

İstanbul'un Akıllı Şehir Endeksi Kapsamında Değerlendirilmesi

Ertan VARLI^{1*}, Güzide GİDER²

¹Mimarlık Bölümü, Mimarlık Fakültesi, Trakya Üniversitesi, Edirne, Türkiye

²Mimarlık Bölümü, Fen Bilimleri Enstitüsü, Trakya Üniversitesi, Edirne, Türkiye

Geliş: 04.06.2025, Kabul: 03.10.2025, Yayınlanma: 05.11.2025

ÖZ

Akıllı şehirler, teknoloji yardımıyla kentlerdeki problemlere akılcı çözümler sunarak, toplum ve teknolojinin uyum içinde bulunduğu sürdürülebilir şehir modeli oluşturmayı amaçlamaktadır. Bu çalışmada, Uluslararası Yönetim Geliştirme Enstitüsü (IMD)'nün yayımladığı 2025 Akıllı Şehir Endeksi verilerine dayanarak, İstanbul ile endekste en yüksek değere sahip ilk on şehir (Zürih, Oslo, Cenevre, Dubai, Abu Dabi, Londra, Kopenhag, Kanberra, Singapur, Lozan) yapı ve teknoloji açısından karşılaştırılmıştır. Karşılaştırma, şehirlerin sağlık ve güvenlik, ulaşılabilirlik, fırsatlar, faaliyetler, yönetim göstergeleri üzerinden yapılmıştır. Değerlendirmeler, İstanbul'un bu alanlarındaki performansının ilk on şehrin gerisinde kaldığını göstermektedir. Sağlık hizmetlerinde dijital uygulamalar güçlü olsa da geri dönüşüm ve hava kalitesi iyileştirilmesi gereken alanlardır. Ulaşımda bilgilendirme sistemleri daha iyi bir performans sergilerken, trafik sıkışıklığı ve toplu taşıma yetersizliği öne çıkmaktadır. Kültürel etkinlikler ve yeşil alanlar yetersiz kalmakta; eğitim ve dijital altyapı eksiklikleri dikkat çekmektedir. Yönetimde, katılımın düşük olması ve şeffaflık ihtiyacı öne çıkan sorunlar olarak belirlenmiştir. Genel olarak, İstanbul'un akıllı şehir olma sürecinde önemli bir potansiyeli bulunmakla birlikte, pek çok alanda gelişime ihtiyaç duyduğu görülmektedir.

Anahtar Kelimeler: Akıllı Şehir; Akıllı Şehir Endeksi; Akıllı Şehir Uygulamaları; İstanbul

Evaluation of Istanbul within the Scope of the Smart City Index

ABSTRACT

Smart cities aim to create a sustainable urban model where society and technology coexist in harmony by providing rational solutions to urban problems through technology. This study, based on the 2025 Smart City Index published by the International Institute for Management Development (IMD), compares Istanbul with the top ten cities in the index (Zurich, Oslo, Geneva, Dubai, Abu Dhabi, London, Copenhagen, Canberra, Singapore, Lausanne) in terms of infrastructure and technology. The comparison covers health and safety, accessibility, opportunities, activities, and governance. Evaluations show that Istanbul lags behind these cities. Digital health applications are strong, but recycling and air quality need improvement. Transportation information systems perform better, while traffic congestion and limited public transport are issues. Cultural events and green spaces are insufficient, and education and digital infrastructure show deficiencies. In governance, low participation and the need for transparency are main concerns. Overall, Istanbul has significant potential to become a smart city, but many areas require development.

Keywords: Smart City; Smart City Index; Smart City Applications; Istanbul

1. GİRİŞ

Kentleşmenin yarattığı nüfus yoğunluğu; barınma sorunu, hava ve çevre kirliliği, trafik yoğunluğu gibi ekonomik, sosyal ve çevresel sorunları beraberinde getirmiştir. Dolayısıyla, daha yaşanabilir şehirler yaratma ihtiyacı duyulmuş ve çeşitli yaklaşımlar geliştirilmiştir. Gelişen teknoloji, bu aşamada sorunların çözümüne önemli ölçüde destekleyici olmuştur. Akıllı şehir kavramı bu süreçte ortaya çıkmış; teknoloji desteği ile akıllıca tasarlanmış sürdürülebilir şehirlerin kurulmasında etkili rol oynamıştır.

Şehirleri akıllı olarak tanımlayabilmek için; meydana gelen problemlerin ortadan kaldırılmasında bilgi ve iletişim teknolojilerinin katkılarından yararlanarak etkili çözümlerin sunulması ve bu çözümlerin vatandaşlar tarafından kabul görüp kullanılması önemlidir. (Uçar ve ark., 2017) Benzer şekilde IMD 2019 Akıllı Şehir Endeksi Akıllı Şehri, teknolojiyi kullanarak kentleşmenin olumlu etkilerini artırıp olumsuz etkilerini hafifleten şehir modeli olarak tanımlanmaktadır.

Akıllı şehirler, teknolojiyi kullanarak şehirlerin yaşanabilir ve sürdürülebilir mekanlara dönüşmesini hedefleyen bir yaklaşımdır. Fakat, teknolojinin bu süreçte bir araç olduğu ve asıl amacın bireyi merkeze alarak onlara yüksek yaşam standartları sunan, akıllıca planlanmış yaşam alanları sağlamak olduğu unutulmamalıdır.

Dünya genelinde giderek önem kazanan akıllı şehir yaklaşımı birçok yenilikçi uygulamanın hayata geçirilmesine öncülük etmiştir. Her yıl yayımlanan akıllı şehir endeksleri, şehirlerin teknolojik gelişimini ve ilerlemelerini ölçmek için önemli bir hale gelmiştir. Türkiye’de de bu alana ilgi artmış ve akıllı şehirler üzerine yapılan çalışmalar zamanla hız kazanmıştır.

Bu çalışma kapsamında, 2054 yılında yayınlanan Akıllı Şehir Endeksi raporunda yer alan ilk on şehir ile İstanbul’un yapı ve teknoloji bakımından karşılaştırılması yapılmıştır. Söz konusu endekste, Ankara ve İstanbul olmak üzere Türkiye’den yalnızca iki şehir yer almaktadır. Ankara üzerine daha önce yapılmış benzer bir değerlendirme bulunması sebebiyle İstanbul ele alınmıştır. Bu yönüyle çalışma, İstanbul’un Akıllı Şehir düzeyine ulaşma sürecindeki potansiyelini ve eksikliklerini değerlendirmek açısından belirleyici olabilir.

1.1. Akıllı şehir kavramı

Akıllı Şehir, kentsel sorunların çözümünde teknolojinin desteğini alarak daha sürdürülebilir şehirler tasarlamak amacıyla ortaya çıkan bir kavramdır (Ulusoy, 2017:1; Canpolat, 2024:30). Henüz ortak bir tanımı bulunmayan Akıllı Şehirler, farklı disiplinlerde çeşitli şekillerde ele alınmış olup; uluslararası çalışmalarda ‘Akıllı Kent’, ‘Dijital Şehir’, ‘Sürdürülebilir Şehir’ gibi ifadeler de kullanılmaktadır. (Ulusoy, 2017:2; Aihemaiti, 2018:2; Barutçu, 2021:3).

Giffenger'a (2007) göre Akıllı şehirler, bağımsız hareket edebilen ve farkındalığı yüksek vatandaşların bir arada olduğu, altı ana bileşen (akıllı yaşam, akıllı yönetim, akıllı insan, akıllı ulaşım, akıllı ekonomi, akıllı çevre) çevresinde şekillenen ve geleceğe dönük yüksek verimlilik ortaya koyan bir oluşumu ifade etmektedir.

PAS 180: Akıllı Şehirler Terminolojisi (2014), akıllı şehirleri; bireylere sürdürülebilir ve yaşam standartları yüksek bir gelecek sunmak hedefiyle; fiziksel, dijital ve beşeri sistemlerin yapıları çevreye entegre edildiği sistemler olarak tanımlamıştır. (İngiliz Standartları Enstitüsü)

Deloitte Raporu'na (2015) göre, akıllı şehirler; insan, sosyal sermaye, teknoloji ve altyapıya yapılan yatırımlar kapsamında, toplumsal katılımı sağlayan, kaynakları verimli kullanan, sürdürülebilir büyümeyi destekleyen ve bu sayede yüksek hayat standartlarına erişen şehirlerdir.

Viyana Teknoloji Üniversitesi'nde Avrupa Akıllı Şehirleri Modelini geliştiren ekip, akıllı şehirleri; şehirleşmenin altı ana başlıkta etkili performans gösteren, farkındalığı yüksek ve bağımsız bireylerin kendi kararlarını verebildiği bir yapı olarak tanımlamaktadır. (Mkrtychev ve ark., 2018)

Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı'na (2020-2023) göre, akıllı şehirler; muhtemel sorunlara teknoloji ile birlikte veri ve uzmanlığa dayalı yenilikçi çözümler sunan, iş birliğiyle hayata geçirilen yaşanabilir ve sürdürülebilir şehirlerdir.

Giffenger (2007), Akıllı şehirleri oluşturan temel bileşenleri akıllı yaşam, akıllı insan, akıllı çevre, akıllı yönetim, akıllı ulaşım ve akıllı ekonomiler olmak üzere altı ana kategoride ele almıştır.

Ulusoy (2017), akıllı bir şehrin akıllı vatandaş, akıllı sağlık, akıllı enerji, akıllı binalar, akıllı teknoloji, akıllı ulaşım, akıllı altyapı, akıllı güvenlik, akıllı yönetim ve eğitim bileşenlerinden en az beşini sağlaması gerektiğini vurgulamaktadır.

Çeşitli kaynaklarda farklı şekillerde ele alınan bileşenler; İstanbul Büyükşehir Belediyesi'nin oluşturduğu "2030 İstanbul Akıllı Şehir Stratejik Planı" kapsamında ise akıllı mobilite, akıllı enerji, akıllı çevre, akıllı insan, akıllı yaşam, akıllı ekonomi, akıllı yönetim, akıllı güvenlik olmak üzere sekiz ana bileşenden oluşmaktadır (2030 İstanbul Akıllı Şehir Stratejik Planı).

1.2. İstanbul'daki akıllı şehir uygulamaları

İstanbul Büyükşehir Belediyesi, 2015 yılı öncesi akıllı şehirlerle ilgili çeşitli çalışmalar yürütmüş olmakla birlikte; bu alandaki faaliyetler 2015 yılı itibarıyla kapsamlı bir boyut kazanmıştır (Çelikyay, 2017). 2016 – 2017 yılları arasında bir proje ile geliştirilen çalışmalar, 2018 yılından itibaren somut uygulamalara dönüşmüştür (Akıllı Şehirler Beyaz Bülteni, 2019). Akıllı şehir çalışmalarını değerlendirmek amacıyla uluslararası kaynaklarda yer alan 880 göstergeden 60 tanesi seçilerek

İstanbul'un Akıllı Şehir Endeksi oluşturulmuş ve bu kapsamda uluslararası alandaki konumu belirlenmiştir. Ayrıca, İstanbul'a özel geliştirilen bir model ile süreç, strateji ve yönetim gibi alanlarda düzenli değerlendirmeler yapılmaktadır. İBB, Akıllı Şehir çalışmaları kapsamında çevre, insan, yaşam, enerji, mobilite, güvenlik, ekonomi ve yönetişimi içeren sekiz temel alan belirleyerek faaliyetlerini bu çerçevede sürdürmüştür (2020-2023 Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı, 2019).

Türkiye'deki akıllı şehir uygulamaları kapsamında İstanbul önde gelen şehirlerden biridir. Literatürde çeşitli kaynaklarda İstanbul'daki çeşitli akıllı şehir uygulamalarından söz edilmiştir. Aşağıda, İstanbul Büyükşehir Belediyesi'nin Akıllı Şehir uygulamaları kapsamında yapmış olduğu çalışmalardan bazı örneklerle yer verilmiştir.

İBB'nin İstanbul'da ulaşım ve toplu taşıma alanındaki faaliyetlerini dijitalleştirmeye yönelik olarak geliştirdiği Trafik Kontrol Merkezi, Araç takip Sistemi, İTaksi, İBB Cep Trafik gibi mobil uygulamalar ile akıllı duraklar, akıllı şehir çözümlerine örnek olarak sıralanabilir (İtü Vakfı Dergisi, 2017).

İstanbul'da, şehirlerin geceleri daha güvenli hale gelmesi ve doğru aydınlatma ile enerji tasarrufu sağlanması amacıyla İstanbul Aydınlatma Master Planı; sanat ve mesleki eğitim veren İSMEK'in (Hayat Boyu Öğrenme Merkezi) mobil uygulaması ile eğitimlere uzaktan erişimin sağlanması; Trafik Yönetim sistemi ile trafik akışının hızlanması; Toplu Ulaşım Bilgi Sistemi (TUBS), Akıllı Durak Teknolojisi gibi uygulama ve araçlar sayesinde vatandaşların ulaşım hakkında bilgilendirmesi gibi örnekler İstanbul'daki Akıllı Şehir faaliyetlerine örnek gösterilebilir (Sürdürülebilir Akıllı Şehirler Çalıştayı, 2017).

İBB, Akıllı Çevre bağlamında suyun verimli kullanımı amacıyla kayıp ve kaçakların minimuma indirilmesi, atık yönetimi kapsamında metallerin ayrıştırılmasına yönelik çeşitli çalışmalar yürütülmektedir. Pet şişelerin geri dönüşüme kazandırılması karşılığında İstanbulkart'a bakiye yüklemesi yapan Akıllı konteynerler da akıllı şehir uygulamalarının kullanımını teşvik eden çözümler arasındadır. (2020-2023 Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı, 2019).

Akıllı Şehirler Beyaz Bülteni'nde (2019) yer alan Hava Kalitesi İzleme Merkezi, Çevre Kontrol Merkezi, Ulaşım Yönetim Merkezi, Adaptif Trafik Yönetim Sistemi (ATAK), Trafik Sinyalizasyon Sistemleri, IoT Taksi Şapkası, İTaksi Yönetim Sistemi, Akıllı Geri Dönüşüm Konteyneri, İBB Cep Trafik Mobil Uygulaması ve Akıllı Park Yönetimi gibi uygulamalar İstanbul'daki akıllı şehir faaliyetlerine örnek olarak gösterilebilir.

Çakıcı ve Kızılboğa Özasan (2021) tarafından yapılan çalışmada, İBB'nin yapmış olduğu 27 tanesi mobil uygulama olmak üzere toplam 47 akıllı şehir uygulamasından bahsedilmiştir. Mobil uygulama örnekleri arasında; İBB İstanbul, İBB CepTrafik, İSPARK, İsbike Akıllı Bisiklet, İTaksi, İstanbul Şehir haritası gibi çeşitli uygulamalar yer almaktadır. Diğer uygulamalara örnek olarak; Hava Kalitesi İzleme Merkezi,

Çevre Kontrol Merkezi, İBB Wi-Fi, Akıllı Konteyner Projesi, Akıllı Şebeke ve Sayaçlar, Açık Veri Portalı gibi çeşitli çalışmalardan bahsedilmiştir.

1.3. Akıllı şehir endeksi

IMD Dünya Rekabet Gücü Merkezi ile SUTD (Singapur Teknoloji ve Tasarım Üniversitesi) arasındaki yakın iş birliğiyle ve yürütülen Akıllı Şehir Endeksi (SCI); giderek daha akıllı hale gelen şehirlerin yaşanabilirliğini, vatandaşlar tarafından kapsamlı şekilde değerlendirilebilmesine olanak tanıyan ve dünya genelinde çok sayıda uzman ile şehir planlayıcısının katkılarından yararlanan bir girişimdir. IMD Akıllı Şehir Endeksi'ni özgün kılan, endekste yer alan şehirlerde yaşayan ve çalışan vatandaşların değerlendirmelerini temel almasıdır, bu yönüyle teknolojiyi merkeze alan mevcut endekslerden ayrılır.

2019 yılında yayınlanan ilk baskı; iki yıllık bir çalışmanın ürünü olup her şehirden 120 kişi ile yapılan değerlendirmeler ışığında 102 şehri küresel ölçekte sıralar. Şehirler, 'yapılar' ve 'teknolojiler' kapsamında; sağlık ve güvenlik, faaliyetler, ulaşılabilirlik, yönetim ve fırsatlar olmak üzere beş temel alanda incelenir. Birleşmiş Milletler İnsani Gelişmişlik Endeksi (HDI) puanına göre ayrılarak; en yüksek grup AAA'dan, en düşük grup D'ye kadar olan bir sıralama ölçeği ile değerlendirilir. Ardından sıralamalar, her iki kategoriye ait derecelendirme ve genel sıralama olarak iki formatta sunulur. (IMD Akıllı Şehir Endeksi, 2019)

Beş yıllık süreçte yeni veriler doğrultusunda gelişmeye devam eden 2024 yılı Akıllı Şehir Endeksi'nde 142 şehrin verilerine yer verilmiştir. Yüksek performans gösteren şehirler önceki yılın listesiyle büyük ölçüde örtüşerek 'SCI şampiyonları' tanımını ortaya koymuştur. Hiç düşüş yaşamayan şehirlere bakıldığında 12 şehir ön plana çıkmıştır. Bu 12 şehir; ilk 20'de yer alan 'süper şampiyonlar' ve 20- 35. sıra aralığında yer alan 'hızlı yükselen adaylar' olmak üzere iki alt gruba ayrılmıştır. Zürih, Oslo, Singapur, Abu Dabi, Pekin ve Seul süper şampiyonlar grubunda yer alırken; Sidney, Hong Kong, Şanghay, Tallinn, Riyad ve Melbourne hızlı yükselen adaylar grubunda yer almıştır. Bu şehirlerin ortak özellikleri arasında, sosyal ve ekonomik açıdan öngörülebilir bölgelerde yer almaları; yaşam kalitesini artırmaya yönelik projelerin ve sürdürülebilir stratejilerin geliştirilmesi gibi unsurlar ön plana çıkmaktadır. (IMD Akıllı Şehir Endeksi, 2024)

2025 Akıllı Şehir Endeksi'ne, AIUla, Astana (eski adıyla Nur Sultan), Caracas, Kuveyt, Manama ve San Juan olmak üzere altı yeni şehir eklenmiştir. Ayrıca, yeterli veri bulunmadığından Tianjin ve Zhuhailiste dışında bırakılmıştır. Böylelikle, anket kapsamındaki şehir sayısı 2025 yılı itibariyle 142'den146'ya yükselmiştir. Sıralama genel olarak üst sıralarda minimal değişiklikler dışında istikrar göstermektedir. İlk yirmi şehirde geçen yıla göre değişiklikler, Taipei'nin 16. sıradan 23. sıraya gerilemesi ve Ljubljana'nın 32. sıradan 16. sıraya yükselmesi olarak gözlemlenmektedir. Zürih, bir kez daha listenin zirvesinde yer

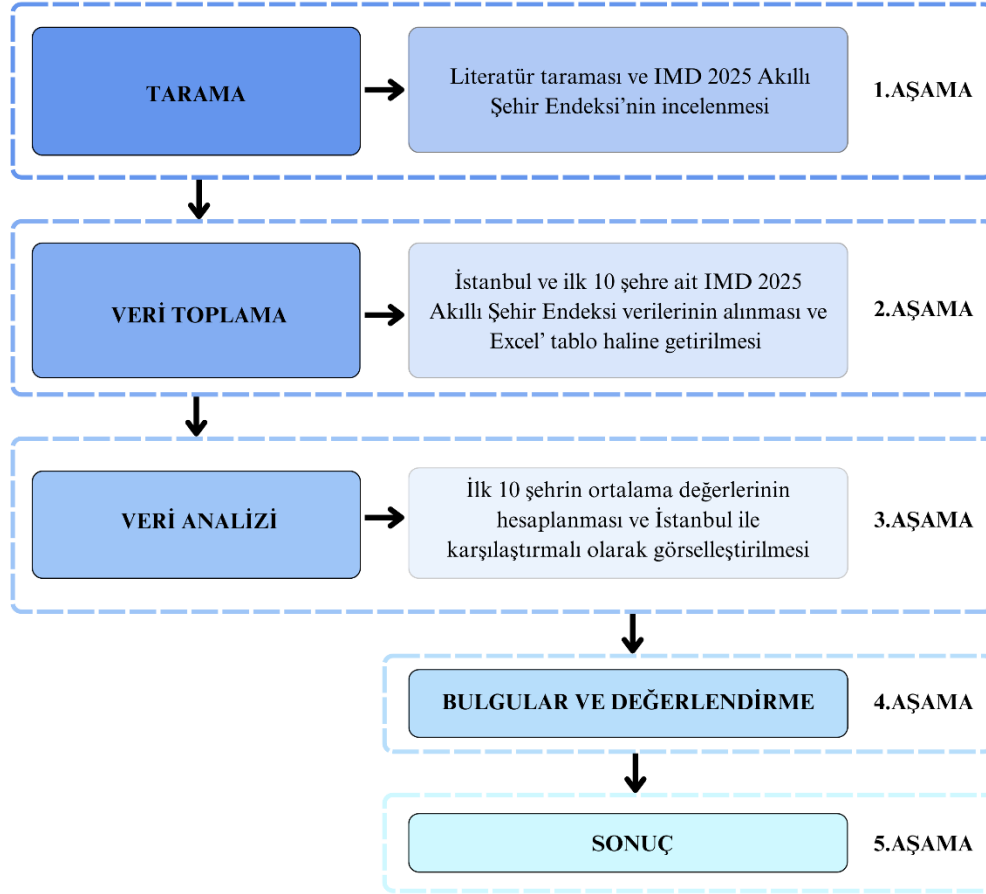
almıştır. Dubai ve Abu Dabi ise önceki yıl yer aldıkları 12. ve 10. sıralardan 4.ve 5. sıraya yükselmiştir. Tablo 1’de İlk 10 şehir ve İstanbul’un bulunduğu gruplar ve şehir sıralamaları gösterilmiştir. (IMD Akıllı Şehir Endeksi, 2025)

Tablo1: 2025 Akıllı Şehir Endeksi İlk 10 Şehir ve İstanbul’un Sıralaması (IMD 2025 Akıllı Şehir Endeksi)

Şehir	Şehir Sıralaması 2025 (146 Şehir)	Grup	Toplam Derece	Yapılar 2025	Teknolojiler 2025	Şehir Sıralaması 2024 (142 Şehir)	Değişim
Zürih	1	1	AAA	AAA	AAA	1	-
Oslo	2	1	AAA	AAA	AAA	2	-
Cenevre	3	1	AAA	AAA	AAA	4	+1
Dubai	4	2	A	A	A	12	+8
Abu Dabi	5	2	A	A	A	10	+5
Londra	6	1	AA	AAA	AAA	8	+2
Kopenhag	7	1	AAA	AAA	AA	6	-1
Kanberra	8	1	AAA	AAA	A	3	-5
Singapur	9	1	AAA	AAA	AAA	5	-4
Lozan	10	1	AAA	AA	AA	7	-3
İstanbul	111	3	C	CC	CCC	110	-1

2. YÖNTEM

Bu çalışmada, IMD Dünya Rekabet Gücü Merkezi tarafından yayımlanan 2025 yılı Akıllı Şehir Endeksi (SCI) temel alınarak İstanbul ile endekste ilk 10 sırada yer alan şehirler karşılaştırılmıştır. Araştırma, nitel bir yöntem olandoküman incelemesi tekniğiyle gerçekleştirilmiştir. Dokümanincelemesi, yazılı belgelerin bir sistem çerçevesinde ayrıntılı bir şekilde incelenmesine dayanan bir yöntemdir (Wach, 2013). Akıllı Şehir Endeksi 2025 raporu detaylı şekilde incelenerek endekste yer alan ilk on şehir ile İstanbul’un performans verileri seçilmiş ve Excel programı kullanılarak tablolar hâline getirilmiştir. İlk on şehrin verileri birleştirilerek ortalama değerleri hesaplanmış, İstanbul verileri ise ayrı bir tabloda tutulmuştur. Elde edilen veriler, karşılaştırmalı olarak görselleştirilmiş ve öne çıkan göstergeler yorumlanmıştır. Bu çalışmada temel yöntem, mevcut rapor verilerinin derlenmesi ve karşılaştırmalı olarak okuyucuya sunulmasıdır.

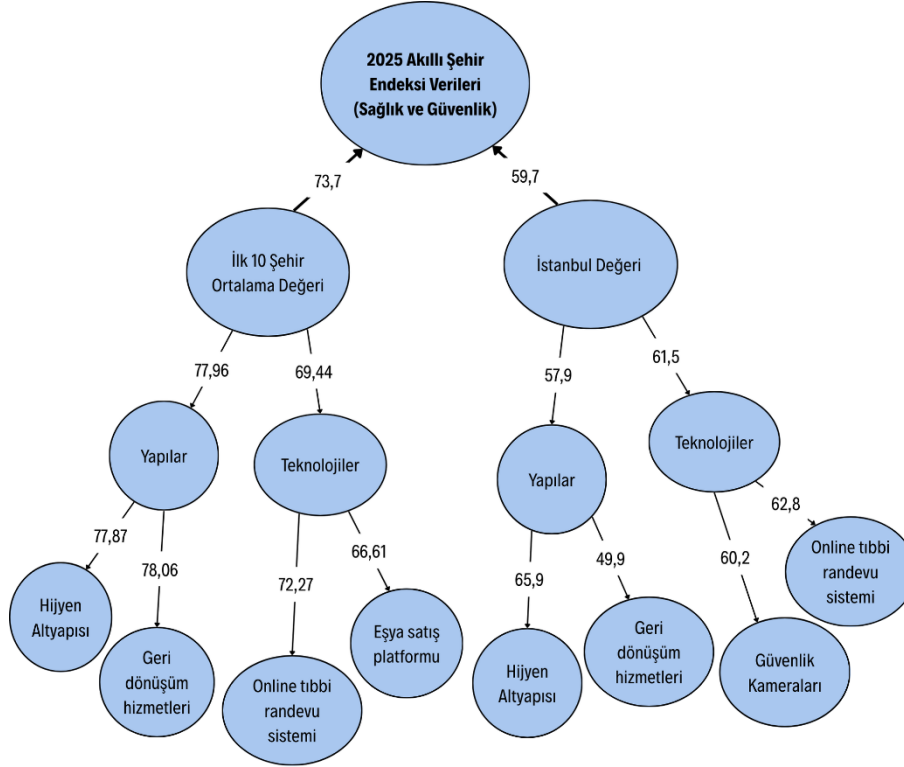


Şekil 1: Yöntem akış diyagramı.

3. BULGULAR

IMD' nin yayınladığı 2025 Akıllı Şehir Endeksi'ndeki akıllı şehir göstergeleri (sağlık ve güvenlik, faaliyetler, ulaşılabilirlik, yönetim ve fırsatlar) kapsamında, ilk on sırada yer alan (Zürih, Oslo, Cenevre, Dubai, Abu Dabi, Londra, Kopenhag, Kanberra, Singapur, Lozan) şehirler ile İstanbul yapılar ve teknolojiler bakımından karşılaştırılmalı olarak ele alınmıştır. Ulaşılan veriler, ilgili değerlendirme kriterleri çerçevesinde okuyucuya sunulmuştur.

3.1. Sağlık ve Güvenlik



Şekil 2: Sağlık ve güvenlik göstergesi değerleri.

2025 Akıllı Şehir Endeksi verileri sağlık ve güvenlik kapsamında değerlendirildiğinde, ilk on şehirdeki ortalama değerlerin İstanbul'un değerlerinden daha yüksek olduğu görülmektedir. Yapılar bakımından incelendiğinde hem ilk on şehirde hem de İstanbul'da "hijyen altyapısı" ve "geri dönüşüm hizmetleri" en yüksek değere sahip göstergeler olarak öne çıkmaktadır. Ancak özellikle geri dönüşüm hizmetleri konusunda İstanbul'un değerinin ilk on şehre kıyasla oldukça düşük kaldığı dikkat çekmektedir. Bu durum, geri dönüşüme yönelik akıllı uygulamaların yetersiz olduğunu düşündürmektedir.

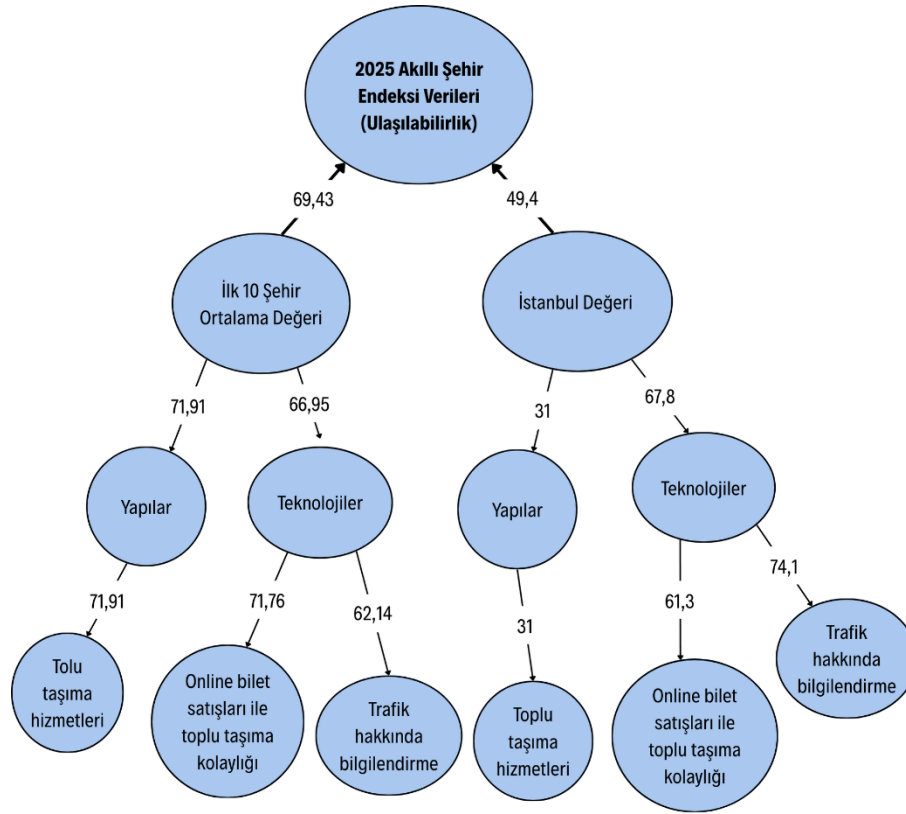
İstanbul'u kendi içinde ele aldığımızda "kiralık konut bulma" ve "hava kirliliği" en düşük değere sahip göstergeler arasında yer almaktadır. Ayrıca kamu güvenliği de İstanbul için önemli bir sorun alanı olarak öne çıkmaktadır. Barınma ve güvenlik gibi temel ihtiyaçların düşük değerlere sahip olması, bu alanlara yönelik çalışmaların artırılması gerektiğini göstermektedir.

Teknolojik göstergeler açısından incelendiğinde "online tıbbi randevu sistemi" hem İstanbul hem de ilk on şehir için en yüksek değere sahip gösterge olarak öne çıkmaktadır. İlk on şehrin ortalamasında "eşya satış platformu" en yüksek değere sahip ikinci gösterge olarak dikkat çekmektedir. İstanbul özelinde ise sağlık hizmetlerine erişimin nispeten yüksek olduğu, e-Nabız ve MHRS gibi uygulamaların bu noktada önemli

katkıları sunduğu söylenebilir. Bununla birlikte, sorunların çevrimiçi bildirilmesi ve hava kirliliğinin etkili şekilde takip edilmesi en düşük değerlere sahip göstergeler arasında yer almaktadır. Bu göstergelerin iyileştirilmesi için yeni web sitesi ve uygulamaların geliştirilmesi ya da mevcut uygulamaların daha işlevsel hale getirilmesi önemlidir.

Önceki bölümlerde 'İstanbul'daki akıllı şehir uygulamaları' başlığı altında da belirtildiği üzere, İstanbul'da hava kalitesi izleme merkezi kurulmuş ve çeşitli mobil uygulamalar geliştirilmiştir. Ancak son endeks verileri, bu konularda hala sorunlar olduğunu göstermektedir. Bu durum, yalnızca bu uygulamaların yapılmasının yeterli olmadığını, aynı zamanda insanların bu uygulamalardan haberdar olmasının da önemli olduğunu ortaya koymaktadır.

3.2. Ulaşılabilirlik



Şekil 3: Ulaşılabilirlik göstergesi değerleri.

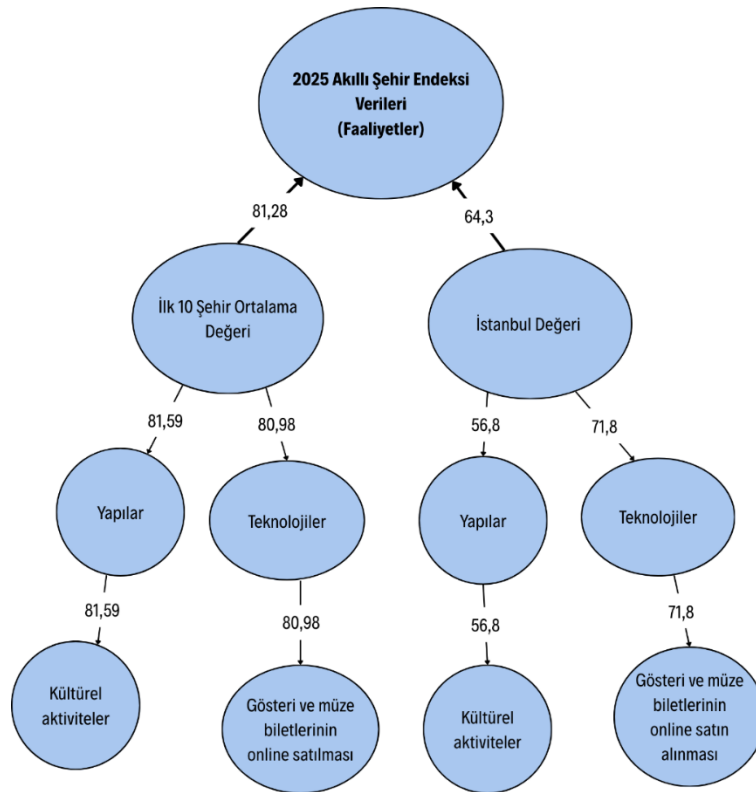
2025 Akıllı Şehir Endeksi verileri ulaşılabilirlik açısından incelendiğinde, İstanbul'un değerinin ilk on şehrin ortalamasının altında kaldığı görülmektedir. Yapılar bakımından değerlendirildiğinde "toplu taşıma hizmetleri" hem İstanbul hem de diğer şehirlerde en yüksek değere sahip gösterge olarak öne çıkmaktadır. Ancak İstanbul'da bu değer diğer şehirlere kıyasla belirgin şekilde düşük olması dikkat çekicidir. İstanbul'un kendi içinde ele alındığında ise "trafik sıkışıklığı" önemli bir sorun olarak öne çıkmaktadır.

Önceki bölümlerde İstanbul'da ulaşım konusunda birçok akıllı şehir uygulamasının hayata geçirildiğinden bahsedilmiştir. Ancak son endeks verileri, bu uygulamalara rağmen sorunların devam ettiğini ve yapılan çalışmaların yetersiz kaldığını göstermektedir. Nüfus yoğunluğundan getirilerinden olan bu sorunlar yalnızca İstanbul'a özgü olmamakla birlikte, diğer şehirlerin yüksek değerlere sahip olması İstanbul için çözüm üretme bakımından umut vericidir.

Teknolojik göstergeler açısından bakıldığında, “online bilet satışları ile toplu taşıma kolaylığı” ve “trafik hakkında bilgilendirme” göstergeleri hem İstanbul hem de diğer şehirlerde en yüksek değerlere sahip göstergeler olarak öne çıkmaktadır. Arada ciddi bir fark olmamakla birlikte, İstanbul'un teknoloji bakımından daha yüksek değere sahip olduğu görülmektedir. Özellikle “trafik hakkında bilgilendirme” göstergesi bu noktada öne çıkmaktadır. İstanbul'da trafik kontrol merkezi, İBB Cep Trafik, Toplu Taşıma Bilgi Sistemi ve akıllı durak gibi uygulamaların bu değerin yükselmesinde etkili olduğu düşünülmektedir.

Bununla birlikte, diğer göstergelere bakıldığında bu uygulamaların trafik sıkışıklığını azaltmadığı ve seyahat sürelerini kısaltmadığı yönünde sorunların devam ettiği görülmektedir. İstanbul, geliştirdiği uygulamalar sayesinde kullanıcıyı bilgilendirme konusunda başarılı olsa da ulaşım sorunlarının azaltılması ve ortadan kaldırılması bakımından hala gelişim potansiyeline sahiptir.

3.3. Faaliyetler

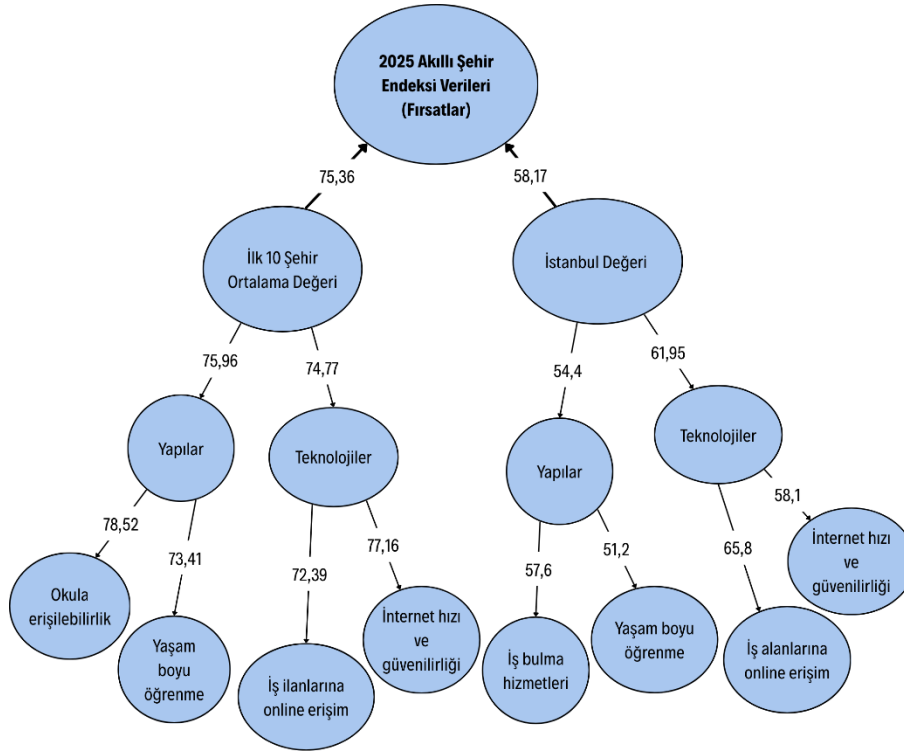


Şekil 4: Faaliyetler göstergesi değerleri.

2025 Akıllı Şehir Endeksi faaliyetler bakımından değerlendirildiğinde, ilk on şehrin ortalama değerlerinin İstanbul'un değerlerinden belirgin şekilde yüksek olduğu görülmektedir. Yapılar açısından incelendiğinde tüm şehirler için “kültürel aktiviteler” en yüksek değere sahip gösterge olarak öne çıkmaktadır. Ancak İstanbul özelinde bu göstergenin yeterince tatmin edici düzeyde olmadığı dikkat çekmektedir. Bununla birlikte “yeşil alanlar” göstergesi İstanbul için en düşük değere sahip unsur olarak öne çıkmaktadır. Artan nüfus ve yoğun yapılaşmanın etkisiyle yeşil alanların yetersizliği, bu göstergenin düşük olmasında önemli bir faktör olarak değerlendirilmektedir.

Teknolojiler bakımından incelendiğinde, “gösteri ve müze biletlerinin online satın alınması” tüm şehirlerde ortak olarak öne çıkan tek gösterge durumundadır. İstanbul'un bu alandaki değeri ise diğer şehirlere kıyasla daha düşük kalmaktadır. Bu bakımdan, geliştirilecek mobil uygulamalar ve web siteleri aracılığıyla gösteri ve müzelere erişimin kolaylaştırılması katılımı daha kolay hale getirerek İstanbul'un söz konusu göstergedeki değerinin yükselmesine katkı sağlayacağı düşünülmektedir.

3.4. Fırsatlar



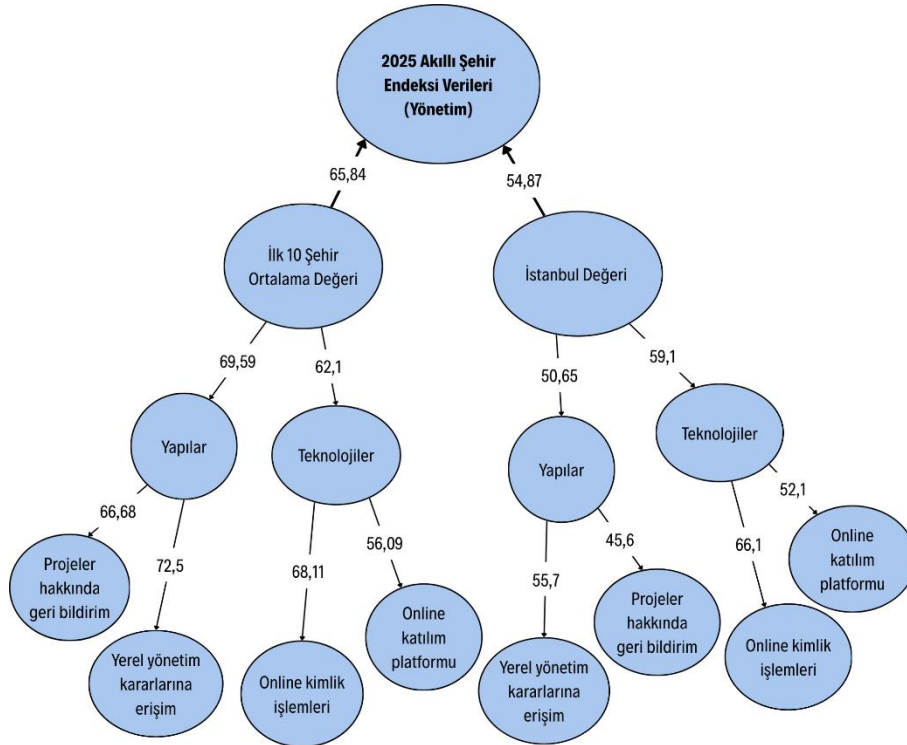
Şekil 5: Fırsatlar göstergesi değerleri.

2025 Akıllı Şehir Endeksi fırsatlar bakımından değerlendirildiğinde, İstanbul'un değerinin ilk on şehrin ortalamasının altında kaldığı görülmektedir. Yapılar açısından incelendiğinde, ilk on şehirde “yaşam boyu öğrenme” ve “okula erişilebilirlik” göstergeleri öne çıkarken, İstanbul'da farklı olarak “iş bulma hizmetleri” göstergesi ön plana çıkmaktadır. İstanbul özelinde değerlendirildiğinde, “okula erişilebilirlik”

en düşük gösterge olarak dikkat çekmektedir. Ayrıca diğer şehirlerle karşılaştırıldığında bu değer oldukça düşük seviyede kalması, İstanbul'un bu alanda önemli derecede gelişime açık olduğunu göstermektedir.

Teknolojik göstergeler açısından bakıldığında 'iş ilanlarına online erişim' ile 'internet hızı ve güvenilirliği', en yüksek değere sahip göstergeler arasında yer almaktadır. Ancak, İstanbul özellikle internet hızı ve güvenilirliği konusunda diğer şehirlere göre düşük bir değere sahiptir. Ayrıca, şehirde okullarda bilgisayar becerilerinin yeterince etkili bir şekilde kazandırılmaması önemli bir sorun olarak öne çıkmaktadır. Bu durum, İstanbul'un hem eğitim hem de teknoloji alanında gelişime açık olduğunu göstermektedir.

3.5. Yönetim



Şekil 6: Yönetim göstergesi değerleri.

2025 Akıllı Şehir Endeksi yönetim boyutu incelendiğinde, şehirler arasında öne çıkan ortak göstergelerin başında “yerel yönetim kararlarına erişim” ve “projeler hakkında geri bildirim” gelmektedir. Bu göstergeler tüm şehirler için en yüksek değerlere sahipken, İstanbul'un söz konusu alanlardaki performansı diğer şehirlere kıyasla daha düşük seviyede kalmaktadır. İstanbul'un özelinde değerlendirildiğinde ise yolsuzluk endişe verici bir sorun olarak öne çıkmakta ve yönetim başlığında en düşük değere sahip göstergesi oluşturmaktadır. Bunu, şehir sakinlerinin yönetimin karar alma süreçlerine yeterince katılım göstermemesi takip etmektedir. Bu durum, İstanbul'un yönetim konusunda oldukça

düşük bir performansa sahip olduğunu ve bu alanda etkili uygulamalara duyulan ihtiyacı açıkça ortaya koymaktadır.

Teknolojik göstergeler açısından bakıldığında ise “çevrimiçi katılım platformları” ve “çevrimiçi kimlik işlemleri” şehirler arasında öne çıkan ortak göstergeler olarak dikkat çekmektedir. İstanbul’da ise teknolojik açıdan en düşük değer “çevrimiçi oylama” göstergesinde görülmektedir. Bu durum, İstanbul’un teknolojiye dayalı yönetimde ilerleme kaydetmesi gerektiğini ifade etmektedir.

4. SONUÇ

İnsana sunduğu imkanlar sayesinde günümüzde şehir yaşamına artan ilgi; ekonomik, sosyal ve çevresel pek çok problemi beraberinde getirmiştir. Şehirlerde düzensizliğe neden olan ve yaşam koşullarını güçleştiren bu problemler; şehirler için daha sürdürülebilir çözümlerin gerekliliğini ortaya koymuştur. Bir yandan gelişmeye devam eden teknoloji, bu süreci destekleyerek çözüm arayışına büyük ölçüde katkıda bulunmuştur. Bu gelişmeler, teknolojiyi merkeze alan akıllıca tasarlanmış şehir modelinin temelini oluşturan Akıllı Şehir kavramını gündeme getirmiştir.

Dünyada ve Türkiye’de akıllı şehirler konusu daha fazla önem kazanmış; buna paralel olarak bu alanda yapılan çalışmalar artmıştır. Yıllık olarak kamuoyuna sunulan akıllı şehir endeksleri, şehirlerin dönüşüm süreçlerini izleme açısından etkili bir değerlendirme aracı olmuştur. Bu doğrultuda, 2025 IMD Akıllı Şehir Endeksi verileri temel alınmış; İstanbul ile diğer şehirlerin karşılaştırılması yapılarak akıllı şehir olma sürecindeki durumunun değerlendirilmesi amaçlanmıştır. 2025 Akıllı Şehir Endeksi’nde 146 şehir içinde yer alan ilk on şehir (Zürih, Oslo, Cenevre, Dubai, Abu Dabi, Londra, Kopenhag, Kanberra, Singapur, Lozan) ile İstanbul, ‘yapılar’ ve ‘teknolojiler’ kapsamında ‘sağlık ve güvenlik, yönetim, faaliyetler, fırsatlar ve ulaşılabilirlik’ göstergeleri bakımından kıyaslanmıştır.

Ortaya çıkan sonuçlar sağlık ve güvenlik bakımından ele alındığında İstanbul, sağlık ve güvenlik alanlarında diğer büyük şehirlerin gerisindedir. Özellikle geri dönüşüm, kiralık konut ve hava kalitesi göstergeleri düşük seviyededir. Teknolojik uygulamalar sağlık hizmetlerine erişimi kolaylaştırırsa da çevrimiçi sorun bildirme ve hava kalitesi takibi hâlâ yeterli değildir. Bu durum, mevcut akıllı şehir uygulamalarının geliştirilmesi, yeni çözümler üretilmesi ve vatandaşların bu uygulamalardan haberdar edilmesinin önemini ortaya koymaktadır.

Ulaşılabilirlik açısından sonuçlar değerlendirildiğinde, İstanbul, ulaşılabilirlik açısından diğer büyük şehirlerin gerisinde olup, özellikle trafik sıkışıklığı önemli bir sorun olarak öne çıkmaktadır. Mevcut teknolojik uygulamalar, bilgilendirme konusunda etkili olsa da trafik ve seyahat sürelerini iyileştirmekte yetersiz kalmaktadır. Bu durum, ulaşım çözümlerinin geliştirilmesi ve daha etkin stratejilerinin uygulanması gerektiğini göstermektedir.

Sonuçlar faaliyetler kapsamında değerlendirildiğinde, İstanbul, kültürel etkinlikler ve yeşil alanlar bakımından diğer büyük şehirlerin gerisindedir. Mevcut teknolojik uygulamalar, katılımı kolaylaştırırsa da yetersiz kalmaktadır. Bu nedenle, İstanbul'da kültürel etkinliklere ve yeşil alanlara erişimi artıracak yeni uygulamalar geliştirilmelidir.

Fırsatlar bakımından sonuçlar değerlendirildiğinde, İstanbul, diğer büyük şehirlerin gerisindedir. Okula erişim ve internet hızı sorunları öne çıkmakta, bilgisayar becerilerinin yeterince kazandırılmaması önemli bir eksiklik olarak görülmektedir. Bu nedenle, eğitim olanaklarını artıracak ve dijital altyapıyı güçlendirecek çözümler geliştirilmelidir.

Son olarak sonuçlar yönetim kapsamında değerlendirildiğinde, İstanbul, yönetim ve katılım açısından diğer şehirlerin gerisindedir. Yolsuzluk endişesi ve düşük vatandaş katılımı öne çıkmaktadır. Teknolojik olarak da çevrimiçi oylama yetersizdir. Bu nedenle, yönetimde şeffaflığı artıracak ve katılımı güçlendirecek uygulamalar geliştirilmelidir.

Bu çalışma, IMD 2025 Akıllı Şehir Endeksi'ndeki ilk on şehir ile İstanbul'u karşılaştırarak, İstanbul'un akıllı şehir olma sürecindeki durumuna dair bir değerlendirme sunmaktadır. Elde edilen sonuçlar doğrultusunda, İstanbul'un genel tabloda geride kaldığı görülmektedir. Bu durum, İstanbul'un akıllı şehir olma sürecinde daha çok adım atılması gerekliliğini ortaya koymaktadır.

Çalışma, gelecekte Akıllı Şehirler ile ilgili yapılacak uygulamalara rehberlik etmesi, İstanbul'un durumunun değerlendirilmesi ve literatüre katkı sağlaması amacıyla hazırlanmıştır. Aynı zamanda, bu alanda yapılan az sayıda karşılaştırmalı araştırmalar arasında yer almaktadır. IMD 2025 Akıllı Şehir Endeksi'nde Türkiye'den Ankara ve İstanbul'un yer alması, çalışma açısından bir sınırlılık oluşturmaktadır. Bununla birlikte, gelecekte yapılacak çalışmalarda, farklı akıllı şehir endekslerinin (örneğin IsiLab Smart Cities Index, Cities in Motion Index vb.) kullanılması, değerlendirmelere çeşitlilik bakımından katkı sağlayabilir.

ÇIKAR ÇATIŞMASI BEYANI

Yazarlar arasında çıkar çatışması bulunmamaktadır.

YAZARLARIN KATKILARI

E.V.: Yöntem, yazılım, doğrulama, araştırma, kaynaklar, yazı yazma - gözden geçirme ve düzenleme.

G.G.: Kavramsallaştırma, yöntem, yazılım, doğrulama, formel analiz, araştırma, kaynaklar, yazı yazma - orijinal taslak hazırlama.

KAYNAKLAR

- Aihemaiti, A. (2018). *Türkiye'deki akıllı şehirlerin sıralama modeli*, Yüksek Lisans Tezi, İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü.
- Barutçu, B. (2021). *Akıllı şehirler üzerine sistemik bir literatür taraması ve akıllı şehirlerde endüstri mühendisliği uygulama alanları*, Yüksek Lisans Tezi, Başkent Üniversitesi Fen Bilimleri Enstitüsü.
- Canpolat, B. Y. (2024). *Geleceğin mimarisinde kentsel dönüşüm: Akıllı şehirler*, Yüksek Lisans Tezi, İstanbul Aydın Üniversitesi Lisansüstü Eğitim Enstitüsü.
- Çakıcı, K., & Özasan Kızılboğa, R. (2021). Birleşmiş Milletler 2030 sürdürülebilir kalkınma amaçlarının akıllı kent uygulamalarındaki karşılığı: İstanbul Büyükşehir Belediyesi örneği. *Gümüşhane Üniversitesi Sosyal Bilimler Enstitüsü Elektronik Dergisi*, 12(2), 209–233.
- Çelikyay, H. H. (2017). İstanbul perspektifinden akıllı şehirlere bakış: Şehirleri akıllı kılan sadece teknoloji mi? *Van Yüzüncü Yıl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 14(4), 505–512.
- Çevre ve Şehircilik Bakanlığı. (2019), *Akıllı Şehirler Beyaz Bülteni*, 7 Nisan 2024 tarihinde Akıllı Şehirler Portalı sitesi: <https://webdosya.csb.gov.tr/db/cbs/akillisehirler/> adresinden alındı.
- Deloitte. (2015), *Smart Cities Report*, 21 Nisan 2025 tarihinde Deloitte sitesi: <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Public-Sector/gx-public-sector-smart-cities-report.pdf> adresinden alındı.
- Designing Buildings Wiki. (2014), *PAS 180: Akıllı Şehirler Terminolojisi*, 24 Aralık 2024 tarihinde Designing Buildings Wiki sitesi: https://www.designingbuildings.co.uk/wiki/PAS_180:2014_Smart_cities_%E2%80%93_Vocabulary adresinden alındı.
- Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler Milanović, N., & Meijers, E. (2007). Smart cities ranking: An effective instrument for the positioning of the cities. Final report. *Centre of Regional Science (SRF)*, Vienna University of Technology.
- IMD Smart City Observatory. (2019), *Akıllı Şehirler Endeksi 2019 Raporu*, 14 Nisan 2025 tarihinde IMD Smart City Observatory sitesi: https://imd.widen.net/view/pdf/vymhxyzlty/smart_city_index_digital.pdf adresinden alındı.
- IMD Smart City Observatory. (2024), *Akıllı Şehirler Endeksi 2024 Raporu*, 26 Şubat 2025 tarihinde IMD Smart City Observatory sitesi: https://imd.widen.net/s/q7flvgtvbs/20240412-smartcityindex-2024-full-report_4 adresinden alındı.
- IMD Smart City Observatory. (2025), *Akıllı Şehirler Endeksi 2025 Raporu*, 8 Eylül 2025 tarihinde IMD Smart City Observatory sitesi: https://imd.widen.net/s/psdrsvpbk7/imd_smart_city_2025_report adresinden alındı.
- Işık, K. C., & Karagöz, Y. (2019). Bankacılık reklamlarının nitel video analizi ile incelenmesi: NVivo 10 örneği. *Elektronik Sosyal Bilimler Dergisi*, 18(72), 1558–1579. doi.org/10.17755/esosder.515569
- İstanbul Büyükşehir Belediyesi. (2017), *Sürdürülebilir Akıllı Şehirler Çalıştayı: Bilgi Kitapçığı*, 21 Nisan 2025 tarihinde İstanbul Büyükşehir Belediyesi sitesi: https://rec.org.tr/wp-content/uploads/2017/08/surdurulebilirakillisehirlercalistayi_kitapcik_vf.pdf adresinden alındı.
- İTÜ Vakfı. (2017), *İTÜ Vakfı Dergisi: Akıllı Şehirler (Sayı: 77)*, 8 Nisan 2025 tarihinde İTÜ Vakfı sitesi: https://www.ituvakif.org.tr/_files/ugd/8db14c_84a319a4ac18478c905e1ff0e2a4f28e.pdf adresinden alındı.
- İstanbul Büyükşehir Belediyesi. (2024), *2030 İstanbul Akıllı Şehir Stratejik Planı*, 13 Nisan 2025 tarihinde İstanbul Büyükşehir Belediyesi sitesi: https://www.akillisehir.istanbul/wp-content/uploads/2024/03/2030_Istanbul_Akilli_Sehir_Stratejik_Plani.pdf adresinden alındı.
- Mkrtychev, O., Starchyk, Y., Yusupova, S., Zaytceva, O., & Starchyk, Y. (2018). Analysis of various definitions for smart city concept. *Materials Science and Engineering*, doi:10.1088/1757-899X/365/2/022065
- Silik, C. E., & Özdemir Akgül, S. (2021). Akıllı Şehir Endeksi kapsamında Ankara'ya ilişkin karşılaştırmalı bir analiz. *Türk Turizm Araştırmaları Dergisi*, 5(1), 542–557. doi:10.26677/tr1010.2021.679

- Uçar, A., Şemşit, S., & Negiz, N. (2017). Avrupa Birliği akıllı kent uygulamaları ve Türkiye'deki yansımaları. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, Özel Sayı, 1785–1798.
- Ulusoy, M. (2017). *Akıllı şehirler*, Yüksek Lisans Tezi, İstanbul Bilgi Üniversitesi Sosyal Bilimler Enstitüsü.
- Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. *Journal of Nursing Education and Practice*, 6(5), 100–110.
- Wach, E., & Ward, R. (2013). Learning about qualitative document analysis. *IDS Practice Paper In Brief*, 13.

Keten Tohumu Müsilajı İçeren Sıcaklık Duyarlı Hiyaluronik Asit Hidrojellerinin Sentezi

Halide Ceren HAZAVİTLİGİL¹ , Burcu OKUTUCU^{1*} 

¹Biyokimya Bölümü, Fen Fakültesi, Ege Üniversitesi, İzmir, Türkiye

Geliş: 30.04.2025, Kabul: 30.09.2025, Yayınlanma: 13.11.2025

ÖZ

Hidrojeller, bir veya daha fazla monomerin reaksiyonu ile sentezlenen çapraz bağlı polimerik yapılardır. Hidrojeller monomer kaynaklarına göre doğal ve sentetik polimerik malzemelerdir. Doğal biyopolimerlerden hazırlanan hidrojeller; biyolojik olarak uyumlu olmaları, biyolojik olarak parçalanabilir olmaları, düşük toksisiteye sahip olmaları, karakteristik dokuya benzer şekilde yüksek derecede esnekliğe sahip olmaları ve modifikasyonlarının kolay olmasından dolayı medikal alanda kullanılmaktadırlar. Doğal biyopolimerler hayvansal, bitkisel ya da atıklardan elde edilen karbohidrat, protein ya da müsilajlar olabilir. Hidrojel bileşeni olarak bitki bazlı biyopolimer (müsilaj) toksik olmaması, düşük maliyeti, yumuşatıcı ve tahriş edici olmaması nedeniyle tercih edilir. Hiyaluronik asitin; kozmetik ürünlerin üretiminde, yara tedavi edici ilaçlarda, doku rejenerasyonu ve tedavisinde ve medikal implant olarak kullanım alanı mevcuttur. Bu çalışmada doğal biyopolimer bileşeni olarak seçilen keten tohumu müsilajı ve sıcaklık duyarlı monomer N-izopropilakrilamid ile hiyaluronik asit hidrojelini sentezlendi. Doğal bileşeni keten tohumu müsilajı olan sıcaklık duyarlı hiyaluronik asit hidrojellerinin elastikiyetini artırarak medikal kullanım olanakları artırılmıştır.

Anahtar Kelimeler: Sıcaklık duyarlı hidrojel; Keten Tohumu; Müsilaj; Hiyaluronik asit

Synthesis of Temperature-Sensitive Hyaluronic Acid Hydrogels Containing Flaxseed Mucilage

ABSTRACT

Hydrogels are cross-linked polymeric structures synthesized through the reaction of one or more monomers. Depending on their monomer sources, hydrogels can be natural or synthetic polymeric materials. Hydrogels prepared from natural biopolymers are used in the medical field due to their biocompatibility, biodegradability, low toxicity, high flexibility resembling natural tissue characteristics, and ease of modification. Natural biopolymers can be carbohydrates, proteins, or mucilage derived from animal, plant, or waste sources. Plant-based biopolymers (mucilage) are preferred as hydrogel components because they are non-toxic, low-cost, non-irritating, and act as natural softening agents. Hyaluronic acid is used in cosmetic products, wound healing drugs, tissue regeneration and treatment, various medical implant applications. In this study, hyaluronic acid hydrogels were synthesized using flaxseed mucilage, (natural biopolymer component), and the temperature-sensitive monomer N-isopropylacrylamide. By incorporating flaxseed mucilage as a natural component, the elasticity of temperature-sensitive hyaluronic acid hydrogels was enhanced, thereby increasing their potential for medical applications.

Keywords: Temperature- sensitive hydrogel; Flaxseed; Mucilage; Hyaluronic acid

1. GİRİŞ

Bitki bazlı biyopolimerler; seyrelticiler, bağlayıcılar, tabletlerdeki parçalayıcılar, oral sıvılardaki koyulaştırıcılar, süspansiyonlardaki koruyucu kolloidler, fitillerde jellerde ve bazlarda jelleştirici maddeler gibi çeşitli farmasötik uygulamalarda, kozmetik, tekstil, boya ve kağıt yapımında son yıllarda yaygın olarak kullanılmaktadır (Haseeb, et al, 2024; Hosseini & Nabid, 2020). Bitki bazlı biyopolimerler, matris kontrollü sistem, film kaplama ajanları, mikro kürecikler, nanopartiküller, süspansiyonlar, implantlar gibi viskoz sıvı formülasyonlar gibi farklı farmasötik dozaj formlarında uygulanmaların etkinliği kanıtlanmıştır (Ahmad et al, 2019; George & Suchithra, 2019; Zhao et al, 2012; Prajapati et al, 2013). Genelde suda çözünür polisakaritler olan ve bitkisel bazlı biyopolimer sınıfına giren sakızlar ve müsilajlar hem gıda hem de gıda dışı endüstrilerde geniş uygulamalara sahiptir. Sakızlar ve müsilajlar, ilaç endüstrilerinde koyulaştırıcı, bağlayıcı, emülsifiye edici, süspanse edici ve stabilize edici ajanlar ve ilaç salınım sistem matriksleri olarak kullanılır. Müsilajlar, sürdürülebilir bir kaynak olmalarının yanı sıra suda şişmeleri, biyoyumlu, toksik olmamaları (insan (göz tahrişi ve cilt) sağlığı veya çevre sağlığı), düşük maliyetleri ve kolay temin edilebilirlikleri nedeniyle farmasötik formülasyonların hazırlanmasında ilgi çekici polimerlerdir (Hacker & Nawaz, 2015). Bu çalışmada müsilaj kaynağı olarak seçilen Keten tohumu (*Linum usitatissimum*) dünyanın farklı ülkelerinde yetiştirilen (Kanada, Amerika, Hindistan, Çin) besin maddesi ve lifsi yağlı tohumdur. Keten tohumunun yaklaşık %55'i yağdır ve çoğunluğunu poli-doymamış yağ asitleri ve trigliseridler oluşturur. Keten tohumu yağı soğuk pres ya da tohumun basınçla ezilmesiyle elde edilir. (Chen et al, 2006; Goyal et al, 2014; Puligundla and Lim, 2022). Keten tohumu karbohidratları genelde sindirilemeyen çözünür ve çözünmez liflerden oluşur. Selüloz ve lignin çözünmez lifsi kısmı oluştururken çözünür lifsi kısım musilajın ana iskeletini oluşturur. Keten tohumu müsilajı total tohumun %3-9 oluşturur. Müsilajın içeriğini %50–80 karbohidratlar, %4–20 proteinler ve %3-9 kül oluşturur. Yapılan kimyasal yapı çalışmaları keten tohumu müsilajının gıda hidrokolloid (hidrojel) olduğunu göstermiştir. Yapısında yağ, karbohidrat ve lifsi kısım içerdiği için oldukça fazla kullanım alanı mevcuttur. Gıdalarda kalınlaştırıcı, jelastasyan ajanı, yapısal düzenleyici, emülsifiye edici, enkapsülasyon ajanı, kaplama materyali, prebiyotik, gıda kaplama materyali, yağ çekici gibi alanlar sayılabilir. (Kaur et al, 2018; Goyal et al, 2014; Rashid et al, 2019; Ziokovska, 2012).

Hidrojeller bileşenlerine göre doğal ve sentetik olmak üzere ikiye ayrılırlar. Doğal hidrojel polimer bileşenleri; dekstran, agaroz, hiyaluronik asit, aljinat, karrajenan, pektin, kitosan, polilizin, kollajen, karboksimetil kitin, fibrindir. Sentetik hidrojel polimer bileşenleri polietilen glikol (PEG), poli laktik-ko-glikolik asit (PLGA), polilaktik asit (PLA), polivinil alkol (PVA) dır. (Ahmed, 2015; Mahinroosta et al, 2018; Ullah et al, 2015). Sentetik hidrojellerin, farmasötik ve biyomedikal uygulamalarda kullanımı yaygındır. Sentetik hidrojeller; istenen uygulamaya bağlı olarak fiziksel veya kimyasal olarak çapraz bağlanma, pH, sıcaklık, elektrik sinyali, ışık veya glukoz gibi çevresel koşullardaki değişikliklere yanıt

verme yeteneğine sahiptir. Bu özellikleri sayesinde sentetik hidrojel farklı uygulamalar için spesifik olarak tasarlanabilirler (Agnihotri & Aminabhvi, 2006; Borzacchiello & Ambrosio, 2009; Lavrador et al, 2018; Lee, 2018; Lee & Mooney, 2001; Mantha et al, 2019; Zhu & Marchant, 2011).

İlaç taşınım sistemlerinde hidrojel yaygın çalışılma nedenleri; biyolojik olarak uyumlu olmaları, parçalanabilir olmaları, düşük toksisiteye sahip olmaları, karakteristik dokuya benzer şekilde yüksek derecede esnekliğe sahip olmaları ve modifikasyonlarının kolay olmasıdır. Ayrıca pH, metabolit konsantrasyonu gibi çevresel değişikliklere yanıt olarak kimyasal ve yapısal değişimleri sonucunda içerdikleri bileşikler serbest bırakmaları farklı kullanım avantajı getirir. Genellikle akıllı hidrojel olarak bilenen uyaranlara duyarlı hidrojel sistemleri, yüksek şişme ve jelleşme kapasiteleri nedeniyle taşınım sistemlerinde, çevre çalışmalarında ve farklı biyomedikal alanlarda kullanılmaktadır. Bu hidrojel; pH, sıcaklık, ışık, elektrik alanı gibi çevresel uyarıcılara ve iyonik konsantrasyon gibi kimyasal uyarıcılara yanıt vererek, tersinir bir değişim ile hacim değiştirerek sol jel formatında yapıları meydana getirirler (Dragan 2014; Fan et al, 2022; Hilmi et al, 2016; Liu and Yin, 2015; Klouda & Mikos, 2008; Li et al, 2013; Lavrador et al, 2021). Poli N-izopropilakrilamid (PNIPAM) sıcaklık duyarlı bir organik polimerdir. Organik polimerik malzemelerin hazırlanmasında etkin olan bağlar; hidrojen bağları ve $\pi-\pi$, van der Waals, ve hidrofobik etkileşimlerdir. Yapısında hidrofilik amid ($-\text{CONH}_2$) ve hidrofobik izopropil ($-\text{CH}(\text{CH}_3)_2$) bulunur. Düşük sıcaklıklarda üç boyutlu yapısında büyük boşluklar bulunur. Bu boşluklar sulu çözeltilerde su ile doldurulmuştur. Su molekülleri amid grupları ile hidrojen bağları oluşturur ve bir su moleküler tabakası oluşur. Sıcaklık yükseldikçe hidrojen bağları kırılır, izopropil dehidratların su içeriği azalır ve sonuç olarak hidrofobik gruplar birleşerek su dışarı atılır ve hidrojel yapısı meydana gelir (Ansari et al, 2022; Brahma et al, 2017; El- Hussein et al, 2022; Haq et al, 2017).

Canlıda kullanılacak materyallerin biyoyumlu, biyodegrade olabilen ve olası parçalanma durumunda canlıda herhangi bir immün yanıtı neden olmaması gerekir. Sentetik monomerler ile hazırlanan hidrojel (akrilamid bazlı) insan sağlığına olumsuz etkileri vardır. Bu yüzden ısı duyarlı hidrojel bileşenlerinden biri sentetik monomerde olsa yapısında doğal bir biyopolimerin olması önemlidir (Atoufi et al, 2019; Vihola et al, 2005). Bu çalışmada; sıcaklık duyarlı hidrojel monomer doğal bileşeni olarak bitki bazlı keten tohumu müsilajı seçildi. Müsilajın kullanılmasıyla meydana gelen kıvam artışı sayesinde yüksek miktarda toksik çapraz bağlayıcı kullanımına da gerek kalmamıştır.

Hyaluronik asit (HA) değişen D-glukuronik asit ve N-asetil glukozamin bileşenlerine sahip lineer yüksek molekül ağırlıklı asidik bir ekstrasellüler matriks polisakkaritidir. Hyaluronik asit; hücre farklılaşması ve proliferasyonu, morfogenez, inflamasyon ve yara iyileştirme, kırık bozuklukları, eklem rahatsızlıkları, osteoartrit, kalp kapak mühendisliği, kardiyak onarım gibi oldukça geniş bir spektrumda biyolojik kullanım özelliklerine sahiptir (Bukhari et al, 2018; Mihajlovic et al, 2021; Papakonstantinou et al, 2012; Perez et al,

2021). Bu çalışmada; keten tohumu müsilajı, ısı duyarlı hidrojeller ve hiyaluronik asit ile birleştirilerek ısı duyarlı elastik jellerin sentezi çalışıldı.

2. MATERİYAL ve METOD

2.1. Kimyasallar

Hidrojellerin sentezinde kullanılan N-İzopropilakrilamid (NIPAM), amonyum persülfat (APS) (% 98), tetrametiletilediamin (TEMED), etanol ($\geq$ % 99.9), kitosan (CS), hiyaluronik asit (HA), karboksimetil selüloz (CMC), sodyum dodesil sülfat (SDS) Sigma-Aldrich'ten satın alınmıştır. Keten tohumu aktardan temin edilmiştir.

2.2. Keten Tohumu Müsilajının Eldesi

Keten tohumu aktardan temin edildi. Keten tohumu (10 g) 24 saat boyunca 50°C'de kurumaya bırakıldı. Keten tohumu yağlı bir tohum olduğu için yağsızlaştırma işlemi yapıldı. Yağsızlaştırma için kurutulan örnek 350 mL etanol kullanılarak Soxhlet ekstraksiyonu yapıldı (üç döngü-90 dak.). Örnek işlem sonrası örnek kurutuldu ve tartıldı.

Müsilaj eldesi için farklı pH, sıcak ve alkol gradienti çalışıldı. Müsilaj oluşumunda yapıdaki karbohidratların yapısına bağlı olarak (asidik polisakkaritler) pH oldukça önemlidir. Bu çalışma da tohumlar (1g) farklı pH' lardaki 1 M Sodyum asetat tamponu (pH 4,0-5,0); Sodyum fosfat tamponu (pH 7,0); Tris-HCl tamponu (pH 8,5); Sodyum karbonat tamponu (pH 9,5)) ve dört farklı sıcaklıkta (oda sıcaklığı; 30°C; 37°C, 45°C) 30 dak. bekletildi. Müsilaj oluşumu için 550 µL %100 etanol eklendi ve 24 saat boyunca +4 °C'de bekletildi. 24 saat sonunda müsilajların birleşmesi için 60°C'de su banyosunda 30 dak bekletildi ve ardından bütün tüpler 4500 rpm'de santrifüjlendi, müsilajlar toplandı.

Alkol gradienti çalışmasında farklı etanol yüzdeleri (%) (30; 50; 60; 80) 1g tohumların üzerine 1mL eklendi ve dört farklı sıcaklıkta (oda sıcaklığı, 30°C, 37°C, 45°C) 30 dak bekletildi. Ardından müsilaj oluşumu için 550 µL %100 etanol eklendi ve 24 saat boyunca +4 °C'de bekletildi. 24 saat sonunda müsilajların birleşmesi için 60°C'de su banyosunda yarım saat bekletildi ve ardından tüpler 4500 rpm'de santrifüjlendi, müsilajlar toplandı. Kurutularak çalışmalarda kullanılmak üzere buzdolabında saklandı.

2.3. PNIPAM hidrojellerinin hazırlanması

Bu çalışmada PNIPAM hidrojellerinin hazırlanmasında elastikiyeti artırmak için sodyum dodesil sülfat (SDS), kitosan (CS) ve karboksimetil selüloz (CMC) kullanıldı

2.3.1. PNIPAM-SDS hidrojellerinin hazırlanması

PNIPAM hidrojelleri serbest radikal polimerizasyonu ile hazırlandı. Değişen miktarlarda NIPAM (50-100-200mg) ve değişen miktarlarda çapraz bağlayıcı N, N-Metilenbisakrilamid (MBA) (20-100mg) ve SDS (20mg) 5 ml distile suda çözüldü. 2 saat 60°C'da karıştırıldı. Homojen karışım elde edildikten sonra polimerizasyonu başlatmak için APS (25mg) ve TEMED (25µL) eklendi ve 37°C'de polimerizasyona bırakıldı.

2.3.2. PNIPAM-CMC hidrojellerinin hazırlanması

Değişen miktarlarda NIPAM (50-100-200mg) ve değişen miktarlarda çapraz bağlayıcı MBA (20-100mg) ve değişen hacimlerde %1'lik CMC (500µL-2,5mL) eklenerek son hacim 5 mL distile su ile tamamlandı. 37°C'de 30 dak. karıştırıldı. Homojen karışım elde edildikten sonra polimerizasyonu başlatmak için APS (25mg) ve TEMED (25µL) eklendi ve 37°C'de polimerizasyona bırakıldı.

2.3.3. PNIPAM-Kitosan hidrojellerinin hazırlanması

Değişen miktarlarda NIPAM (25-50mg) ve değişen miktarlarda çapraz bağlayıcı MBA (20-100mg) ve değişen hacimlerde CS (%1) (1,5mL- 2,5mL) eklenerek son hacim 5 mL distile su ile tamamlandı 37°C'de 30 dak. karıştırıldı. Homojen karışım elde edildikten sonra polimerizasyonu başlatmak için APS (25mg) ve TEMED (25µL) eklendi ve 37°C'de polimerizasyona bırakıldı.

2.4. PNIPAM Müsilaj- Hiyaluronik asit (HA) jellerinin hazırlanması

PNIPAM polimerleri: keten tohumu müsilajları 1:1, 1:2.5, 1:5 ve 1:10 olacak şekilde birleştirildi. PNIPAM hidrojeli hazırlanırken 200mg NIPAM 5mL distile suda 37°C'de çözüldü. Çözeltiye 20mg SDS ve 40mg MBA eklendi ve karıştırıldı. Belirlenen oranlarda müsilajlar eklenip ardından 25mg APS ve 25µL TEMED eklendi. 37 °C'de polimerizasyona bırakıldı. Hazırlanan PNIPAM: müsilaj hidrojellerine 2500ppm HA eklendi. Oda sıcaklığında 30 dak. karıştırıldı. PNIPAM: müsilaj hidrojellerinin yapısına katılan HA miktarı karbozol yöntemiyle tayin edildi.

Hazırlanan polimerin cilt ile uyumunu test etmek hidrojel miktarının (PNIPAM: müsilaj: HA) 2 katı hacimde 1M Sodyum asetat tamponu (pH 5,50) içerisinde farklı süreler (1-5-10-20-30) dakika bekletildi ve üst fazdan örnek alındı. Hazırlanan bu hidrojelden zamana bağlı HA salınımı izlendi.

3. SONUÇLAR

3.1. Keten tohumu müsilaj eldesinin optimizasyonu

Farmasötik amaçlı kullanılacak biyomalzemelerin en temel özelliklerinden biri de pH ile olan etkileşim ve pH değişimindeki fizikokimyasal değişimleridir. Müsilaj eldesinde kullanılan bitkinin cinsi, yetiştiği bölge kadar etkili olan diğer faktör de ekstraksiyon ve eldesi esnasındaki ortamın pH'sıdır. Tohumlardaki müsilajın eldesinde farklı teknikler, çözgenler ve ortamlar çalışılmaktadır. Müsilaj eldesinde ekstraksiyon

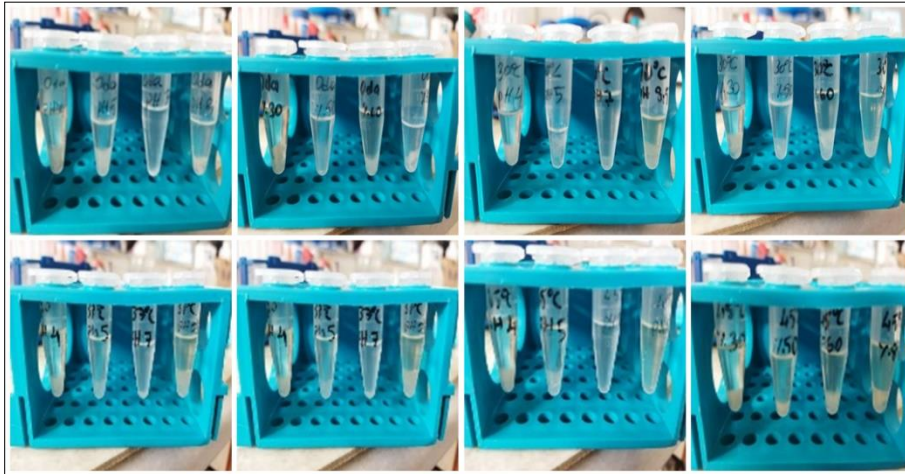
faktörleri müsilajın yapısındaki polisakkaritlerin denatüre olmadığı koşul olarak yöntem tasarlanır (Haseeb et al, 2024; Ramteke et al, 2022).

Çizelge 1: Farklı sıcaklıklar ve tamponların müsilaj oluşumu üzerine etkisi

Sıcaklık	Tamponlar				
	Sodyum Asetat ¹	NaAsetat ²	Sodyum Fosfat ³	Tris-HCl ⁴	Sodyum Karbonat ⁵
Oda Sıc.	+	+	-	+	+
30°C	+	-	-	+	+
37°C	+	+	-	+	+
45°C	+	+	-	-	+

Bu çalışma da müsilaj oluşumu üzerine pH etkisi (farklı tamponlar) Çizelge 1 'de verilmiştir. Müsilaj oluşumunu analizlemek için değişen pH 4,0-10,0 aralıklarında tamponlar denendi (Müsilaj oluşumu (+), Müsilaj oluşumu yok (-)). Keten tohumunun yapısındaki kompleks polisakkaritler asidik ve nötral yapıdadır. Bu sebeple alkali koşullarda çözünmeyen yapıdaki polisakkaritler çözünür forma geçebilir, bu sebeple pH 7,0-8,0'de müsilaj oluşumu gözlenmedi. Düşük pH'larda ise (pH 4,0-5,0) müsilaj oluşumunun sıcaklık ile bağlantılı olduğu bulundu. Bunun da sebebi polisakkarit zincirindeki grupların iyonlaşmasıdır. İyonik bağ oluşumu molekülün polaritesini artırır ve sudaki çözünürlüğü düşer ancak oluşan müsilaj sıcaklık koşulları oda sıcaklığına döndüğünde çözünmüştür. Yeterli mekanik dayanıklılıkta 1 M Sodyum karbonat tamponunda (pH 9,5) 45°C'de elde edilen müsilajlar çalışmanın kalan kısmında kullanıldı.

Müsilaj oluşumunda değerlendirilen bir diğer faktör de alkol çöktürmesidir. Bu çalışmada müsilaj oluşumu üzerine alkol etkisi için farklı yüzdelerdeki etanol ve farklı sıcaklıklar çalışıldı. Farklı sıcaklıklardaki (Oda sıcaklığı; 30°C; 37°C; 45°C) müsilaj oluşumları (farklı tampon ve farklı etanol yüzdelerindeki) Şekil 1'de sıcaklık sırasıyla verilmiştir. Elde edilen sonuçlara göre alkolün protein yapılarının müsilajdan uzaklaştıramamasına bağlı olarak müsilajlar oluşsa bile çok kısa sürede dağıldığı için keten tohumundan müsilaj eldesi için alkol gradientinin uygun olmadığı sonucuna varıldı.



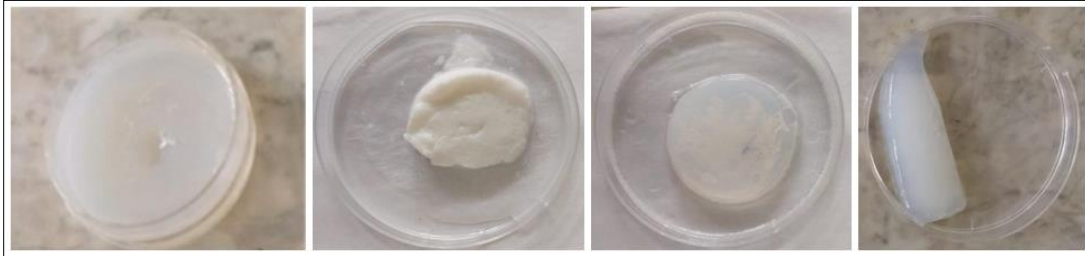
Şekil 1: Farklı sıcaklıklar (Oda sıcaklığı; 30°C; 37°C; 45°C); farklı pH (tampon) ve farklı yüzdelerdeki etanolde elde edilen müsilajlar.

3.2. PNIPAM hidrojellerinin hazırlanması ve optimizasyonu

Vücut sıcaklığına yakın kritik faz değıştirme özelliğine sahip sıcaklık duyarlı hidrojeli PNIPAM medikal alanda en çok kullanılan hidrojeldir. Düşük kritik sıcaklık altında (LCST), faz değışiminin gerçekleştiği PNIPAM hidrojelleri suda şişmelerine neden olan hidrojen bağlarını oluşturur, sıcaklık arttırıldığında ise hidrojen bağları kopar ve hidrojeller büzülür. Bu özellik sayesinde kontrollü bir şekilde içine aldığı ilaç ya da bileşikler bırakır. PNIPAM hidrojelin şişme/büzülme özelliğini çapraz bağlayıcı tipi, çapraz bağlayıcı miktarı, hidrofobik/ hidrofilik grup dengesi ve insiyatör tipi etkiler (Balan et al, 2022; Huang et al, 2006; Cui et al, 2014; Warrena et al, 2020).

Genelde 4 farklı çapraz bağlayıcı PNIPAM hidrojellerinin hazırlanmasında kullanılır. Etilen glikoldimetakrilat (EGDMA), bütandiol diakrilate (BDDA), N,N'-metilenbisakrilamid (MBA) ve trimetilolpropan triakrilattır (TMPTA) (Maitra & Shukla, 2014; Liu et al, 2024). Bu çalışmada biyolojik uyumu daha iyi olan ve suda kolay çözülen MBA çapraz bağlayıcı ve TEMED insiyatör sistemi olarak çalışıldı.

Bu çalışmada PNIPAM hidrofilik zincirleri içerisinde hidrofobik etkileşimi arttırmak için kuvvetli hidrofobik etkileşimleri gerçekleştiren SDS olarak sisteme eklendi.



Şekil 2: PNIPAM-SDS değışen MBA ile hazırlanan hidrojeller

Şekil 2’te görüldüğü gibi PNIPAM-SDS-MBA hidrojellerinde çapraz bağlayıcı arttıkça yapıda çözümler başlamıştır. Yapının elastikiyetini ve suda dağılmadan kalmasını SDS’nin sağladığı görülmektedir. Oluşan jeller hidrofobik bloklarından dolayı suda çözünmez ve jelde fiziksel çapraz bağları meydana getirirler. Bu fiziksel çapraz bağların geçici olarak kırılıp tekrar oluşması sayesinde oda sıcaklığında hidrojeller hızlıca eski hallerine dönebilirler. Yapıdaki kimyasal çapraz bağlayıcı (MBA) hidrojellere suda dağılmamayı sağlar. Bu durumun mekanik dayanıma katkısı fazladır.

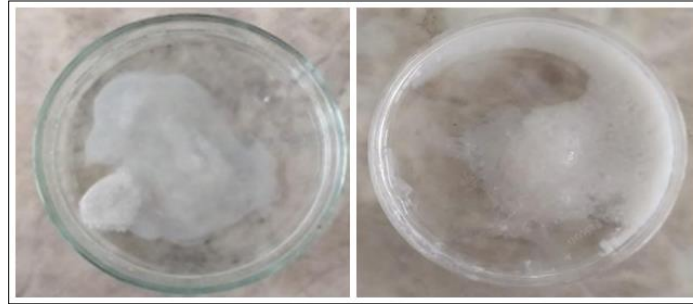
Karboksimetil selüloz (CMC), biyomedikal hidrojel hazırlanmasında en yaygın kullanılan selüloz türevlerindendir. Karboksimetil selüloz; kokusuz, renksiz, suda çözünürlüğü yüksek, toksik olmayan, biyobozunur ve biyouyumlu bir biyopolimerdir. Yapısındaki yüksek orandaki karboksilik gruplarından dolayı anyonik bir polisakkarittir. Bu çalışmada sıcaklık duyarlı hidrofobik karakterli PNIPAM hidrojelinin hidrofilik CMC (iyonize olabilen karboksilik asit gruplarıyla) ile kullanılması daha esnek bir iyonik hidrojel eldesini sağlamaktır (Liu et al, 2022; Chen et al, 2024).



Şekil 3: PNIPAM- değişen CMC oranlarındaki hidrojeller

Şekil 3’de PNIPAM–CMC hidrojellerinin en iyi olanları görülmektedir (%1-2,5). Farklı oranlar denenmesine rağmen (PNIPAM: CMC) oluşan hidrojel polimerlerin çok kısa sürede sertleştiği ve çok kırılgan olduğu bulundu. Kırılganlık suda çözünmenin ve aynı zamanda mekanik dayanıksızlığı gösterdiği için bu hidrojeller enkapsülasyonda kullanılmadı.

Kitosan toksik olmaması, biyoyumlu ve biyobozunur olması, amin ve hidroksil grupları sayesinde modifiye edilebilir olmasından dolayı medikal, kozmetik, doku mühendisliği, farmakoloji gibi birçok alanda uygulamaya sahiptir. Sıcaklığa duyarlı hidrojel eldesinde sıklıkla kullanılan doğal polimerlerden biri kitosandır (Rasiba et al, 2016; Wu et al, 2018). PNIPAM hem hidrofilik hem de hidrofobik kısımlara sahip olan polimer zincirinden meydana gelmektedir. Kitosan ile uyumlu çalışmasının temel nedeni grup etkileşimini sağlamış olmasıdır.



Şekil 4: PNIPAM- Kitosan hidrojelleri

Şekil 4’ de elde edilen PNIPAM–kitosan hidrojellerinin en iyi olanları görülmektedir (%1-3). Farklı oranlar denenmesine rağmen PNIPAM: kitosan hidrojel polimerlerin sertleştiği ve çok kırılgan olduğu ve hidrojel formunda olmadığı bulundu. Kitosan ile istenen hidrojel yapılarının oluşmamasında birincil etkenin pH olduğu düşünülmektedir. Kitosan çözeltilerinde pH değeri 6,2’yi aştığında jel benzeri bir çökelek oluşmaktadır ve PNIPAM ile etkileşim bu pH’da olduğu için istenen hidrojel formu meydana gelmemiştir. Bu çalışmada eklenen çapraz bağlayıcı miktarı ve PNIPAM oranı ve insiyatör sisteminin polimerizasyon aşamasında kitosan ile uyumsuzluğundan kaynaklı olarak oluşmadığı sonucuna varıldı.

3.3. PNIPAM-Müsilaj: Hiyaluronik asit hidrojjellerinin optimizasyonu

Sol-jel geçişinin tersinirliğı, sentezlenen hidrojjelin biyomedikal alanda kullanımı ile oldukça paralel bir özelliktir. Tipik hücre tedavisi protokollerinde hücrelerin matris içerisinde ve fizyolojik sıcaklıkta kolayca karıştırılmasının sağlanması gerekmektedir. (Oda sıcaklığında sol durumunda, fizyolojik sıcaklıkta jel durumunda bulunabilirlik.) Sıcaklık bağımlı jelleşme hücre canlılığını korumak için yüksek mekanik şiddet olmadan kolay bir enjeksiyona, hedef dokulara kolay entegrasyona izin verir. (Papakonstantinou, Roth & Karakioulakis, 2012) Cilt üzerine olan etkileri bilinen HA, sentetik polimerler yerine bitki bazlı müsilaj kullanarak sentezlenen sıcaklık duyarlı hidrojele kapsüle edilerek biyomedikal kullanımını çalışılabilirliğı amacıdır.

Çalışma da PNIPAM-müsilaj:SDS oranlarının optimizasyonu için farklı sıcaklıklar denendi.

30°C’de PNIPAM: müsilaj: SDS çalışılan 1:1:1-1:2,5:1-1:5:1 ve 1:10:1 oranlarının tamamında hidrojel oluşumu gözlemlendi. Ancak 30°C hem müsilajın eldesi koşullarına yakın olması ve aynı zamanda PNIPAM için sınır LCST aralığında olduğu için oluşan yapılarda kısa sürede bozulmalar gözlemlendi.

37 °C’de PNIPAM: müsilaj: SDS çalışılan 1:1:1-1:2,5:1-1:5:1 ve 1:10:1 oranlarının tamamında hidrojel oluşumu gözlemlendi. Bu sıcaklık hem müsilajın eldesi koşullarına yakın olması ve aynı zamanda PNIPAM için LCST değeri olduğu için olumlu sonuçlar elde edildi ve bozulmadan kalan en stabil oluşum 1:5:1 oranı bulundu (Şekil 5).

45 °C’de PNIPAM: müsilaj: SDS oranlarından sadece 1:5:1 ve 1:10:1 oranlarında hidrojel oluşumu gözlemlendi. Bu sıcaklıkta PNIPAM:müsilaj ile SDS arasında birleşme gerçekleşmedi.

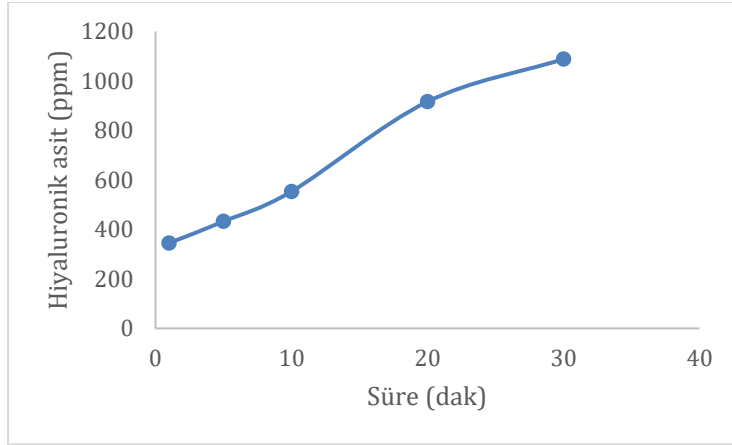


Şekil 5: 37°C’de PNIPAM:müsilaj:SDS (1:5:1)

3.3.1. Hiyaluronik asit Salınımının İzlenmesi

Hiyaluronik asit (2500ppm) çözeltisi, PNIPAM: müsilaj: SDS(1:5:1) (37°C) hidrojjeline eklenerek belli aralıklarda örnekler alındı. Yapılan bu salınım testi için seçilen süre aralıkları 0-5-10-20-30 dak olarak belirlendi. Şekil 6’da görüldüğü gibi ilk 10 dak. salınım miktarı oldukça düşüktür. Bunun sebebi polimerin fiziksel olarak ortama ve sıcaklığa uyum sağlaması için geçen süredir. Bu süre boyunca mekanik kararlılığı ve dayanıklılığı devam etmektedir. 10-30 dak arasında ise hızlı bir salınım gözlenmiştir. Bu durumda hidrojjelin gevşeme ve içeriğın salınımını gerçekleştirdiği zaman olarak tanımlanabilir. Yapılan salınım

çalışması sonuçlarına göre başlangıç hyaluronik asiti yaklaşık %50 kısmının polimerin dışına çıktığı gözlemlendi. Geri kalan kısmında 24 saat içerisinde çıkışı gözlemlendi. Sonuçlara göre PNIPAM-müsilaj-HA hidrojelinin HA enkapsülasyonuna uygun olduğu gözlemlendi (Varshosaz & Falamarzian, 2001).



Şekil 6: Hyaluronik asitin zamana bağlı salınımı

4. SONUÇ

Bitkisel müsilajların biyomedikal alandaki kullanımları son yıllarda biyoyumlu olmaları ve biyotoksitelerinin olmaması sebebiyle gittikçe artmaktadır. Oldukça fazla bitki tohumu müsilaj kaynağı olarak kullanılmaktadır. Bunlar içerisinde en fazla ilgi çekenler keten tohumu ve bamyadır. Bamyaya tohum müsilajı yeterince elastikliğe sahip olmadığı için bu çalışmada keten tohumu müsilajı kullanılmıştır. Keten tohumu müsilajı yağsızlaştırıldıktan sonra farklı sıcaklık, pH ve alkol yüzdelerinde en iyi müsilaj verimi için optimizasyon çalışmaları yapılmıştır. Müsilajın tek başına vücutta taşınım sistemi olarak kullanımı mümkün değildir. Bu sebeple sıcaklık duyarlı bir monomer olan PNIPAM ile hazırlanan hidrojel ile birleştirilmiştir. Çalışmanın insan sağlığı için en önemli katma değeri, insan sağlığına zararlı sentetik monomerler ve toksik çapraz bağlayıcılar yerine bitki bazlı biyopolimer olan keten tohumu müsilajının kullanılmasıdır. Hyaluronik Asit ile keten tohumu müsilajıyla hazırlanan hidrojeller bir çok farklı alanda kullanılabilecektir. Tamamen sürdürülebilir, biyoyumlu ve doğal kaynaklardan elde edilerek hazırlanan, sıcaklık ile fiziksel özellikleri değişebilen hidrojellerin HA birleştirilmesiyle sentetik polimerlerin toksik etkilerini ortadan kaldırarak biyomedikal kullanımına olanak sağlayacaktır.

ÇIKAR ÇATIŞMASI BEYANI

Yazarlar arasında çıkar çatışması bulunmamaktadır.

TEŞEKKÜR

Bu çalışmadaki katkılarından dolayı Yüksek Biyokimyager Hüseyin Özyay'a ve Ege Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimine (22937) teşekkür ederiz.

YAZARLARIN KATKILARI

H.H.: Yöntem, araştırma, kaynaklar, yazı yazma-orijinal taslak hazırlama.

B.O.: Araştırma, kaynaklar, yazı yazma-gözden geçirme ve düzenleme.

KAYNAKLAR

- Agnihotri, S.A., & Aminabhavi, T.M. (2006). Novel interpenetrating network chitosan- poly(ethylene oxide-g-acrylamide) hydrogel microspheres for the controlled release of capecitabine. *International Journal of Pharmaceutics*, 324(2):103-115. doi: 10.1016/j.pharm.2006.05.061.
- Ahmad S., Ahmad, M., Manzoor, K., Purwar, R., & Ikram, S. (2019). A review on latest innovations in natural gums based hydrogels: Preparations & applications. *International Journal of Biological Macromolecules*, 136, 870–890. doi:10.1016/j.ijbiomac.2019.06.113.
- Ahmed, E. M. (2015). Hydrogel: Preparation, characterization, and applications: *A review. Journal of Advanced Research*, 6(2): 105-121 Doi:10.1016/j.jare.2013.07.006
- Ansari, M., J., Rajendran, R.R., Mohanto, S., Panda, U.A.K., Dhotre, K., Manne, R., Deepak, A., Zafar, A., Yasir, M., Pramanik, S. (2022). Poly(N-isopropylacrylamide)-based hydrogels for biomedical applications: A review of the state-of-the-art. *Gels*, 8: 454-499. doi:10.3390/gels8070454.
- Atoufi, Z., Kamrava, S., K., Davachi, S., M., Hassanabadi, M., Garakani, S., S., Alizadeh, R., Farhadi, M., Tavakol, S., Bagher, Z., Motlagh, G., H. (2019). Injectable PNIPAM/Hyaluronic acid hydrogels containing multipurpose modified particles for cartilage tissue engineering: Synthesis, characterization, drug release and cell culture study. *International Journal of Biological Macromolecules*, 139:1168–1181. doi:10.1016/j.ijbiomac.2019.08.101.
- Balan, K., E., Boztepe, C., Künkül, A. (2022). Modeling the effect of physical crosslinking degree of pH and temperature responsive poly(NIPAAm-co-VSA)/alginate IPN hydrogels on drug release behavior. *Journal of Drug Delivery Science and Technology*, 75:103671. doi:10.1016/j.jddst.2022.103671.
- Borzacchiello, A., Ambrosio L. (2009). Structure-property relationships in hydrogels, *Hydrogels*, Springer, 9-20 pp.
- Brahima, S., Boztepe, C., Kunkul, A., Yuceer, M. (2017). Modeling of drug release behavior of pH and temperature sensitive poly(NIPAAm-co-AAc) IPN hydrogels using response surface methodology and artificial neural networks. *Materials Science and Engineering C*, 7(5): 425–432. doi:10.1016/j.msec.2017.02.081.
- Bukhari, S. N. A., Roswandi, N. L., Waqas, M., Habib, H., Hussain, F., Khan, S., Sohail, M., Ramli, N. A., Thu, H. E., & Hussain, Z. (2018). Hyaluronic acid, a promising skin rejuvenating biomedicine: A review of recent updates and pre-clinical and clinical investigations on cosmetic and nutricosmetic effects. *International Journal of Biological Macromolecules*, 120:1682–1695. doi:10.1016/j.ijbiomac.2018.09.188.
- Chen, H. H., Xu, S. Y., Wang, Z. (2006). Gelation properties of flaxseed gum. *Journal of Food Engineering*, 77(2):295–303. doi:10.1016/j.jfoodeng.2005.06.033
- Chen, G., Ma, F., Li, J., Yang, P., Wang, Y., Li, Z., Meng, Y. (2024). Preparation of CMC-poly(N-isopropylacrylamide) semi-interpenetrating hydrogel with temperature-sensitivity for water retention. *International Journal of Biological Macromolecules*, 268 :131735. doi:10.1016/j.ijbiomac.2024.131735
- Cui L., Jia, J., Guo, Y., Liu, Y., Zhu, P. (2014). Preparation and characterization of IPN hydrogels composed of chitosan and gelatin cross-linked by genipin. *Carbohydrate Polymers*, 99:31-38. Doi:10.1016/j.carpol.2013.08.048
- Dragan, S.D. (2014). Design and applications of interpenetrating polymer network hydrogels. A review. *Chemical Engineering Journal*, 243:572-590. doi:10.1016/j.cej.2014.01.065
- El-Husseiny, H., M., Mady, E., A., Hamabe L., Abugomaa, A., Shimada K., Yoshida, T., Tanaka, T., Yokoi, A., Elbadawy, M., Tanaka, R. (2022). Smart/stimuli-responsive hydrogels: Cutting-edge platforms for tissue engineering and other biomedical applications. *Materials Today Bio*, 13: 100186. doi:10.1016/j.mtbio.2021.100186.
- Fan, R., Cheng, Y., Wang, R., Zhang, T., Zhang, H., Li, J., Song, S., Zheng, A. (2022). Thermosensitive hydrogels and advances in their applicationin disease therapy. *Polymers*, 14: 2379-2400. doi:10.3390/polym14122379.

- George, B., & Suchithra T.V. (2019). Plant-derived bioadhesives for wound dressing and drug delivery system. *Fitoterapia*, 137:104241. doi:10.1016/j.fitote.2019.104241
- Goyal, A., Sharma, V., Upadhyay, N., Gill, S., Sihag, M. (2014). Flax and flaxseed oil: an ancient medicine & modern functional food. *Journal of Food Science and Technology*, 51(9):1633–1653. doi:10.1007/s13197-013-1247-9.
- Hacker, M., C., Nawaz, H., A. (2015). Multi-functional macromers for hydrogel design in biomedical engineering and regenerative medicine. *International Journal of Molecular Science*, 16(11): 27677–27706. doi:10.3390/ijms161126056.
- Haseeb, M., T., Muhammad, G., Hussain, M., A., Bukhari, S., N., A., Sheikh, F., A. (2024). Flaxseed (*Linum usitatissimum*) mucilage: A versatile stimuli–responsive functional biomaterial for pharmaceuticals and healthcare. *International Journal of Biological Macromolecules*. 278: 134817. doi:10.1016/j.ijbiomac.2024.134817.
- Haq, M., A., Su, Y., Wang, D. (2017). Mechanical properties of PNIPAM based hydrogels: A review. *Materials Science and Engineering C*, 70: 842–855. doi:10.1016/j.msec.2016.09.081.
- Hilmi, B., Hamid, A., Akil, H., Yahay, B., H. (2016). The characteristics of the smart polymer as temperature or pH responsive hydrogel. *Procedia Chemistry*, 19:406–409. doi:10.1016/j.proche.2016.03.031.
- Hosseini, M., S., & Nabid, M., R. (2020). Synthesis of chemically cross-linked hydrogel films based on basil seed (*Ocimum basilicum* L.) mucilage for wound dressing drug delivery applications. *International Journal of Biological Macromolecules*, 163:336–347. doi:10.1016/j.ijbiomac.2020.06.252.
- Huang, J., Wang, X., Yu, X. (2006). Solute permeation through the polyurethane- NIPAAm hydrogel membranes with various cross-linking densities. *Desalination*, 192: 125– 131. doi:10.1016/j.desal.2005.04.133.
- Kaur, M., Kaur, R., Punia, S. (2018). Characterization of mucilages extracted from different flaxseed (*Linum usitatissimum* L.) cultivars: A heteropolysaccharide with desirable functional and rheological properties. *International Journal of Biological Macromolecules*, 117:919–927. doi:10.1016/j.ijbiomac.2018.06.010.
- Klouda, L., & Mikos, A., G. (2008). Thermoresponsive hydrogels in biomedical applications. *European Journal of Pharmaceutics and Biopharmaceutics*, 68:34–45. doi:10.1016/j.ejpb.2007.02.025.
- Lavrador, P., Esteves, M.R., Gaspar, V.M., Mano, J.F. (2021). Stimuli-responsive nanocomposite hydrogels for biomedical applications. *Advanced Functional Materials*, 31(8): 2005941. doi:10.1002/adfm.202005941.
- Lavrador, P., Gaspar, V.M., Mano, J.F. (2018). Stimuli-responsive nanocarriers for delivery of bone therapeutics - barriers and progresses. *Journal of Controlled Release*, 273: 51–67. doi:10.1016/j.jconrel.2018.01.021.
- Lee, J.H. (2018). Injectable hydrogels delivering therapeutic agents for disease treatment and tissue engineering. *Biomaterial Research*, 22(1):1–14. doi:10.1186/s40824-018-0138-6.
- Lee, K.Y., & Mooney, D.J. (2001). Hydrogels for tissue engineering. *Chemical Reviews*, 101(7):1869–1880. doi:10.1021/cr000108x.
- Li, P., Xua, K., Tan, Y., Lu, C., Li, Y., Wang, P. (2013). A novel fabrication method of temperature-responsive poly(acrylamide) composite hydrogel with high mechanical strength. *Polymer*, 54: 5830–5838. doi:10.1016/j.polymer.2013.08.019.
- Liu, J., Yin, Y. (2015). Temperature responsive hydrogels: construction and applications, *Polymer Science*, 1(13):1–6. doi:10.21767/2471-9935.100003.
- Liu, Z., Zhang, S., Gao, C., Meng, X., Wang, S., Kong, F. (2022). Temperature/ pH-Responsive Carboxymethyl Cellulose/ Poly (N-isopropyl acrylamide) Interpenetrating Polymer Network Aerogels for Drug Delivery Systems. *Polymers*, 14: 1578. doi: 10.3390/polym14081578.
- Liu, L., Liu, H., Wang, R., Zhou, J., Zhao, L., Li, Q., Liu, Z. (2024). Preparation and application of environmentally-responsive hydrogels in tissue engineering. *Materials Today Communications*, 40: 109493. doi:10.1016/j.mtcomm.2024.109493.
- Mahinroosta, M., Farsangi, Z.J., Allahverdi, A., Shakoori, Z. (2018). Hydrogels a intelligent materials: A brief review of synthesis, properties and applications, *Materials Today Chemistry*, 8:42–55. Doi:10.1016/j.mtchem.2018.02.004

- Maitra, J., & Shukla, V.K. (2014). Cross-linking in Hydrogels - A Review, *American Journal of Polymer Science*, 4(2): 25-31. doi:10.5923/j.ajps.20140402.01.
- Mantha, S., Pillai, S., Khayambashi, P., Upadhyay, A., Zhang, Y., Tao, O., Pham, H.M., Tran, S.D. (2019). Smart hydrogels in tissue engineering and regenerative medicine. *Material (Basel)*, 12(20): 3323. doi:10.3390/ma12203323.
- Mihajlovic M., Fermin, M., M., L., Ito, K., Nostrum, C., F., Vermonden, T. (2021). Hyaluronic acid-based supramolecular hydrogels for biomedical Applications. *Multifunctional Materials*, 4(3): 03200. doi:10.1088/2399-7532/ac1c8a.
- Papakonstantinou, E., Roth, M., Karakiulakis, G. (2012). Hyaluronic acid: A key molecule in skin aging. *Dermato-Endocrinology*, 4(3):37–41. doi:10.4161/derm.21923.
- Pérez, L., A., Hernández, R., Alonso, J., M., Pérez-González, R., Sáez-Martínez, V. (2021). Hyaluronic Acid Hydrogels Crosslinked in Physiological Conditions: *Synthesis and Biomedical Applications. Biomedicines*, 9:1113-1134. doi:10.3390/biomedicines9091113.
- Prajapati, V. D., Jani, G. K., Moradiya, N. G., Randeria, N. P. (2013). Pharmaceutical applications of various natural gums, mucilages and their modified forms. *Carbohydrate Polymers*, 92(2): 1685–1699. doi:10.1016/j.carbpol.2012.11.021.
- Puligundla, P., & Lim, S. (2022). A Review of Extraction Techniques and Food Applications of Flaxseed Mucilage. *Foods*, 11:1677-1697. doi:10.3390/foods11121677.
- Ramteke, S., Haigune, N., More, S., Pise, S., Pise, A., Kharwade, R. (2022). Flaxseed Mucilage Hydrogel based Floating Drug Delivery System: Design and Evaluation. *Research Journal of Pharmacology. and Technology*, 15(4):1549- 1554. doi:10.52711/0974-360X.2022.00258.
- Rashid, F., Ahmed, Z., Hussain, S., Huang, J. Y., Ahmad, A. (2019). Linum usitatissimum L. seeds: Flax gum extraction, physicochemical and functional characterization. *Carbohydrate Polymers*, 215:29–38. doi:10.1016/j.carbpol.2019.03.054.
- Rasiba, S., Z., M., Akil, H., Yahya, A., S. (2016). Effect of Different Composition on Particle Size Chitosan-PMAA-PNIPAM Hydrogel. *Procedia Chemistry*, 19:388–39. doi:10.1016/j.proche.2016.03.028
- Ullah, F., Othman, M.,B.,H., Javed, F., Ahmad, Z., Akil, H. (2015). Classification, processing and application of hydrogels: A review. *Materials Sciences and Engineering: C*, 57:414-443. doi:10.1016/j.msec.2015.07.053.
- Varshosaz, J., & Falamarzian, M. (2001). Drug diffusion mechanism through pH- sensitive hydrophobic/polyelectrolyte hydrogel membranes. *European Journal of Pharmaceutics and Biopharmaceutics*, 51(3):235-240. doi:10.1016/S0939-6411(01)00126-6
- Vihola, H., Laukkanen, A., Valtola, L., Tenhu, H., Hirvonen, J. (2005). Cytotoxicity of thermosensitive polymers poly(N-isopropylacrylamide), poly(N-vinylcaprolactam) and amphiphilically modified poly(N-vinylcaprolactam). *Biomaterials*, 26(16): 3055–3064. doi:10.1016/j.biomaterials.2004.09.008.
- Warrena, H., Shepherd, D., J., Panhuis, M., Officera, D., L., Spinks, G., M. (2020). Porous PNIPAm hydrogels: Overcoming diffusion-governed hydrogel actuation. *Sensors and Actuators A*, 301: 111784. doi:10.1016/j.sna.2019.111784.
- Wu, S., Liua, X., Miller II, L., A., Cheng, Y., Yeh, M., Lu, L. (2018). Strengthening injectable thermo-sensitive NIPAAm-g-chitosan hydrogels using chemical cross-linking of disulfide bonds as scaffolds for tissue engineering. *Carbohydrate Polymers*, 192:308–316. doi: 10.1016/j.carbpol.2018.03.047.
- Zhao, W., Jin, X., Cong, Y., Liud, Y. and Fue, J. (2012). Degradable natural polymer hydrogels for articular cartilage tissue engineering. *Journal of Chemical Technology and Biotechnology*, 88(3):327-339. doi:10.1002/jctb.3970.
- Zhu, J., & Marchant, R.E. (2011). Design properties of hydrogel tissue-engineering scaffolds. *Expert Review of Medical Devices.*, 8(5): 607–626. doi:10.1586/erd.11.27.
- Ziolkovska, A. (2012). Laws of flaxseed mucilage extraction. *Food Hydrocolloids*, 26(1):197– 204. doi:10.1016/j.foodhyd.2011.04.022.

IoU-Based Anchor Box Estimation for Enhanced Lung Region Localization in Chest X-rays Using YOLO v4

Serdar Abut^{1*} 

¹Department of Artificial Intelligence and Machine Learning, Kayseri University, Kayseri, Türkiye

Received: 15.10.2025, Accepted: 08.12.2025, Published: 20.12.2025

ABSTRACT

Precise lung region detection in chest radiographs is an essential preprocessing step for computer-aided diagnostics. This study presents a YOLO v4-based framework to automatically localize lung regions in posteroanterior (PA) chest X-rays. A subset of the CheXpert dataset, containing 456 manually annotated PA radiographs, was used. Anchor boxes were estimated via an Intersection-over-Union (IoU)-based clustering method, improving scale invariance and shape alignment over Euclidean metrics. Empirical evaluation showed that six anchor boxes achieved the best balance between mean IoU (0.883) and computational efficiency. The trained model was tested on 144 images, yielding Average Precision (AP) of 0.9043 for the lung_region class, which represents only the anatomical lung area and not any specific pathology. The precision-recall curve indicated high precision across most recall values, and the confusion matrix showed 124 true positives, 13 false positives, and 7 false negatives. These results demonstrate that YOLO v4 with optimized anchor box estimation enables accurate, efficient lung region localization, supporting automated radiology workflows.

Keywords: Lung Region Detection; Image Processing; YOLO v4; Anchor Box Optimization; Object Detection

YOLO v4 Kullanarak Göğüs Röntgenlerinde Geliştirilmiş Akciğer Bölgesi Yerelleştirme için IoU Tabanlı Çapa Kutusu Tahmini

ÖZ

Göğüs röntgenlerinde akciğer bölgesinin doğru tespiti, bilgisayar destekli tanı sistemleri için kritik bir ön işleme adımdır. Bu çalışmada, posteroanterior (PA) göğüs röntgenlerinde akciğer bölgelerini otomatik olarak yerelleştirmek için YOLO v4 tabanlı bir çerçeve sunulmuştur. CheXpert veri kümesinden 456 adet elle anotlanmış PA röntgen kullanılmıştır. Çapa kutuları, ölçekten bağımsız mesafe ölçümü ve şekil hizalamasında iyileşme sağlayan IoU (Intersection-over-Union) tabanlı kümeleme yöntemiyle tahmin edilmiştir. Deneyisel değerlendirmeler, altı çapa kutusunun ortalama IoU (0,883) ve hesaplama verimliliği açısından en iyi dengeyi sunduğunu göstermiştir. Eğitilen model, 144 görüntüden oluşan test kümesinde çalıştırılmış ve lung_region sınıfı, yalnızca anatomik akciğer bölgesini temsil etmekte olup herhangi bir patolojiyi göstermemektedir; Doğruluk (AP) değeri 0,9043 elde edilmiştir. Kesinlik-duyarlılık eğrisi, çoğu duyarlılık değerinde yüksek kesinlik göstermiştir. Karmaşıklık matrisi ise 124 doğru pozitif, 13 yanlış pozitif ve 7 yanlış negatif tespit etmiştir. Sonuçlar, optimize edilmiş çapa kutusu tahmini ile YOLO v4'ün doğru ve verimli akciğer bölgesi yerelleştirmesi sağlayabildiğini göstermektedir.

Anahtar Kelimeler: Akciğer Bölgesi Tespiti; Görüntü İşleme; YOLO v4; Çapa Kutusu Optimizasyonu; Nesne Tespiti

*Corresponding Author: Serdar Abut
E-mail: serdarabut@kayseri.edu.tr

1. INTRODUCTION

The increasing use of Artificial Intelligence (AI) in the medical domain, particularly in medical image analysis, has been propelled by a paradigm shift from traditional Artificial Neural Networks (ANNs) to deep Convolutional Neural Networks (DCNNs) (Abut et al., 2024). These models have shown remarkable performance in tasks such as disease classification from chest X-rays, facial image-based gender recognition (Ergen & Abut, 2013), and broader applications in medical decision-making (Abut & Okut, 2024). Despite these successes, explainability and reliability remain critical challenges. A promising solution is to apply segmentation techniques to isolate the lung region prior to classification. By enforcing the model to learn only from relevant anatomical structures, segmentation improves both accuracy and interpretability. For example, COPD-GradeNet is a model proposal that integrates lung-region-focused learning to predict the severity of chronic obstructive pulmonary disease (COPD) (Abut, 2024).

Recent studies emphasize the importance of considering anatomical structures in medical imaging. This attention spans multiple domains, including nuclear marker quantification in tumor tissues (Yilmaz et al., 2025b), AI-based segmentation of kidney biopsies (Yilmaz et al., 2025a; Yilmaz et al. 2025c), and emergency care applications (Berikol et al., 2025). These approaches collectively underline the importance of guiding AI models toward relevant visual features to improve diagnostic reliability.

Although deep learning-based models have shown promising results in the classification of lung diseases, careful visual analyses are required to explain their decision-making processes and improve their accuracy. Visualization techniques such as Grad-CAM make it possible to analyze which regions the model focuses on when making decisions. As shown in Figure 1, the classification model focused not on the lung image itself but on irrelevant areas in the corner of the image. Such deviations indicate that the model may lead to incorrect generalizations and pose reliability issues. In this context, applying lung region segmentation to ensure that the model considers only the relevant anatomical areas (e.g., lung lobes) during decision-making is of critical importance. Performing segmentation as a preprocessing step helps guide the model's attention to the correct region and improves classification performance (Irvin et al., 2019; Rajpurkar et al., 2017).



Figure 1: Grad-CAM visualization applied to a chest X-ray image

In Figure 1, the left image shows the attention map overlaid on the original X-ray, where red areas indicate higher attention weights of the model. It can be observed that the classifier has focused on irrelevant regions such as the top-left corner of the image rather than the lung area, which is the anatomically meaningful region for diagnosis. This highlights the necessity of restricting the model's attention to the lung region via segmentation-based preprocessing for more accurate disease classification.

Recent advances in deep learning, particularly convolutional neural networks (CNNs), have revolutionized medical image analysis by enabling hierarchical feature learning directly from raw data, leading to improved accuracy in various tasks including classification, segmentation, and localization of diseases in CXR images. Deep learning models have demonstrated superior performance in detecting lung infections, cancers, and other thoracic abnormalities by effectively learning spatial-contextual patterns that are not readily captured by traditional approaches.

Given the importance of correctly identifying the lung region to reduce irrelevant information and improve model performance, this study employs a YOLOv4-based deep learning framework for automatic localization and rectangular cropping of lung areas in posteroanterior chest X-rays. Unlike methods that perform precise boundary segmentation, the proposed approach focuses on bounding-box localization rather than pixel-level delineation, thereby preserving peripheral contextual cues that may contribute to disease classification. This design choice also prevents potential loss of diagnostically relevant patterns beyond the strict anatomical lung borders.

Prior studies have explored various segmentation techniques, including texture-based filtering combined with morphological operations to generate lung masks, which successfully excluded non-lung areas such as the shoulder and diaphragm regions (Matsuyama, 2021). Other research leveraged deep CNN architectures

like DenseNet-121 with transfer learning to classify lung nodules and cancer from chest X-rays, highlighting the role of precise lung region localization in improving diagnostic accuracy (Ausawalaithong et al., 2018). More recent work has integrated segmentation and classification networks, combining encoder-decoder frameworks such as UNet++ with classification modules, achieving high precision in lung infection localization and disease detection with efficient model designs suitable for clinical applications (Miah et al., 2024).

Moreover, deep Siamese networks have been employed to compare symmetrical lung segments for pneumonia classification, illustrating the utility of lung segmentation as a preprocessing step for enhanced disease detection (Acharya & Satapathy, 2020). Surveys and reviews indicate a growing trend toward leveraging advanced deep learning models for diverse chest X-ray analysis tasks, emphasizing the need for robust lung localization to support downstream applications (Ait Nasser & Akhloufi, 2023; Çallı et al., 2021).

In this context, the present study introduces an automated deep learning pipeline focusing on accurate lung region detection and bounding box-based cropping in chest X-rays, intended to facilitate and improve the performance of various CAD systems for thoracic disease diagnosis. While object detection networks such as YOLO have been widely adopted in natural image analysis, their systematic application for lung region localization within MATLAB-based medical imaging workflows has been limited. This study addresses this practical gap by implementing and optimizing YOLOv4 for robust, efficient localization of lung fields, mitigating the confounding influence of non-anatomical artifacts (e.g., labels or external patterns) commonly present in chest radiographs. The resulting framework offers a reproducible preprocessing solution that enhances downstream diagnostic modeling by providing clean, context-preserving input regions.

2. MATERIAL AND METHOD

2.1. Dataset

In this study, a subset of the publicly available CheXpert dataset (Saporta et al., 2022) was utilized to develop and evaluate the lung localization framework. Only posteroanterior (PA) view radiographs were included to maintain consistency in image orientation and diagnostic relevance, while lateral views were excluded. From the CheXpert collection, 456 images were selected based on the presence of clear lung regions and sufficient image quality, following a semi-random sampling procedure that ensured diversity across different thoracic pathologies. This subset was primarily used for lung region localization; a related

but distinct study is being conducted to investigate Atelectasis classification using the same source dataset but with a different methodological focus.

2.2. Manual Annotation

Prior to training the YOLO model, each of the 456 selected chest radiographs underwent manual annotation to identify the lung regions. Manual annotation of lung regions was carried out to generate the ground-truth bounding boxes used for training and evaluation. Each chest X-ray was visually inspected, and the lung boundaries were marked to define the minimal rectangular area covering both lungs. The manual annotations were performed by a biomedical engineer experienced in medical image processing, following consistent visual criteria to ensure reproducibility and minimize subjectivity across all samples. Figure 2 illustrates an example of the annotated regions used in the experiments. A rectangular bounding box was drawn around the visible lung fields in each image, ensuring that the entire lung area was enclosed within the frame. The corner coordinates (top-left and bottom-right) of each bounding box were manually recorded for every image. These coordinates served as ground truth data for the subsequent training of the object detection model (Figure 2).

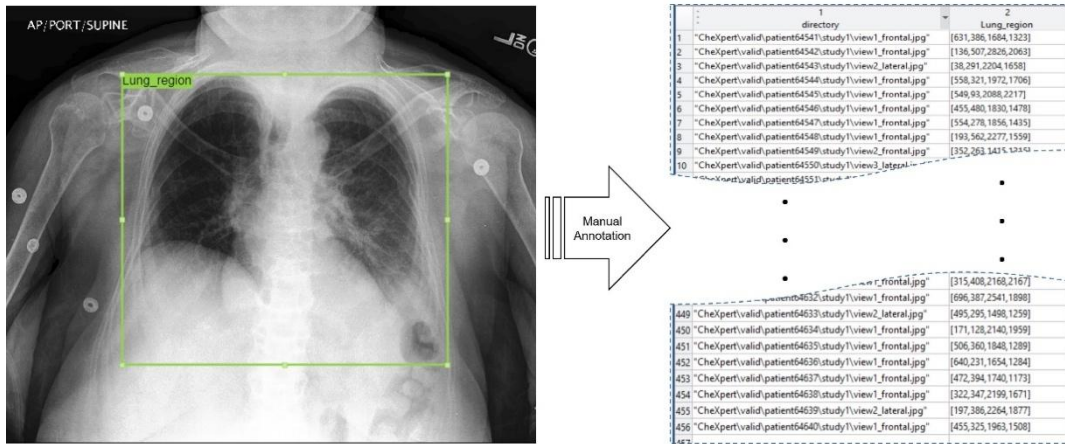


Figure 2: Manual annotation of the lung region

2.3. Implementation of YOLO v4 for Object Detection and Anchor Box Estimation

In this study, the YOLO v4 object detection network was utilized for lung region detection. YOLO v4 is a one-stage object detector that provides a balanced trade-off between detection speed and accuracy (Bochkovskiy et al., 2020).

During model training, the input image size was fixed at 416×416 pixels. Based on the locations and sizes of lung regions in the training data, suitable predefined anchor boxes were estimated. The YOLOv4 model was trained using the Adam optimizer with gradient decay factors of 0.9 and 0.999, an initial learning rate of 0.001, and an L2 regularization factor of 0.0005. The training was performed for 80 epochs with a mini-

batch size of 4, and the best-performing model was selected based on the lowest validation loss. Training and validation data were shuffled at each epoch to prevent overfitting, and checkpointing was applied to save intermediate models. The total training process was completed in approximately 45 minutes.

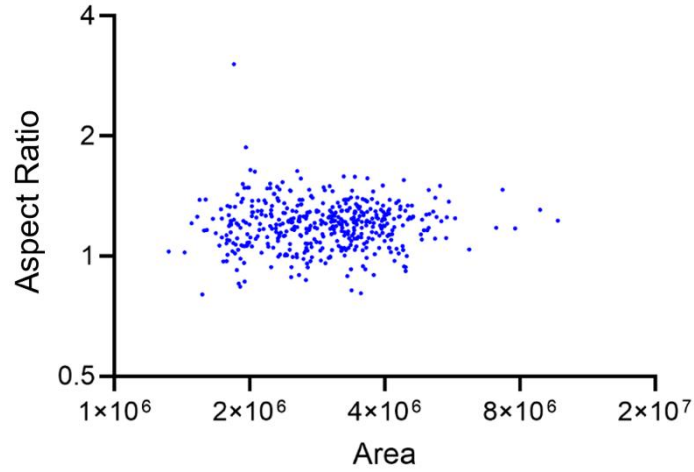


Figure 3: Distribution of anchor box areas (in pixels) with respect to their aspect ratios

Figure 3 shows the distribution of aspect ratios of the manually annotated rectangles during training, plotted against their corresponding areas within the images. It can be observed from the figure that the distribution appears approximately uniform on average. Anchor boxes were estimated from the training data using a method based on the intersection-over-union (IoU) distance metric (Figure 4). The reason for using an IoU-based distance metric is its invariance to box sizes, unlike the Euclidean distance metric, which tends to produce larger errors as box dimensions increase (Redmon & Farhadi, 2017). In addition, it was considered that the IoU metric would cause boxes corresponding to the lung region—generally having similar aspect ratios and sizes across the images—to be clustered together, thus resulting in anchor boxes that better fit the data.

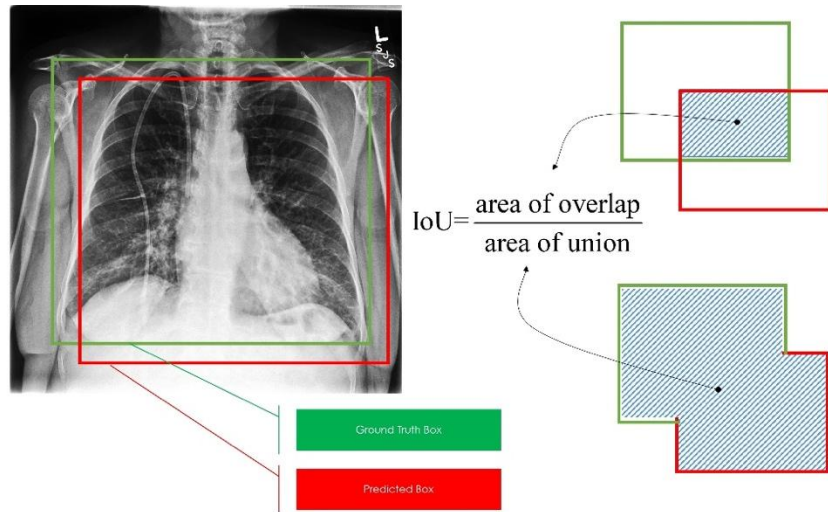


Figure 4: Mean IoU values obtained by using different numbers of anchor boxes.

A mean IoU value of approximately 0.83 is achieved using two anchor boxes, and the mean IoU is improved up to about 0.88 by increasing the number of anchor boxes to six. Beyond six anchor boxes, only marginal improvements in mean IoU are observed. Given these results, multiple object detectors are planned to be trained and evaluated using anchor box counts between 2 and 6. This empirical analysis is conducted to determine the optimal number of anchor boxes required to balance detection accuracy and computational efficiency. The IoU value improves significantly up to 6 anchor boxes (0.883), after which the gains become marginal (Figure 5). Based on this, 6 anchor boxes were selected as a balance between performance and model complexity for further training and evaluation.

Six anchor boxes were estimated and sorted in descending order based on their areas. Thus, larger anchor boxes were assigned to lower-dimensional feature map layers of the network, while smaller anchor boxes were assigned to higher-dimensional feature map layers. This approach enabled more effective detection of lung regions at different scales.

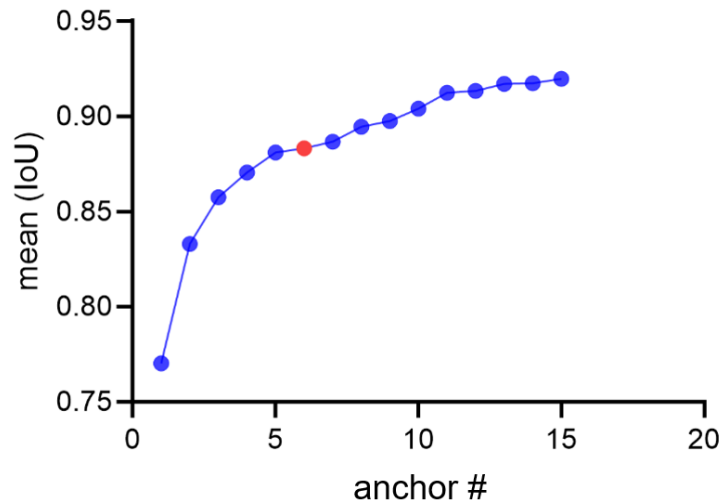


Figure 5: Mean IoU values obtained by using different numbers of anchor boxes.

2.4. Evaluation Metric (Average Precision)

For performance evaluation, the Average Precision (AP) metric was employed since the model detects a single class (lung_region). The AP was computed at an Intersection-over-Union (IoU) threshold of 0.5, corresponding to the conventional mAP@0.5 criterion used in object detection tasks. This threshold ensures that the predicted bounding boxes have at least 50% overlap with the ground-truth annotations, which is considered sufficient for reliable localization. Additionally, the mean IoU between predicted and annotated boxes was analyzed as a complementary measure to assess the spatial alignment quality. The relationship between the number of anchor boxes and the mean IoU was also examined to balance detection accuracy and computational efficiency.

Regardless of the interpolation strategy employed, the Average Precision (AP) is calculated separately for each object class. In datasets containing a large number of classes, it is often desirable to summarize detection performance with a single, representative metric.

3. RESULTS AND DISCUSSIONS

3.1. Results

The performance of the trained object detection model was evaluated using a dedicated test set consisting of 144 chest X-ray images. It should be noted that the test dataset consists of 144 images, which constitutes a relatively small sample. While the results demonstrate promising lung region detection performance, larger-scale evaluation on diverse datasets would be necessary to fully validate the generalizability of the method. The evaluation focused on assessing the model’s ability to accurately detect the lung_region class.

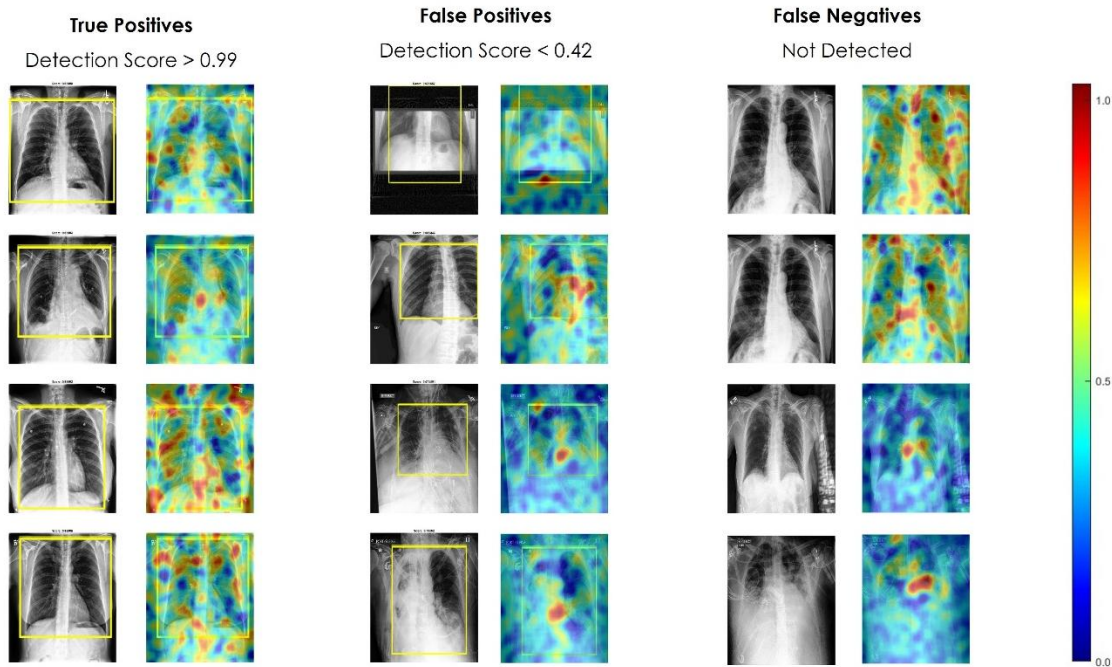


Figure 6: Representative detection outcomes from the test set with Grad-CAM visualizations using the D-RISE method (Petsiuk et al., 2021): True positive, false positive, and false negative cases are shown; detailed column descriptions are provided in the main text.

It should be clarified that the detected single class, lung_region, does not correspond to a pathology. The aim is to isolate the lung area in chest X-rays to reduce irrelevant information and improve model attention, rather than to detect diseases. Hence, the single-class detection serves as a preprocessing step for downstream tasks and does not limit the framework’s potential applicability to disease detection with additional data.

Figure 6 illustrates representative detection outcomes from the test set, complemented by Grad-CAM visualizations generated using the D-RISE method (Petsiuk et al., 2021). The first column shows true positive cases where the lung region is correctly detected by the model. The second column presents false positive cases in which non-lung areas were mistakenly identified as lung regions. The third column displays false negative cases, where actual lung regions were present but not detected by the model. This breakdown highlights both the strengths and limitations of the YOLOv4-based localization approach.

A low detection threshold of 0.01 was used during inference to ensure that a broad range of detections could be captured. This strategy facilitated the analysis of the model's precision over a wide spectrum of recall values, thereby enabling a more comprehensive performance characterization.

The evaluation was conducted based on the average precision (AP) metric, which integrates both the ability of the model to correctly classify detected objects (precision) and its ability to retrieve all relevant instances (recall). The computed AP value for the lung_region class was 0.9043, indicating that the detector achieved high precision while maintaining a substantial level of recall (Figure 7).

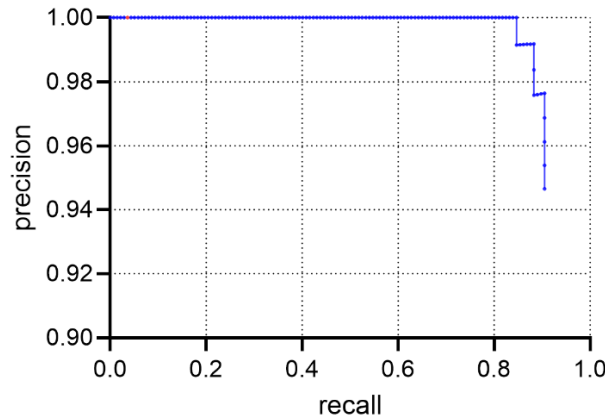


Figure 7: Precision–recall (PR) curve for the lung_region class obtained from the test dataset.

The corresponding precision-recall (PR) curve, shown in Figure 7, revealed that precision remained consistently at 1.0 throughout a large portion of the recall range, with a gradual decline observed only at higher recall levels. This behavior suggests that the model was able to detect most of the relevant lung regions without introducing a significant number of false positives.

Figure 6 presents representative detection outcomes from the test set, along with Grad-CAM visualizations generated using the D-RISE method (Petsiuk et al., 2021). In the true positive cases (left column), the YOLO v4 model achieved high detection scores (> 0.99) and the activation maps show strong spatial correspondence between high-intensity regions and the annotated lung fields, indicating that the network effectively focused on relevant anatomical structures. In contrast, false positives (middle column) primarily

occurred when non-lung anatomical regions such as the clavicle, shoulder, or mediastinal area were misclassified as lung regions. The associated Grad-CAM maps reveal attention leakage toward surrounding high-contrast structures, often outside the intended lung boundaries. False negatives (right column) illustrate instances where lung regions were present but not detected; although activation maps occasionally showed partial focus on the lungs, this attention was insufficient to trigger a detection output. These failure cases suggest potential improvements through enhanced anchor box tuning, inclusion of more challenging training examples, and augmentation strategies that increase robustness to anatomical variability and low-contrast imaging conditions.

Overall, the D-RISE visualizations provide valuable insight into the model's decision-making process, highlighting both strengths in accurate lung localization and limitations leading to missed or spurious detections.

In addition, a confusion matrix in Equation (1) was generated to provide further insight into the classification performance. As shown below, the model produced 124 true positives, 13 false positives, 7 false negatives, and 0 true negatives:

$$\begin{bmatrix} \text{TP} & \text{FN} \\ \text{FP} & \text{TN} \end{bmatrix} = \begin{bmatrix} 124 & 7 \\ 13 & 0 \end{bmatrix} \quad (1)$$

This matrix confirms the high sensitivity of the detector, although it also highlights a modest false positive rate. Nevertheless, given the high AP score and the favorable shape of the PR curve, it can be concluded that the model performs reliably in detecting lung regions in chest X-ray images.

Based on these results, the lung region was correctly detected in 124 out of the 144 images in the test group. In 7 cases, an area outside the actual lung region was incorrectly detected as a lung region, while in 13 chest X-ray images, the detector failed to identify a region despite the presence of an actual lung region. Figure 6 presents five representative examples each of true positives, false positives, and false negatives.

In comparison to previous works using YOLO-based object detection in chest imaging, our approach focuses specifically on bounding-box localization and cropping of the lung region rather than pathology classification. For instance, Detection Method of Chest X-ray Lung Nodules Based on Improved YOLOv3 Model(Yan, 2020) used an improved YOLOv3 architecture to detect lung nodules on chest X-rays and reported a detection accuracy of approximately 61% on nodules.

Meanwhile, Anomalies Detection in Chest X-Ray Images Using Faster R-CNN and YOLO (2025) applied YOLOv5s to identify five different thoracic abnormalities and reported AP@0.5 and AP@[0.5:0.95] values of roughly 0.616 and 0.322 respectively on 2,500 X-ray images (Nguyen et al., 2023).

By contrast, our method achieves a mean IoU of 0.883 and an AP of 0.9043 for the lung_region class in a 144-image test set. Although direct comparisons are limited due to differing tasks and datasets, the strong localization performance of our method suggests that bounding-box cropping of the lung region may provide a robust preprocessing step for downstream diagnostic models.

3.2. Discussion

Accurate localization of lung regions is a critical prerequisite for computer-aided diagnosis (CAD) systems targeting thoracic pathologies such as Atelectasis, Cardiomegaly, and Pleural Effusion. In this study, YOLO v4 was selected due to its proven balance between detection speed and accuracy, aligning with prior evaluations demonstrating its superior real-time performance even on moderate GPUs % (Bochkovskiy et al., 2020) and other benchmarks in medical contexts (Ragab et al., 2024).

The relatively small size of the test dataset may limit the generalizability of the model and increase the risk of overfitting. Future studies with larger and more diverse datasets are needed to validate and further improve the robustness of the proposed approach.

Future work could explore strategies to further enhance model robustness and generalizability. Approaches such as data augmentation, transfer learning, and domain adaptation may help address the limitations imposed by the relatively small dataset, reduce overfitting, and improve performance across diverse patient populations and imaging conditions. Implementing these techniques could facilitate the extension of the current framework to multi-class detection tasks and broader clinical applications.

The high AP score (0.9043) and the PR curve characteristics suggest that the proposed model can reliably detect lung regions in chest X-rays while maintaining low false-positive rates. The IoU-based anchor box estimation played a pivotal role in optimizing detection accuracy (Redmon & Farhadi, 2016), consistent with prior reports that IoU-based clustering outperforms Euclidean distance for object detection tasks in medical imaging (Redmon & Farhadi, 2016). Despite the encouraging results, some false positives and false negatives were observed. The false positives often occurred in regions adjacent to the diaphragm or mediastinum, potentially due to overlapping soft tissue densities. False negatives were mainly associated with atypical lung shapes or reduced contrast in diseased lungs, highlighting a limitation in model generalization to highly variable anatomical presentations.

IoU-based anchor box optimization helps the model focus on actual lung regions by excluding irrelevant areas such as image borders or markers. These non-lung regions can distract deep networks, so aligning anchors with lung fields reduces such noise and conceptually improves feature learning and downstream performance.

Future research should address these limitations by integrating contrast enhancement techniques, multi-view learning, and domain adaptation to handle variations across datasets. Additionally, incorporating explainability methods such as Grad-CAM could aid in model interpretability, thereby increasing clinician trust in AI-assisted diagnostic workflows (Han et al., 2022; Holzinger et al., 2020)

The Grad-CAM and D-RISE visualizations provide insights into the regions of the chest X-rays that the model focuses on during lung region detection, enhancing the clinical interpretability of the results. In this study, we plan to incorporate these visualizations in ablation experiments to systematically evaluate how excluding or modifying certain image regions affects model attention and performance. This approach will help demonstrate the reliability of the model's focus and support the interpretability of automated detection outcomes.

4. CONCLUSION

This study demonstrated that YOLO v4, combined with IoU-based anchor box optimization, provides an effective and computationally efficient solution for lung region detection in chest X-ray images. Using six optimized anchor boxes yielded a high AP score of 0.9043 and robust detection performance across the test set, confirming the suitability of this approach for clinical pre-processing tasks in computer-aided diagnosis (Bochkovskiy et al., 2020; Redmon & Farhadi, 2017).

The accurate identification of lung regions facilitates subsequent diagnostic tasks, including the detection and classification of Atelectasis and other thoracic pathologies. While some false positives and false negatives persist, these can be mitigated through enhanced preprocessing, increased dataset diversity, and the integration of advanced explainability techniques (Bochkovskiy et al., 2020; Holzinger et al., 2020; Redmon & Farhadi, 2016).

Given its strong performance and scalability, the proposed method holds significant potential for deployment in automated radiology pipelines, improving workflow efficiency and diagnostic consistency in clinical settings.

DATA AVAILABILITY

The MATLAB implementation of the YOLOv4-based lung region detection framework, including functions for annotation and cropping, is publicly available at:

- <https://www.mathworks.com/matlabcentral/fileexchange/182683-chestx-lungcrop>

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest among the authors.

REFERENCES

- Abut, S. (2024). AI-based model design for prediction of COPD grade from chest X-ray images: a model proposal (COPD-GradeNet). *Cukurova University Journal of the Faculty of Engineering*, 39(2), 325-338.
- Abut, S., & Okut, H. (2024). The Importance of Artificial Neural Networks in Decision Making for the Field of Medicine. In G. A. Indrajit, Mittal; Hemlata, Jain (Ed.), *The Future of Artificial Neural Networks* (pp. 1-24). Nova Science. <https://doi.org/10.52305/YUZX7201>
- Abut, S., Okut, H., & Kallail, K. J. (2024). Paradigm shift from Artificial Neural Networks (ANNs) to deep Convolutional Neural Networks (DCNNs) in the field of medical image processing. *Expert Systems with Applications*, 244, 122983.
- Acharya, A. K., & Satapathy, R. (2020). A deep learning based approach towards the automatic diagnosis of pneumonia from chest radio-graphs. *Biomedical and Pharmacology Journal*, 13(1), 449-455.
- Ait Nasser, A., & Akhloufi, M. A. (2023). A review of recent advances in deep learning models for chest disease detection using radiography. *Diagnostics*, 13(1), 159.
- Ausawalaithong, W., Thirach, A., Marukatat, S., & Wilaiprasitporn, T. (2018, 21-24 Nov. 2018). Automatic Lung Cancer Prediction from Chest X-ray Images Using the Deep Learning Approach. 2018 11th Biomedical Engineering International Conference (BMEiCON),
- Berikol, G. B., Kanbakan, A., Ilhan, B., & Doğanay, F. (2025). Mapping artificial intelligence models in emergency medicine: A scoping review on artificial intelligence performance in emergency care and education. *Turkish Journal of Emergency Medicine*, 25(2), 67-91.
- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- Çallı, E., Sogancioglu, E., Van Ginneken, B., van Leeuwen, K. G., & Murphy, K. (2021). Deep learning for chest X-ray analysis: A survey. *Medical image analysis*, 72, 102125.
- Ergen, B., & Abut, S. (2013). Gender recognition using facial images. *International Proceedings of Chemical, Biological & Environmental Engineering*, 60(22), 112-117. <https://doi.org/10.7763/IPCBE.2013.V60.22>
- Han, Y., Chen, C., Tewfik, A., Glicksberg, B., Ding, Y., Peng, Y., & Wang, Z. (2022). Knowledge-augmented contrastive learning for abnormality classification and localization in chest X-rays with radiomics using a feedback loop. *Proceedings of the IEEE/CVF winter conference on applications of computer vision*,
- Holzinger, A., Saranti, A., Molnar, C., Biecek, P., & Samek, W. (2020). Explainable AI methods-a brief overview. *International workshop on extending explainable AI beyond deep models and classifiers*,
- Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Illcus, S., Chute, C., Marklund, H., Haghighi, B., Ball, R., & Shpanskaya, K. (2019). Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *Proceedings of the AAAI conference on artificial intelligence*,
- Matsuyama, E. (2021). A novel method for automated lung region segmentation in chest X-ray images. *Journal of biomedical science and engineering*, 14(6), 288-299.
- Miah, M. A. I., Paul, S., Das, S., & Hashem, M. (2024). Inflocnet: Enhanced lung infection localization and disease detection from chest x-ray images using lightweight deep learning. *arXiv preprint arXiv:2408.06459*.
- Nguyen, H. T., Nguyen, M. N., Phung, L. D., & Pham, L. T. T. (2023). Anomalies Detection in Chest X-Rays Images Using Faster R-CNN and YOLO. *Vietnam Journal of Computer Science*, 10(04), 499-515. <https://doi.org/10.1142/s2196888823500094>
- Petsiuk, V., Jain, R., Manjunatha, V., Morariu, V. I., Mehra, A., Ordonez, V., & Saenko, K. (2021). Black-box explanation of object detectors via saliency maps. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*,
- Ragab, M., Jadid Abdulkadir, S., Muneer, A., Alqushaibi, A., Sumiea, E., Qureshi, R., Al-Selwi, S., & Alhussian, H. (2024). A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023). *IEEE Access*, PP, 1-1. <https://doi.org/10.1109/ACCESS.2024.3386826>

- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., & Shpanskaya, K. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- Redmon, J., & Farhadi, A. (2016). YOLO9000: Better, Faster, Stronger. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6517-6525.
- Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. *Proceedings of the IEEE conference on computer vision and pattern recognition*,
- Saporta, A., Gui, X., Agrawal, A., Pareek, A., Truong, S. Q. H., Nguyen, C. D. T., Ngo, V.-D., Seekins, J., Blankenberg, F. G., Ng, A. Y., Lungren, M. P., & Rajpurkar, P. (2022). Benchmarking saliency methods for chest X-ray interpretation. *Nature Machine Intelligence*, 4(10), 867-878. <https://doi.org/10.1038/s42256-022-00536-x>
- Yan, S. (2020). Detection Method of Chest X-ray Lung Nodules Based on Improved YOLOV3 Model. *Science Discovery*, 8(1), 18-23. <https://doi.org/10.11648/j.sd.20200801.15>
- Yilmaz, I., Alazab, H., Doganay, F., Dangott, B., Albadri, S., Nassar, A., . . . Salem, F. (2025). 1492 SegRenal: AI-Powered Segmentation and Quantification of Frozen Section Histopathology of Preimplantation Kidney Biopsies. *Laboratory Investigation*, 105(3). doi:10.1016/j.labinv.2024.103732
- Yilmaz, I., Alazab, H., Doganay, F., Dangott, B., Nassar, A., & Akkus, Z. (2025). 1398 AI-Powered Detection and Quantification of General-Purpose Nuclear Marker in Immunohistochemically Stained Tumor Tissues. *Laboratory Investigation*, 105(3).
- Yilmaz, I., Alazab, H. M., Doganay, F., Dangott, B., Albadri, S., Nassar, A., . . . Akkus, Z. SegRenal: AI-Driven Segmentation of Frozen Sections in Transplant Kidney Biopsies-A Comparative Analysis of Deep Learning Models. *Authorea Preprints*.

Comparative Evaluation of Regression Models for Building Energy Efficiency Assessment Based on Heating and Cooling Load Requirements

Sinan ATICI^{1*}, Gürkan TUNA²

¹Sechard Information Technologies, Istanbul, Türkiye

²Department of Computer Technologies, Vocational School of Technical Sciences, Trakya University, Edirne, Türkiye

Received: 15.10.2025, Accepted: 05.12.2025, Published: 20.12.2025

ABSTRACT

The accurate prediction of heating and cooling loads is a critical prerequisite for designing energy-efficient buildings and reducing their environmental footprint. This study presents a comprehensive comparative analysis of multiple regression models for estimating the energy efficiency of residential buildings based on their architectural parameters. Using the Energy Efficiency dataset, we evaluated the performance of seven distinct modelling approaches: Linear Regression, Decision Tree, Random Forest, Support Vector Regression with a Radial Basis Function kernel, K-Nearest Neighbours, Multi-Layer Perceptron, and Deep Neural Networks. Models were rigorously assessed using Root Mean Square Error, Mean Absolute Error, and the coefficient of determination (R^2). The results demonstrate that non-linear machine learning methods significantly outperform traditional linear models. Specifically, the Random Forest and Support Vector Regression models achieved superior predictive accuracy, with RMSE values as low as 0.46 for heating load and 1.53 for cooling load, and R^2 scores exceeding 0.97. Furthermore, feature importance analysis identified Overall Height and Relative Compactness as the most influential parameters for heating and cooling load predictions, respectively, providing actionable insights for architectural design. This research shows that advanced machine learning models, particularly Random Forest and Support Vector Regression, offer a robust and accurate framework for building energy assessment.

Keywords: Energy efficiency; Load prediction; Machine learning; Regression models; Sustainable building design

Isıtma ve Soğutma Yüğü Gereksinimlerine Dayalı Bina Enerji Verimliliğı Değerlendirmesi için Regresyon Modellerinin Karşılaştırmalı Değerlendirmesi

ÖZ

Isıtma ve soğutma yüklerinin doğru tahmini, enerji verimli binalar tasarlamak ve çevresel ayak izlerini azaltmak için kritik bir ön koşuldur. Bu çalışma, mimari parametrelerine dayalı olarak konut binalarının enerji verimliliğini tahmin etmek için çoklu regresyon modellerinin kapsamlı bir karşılaştırmalı analizini sunmaktadır. Enerji Verimliliğı veri setini kullanarak, yedi farklı modelleme yaklaşımının performansını değerlendirdik: Doğrusal Regresyon, Karar Ağacı, Rastgele Orman, Radyal Taban Fonksiyonu çekirdeğine sahip Destek Vektör Regresyonu, K-En Yakın Komşular, Çok Katmanlı Algılayıcı ve Derin Sinir Ağları. Modeller, Kök Ortalama Karesel Hata, Ortalama Mutlak Hata ve belirleme katsayısı (R^2) kullanılarak titizlikle değerlendirildi. Sonuçlar, doğrusal olmayan makine öğrenmesi yöntemlerinin

geleneksel doğrusal modellerden önemli ölçüde daha iyi performans gösterdiğini göstermektedir. Özellikle, Rastgele Orman ve Destek Vektör Regresyonu modelleri, ısıtma yükü için 0,46 ve soğutma yükü için 1,53 kadar düşük RMSE değerleri ve 0,97'yi aşan R^2 puanlarıyla üstün tahmin doğruluğu elde etti. Ayrıca, özellik önem analizi, ısıtma ve soğutma yükü tahminleri için sırasıyla Toplam Yükseklik ve Göreceli Kompaktlığı en etkili parametreler olarak belirleyerek mimari tasarım için uygulanabilir içgörüler sağlamıştır. Bu araştırma, özellikle Rastgele Orman ve Destek Vektör Regresyonu olmak üzere gelişmiş makine öğrenimi modellerinin, bina enerji değerlendirmesi için sağlam ve doğru bir çerçeve sunduğunu göstermiştir.

Anahtar Kelimeler: Enerji verimliliği; Yük tahmini; Makine öğrenimi; Regresyon modelleri; Sürdürülebilir bina tasarımı

1. INTRODUCTION

Energy efficiency policy remains the most influential driver in Europe. The European Green Deal and Renovation Wave Strategy (European Commission, 2023) aim to double annual energy renovation rates by 2030. Member States have implemented building codes and energy performance certificates under the EPBD. However, enforcement varies across regions (IEA, 2024). Financial incentives such as tax reductions and green loans have supported adoption, though split incentives and limited awareness remain challenges (Gou et al., 2022). The building sector accounts for approximately 40% of total energy consumption and 36% of greenhouse gas emissions in the European Union (European Commission, 2024). Recent directives, including the recast of the Energy Performance of Buildings Directive (EPBD), aim to accelerate renovation rates and promote nearly zero-energy buildings.

Retrofitting existing buildings offers the greatest opportunity for energy reduction. Integrated retrofits such as combining envelope upgrades, HVAC replacement, and renewable energy achieve up to 70% reductions in operational energy (Guerra Santin & Tweed, 2022). However, financing and tenant-owner split incentives remain barriers. Deep renovation packages are often limited to publicly funded projects or social housing. Envelope improvements are among the most cost-effective energy-saving strategies. Studies demonstrate that enhanced insulation, airtightness, and high-performance glazing can reduce heating energy use by 20–40% in cold climates (Ascione et al., 2023). Emerging materials such as phase change materials, aerogels, and dynamic façades offer additional benefits. For Mediterranean climates, adaptive shading and ventilated façades are particularly effective (Palermo et al., 2024).

Heating, ventilation, and air-conditioning (HVAC) systems represent the largest end-use in European buildings. Research on variable refrigerant flow systems, demand-controlled ventilation, and heat pumps shows potential energy savings of 30–60% (Bertoldi et al., 2023). Heat pump adoption is increasing due to electrification policies and the decarbonisation of electricity grids (IEA, 2024). Proper commissioning and system integration are crucial to realise expected performance gains.

Occupant behaviour significantly affects actual energy performance. Behavioural interventions, feedback systems, and smart meters can reduce household energy use by 5–15% (Andersen et al., 2022). Comfort preferences, cultural norms, and income levels influence the effectiveness of efficiency measures. User engagement is increasingly recognised as essential for closing the performance gap.

Estimating heating load and cooling load accurately is essential for building design, HVAC system sizing, energy efficiency, operational cost reduction, and meeting regulatory/green building standards. Traditional methods include physics-based simulation and empirical regression. More recently, machine learning methods have been applied to predict heating load/cooling load, both in design-stage and in operational

forecasting (Tsanas & Xifara, 2012a). The rise of large and synthetic datasets, improved computational resources, and demand for rapid, cost-efficient decision support have driven this trend. Building energy management systems, Internet of Things sensors, and artificial intelligence enable predictive maintenance and adaptive control. Studies indicate that artificial intelligence-based controls can reduce building energy use by 15–30% while maintaining occupant comfort (Zhou et al., 2024). Data interoperability, cybersecurity, and privacy are major challenges for large-scale deployment (D’Oca et al., 2023).

This study contributes to the existing literature on the Energy Efficiency dataset by conducting a comprehensive and systematic comparison of seven widely used regression algorithms, Linear Regression, Decision Tree, Random Forest, SVR with RBF kernel, K-Nearest Neighbours, Multi-Layer Perceptron, and Deep Neural Networks, evaluated simultaneously on both heating and cooling loads under identical experimental conditions. This analysis extends the original six-model benchmark of Tsanas and Xifara (2012a) through the inclusion of Deep Neural Networks and a broader performance assessment framework. Most importantly, the study provides a unified feature importance ranking derived from multiple models, clearly demonstrating that Orientation and Glazing Area Distribution exert negligible influence on both heating and cooling loads. These findings offer architects substantially clearer and more actionable engineering insights than those typically reported in prior work, emphasising greater design flexibility in façade orientation and glazing distribution without compromising energy performance. Supporting analyses, including detailed multicollinearity diagnostics and extensive visual exploration of the dataset, further strengthen the interpretability and practical value of the results.

The rest of this paper is as follows. First, it presents a literature review that provides a synthesis of recent studies (2020–2025) to evaluate the state of knowledge, identify effective strategies, and outline remaining challenges. Then, our approach to assess the energy efficiency of a given building using heating and cooling load requirements and machine learning methods is presented. Afterwards, the results of the simulation study are reported. Finally, the paper concludes with recommendations for future research on embodied carbon, post-occupancy monitoring, and equitable energy transitions.

2. RELATED WORK

Considering building energy efficiency, despite technological advances, multiple barriers remain. These include high upfront costs, limited skilled workforce, lack of performance monitoring, and fragmented regulations across EU member states (Kwak et al., 2023). The European Renovation Wave seeks to address these through funding, harmonised standards, and workforce training. Still, achieving the 2050 carbon neutrality target will require accelerated action. Future research in Europe should focus on integrating digitalisation with retrofit strategies, developing consistent embodied carbon databases, and

promoting equitable access to energy-efficient housing. Long-term monitoring and post-occupancy evaluation are needed to validate performance claims. Cross-disciplinary collaboration between engineers, architects, policymakers, and social scientists will be essential.

This literature review examines recent studies that focus on machine learning methods for heating load/cooling load prediction: what algorithms are used, what features, how models are validated, interpretability, and what gaps remain. Recent research shows significant advances in applying machine learning to predict building thermal loads. For example, Abdel-Jaber and Dirks (2024) provided a comprehensive review of data-driven prediction techniques for heating and cooling loads, highlighting the dominance of ensemble and artificial neural network models. Chaganti et al. (2022) demonstrated that ensemble models such as Random Forest and Gradient Boosting achieved R-Squared (R^2) values above 0.99 on the University of California, Irvine (UCI) Energy Efficiency dataset. Lu et al. (2023) introduced an AutoML framework that achieved R^2 scores above 0.98 while also incorporating interpretability features. Other studies, such as Miskolc et al. (2023), examined neural network optimisation for different building geometries and materials, showing that thermal properties of envelope layers significantly affect heating and cooling loads.

Ensemble tree-based methods such as Random Forest, Gradient Boosting, and Cubist consistently perform best for design datasets due to their ability to model non-linear relationships between geometry and thermal loads. Neural networks, including multilayer perceptrons and radial basis functions, achieve comparable results but require larger datasets and careful tuning. AutoML has emerged as a convenient alternative for practitioners. Interpretability remains limited, although SHapley Additive exPlanations (SHAP) and feature importance analyses are increasingly used.

Most studies rely on the UCI ENB2012 dataset (UCI, 2025) with 80/20 training-testing splits and cross-validation. Common evaluation metrics include R^2 , Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Reported accuracies often exceed 0.99 R^2 for heating load and 0.98 R^2 for cooling load. However, generalisation to real-world data remains a challenge due to synthetic nature of most datasets.

Despite excellent accuracy, machine learning models for heating load/cooling load prediction face key challenges: (1) limited generalisation to real buildings, (2) scarcity of high-quality, labelled datasets, (3) limited interpretability of complex models, and (4) lack of integration with physical/thermal simulation approaches. Hybrid models combining physics-based knowledge and machine learning show potential to improve transferability across climates and building types.

To ensure robust and useful machine learning models for heating and cooling load prediction, studies recommend: (1) using ensemble or AutoML methods, (2) integrating interpretability tools such as SHAP, (3) validating across climates and datasets, and (4) reporting multiple error metrics. Future work should emphasise real-building datasets and hybrid physics-informed machine learning.

3. METHODOLOGY

3.1. Dataset Description and Characteristics

This study utilises the Energy Efficiency dataset, which consists of 768 residential building configurations generated through controlled simulation experiments (Tsanas & Xifara, 2012a, 2012b). The dataset was created using 12 fundamental building shapes differing in relative compactness, aspect ratio, and height. For each of these 12 base shapes, the remaining design variables, orientation, glazing area, and glazing area distribution, were systematically varied across 18 combinations, yielding a total of 768 unique instances.

All energy performance values were computed using the building energy simulation software Ecotect (version 2011), which served as the ground truth for both heating load (Y_1) and cooling load (Y_2) under standardised operating conditions and assumptions (Tsanas & Xifara, 2012a). The eight input features were specifically chosen to capture the most influential architectural parameters affecting thermal performance. The input features are defined as follows:

- X_1 : Relative Compactness
- X_2 : Surface Area
- X_3 : Wall Area
- X_4 : Roof Area
- X_5 : Overall Height
- X_6 : Orientation
- X_7 : Glazing Area
- X_8 : Glazing Area Distribution

The primary objective of this study is to predict the two continuous target variables, heating load (Y_1) and cooling load (Y_2), using the eight input features described above, following the experimental framework originally established by Tsanas and Xifara (2012a). The dataset is fully deterministic and perfectly balanced, comprising exactly 768 building simulations. According to the original authors (Tsanas & Xifara, 2012a), the 768 instances were generated in two carefully designed parts:

- **720 buildings with glazing:** 12 distinct building geometries $\times$ 3 non-zero glazing area ratios (10%, 25%, 40%) $\times$ 5 glazing area distributions (0 = uniform, 1–5 = non-uniform) $\times$ 4 orientations $\rightarrow 12 \times 3 \times 5 \times 4 = 720$
- **48 buildings without glazing (glazing area = 0%):** 12 building geometries $\times$ 4 orientations $\rightarrow 12 \times 4 = 48$
- **Total:** $720 + 48 = 768$ unique configurations.

This deliberate construction ensures perfect balance across all design variables while explicitly including the physically realistic case of fully opaque façades (0% glazing). Each of the 768 rows represents a unique combination with no duplicates, no missing cells, and no random sampling bias. Although the target variables, heating load Y_1 and cooling load Y_2 , appear continuous, they take only 587 and 636 distinct values respectively, due to the discrete resolution of the Ecotect simulation engine. Descriptive statistics of the eight input features and two target variables are presented in Table 1.

Table 1: Descriptive statistics of the input features and target variables

Feature	Count	Mean $\pm$ Std	Min	25%	Median	75%	Max	Unique
Relative Compactness (X1)	768	0.764 ± 0.106	0.620	0.682	0.750	0.830	0.980	12
Surface Area (X2) (m^2)	768	671.71 ± 88.09	514.50	606.38	673.75	741.12	808.50	12
Wall Area (X3) (m^2)	768	318.50 ± 43.63	245.00	294.00	318.50	343.00	416.50	7
Roof Area (X4) (m^2)	768	176.60 ± 45.17	110.25	140.88	183.75	220.50	245.00	12
Overall Height (X5) (m)	768	5.25 ± 1.75	3.50	3.50	5.25	7.00	7.00	2
Orientation (X6)	768	3.50 ± 1.12	2	2.75	3.50	4.25	5	4
Glazing Area (X7) (fraction)	768	0.234 ± 0.133	0	0.10	0.25	0.40	0.40	4
Glazing Area Distribution (X8)	768	2.81 ± 1.55	0	1.75	3.00	4.00	5	6
Heating Load (Y1) (kW)	768	22.31 ± 10.09	6.01	12.99	18.95	31.67	43.10	587
Cooling Load (Y2) (kW)	768	24.59 ± 9.51	10.90	15.62	22.08	33.13	48.03	636

3.1.1. Multicollinearity Analysis

A critical characteristic of the dataset is the presence of perfect collinearity among three geometric features due to the fixed rectangular prism geometry assumed in the simulations. Specifically, the following exact linear relationship holds for all 768 instances: *Surface Area* = *Wall Area* + $2 \times$ *Roof Area*.

This physical constraint renders the design matrix singular when all three variables are included simultaneously. Consequently, Variance Inflation Factor (VIF) values for Surface Area, Wall Area, and Roof Area are infinite (Table 2).

Table 2: Descriptive statistics of the input features and target variables Variance Inflation Factors (VIF) for numeric features (standardised prior to computation). Infinite values arise from exact linear dependence.

Feature	VIF
Surface Area	∞
Wall Area	∞
Roof Area	∞
Relative Compactness	73.06
Overall Height	31.21
Glazing Area Distribution	1.05
Glazing Area	1.05

This perfect collinearity is not a data quality issue but an inherent property of the simulation model stemming from the fixed building geometry. High but finite VIF values for Relative Compactness and Overall Height are also expected due to their strong physical relationship. For linear models sensitive to multicollinearity, e.g., ordinary least squares regression, we additionally report performance using a reduced feature set comprising only Relative Compactness, Overall Height, Glazing Area, Glazing Area Distribution, and Orientation (all VIF < 10).

The Pearson correlation matrix is presented in Figure 1. It clearly confirms that Overall Height ($r = 0.89$) and Relative Compactness ($r = 0.62$) are the strongest predictors of heating load (Y_1), while Glazing Area exhibits an extremely strong correlation with cooling load (Y_2) ($r = 0.98$), which is fully consistent with established building physics principles. Additionally, the near-perfect correlations among Surface Area,

Wall Area, Roof Area, and Overall Height reflect the inherent geometric constraints and fully support the observed perfect multicollinearity discussed earlier.

Figure 1: Pearson Correlation Matrix

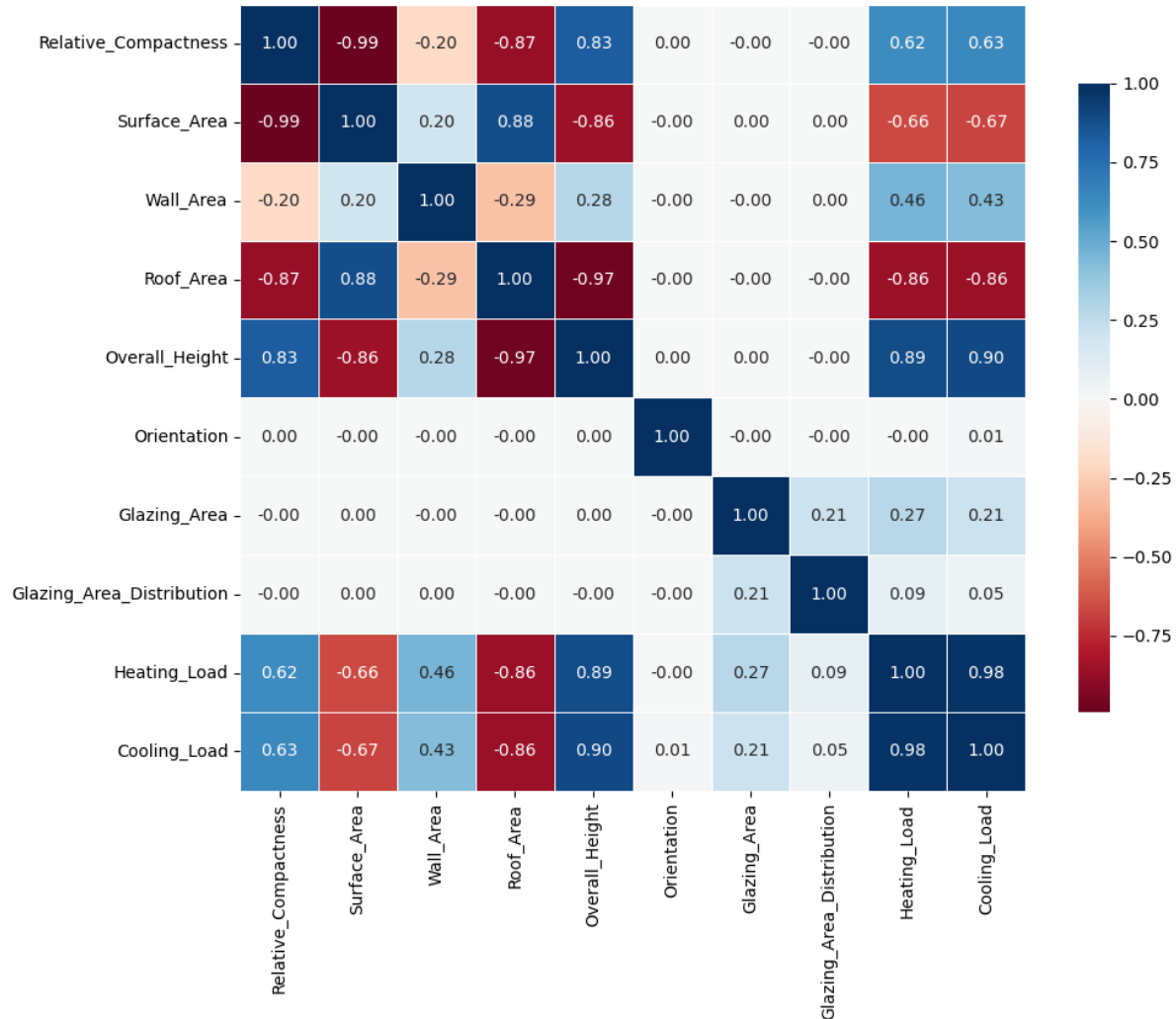


Figure 1: Pearson correlation matrix of all features and targets. Strong positive/negative correlations among geometric variables are clearly visible due to physical constraints.

This comprehensive characterization responds to concerns raised regarding dataset size and potential overfitting. Although the sample size ($n = 768$) may appear modest compared to contemporary large-scale datasets, the perfectly balanced, noise-free, full-factorial experimental design, with deliberate inclusion of the zero-glazing baseline, provides high statistical efficiency and ensures strong representativeness and reliable generalization within the defined architectural design space.

To further illustrate the characteristics of the Energy Efficiency dataset and support the findings from the correlation analysis, selected visualizations are presented in Figure 2. These plots highlight the distributional properties of the target variables, the dominant influence of key design parameters, and the negligible effect of building orientation, all of which are consistent with building physics expectations.

Figure 2: Energy Efficiency Dataset - Selected Visualizations

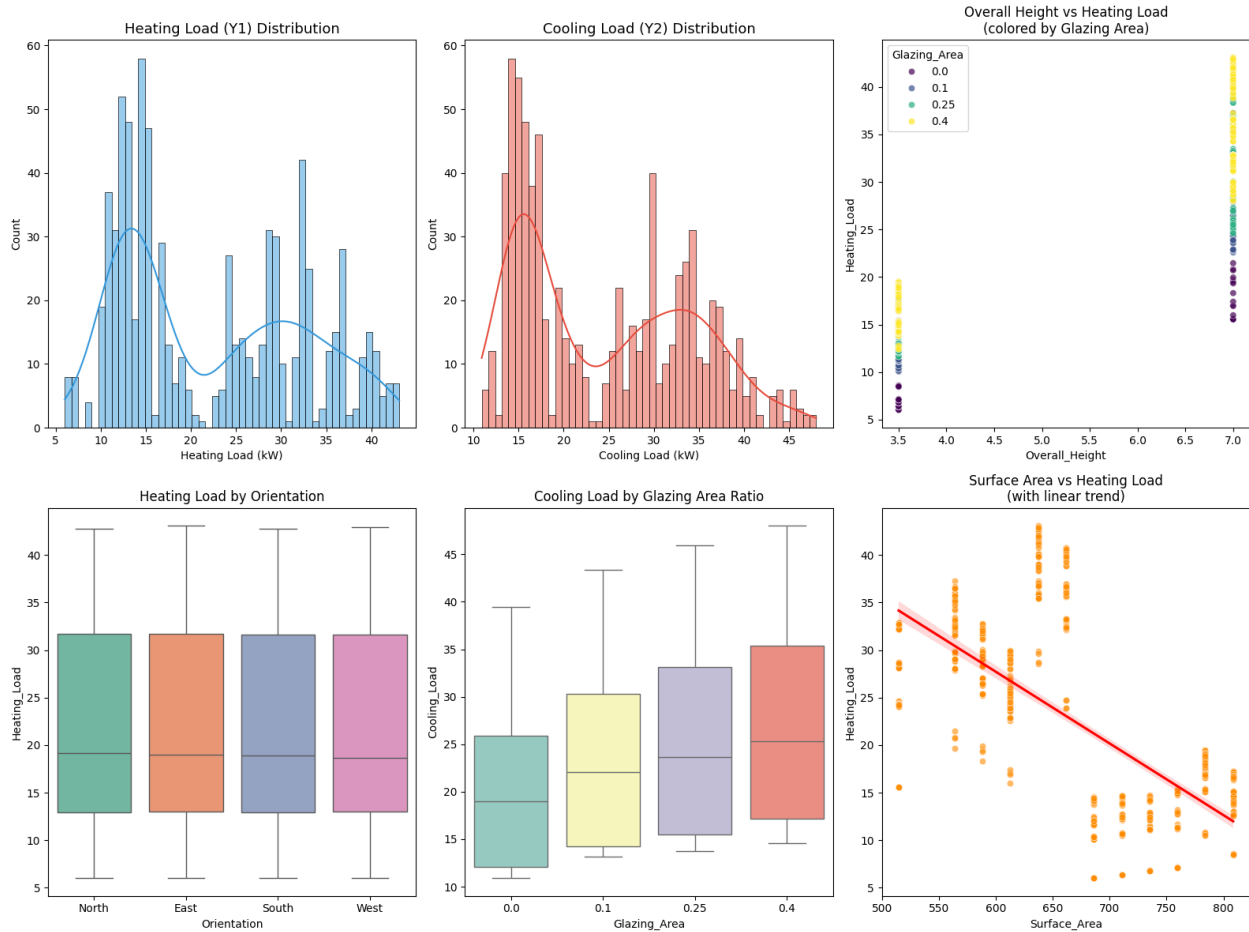


Figure 2: Energy Efficiency Dataset – Selected Visualizations: (a) Heating Load (Y_1) distribution with overlaid kernel density estimate, (b) Cooling Load (Y_2) distribution, (c) Overall Height vs. Heating Load (coloured by Glazing Area ratio), (d) Heating Load by Orientation, (e) Cooling Load by Glazing Area ratio, (f) Surface Area vs. Heating Load with linear trend line.

Figure 2 provides several important insights that reinforce the earlier statistical analyses. The distributions of both heating and cooling loads exhibit clear multimodality (a, b), reflecting the discrete combinations of the 12 underlying building geometries and two height levels (3.5 m and 7.0 m). The strong linear relationship between Overall Height and Heating Load (c) and the inverse relationship between Surface Area and Heating Load (f) are immediately apparent, while the boxplot of Heating Load by Orientation (d) confirms its negligible impact (nearly identical medians and distributions across North, East, South,

and West). In contrast, Cooling Load increases markedly and systematically with Glazing Area ratio (e), further explaining the very high correlation ($r = 0.98$) observed in Figure 1. Collectively, these visualizations, combined with the perfectly balanced experimental design and absence of measurement noise, help address concerns regarding the modest sample size ($n = 768$) and support the reliability and interpretability of the subsequent modelling results.

3.2. Mathematical formulations

3.2.1. Linear regression

It models the relationship between features x and target y as given by Eq.(1):

$$y = w^T x + b + \epsilon \quad (1)$$

where w is the weight vector, b is the bias, and ϵ is the error term. The objective is to minimise the mean squared error (Eq.(2)):

$$\min_{w,b} \sum_{i=1}^n (y_i - (w^T x_i + b))^2 \quad (2)$$

3.2.2. Decision tree regressor

It partitions the feature space into regions based on feature thresholds and predicts the mean target value in each region. The tree is built by minimising (Eq.(3)):

$$\sum_{i \in R_m} (y_i - \bar{y}_m)^2 \quad (3)$$

where R_m is the m -th region and $\bar{y}_m$ is the mean target in that region.

3.2.3. Random forest regressor

It is an ensemble of decision trees, where each tree is trained on a random subset of data and features. The prediction is the average (Eq.(4)):

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(x) \quad (4)$$

where $h_t(x)$ is the prediction of the t -th tree and T is the number of trees.

3.2.4. Support vector regression (SVR)

It fits a function $f(x) = w^T \phi(x) + b$ within an ϵ -insensitive tube as given by Eq.(5):

$$\min_{w,b} \frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^n \max(0, |y_i - f(x_i)| - \epsilon) \quad (5)$$

where $\phi(x)$ maps features to a higher-dimensional space and C controls the trade-off.

3.2.5. K-nearest neighbours regressor

Predicts the target as the average of the k nearest neighbours' targets as given by Eq.(6):

$$\hat{y} = \frac{1}{k} \sum_{i \in N_k(x)} y_i \quad (6)$$

where $N_k(x)$ is the set of k nearest neighbours based on a distance metric (e.g., Euclidean).

3.2.6. Multilayer perceptron (MLP) regressor

It is a feedforward neural network with multiple layers can be formulated by Eq.(7):

$$\hat{y} = f_L(W_L f_{L-1}(W_{L-1} \dots f_1(W_1 x + b_1)) + b_L) \quad (7)$$

where f_L is the activation function, W_L and b_L are weights and biases, and the loss is typically (Eq.(8)):

$$\min \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (8)$$

4. RESULTS

This section presents a comprehensive evaluation of multiple predictive models for estimating heating and cooling loads in building energy efficiency analysis. The models assessed include Linear Regression, Decision Tree, Random Forest, Support Vector Regression (SVR) with Radial Basis Function (RBF) kernel, K-Nearest Neighbours (K-NN) with Euclidean distance, Multi-Layer Perceptron (MLP), and Deep Neural Network (DNN). Performance metrics, including RMSE, MAE, coefficient of determination (R^2), and Cross-Validation RMSE (CV-RMSE), are reported for both heating load (Y_1) and cooling load (Y_2). Additionally, feature importance rankings are provided to elucidate the influence of architectural parameters on energy efficiency predictions.

4.1. Model performance

The performance of the models is summarised in Table 3, which reports Root Mean Square Error (RMSE), Mean Absolute Error (MAE), coefficient of determination (R^2), and 5-fold cross-validated RMSE (CV-RMSE) for both heating and cooling load predictions. All models were trained on an 80/20 train–test split (random_state = 42). Hyperparameters were systematically optimised using GridSearchCV (5-fold CV) for traditional models and RandomizedSearchCV/custom validation for the neural networks. Only the single best-performing configuration of each algorithm was retained for final reporting. The selected hyperparameters are detailed in Table 4.

Table 3: Model performance comparison

Model	Target	RMSE	MAE	R^2	CV-RMSE
Linear Regression	Heating	3.03	2.18	0.91	3.20
	Cooling	3.15	2.20	0.89	3.29
Decision Tree	Heating	0.57	0.40	0.99	1.70
	Cooling	1.92	1.21	0.96	2.10
Random Forest	Heating	0.50	0.36	0.99	1.70
	Cooling	1.78	1.10	0.97	1.89
SVR (RBF)	Heating	0.46	0.35	0.99	1.89
	Cooling	1.53	0.91	0.97	2.23
K-NN (Euclidean)	Heating	2.33	1.50	0.95	3.02
	Cooling	2.63	1.72	0.93	2.74
MLP	Heating	0.61	0.46	0.99	2.00
	Cooling	1.31	0.91	0.98	2.37
DNN	Heating	1.11	0.88	0.99	1.11
	Cooling	1.64	1.22	0.97	1.64

The comparative analysis of model performance, as delineated in **Table 3** and visually corroborated in **Figures 4-9**, reveals a clear hierarchy in predictive capability for building energy load estimation. While linear regression provides a competent baseline (**Figure 3**), explaining over 89% of the variance in both heating and cooling loads ($R^2 > 0.89$), its relatively higher RMSE and MAE values indicate a significant margin for improvement. Non-linear ensemble and neural network methods demonstrably surpass this baseline, with tree-based models such as Random Forest (**Figure 5**) and support vector regression with a radial basis function kernel (**Figure 6**) achieving superior performance, as evidenced by their substantially lower RMSE (e.g., 0.50 for Heating Load) and near-perfect R^2 scores (≥ 0.99). The Multi-Layer Perceptron (**Figure 8**) also exhibits strong performance, though its higher cross-validation RMSE suggests slightly less generalisability compared to the best tree-based models. A notable observation is the performance of the Deep Neural Network (**Figure 9**), which, despite a strong R^2 , shows a marked discrepancy between its test RMSE and that of the most accurate models, potentially indicating underfitting or suboptimal hyperparameter tuning. Conversely, the simplistic K-Nearest Neighbours

model (Figure 7) proved to be the least effective of the non-linear approaches, as visually apparent from the greater dispersion in its prediction plots. The consistent diagonal alignment of data points across all high-performing models (Figures 4-6, 8) confirms their robust predictive capability across the entire range of energy load values.

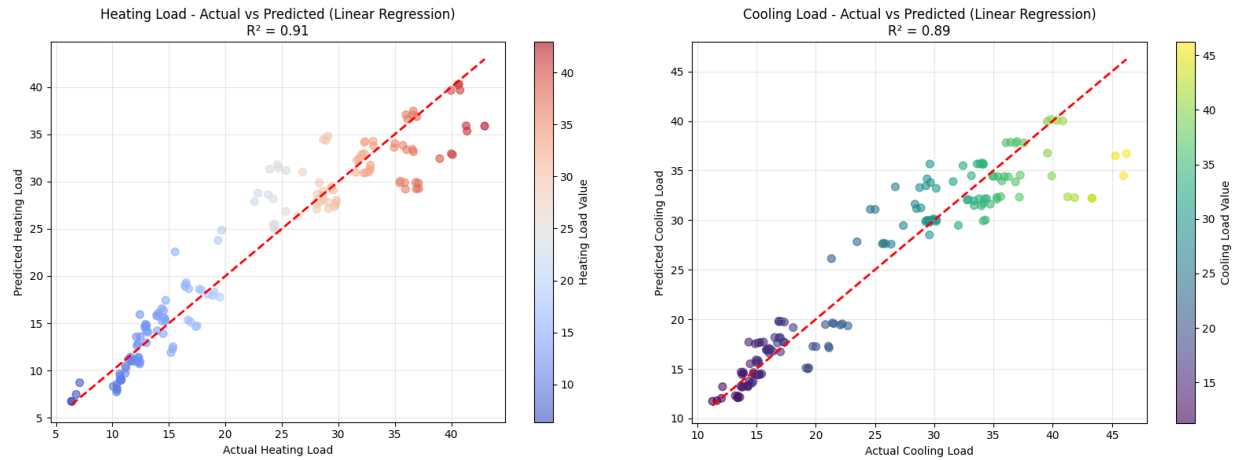


Figure 3: Actual versus predicted heating and cooling loads using linear regression model demonstrating strong linear relationship ($R^2 = 0.91$ for heating, $R^2 = 0.89$ for cooling).

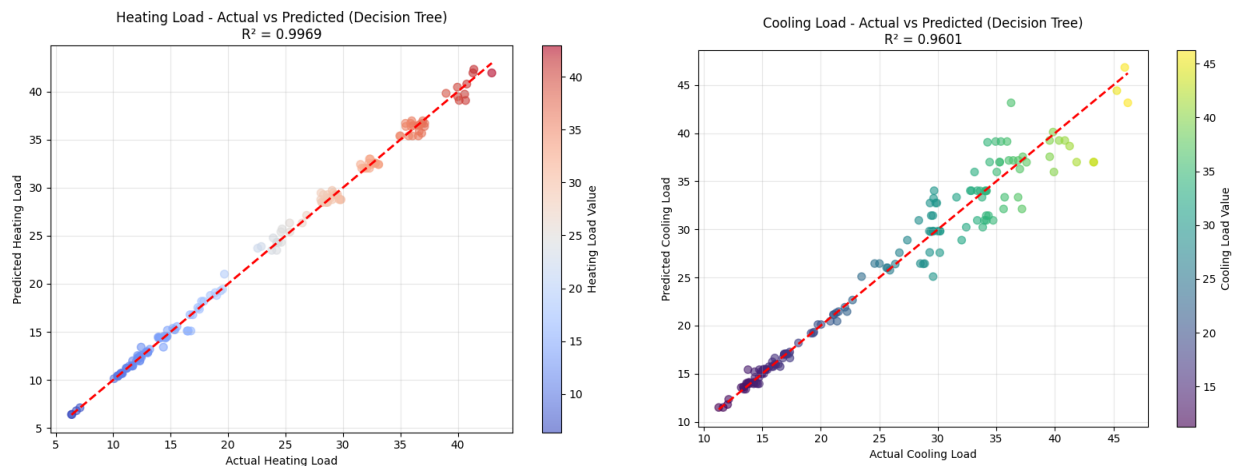


Figure 4: Decision tree regression performance showing near-perfect fit for heating load ($R^2 = 0.99$) and cooling load ($R^2 = 0.96$) with minimal prediction error.

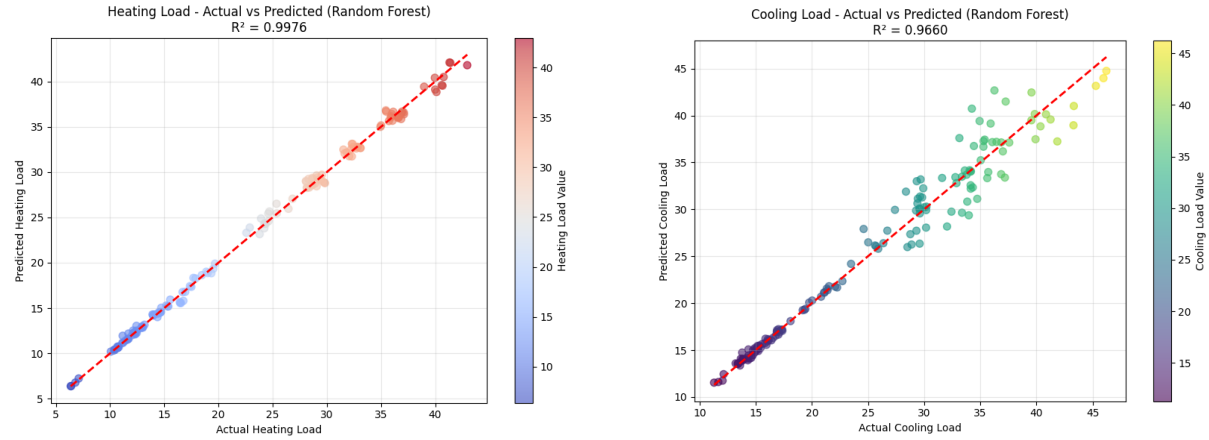


Figure 5: Random forest ensemble predictions illustrating enhanced accuracy and robust performance for both heating and cooling load estimation ($R^2 = 0.99/0.97$).

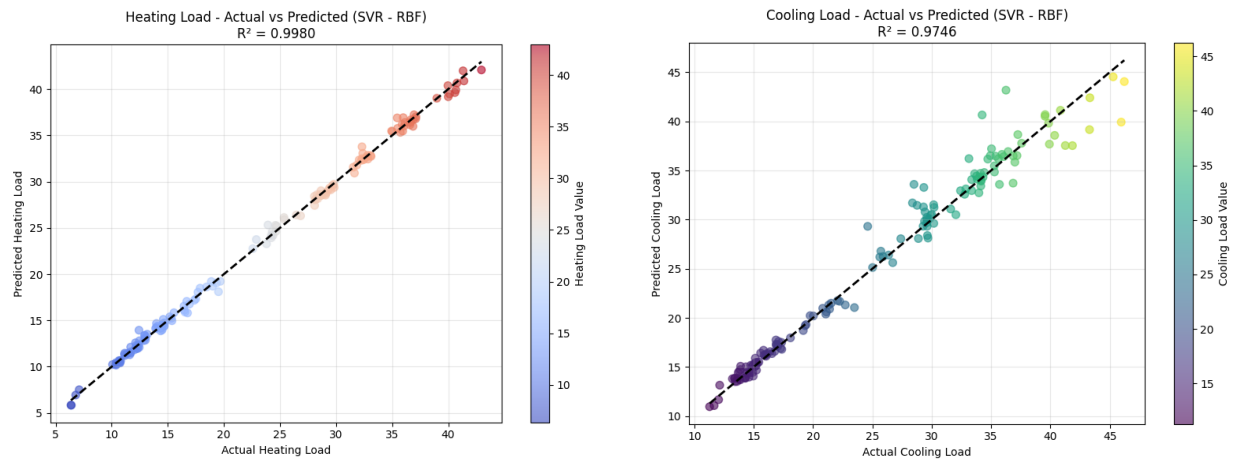


Figure 6: Support vector regression with RBF kernel capturing complex non-linear patterns in building energy load data ($R^2 = 0.99/0.97$).

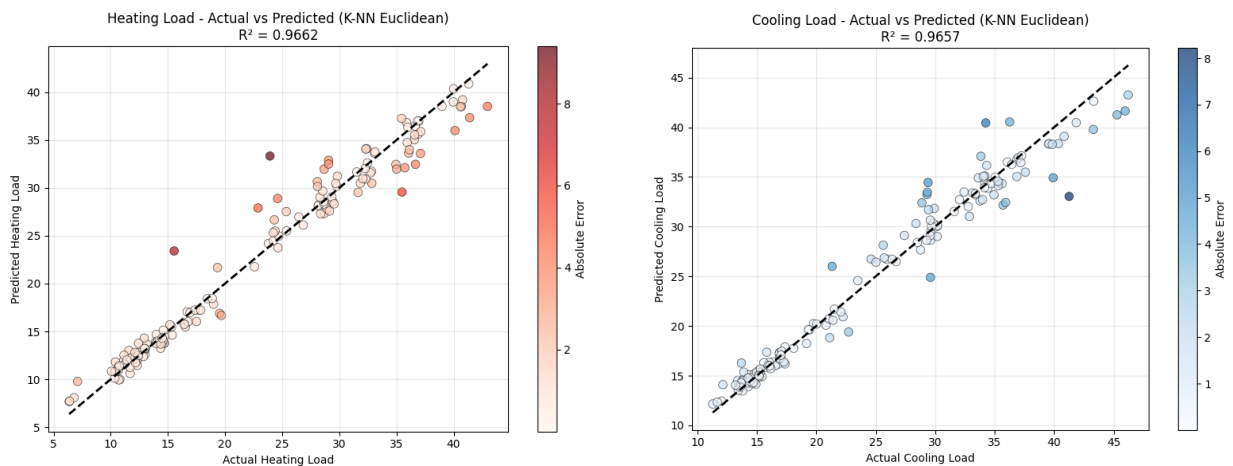


Figure 7: K-nearest neighbours regression performance demonstrating moderate predictive capability with visible variance in cooling load predictions ($R^2 = 0.95/0.93$).

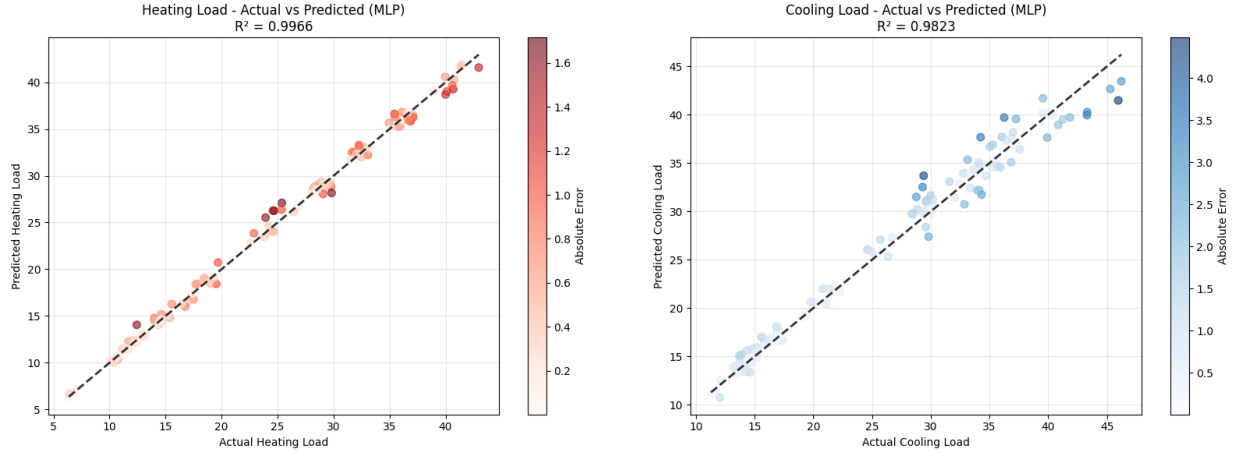


Figure 8: Multi-layer perceptron predictions showcasing strong non-linear modelling capability with high accuracy for both energy loads ($R^2 = 0.99/0.98$).

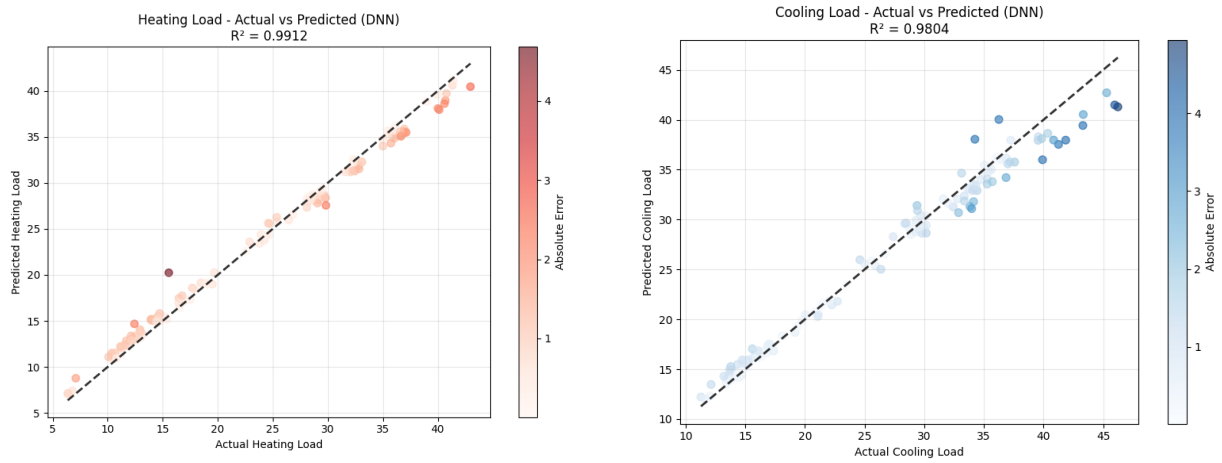


Figure 9: Deep neural network performance displaying competent predictive power with slight underfitting tendencies compared to ensemble methods ($R^2 = 0.99/0.97$).

To ensure reproducibility and optimal performance, a systematic hyperparameter optimisation procedure was applied to all models. An 80/20 train-test split was performed once (random_state = 42) and kept identical across all experiments. For Linear Regression, implemented as Ridge, Decision Tree, Random Forest, SVR with RBF kernel, and K-Nearest Neighbours, an exhaustive GridSearchCV with 5-fold cross-validation was employed on the training set. For the Multi-Layer Perceptron (scikit-learn) and the Deep Neural Network (custom PyTorch implementation), a combination of grid/random search and validation-based early stopping was used to prevent overfitting while maximizing predictive accuracy. In every case, only the single best-performing configuration, selected according to negative mean squared error, was retained for final evaluation on the fixed hold-out test set and for reporting cross-validated metrics. Feature scaling via StandardScaler was uniformly applied to K-NN, SVR, MLP, and DNN models, as these algorithms are sensitive to feature magnitude.

Table 4: Final optimised hyperparameters of the evaluated models

Model	Selected Hyperparameters / Architecture	Selection / Training Details
Linear Regression	$\alpha = 1.0$	GridSearchCV (5-fold CV)
Decision Tree	max_depth = None, min_samples_split = 2, min_samples_leaf = 1	GridSearchCV (5-fold CV)
Random Forest	n_estimators = 200, max_depth = None, min_samples_split = 2, min_samples_leaf = 1, max_features = 'sqrt', random_state = 42	GridSearchCV (5-fold CV)
SVR (RBF)	kernel = 'rbf', C = 100, γ = 'scale', ϵ = 0.1	GridSearchCV (5-fold CV)
K-NN (Euclidean)	n_neighbors = 7, weights = 'uniform', metric = 'euclidean'	GridSearchCV (5-fold CV)
MLP	hidden_layer_sizes = (100, 50), activation = 'relu', solver = 'adam', α = 0.0001, learning_rate = 'adaptive', max_iter = 1000, random_state = 42	Fixed architecture with early stopping disabled
DNN	Architecture: 8 $\rightarrow$ 128 $\rightarrow$ 64 $\rightarrow$ 32 $\rightarrow$ 1 (ReLU + BatchNorm + Dropout 0.2) Optimizer: Adam (lr = 0.001), weight_decay = 10^{-5} , batch_size = 32 Early stopping (patience = 50), maximum 500 epochs	Custom training with validation-based monitoring

Common experimental settings (applied uniformly to all models)

- Train–test split: 80% training, 20% test (random_state = 42)
- Feature scaling: StandardScaler applied for K-NN, SVR, MLP, and DNN
- Reproducibility: Fixed random_state = 42
- Final model selection: Best configuration chosen by negative mean squared error on the training set

The hyperparameters and architectural choices summarized in Table 4 reflect the outcome of the optimisation process described above. This transparent reporting eliminates any ambiguity regarding model configuration and enables exact replication of the reported results. The consistency of the experimental protocol (identical data split, random seed, and preprocessing) combined with rigorous hyperparameter tuning further strengthens the validity of the comparative analysis presented in Table 3.

4.2. Feature importance analysis

Feature importance rankings, derived from the models, highlight the relative influence of architectural parameters on heating and cooling load predictions. **Table 5** provides a simplified overview of the top three features for selected models (Linear Regression, Random Forest, and SVR (RBF)) for both heating and cooling loads. For heating load, Overall Height and Relative Compactness consistently emerged as the most influential features across models, with Linear Regression assigning the highest importance to Overall Height (7.22) and Random Forest to Relative Compactness (0.40). For cooling load, Relative

Compactness was the dominant feature in Linear Regression (7.49) and SVR (RBF) (8.81), while Random Forest prioritised Overall Height (0.39). Orientation and Glazing Area Distribution consistently showed minimal influence across all models, indicating greater design flexibility in these parameters.

Table 5: Top three feature importance rankings for selected models

Load	Model	Feature	Importance
Heating Load (Y_1)	Linear Regression	Overall Height	7.22
		Relative Compactness	6.52
		Roof Area	3.92
	Random Forest	Relative Compactness	0.40
		Surface Area	0.21
		Overall Height	0.14
	SVR (RBF)	Relative Compactness	4.95
		Wall Area	1.72
		Surface Area	0.78
Cooling Load (Y_2)	Linear Regression	Relative Compactness	7.49
		Overall Height	7.08
		Surface Area	4.14
	Random Forest	Overall Height	0.39
		Relative Compactness	0.29
		Surface Area	0.11
	SVR (RBF)	Relative Compactness	8.81
		Wall Area	2.19
		Surface Area	1.65

5. DISCUSSION AND CONCLUSION

The results reveal a clear performance hierarchy among the seven regression models evaluated on the Energy Efficiency dataset. Random Forest and Support Vector Regression with RBF kernel consistently delivered the highest predictive accuracy, achieving RMSE values as low as 0.46–0.50 kW for heating load and 1.53–1.78 kW for cooling load, with R^2 scores exceeding 0.99 and 0.97, respectively. This superior performance can be attributed to their robust ability to model complex non-linear relationships and interaction effects inherent in building physics. Random Forest effectively mitigates overfitting through bootstrap aggregation and feature randomness, enabling stable generalisation even within the constrained, deterministic design space of 768 samples. Similarly, SVR constructs a flexible, high-dimensional decision boundary via the RBF kernel, making it particularly well-suited to capture the sharp transitions induced by discrete geometric and glazing parameters.

In contrast, Linear Regression and K-Nearest Neighbours exhibited the poorest performance. The fundamentally limited expressive power of linear models is severely compromised by the exact collinearity among Surface Area, Wall Area, and Roof Area ($\text{Surface Area} = \text{Wall Area} + 2 \times \text{Roof Area}$), which produces infinite Variance Inflation Factors and an ill-conditioned design matrix. Although ridge regularisation was applied, it cannot fully compensate for perfect multicollinearity. K-NN, despite its non-

parametric nature, suffers from the curse of dimensionality and the high correlation structure of the features; local averaging becomes unreliable when nearest neighbours lie in regions distorted by collinear geometric variables, resulting in substantially higher prediction errors and poor cross-validation stability.

The Decision Tree and single-hidden-layer MLP occupy an intermediate position, demonstrating reasonable accuracy but remaining clearly outperformed by ensemble and kernel-based methods. The Deep Neural Network, while theoretically capable of universal approximation, showed signs of slight underfitting in this specific setting, likely due to the relatively small sample size and the absence of measurement noise, conditions under which simpler, well-regularised non-linear models often prevail.

A key finding with direct practical relevance is the consistent ranking of feature importance across models. Overall Height and Relative Compactness emerged as the dominant drivers of heating load, whereas Glazing Area overwhelmingly governs cooling load (Pearson $r = 0.98$). Crucially, Orientation and Glazing Area Distribution were ranked lowest by all models, with near-zero contribution in both Random Forest permutation importance and SVR coefficient analysis. This robust, multi-model consensus provides architects with clear and actionable design guidance: façade orientation and the spatial distribution of glazing can be treated as flexible aesthetic parameters without meaningful energy penalty, provided total glazing area, building height, and compactness are appropriately optimised.

Although the present models were developed and validated on the widely used synthetic Energy Efficiency dataset (Tsanas & Xifara, 2012), which remains a standard benchmark in the field, we acknowledge that direct extrapolation to real-world buildings, where measurement noise, occupant behaviour, and climatic variability are present, requires additional validation on empirical datasets. Nevertheless, the physically interpretable patterns uncovered, e.g., dominance of height, compactness, and glazing area, align closely with established building science principles and are therefore expected to retain substantial validity across diverse residential contexts.

In conclusion, Random Forest and SVR (RBF) are recommended as the preferred algorithms for rapid and accurate heating and cooling load prediction during early-stage building design. The identified key drivers and negligible parameters offer practical, evidence-based guidance that can immediately inform architectural decision-making and support the development of user-friendly, data-driven energy assessment tools.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest among the authors.

ACKNOWLEDGEMENT

Note: The consequences/responsibilities of unwritten/unmentioned fundings/supports belong entirely to the authors. The Kırklareli University Journal of Engineering and Science does not accept any responsibility in this issue (regard).

CONTRIBUTIONS OF AUTHORS

S.A.: Methodology, simulation design, analysis of results, writing—original draft preparation.

G.T.: Methodology, validation, discussion of results, writing—review and editing.

REFERENCES

- Abdel-Jaber, F., & Dirks, K. N. (2024). A Review of Cooling and Heating Loads Predictions of Residential Buildings Using Data-Driven Techniques. *Buildings*, 14(3), 752. <https://doi.org/10.3390/buildings14030752>
- Andersen, R., Fabi, V., & Corgnati, S. P. (2022). Human behavior and building performance: Understanding occupant-driven energy use. *Energy and Buildings*, 259, 111931.
- Ascione, F., Bianco, N., Mauro, G. M., & Vanoli, G. P. (2023). Thermal envelope optimization for European climates. *Renewable Energy*, 214, 1363–1378.
- Bertoldi, P., Mosconi, R., & Serrenho, T. (2023). Efficient heat pump systems for decarbonizing EU buildings. *Energy Policy*, 182, 113976.
- Chaganti, R., Rustam, F., Daghriri, T., Díez, I. d. l. T., Mazón, J. L. V., Rodríguez, C. L., & Ashraf, I. (2022). Building Heating and Cooling Load Prediction Using Ensemble Machine Learning Model. *Sensors*, 22(19), 7692. <https://doi.org/10.3390/s22197692>
- D'Oca, S., Hong, T., & Langevin, J. (2023). Smart building technologies for energy management in Europe. *Building and Environment*, 233, 110161.
- European Commission. (2023). *Renovation Wave for Europe: Greening our buildings, creating jobs, improving lives*. Brussels: EC Publications.
- European Commission. (2024). *Energy Performance of Buildings Directive (Recast)*. Brussels: EC Publications.
- Gou, S., Zhao, J., & Pan, W. (2022). Evaluating energy policy impacts on building retrofitting. *Energy Research & Social Science*, 87, 102452.
- Guerra Santin, O., & Tweed, C. (2022). Deep renovation strategies in European housing. *Energy and Buildings*, 259, 111912.
- IEA. (2024). *Energy Efficiency 2024: Global Progress and European Insights*. Paris: International Energy Agency.
- Kwak, S. Y., Bertoldi, P., & Economidou, M. (2023). Barriers and drivers for energy renovation in the EU. *Energy Policy*, 182, 113985.
- Lu, C., Li, S., Penaka, S., & Olofsson, T. (2023). Automated Machine Learning-Based Framework of Heating and Cooling Load Prediction for Quick Residential Building Design. *Energy*, 274, 127334. <https://doi.org/10.1016/j.energy.2023.127334>
- Miskolc, M., et al. (2023). Prediction and Optimization of Thermal Loads in Buildings with Different Shapes by Neural Networks and Recent Finite Difference Methods. *Buildings*, 13(11), 2862. <https://doi.org/10.3390/buildings13112862>
- Tsanas, A., & Xifara, A. (2012a). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49, 560–567. <https://doi.org/10.1016/j.enbuild.2012.03.003>

- Tsanas, A., & Xifara, A. (2012b). Energy Efficiency [Data set]. UCI Machine Learning Repository. <https://doi.org/10.24432/C51307>
- Palermo, S. A., Danza, L., & Nocera, F. (2024). Adaptive façades for Mediterranean climates. *Sustainable Cities and Society*, 104, 104239.
- UCI. (2025). Energy efficiency Data Set. <https://archive.ics.uci.edu/ml/datasets/energy+efficiency>
- Zhou, T., Fang, Y., & Li, Z. (2024). AI-based optimization of building energy systems. *Applied Energy*, 351, 121550.

Influence of Graft Copolymerization Initiators on the Morphology and Wettability of Poly(vinyl alcohol)/(κ-carrageenan-graft-ethylene glycol dimethacrylate) Membranes

Gülcan GEYİK¹, Fatma KURŞUN BAYSAK^{2*}

¹Hitit University, Alaca Avni Çelik Vocational School, Department of Property Protection and Security, Çorum, Türkiye

²Kırklareli University, Faculty of Arts and Sciences, Department of Chemistry, Kırklareli, Türkiye

Received: 22.10.2025, Accepted: 08.12.2025, Published: 21.12.2025

ABSTRACT

The development of hybrid polymeric membranes with tunable surface properties is crucial for advanced separation and controlled release technologies. In this study, poly(vinyl alcohol)/(κ-carrageenan-graft-ethylene glycol dimethacrylate) (VA/CG-g-EGD) membranes were prepared by blending poly(vinyl alcohol) (VA) with κ-carrageenan-graft-ethylene glycol dimethacrylate (CG-g-EGD) copolymers synthesized using different initiators: cerium ammonium nitrate (CAN), ammonium persulfate (APS), and benzoyl peroxide (BPO). The aim was to examine how the initiator type influences grafting efficiency and, consequently, the morphology and surface characteristics of the membranes.

AFM and contact angle analyses revealed that the initiator type significantly affected surface roughness and wettability. Among the initiators, BPO produced the most uniform surface, whereas CAN yielded a rougher morphology with lower grafting efficiency. Compared to pristine VA membranes, the VA/CG-g-EGD membranes exhibited smoother, more stable surfaces with a tunable hydrophilic–hydrophobic balance, confirming the successful grafting.

These findings demonstrate that controlling grafting parameters effectively enhances membrane surface performance. The improved surface integrity and morphological stability make VA/CG-g-EGD membranes promising candidates for pervaporation-based separations and drug-delivery applications requiring selective permeability and mechanical durability.

Keywords: Atomic force microscopy; Contact angle; Poly(vinyl alcohol); Initiator.

Aşı Kopolimerizasyon Başlatıcılarının Poli(vinil alkol)/(κ-karragenan-greft-etilen glikol dimetakrilat) Membranlarının Morfolojisi ve İslatılabilirliği Üzerindeki Etkisi

ÖZ

Ayarlanabilir yüzey özelliklerine sahip hibrit polimerik membranların geliştirilmesi, ileri düzey ayırma ve kontrollü salım teknolojileri açısından büyük önem taşımaktadır. Bu çalışmada, farklı başlatıcılar olan seryum amonyum nitrat (CAN), amonyum persülfat (APS) ve benzoil peroksit (BPO) kullanılarak κ-karragenan (CG) üzerine sentezlenen κ-karragenan-aşı-etilen glikol dimetakrilat (CG-g-EGD) kopolimerleri, poli(vinil alkol) (VA) ile karıştırılarak poli(vinil alkol)/(κ-karragenan-greft-etilen glikol dimetakrilat (VA/CG-g-EGD) membranları hazırlanmıştır. Çalışmanın amacı, başlatıcı tipinin aşı

verimliliğini ve buna bağlı olarak elde edilen membranların morfolojik ve yüzey özelliklerini nasıl etkilediğini incelemektir.

AFM ve temas açısı analizleri, başlatıcı türünün yüzey pürüzlülüğü ve ıslanabilirlik üzerinde önemli bir etkiye sahip olduğunu göstermiştir. Başlatıcılar arasında, BPO en homojen yüzey yapısını oluştururken, CAN daha düşük aşı verimliliğine sahip daha pürüzlü bir morfoloji üretmiştir. Saf VA membranlara kıyasla, sentezlenen VA/CG-g-EGD membranları daha düzgün, daha kararlı yüzeyler ve ayarlanabilir hidrofilik–hidrofobik denge sergileyerek başarılı aşı polimerizasyonunu doğrulamıştır.

Elde edilen sonuçlar, aşı polimerizasyon parametrelerinin kontrolünün membran yüzey performansını etkili biçimde geliştirdiğini göstermektedir. Gelişmiş yüzey bütünlüğü ve morfolojik kararlılık, VA/CG-g-EGD membranlarını pervaporasyon temelli ayırma ve seçici geçirgenlik ile mekanik dayanım gerektiren ilaç salım uygulamaları için umut vadeden adaylar haline getirmektedir.

Anahtar Kelimeler: Atomik kuvvet mikroskobu; Temas açısı; Poli(vinil alkol); Başlatıcı.

1. INTRODUCTION

Atomic Force Microscopy (AFM), first introduced in 1986, has become one of the most powerful and versatile analytical techniques for investigating the surface morphology of solid materials, including polymers and insulators. With the advancement of AFM technology, it has evolved into an indispensable tool for examining membrane surface characteristics and fouling phenomena (Wei et al., 2024). The increasing number of studies on AFM applications in membrane fouling since 2000 highlights its growing importance in membrane science. Compared with Scanning Electron Microscopy (SEM), AFM offers superior imaging resolution and enables nanoscale three-dimensional (3D) surface visualization. While SEM requires complex sample preparation steps, such as metal coating, which can distort soft or gel-like materials, AFM typically allows *in situ* measurements in liquid environments under ambient conditions. Furthermore, AFM is suitable for both conductive and non-conductive materials, making it particularly advantageous for polymeric membrane characterization (Wei et al., 2024).

AFM operates by scanning a sharp-tipped cantilever across the sample surface to detect atomic-level interactions, such as van der Waals forces and electrostatic charges, between the probe and the sample. Depending on the nature of the interaction, AFM can be performed in three principal modes: contact, non-contact, and tapping. Since contact mode may damage soft polymeric surfaces, tapping mode is commonly preferred for membrane surface characterization (Kailash Chandra Khulbe & Matsuura, 2021). Using this approach, AFM provides detailed information on surface roughness, pore size and distribution, and phase separation, while also enabling qualitative evaluation of antifouling and fouling tendencies. Membranes with smoother, more homogeneous surfaces typically exhibit lower contact angles and greater hydrophilicity, allowing water molecules to permeate more easily and thereby reducing fouling while improving rejection performance (Rana & Matsuura, 2010).

Surface morphology plays a decisive role in determining membrane–solution interactions and fouling behavior. AFM analysis directly measures height variations, average roughness (R_a), root-mean-square roughness (R_q), and peak-to-valley differences (R_{max}), parameters that can be correlated with surface energy and wettability characteristics (K. C. Khulbe, Feng, & Matsuura, 2008). Over the past decade, AFM has also been used to quantify adhesion forces between membrane surfaces and foulants via force spectroscopy. Using the colloidal probe technique, in which micrometer-sized particles are attached to a cantilever, interaction forces between the probe and the surface can be measured. A hydrophilic surface generally exhibits stronger interaction with a hydrophilic probe, indicated by a larger phase shift, whereas a hydrophobic surface exhibits a lower phase shift (Rana & Matsuura, 2010). These measurements provide valuable insight into the mechanisms governing membrane fouling and cleaning processes.

AFM is a nanoscale imaging and force analysis tool that offers high resolution, nondestructive detection, strong adaptability, and real-time *in situ* observation. It has found widespread application across disciplines such as materials science, chemistry, physics, and life sciences. (Liu, Zhu, Chen, & Zhu, 2024). In membrane research, AFM contributes to the understanding of surface structure, pore distribution, roughness, hydrophobicity, mechanical properties, and fouling mechanisms. Furthermore, AFM data serves as a basis for optimizing membrane fabrication, modification, and performance evaluation. It allows for both qualitative and quantitative 3D visualization of physical and chemical features, including size, surface smoothness, and cross-sectional structure, across a wide range of particle sizes (1 nm to 8 μm) (Fathiraja et al., 2022).

Surface wettability, another critical factor in membrane performance, is often assessed by contact angle measurements. Contact angle analysis provides a direct measure of the surface's hydrophilicity or hydrophobicity, which significantly influences permeation flux, fouling resistance, and selectivity. Standard experimental techniques for determining contact angles include the sessile drop, captive bubble, and Wilhelmy plate methods (Ismail, Islam, Khorshidi, Tehrani-Bagha, & Sadrzadeh, 2022). The sessile drop method can be applied to both dry and wet samples, whereas the captive bubble method is suitable for submerged surfaces. The Wilhelmy plate method, although precise for homogeneous samples, is less frequently used due to its experimental complexity and sensitivity to surface uniformity. Based on contact angle values, membranes are classified as superhydrophilic (0° – 5°), hydrophilic ($<90^\circ$), hydrophobic (90° – 150°), or superhydrophobic (150° – 180°) (Lusiana et al., 2025).

When a liquid droplet is placed on a solid surface, the equilibrium contact angle (θ) formed reflects the balance of interfacial tensions between solid, liquid, and vapor phases. On an ideally smooth and chemically homogeneous surface, this relationship follows Young's equation. However, in practical systems, deviations occur due to surface heterogeneity and roughness, leading to contact-angle hysteresis and multiple metastable states. These deviations complicate the interpretation of wetting phenomena and underscore the need to combine AFM and contact angle analyses to characterize membrane surfaces and accurately understand their interactions with liquids.

Several studies in the literature have demonstrated the effectiveness of atomic force microscopy (AFM) and contact angle measurements in characterizing membrane surface morphology and wettability. Çaykara *et al.* (Çaykara, Demirci, Eroğlu, & Güven, 2006) prepared poly(N-vinyl-2-pyrrolidone)/sodium alginate (PVP/NaAlg) and poly(vinyl alcohol)/sodium alginate (PVA/NaAlg) blend membranes and analyzed their surface properties using AFM and contact angle measurements. Their findings highlighted the relationship between surface topography and hydrophilicity, emphasizing the influence of polymer composition on membrane performance. Similarly, Lusiana *et al.* (Lusiana et al., 2025) investigated the salt rejection

performance of chitosan/poly(vinyl alcohol) blend membranes. FTIR, XRD, AFM, and contact angle analyses characterized the membranes. The authors reported that increasing the PVA content led to a higher contact angle, indicating reduced surface hydrophilicity, while root-mean-square (RMS) roughness values obtained by AFM decreased, suggesting smoother membrane surfaces at higher PVA ratios. In another study, Suganthi *et al.* (Suganthi et al., 2024) fabricated sodium alginate (SA) membranes blended with poly(vinyl alcohol) (PVA) and various concentrations of histidine-capped silver nanoparticles (H-AgNPs). The incorporation of H-AgNPs resulted in an increase in contact angle compared to pristine PVA/SA membranes, implying an enhancement in surface hydrophobicity due to the presence of the nanoparticles.

Collectively, these studies demonstrate that AFM and contact angle analyses are indispensable techniques for understanding the correlation between membrane surface structure and wettability. Insights from such analyses are crucial for guiding membrane design, modification, and optimization across various separation applications.

In the literature, various κ -carrageenan-based graft copolymers using different monomers have been reported. For example, in a study conducted by Geyik and Işıklan, DMAEMA was grafted onto κ -carrageenan, and the resulting graft copolymer was reported to exhibit enhanced thermal stability and biocompatibility (Geyik & Işıklan, 2020b). In another study by the same research group, DMAEMA and acrylic acid (AA) were simultaneously grafted onto κ -carrageenan, and it was shown that the grafting process improved the mechanical strength of the copolymer membrane (Geyik & Işıklan, 2020a). Similarly, in a study where HEMA was grafted onto κ -carrageenan, the grafting procedure increased the thermal resistance of the material and transformed its structure into a more regular, homogeneous, sponge-like, and porous morphology, while the copolymer remained biocompatible. These findings indicate that the type of monomer and grafting conditions have a decisive effect on the thermal, mechanical, and morphological properties of κ -carrageenan-based copolymers (Geyik & Işıklan, 2023). In another study, PVA was grafted onto carrageenan, and it was reported that increasing the carrageenan content led to higher contact angles, making the surface more hydrophobic (Sukhlaiied & Riyajan, 2013). In studies on iota-carrageenan-g-PVA polyelectrolyte membranes, the grafting process was shown to increase the contact angle, resulting in a more hydrophobic surface (Chavez-Baldovino, Malca-Reyes, Masso, Feng, & Díaz-Vázquez, 2025).

In the present study, κ -carrageenan (CG) was modified separately by graft copolymerization with ethylene glycol dimethacrylate (EGD) using three different radical initiators: cerium ammonium nitrate (CAN), ammonium persulfate (APS), and benzoyl peroxide (BPO). The resulting CG-g-EGD graft copolymers were mixed with VA at 1/1 (w/w) ratio to obtain VA/CG-g-EGD composite membranes. The structural and surface properties of the resulting membranes were characterized by Atomic Force Microscopy (AFM) and contact angle measurements to evaluate their morphology and hydrophilicity. Therefore, this study aims to

investigate how the incorporation of different initiators in graft copolymerization affects membrane surface roughness and hydrophilicity, thereby providing a basis for understanding their potential to enhance separation performance and fouling resistance.

2. MATERIALS AND METHODS

2.1. Materials

VA, CG, CAN, APS, BPO, TEMED, potassium dihydrogen phosphate, sodium hydroxide, acetic acid, acetone, ethanol, and methanol were supplied by Sigma.

2.2. Synthesis of Graft Copolymers

The graft copolymer was synthesized using methods reported in the literature (Baysak & Geyik, 2024; Geyik & Işıklan, 2020b; Işıklan & Kurşun, 2013; Işıklan, Kurşun, & İnal, 2010). Briefly, separate solutions of CG were prepared. Under a nitrogen atmosphere and at a constant temperature, graft copolymerization was carried out by adding EGD monomer, CAN, APS, or BPO initiator, and TEMED accelerator. The resulting graft copolymers were precipitated in cold acetone, washed with ethanol, and then dried in a vacuum oven for storage. The grafting percentage and grafting efficiency were calculated using the following equations;

$$\text{Grafting Percentage (\%)} = \frac{(m_g - m_p)}{m_p} \times 100 \quad (1)$$

$$\text{Grafting Efficiency (\%)} = \frac{(m_g - m_p)}{(m_g - m_p) + m_h} \times 100 \quad (2)$$

where m_g represents the mass of the graft copolymer, m_p denotes the mass of the polymer, and m_h is the mass of the homopolymer.

2.3. Preparation of Membranes

The CG-g-EGD graft copolymer solutions were blended with a 50% (v/v) VA solution and cast into Petri dishes. The resulting films were dried in a vacuum oven at 30 °C and subsequently thermally crosslinked at 150 °C for one hour. The membrane thicknesses were measured using a micrometer (Dasqua, 4210). Among the VA/CG-g-EGD membranes obtained, those with an average thickness of approximately 50 µm were selected for this study.

2.4. Membrane characterization

The surfaces of the synthesized VA and VA/CG-g-EGD membranes were characterized using NT-MDT/Ntegra Solaris Atomic Force Microscopy (AFM) and DATAPHYSICS/OCA 50Micro contact angle (CA) measurements.

2.5. Statistical analysis

All experiments related to vaccination percentage and grafting efficiency were conducted in at least three replicates ($n \geq 3$). The results were expressed as mean $\pm$ standard deviation (SD). Statistical evaluations were performed using SPSS version 16.0 through one-way analysis of variance (ANOVA). A P value below 0.05 was considered statistically significant.

3. RESULTS AND DISCUSSION

In this study, the effect of copolymerization of CG-g-EGD on the VA/CG-g-EGD membrane was investigated. For this purpose, it has previously been reported (Baysak & Geyik, 2024; Geyik, 2025a; Geyik & Işıklan, 2021, 2022; Işıklan & Kurşun, 2013; Işıklan et al., 2010) that the type of initiator influences grafting yield. The grafting yields of the synthesized CG-g-EGD copolymers and the abbreviations of the corresponding VA/CG-g-EGD membranes are presented in Table 1.

Table 1: Abbreviation and properties of the VA/CG-g-EGD. Results are stated as the mean $\pm$ SD ($n=3$) ($P < 0.05$).

Abbreviation	Initiator	Grafting Percentage (%)	Grafting Efficiency (%)
VA/CG-g-EGD ₁	CAN	38.6 $\pm$ 2.6	16 $\pm$ 1.3
VA/CG-g-EGD ₂	APS	55.1 $\pm$ 2.1	25 $\pm$ 2.4
VA/CG-g-EGD ₃	BPO	88.9 $\pm$ 4.1	32 $\pm$ 3.1

In this study, CG-g-EGD copolymers with varying grafting ratios were synthesized via a conventional free-radical graft copolymerization at 70 °C. Ceric ammonium nitrate (CAN), benzoyl peroxide (BPO), and ammonium persulfate (APS) were employed as radical initiators. Under suitable thermal conditions, these initiators decompose to generate reactive radicals that abstract hydrogen from hydroxyl (–OH) groups on CG chains, forming CG macroradicals (Geyik & Işıklan, 2023).

The resulting macroradicals then react with the ethylene glycol dimethacrylate (EGD) monomer by opening its double bonds and generating new active centers. During the propagation stage, EGD monomers are successively added to these active sites, enabling continuous chain growth. In the termination step, active macroradicals and poly(EGD) radicals combine, yielding CG-g-EGD copolymer along with poly(EGD) homopolymer (Geyik, 2025b).

The proposed structural representation and grafting mechanism of EGD onto CG are illustrated in Figures 1 and 2, respectively, demonstrating the radical formation pathway and the bonding of the monomer to the CG backbone.

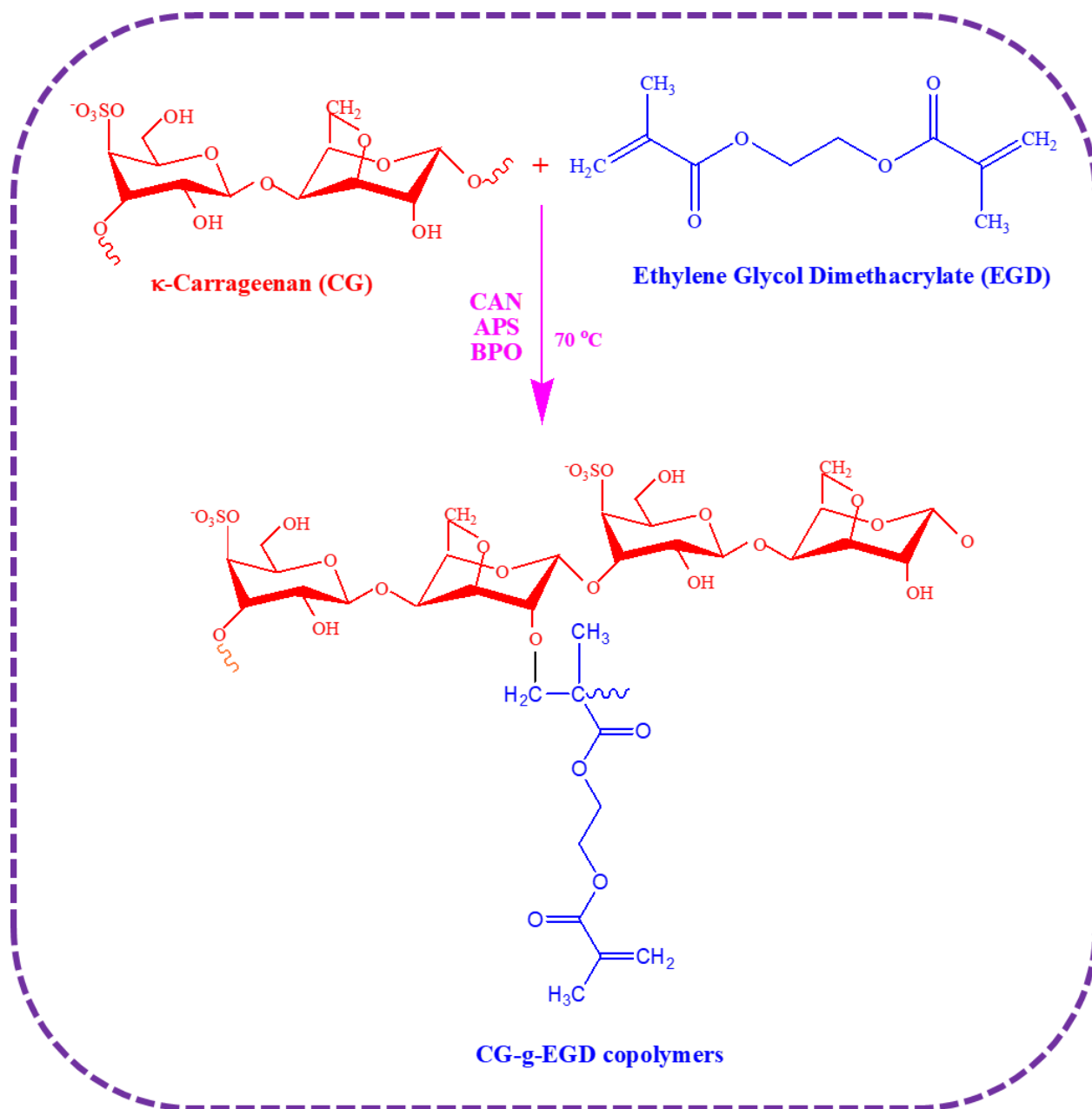


Figure 1. Schematic representation of the synthesis of CR-g-EGD copolymer.

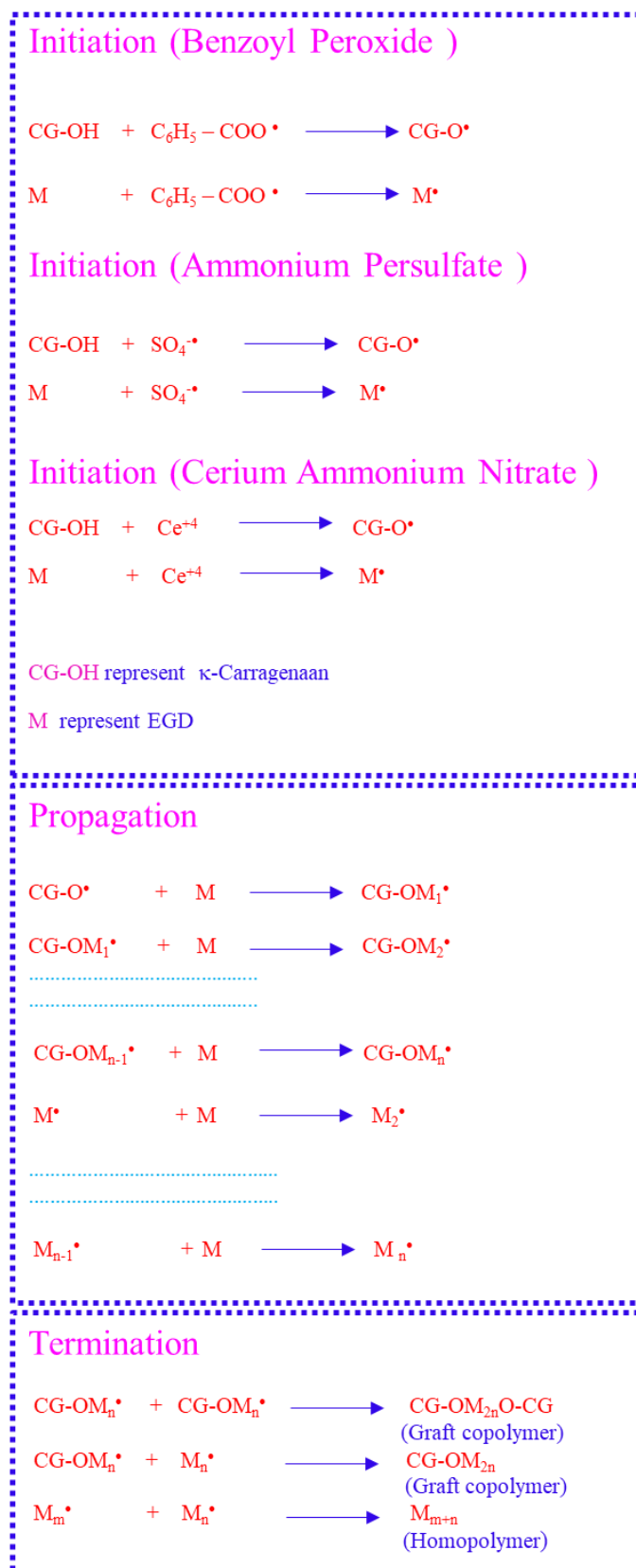


Figure 2: Proposed mechanism for graft copolymerization of ethylene glycol dimethacrylate onto κ -carrageenan.

To examine the influence of initiator-dependent grafting yield variations on surface properties, surface properties, VA and VA/CG-g-EGD membranes were characterized using AFM and CA analyses, and the results are presented in Figures 3 and 4.

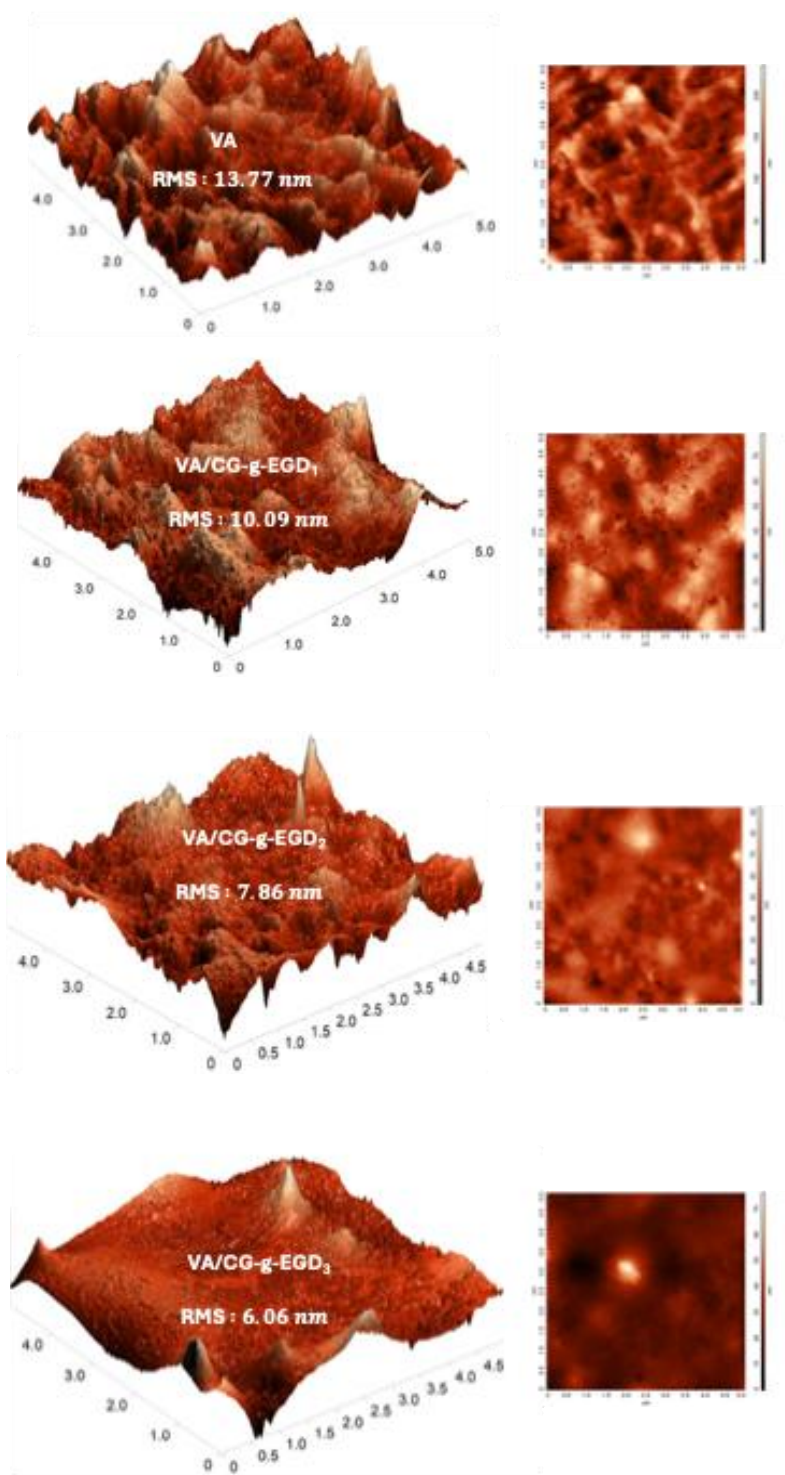


Figure 3: 2D and 3D AFM images of VA and VA/CG-g-EGD membranes

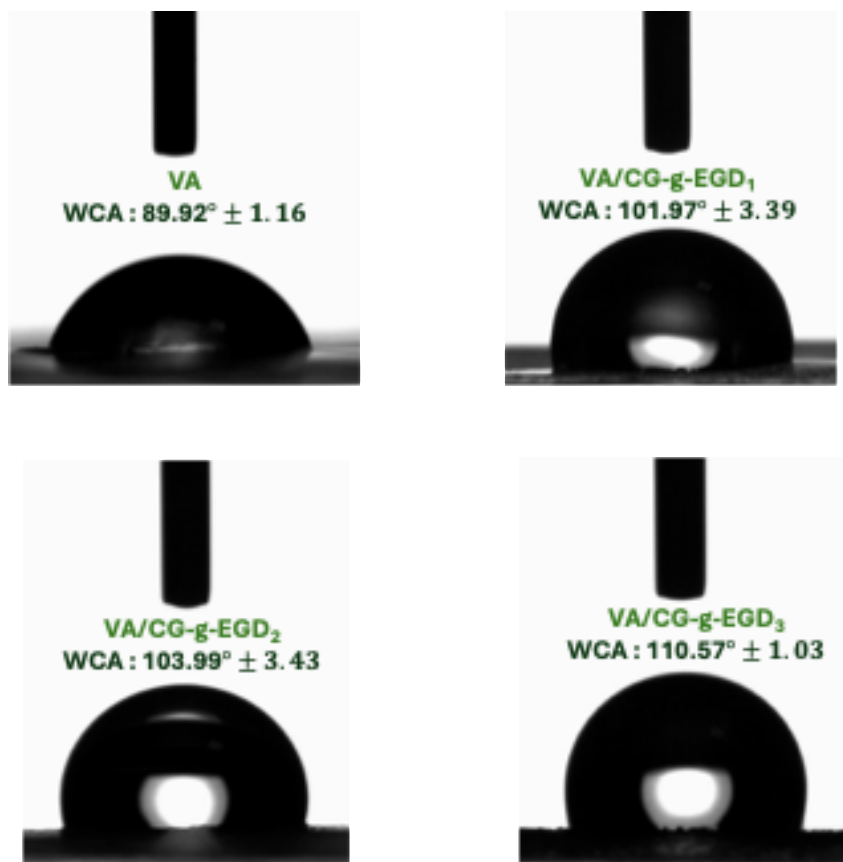


Figure 4: CA images of VA and VA/CG-g-EGD membranes

Table 2: RMS roughness and contact angle (CA) values for VA and VA/CG-g-EGD membranes. Results are stated as the mean $\pm$ SD (n =3) ($P < 0.05$).

Membrane	RMS (nm)	Contact Angle (°)
VA	13.77	89.92 $\pm$ 1.16
VA/CG-g-EGD ₁	10.09	101.97 $\pm$ 3.39
VA/CG-g-EGD ₂	7.86	103.99 $\pm$ 3.43
VA/CG-g-EGD ₃	6.06	110.57 $\pm$ 1.03

Figure 1 displays the 2D and 3D AFM micrographs of pristine VA and VA/CG-g-EGD membranes, each acquired over a scan area of $5 \times 5 \mu\text{m}$. The AFM analyses reveal distinct variations in surface topography, where brighter regions represent elevated surface features, while darker regions correspond to areas of reduced surface height, typically associated with membrane pores. The observations demonstrate that the VA/CG-g-EGD membrane exhibits enhanced structural homogeneity, which can be attributed to improved compatibility and reinforced intermolecular interactions between the grafted segments and the VA matrix.

Since the use of three different initiators in the graft copolymer synthesis affects the CG-g-EGD copolymer structure, the VA/CG-g-EGD membrane surface is also influenced by this variation. The graft yields of graft copolymers synthesized using CAN, APS, and BPO initiators were listed as BPO > APS > CAN. The fact that CAN has the lowest graft yield can be attributed to the limited electron delocalization range in its structure compared to APS and BPO. On the other hand, BPO provides greater resonance stability, resulting in better structural integrity of the CG-g-EGD graft copolymer during copolymerization.

The initiator used in graft copolymerization directly determines the progression of the reaction as well as the fundamental characteristics of the resulting copolymer. Different initiators lead to variations in chain initiation and propagation rates, thereby influencing the polymerization kinetics, reaction controllability, and structural integrity of the final product. Since initiator selection governs not only the onset and uniformity of the reaction but also the morphological and physical properties of the graft copolymer, it plays a critical role in achieving the targeted copolymer architecture (Geyik, 2025a; Tripathy & Singh, 2000). The decisive effect of initiator type on grafting yield has been demonstrated in the literature across various systems. In the study by Geyik (2025), among four initiators, CAN, BPO, AIBN, and APS, the highest grafting yield was obtained using BPO. The grafting percentage (151.12%) and grafting efficiency (85.53%) achieved at 80 °C were attributed to the strong thermal decomposition capacity of BPO and its ability to generate a high number of benzoyloxy and phenyl radicals (Geyik, 2025b). In another study, the influence of initiator type on grafting yield was systematically investigated during the synthesis of ALG-g-PHEMA using four different initiators: cerium ammonium nitrate (CAN), azobisisobutyronitrile (AIBN), benzoyl peroxide (BPO), and ammonium persulfate (APS). Each initiator successfully promoted the attachment of HEMA monomers onto the alginate backbone through a free-radical mechanism, although optimal reaction conditions varied among them. The highest performance was obtained with AIBN, yielding a grafting percentage of 59.12% and a grafting efficiency of 32.55%. While CAN, BPO, and APS also initiated grafting effectively, they produced lower grafting yields due to differences in radical stability, monomer–radical interaction, and chain-initiation kinetics. These findings collectively underscore that initiator selection is a key parameter in the chemical modification of natural polymers and directly governs the grafting yield (Cebeci & Geyik, 2025).

The effects of using different initiators in the graft copolymer synthesis on the structure of the CG-g-EGD copolymer were confirmed by the AFM and CA analysis results. The membrane prepared from the graft copolymer synthesized using the CAN initiator exhibited the highest surface roughness (RMS) value, while the lowest roughness was obtained for the BPO-based membrane (Amiri, Vatanpour, & He, 2023; Kashyap, Pratihari, & Behera, 2016). The contact angle (CA) results, which describe the rate and pattern of liquid penetration, were consistent with the AFM results. As shown in Table 2, the RMS values of the membranes decreased in the order VA > VA/CG-g-EGD₁ > VA/CG-g-EGD₂ > VA/CG-g-EGD₃, while the CA values

increased in the reverse order $VA/CG-g-EGD_3 > VA/CG-g-EGD_2 > VA/CG-g-EGD_1 > VA$ (Danish, 2022; Krainer & Hirn, 2021; Megias-Alguacil, Tervoort, Cattin, & Gauckler, 2011).

When the results are examined, it is evident that the VA membrane exhibits the highest surface roughness and the lowest water contact angle, indicating lower surface uniformity and greater hydrophilicity than the modified membranes. In contrast, the newly synthesized VA/CG-g-EGD membranes exhibit smoother surfaces, higher contact angles, and improved morphological stability, suggesting that grafting enhances structural regularity and reduces surface swelling. These findings indicate that the introduction of CG-g-EGD copolymers effectively improves the overall surface characteristics of the membranes relative to neat VA.

In their study, Yoshida and Cohen (2003) examined the topography of porous silica membranes (500 Å native pore size) modified via silylation with vinyl trimethoxysilane (VTMS) and subsequent graft polymerization with poly(vinyl acetate) (PVAc) using AFM. They observed a reduction in RMS roughness from 19.1 to 13.9 nm, indicating that surface grafting effectively smoothed the membrane surface. Zhuikova et al. (Zhuikova et al., 2020) investigated thin films composed of κ -carrageenan and chitosan prepared by layer-by-layer deposition. The surface morphology of these films was analyzed using AFM, revealing that the formation of polysaccharide coatings on nitinol (NiTi) substrates led to smoother surfaces and enhanced hydrophilicity. The RMS roughness values decreased markedly from 123 nm for bare NiTi to 23 nm for multilayered Chit/Carr/Chit/Carr coatings, indicating a significant improvement in surface homogeneity and film compactness. Furthermore, Guo *et al.* (Guo et al., 2022) prepared κ -carrageenan/hydroxypropyl methylcellulose (HPMC) composite membranes by adding different proportions of HPMC (0%, 3%, 6%, 9%, and 12% w/w) to a 1% κ -carrageenan solution, yielding κ CHM-0, κ CHM-3, κ CHM-6, κ CHM-9, and κ CHM-12 membranes. Contact angle analysis revealed that pure κ -carrageenan membranes exhibited the lowest contact angle, while increasing the HPMC content progressively increased the angle. This rise in water contact angle was attributed to the enhanced hydrophobicity of the membrane surface, which the authors associated with hydrogen-bond formation between HPMC and κ -carrageenan, thereby reducing the availability of hydrophilic hydroxyl groups.

4. CONCLUSION

This study demonstrated that the type of initiator used in the graft copolymerization of CG-g-EGD significantly influences both the grafting efficiency and the surface morphology of the resulting VA/CG-g-EGD membranes. Among the tested initiators, BPO yielded the highest graft efficiency and produced membranes with the most homogeneous and stable surfaces, while CAN generated lower-quality surfaces with greater roughness.

AFM and CA measurements confirmed that incorporating CG-g-EGD copolymers into VA matrices yields membranes with smoother, more uniform, and more stable surface morphologies than pure VA membranes. The reduced surface roughness and improved hydrophobic balance of the VA/CG-g-EGD membranes suggest better resistance to swelling and enhanced structural durability.

Overall, these findings indicate that the synthesis of VA/CG-g-EGD membranes through optimized graft copolymerization provides a promising route to developing high-performance membranes. The improved surface regularity, controlled wettability, and morphological stability make these membranes attractive for future pervaporation-based separation systems and drug-delivery applications, where controlled diffusion, chemical stability, and selective permeability are critical.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest among the authors.

ACKNOWLEDGEMENT

The authors express their gratitude to the Kırklareli University Advanced Technologies Application and Research Center (ITUAM) for granting access to the facilities utilized in this work. Appreciation is also extended to the Kırklareli University Scientific Research Projects Coordination Unit for their financial support under project numbers DPT-2010K121120, KLUBAP-204, and KLUBAP-264.

CONTRIBUTIONS OF AUTHORS

G.G.: Analysis, research, sources, writing - review and editing.

F.K.B.: Method, analysis, research, sources, writing - original draft preparation.

REFERENCES

- Amiri, S., Vatanpour, V., & He, T. (2023). Antifouling thin-film nanocomposite NF membrane with polyvinyl alcohol-sodium alginate-graphene oxide nanocomposite hydrogel coated layer for As(III) removal. *Chemosphere*, 322, 138159. <https://doi.org/10.1016/j.chemosphere.2023.138159>
- Baysak, F. K., & Geyik, G. (2024). Hidrofilik Aşı Kopolimer Membranların Şişme Davranışının Kinetik İncelenmesi. *Kırklareli University Journal of Engineering and Science*, 10(2), 191-206. <https://doi.org/10.34186/klujes.1562185>
- Cebeci, B., & Geyik, G. (2025). Graft copolymerization and characterization of 2-hydroxyethyl methacrylate (HEMA) onto sodium alginate using different initiators. *International Journal of Biological Macromolecules*, 328, 147699. <https://doi.org/10.1016/j.ijbiomac.2025.147699>
- Chavez-Baldovino, E., Malca-Reyes, C. A., Masso, R., Feng, P., & Díaz-Vázquez, L. M. (2025). Development and Characterization of κ -Carrageenan and Boron Nitride Nanoparticle Membranes for Improved Ionic Conductivity in Fuel Cells. *Fuels*, 6(1), 15. <https://doi.org/10.3390/fuels6010015>

- Çaykara, T., Demirci, S., Eroğlu, M. S., & Güven, O. (2006). Surface properties of binary blend films of poly(*N* -vinyl-2-pyrrolidone) and poly(vinyl alcohol) with sodium alginate. *Journal of Polymer Science Part B: Polymer Physics*, 44(2), 426-430. <https://doi.org/10.1002/polb.20712>
- Danish, M. (2022). Contact angle studies of hydrophobic and hydrophilic surfaces. İçinde *In Handbook of Magnetic Hybrid Nanoalloys and Their Nanocomposites* (Pp. 1-22). Cham: Springer International Publishing. *Handbook of Magnetic Hybrid Nanoalloys and their Nanocomposites* (ss. 1-22). Springer, Cham. Geliş tarihi gönderen https://doi.org/10.1007/978-3-030-34007-0_24-1
- Fathiraja, P., Gopalrajan, S., Karunanithi, M., Nagarajan, M., Obaiah, M. C., Durairaj, S., & Neethirajan, N. (2022). Response surface methodology model to optimize concentration of agar, alginate and carrageenan for the improved properties of biopolymer film. *Polymer Bulletin*, 79(8), 6211-6237. <https://doi.org/10.1007/s00289-021-03797-5>
- Geyik, G. (2025a). Etilen Glikol Dimetakrilatın, Sodyum Aljinat, Kitosan ve Polivinil Alkol Üzerine Farklı Başlatıcılar Kullanılarak Aşılması ve Karakterizasyonu. *Afyon Kocatepe University Journal of Sciences and Engineering*, 25(1), 59-69. <https://doi.org/10.35414/akufemubid.1479986>
- Geyik, G. (2025b). Synthesis, Characterization, and Swelling Performance of Chitosan-Graft-Poly(ethylene glycol dimethacrylate). *ChemistrySelect*, 10(37), e03204. <https://doi.org/10.1002/slct.202503204>
- Geyik, G., & Işıklan, N. (2020a). pH /temperature-responsive poly(dimethylaminoethyl methacrylate) grafted κ-carrageenan copolymer: Synthesis and physicochemical properties. *Journal of Applied Polymer Science*, 137(48), 49596. <https://doi.org/10.1002/app.49596>
- Geyik, G., & Işıklan, N. (2020b). Synthesis, characterization and swelling performance of a temperature/pH-sensitive κ-carrageenan graft copolymer. *International Journal of Biological Macromolecules*, 152, 359-370. <https://doi.org/10.1016/j.ijbiomac.2020.02.129>
- Geyik, G., & Işıklan, N. (2023). Chemical modification of κ-carrageenan with poly(2-hydroxypropylmethacrylamide) through microwave induced graft copolymerization: Characterization and swelling features. *International Journal of Biological Macromolecules*, 235, 123888. <https://doi.org/10.1016/j.ijbiomac.2023.123888>
- Guo, J., Dong, S., Ye, M., Wu, X., Lv, X., Xu, H., & Li, M. (2022). Effects of Hydroxypropyl Methylcellulose on Physicochemical Properties and Microstructure of κ-Carrageenan Film. *Foods*, 11(19), 3023. <https://doi.org/10.3390/foods11193023>
- Ismail, M. F., Islam, M. A., Khorshidi, B., Tehrani-Bagha, A., & Sadrzadeh, M. (2022). Surface characterization of thin-film composite membranes using contact angle technique: Review of quantification strategies and applications. *Advances in Colloid and Interface Science*, 299, 102524. <https://doi.org/10.1016/j.cis.2021.102524>
- Işıklan, N., & Kurşun, F. (2013). Synthesis and characterization of graft copolymer of sodium alginate and poly(itaconic acid) by the redox system. *Polymer Bulletin*, 70(3), 1065-1084. <https://doi.org/10.1007/s00289-012-0876-x>
- Işıklan, N., Kurşun, F., & İnal, M. (2010). Graft copolymerization of itaconic acid onto sodium alginate using benzoyl peroxide. *Carbohydrate Polymers*, 79(3), 665-672. <https://doi.org/10.1016/j.carbpol.2009.09.021>
- Kashyap, S., Pratihari, S. K., & Behera, S. K. (2016). Strong and ductile graphene oxide reinforced PVA nanocomposites. *Journal of Alloys and Compounds*, 684, 254-260. <https://doi.org/10.1016/j.jallcom.2016.05.162>
- Khulbe, K. C., Feng, C. Y., & Matsuura, T. (Ed.). (2008). Atomic Force Microscopy. İçinde *Synthetic Polymeric Membranes: Characterization by Atomic Force Microscopy* (ss. 19-45). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-73994-4_3
- Khulbe, Kailash Chandra, & Matsuura, T. (2021). Membrane Characterization. İçinde Kailash Chandra Khulbe & T. Matsuura (Ed.), *Nanotechnology in Membrane Processes* (ss. 89-133). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-64183-2_3
- Krainer, S., & Hirn, U. (2021). Contact angle measurement on porous substrates: Effect of liquid absorption and drop size. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 619, 126503. <https://doi.org/10.1016/j.colsurfa.2021.126503>

- Liu, Q., Zhu, X., Chen, B., & Zhu, X. (2024). Applications of AFM in Membrane Characterization and Fouling Analysis. *ACS ES&T Engineering*, 4(8), 1805-1838. <https://doi.org/10.1021/acsestengg.4c00111>
- Lusiana, R. A., Nuryanto, R., Dayanti, D., Muna, N., Safitri, T., & Sasongko, N. A. (2025). Role of Polyvinyl Alcohol in Enhancing Chitosan Membranes for Desalination: An Experimental and Theoretical Study. *ChemistrySelect*, 10(22), e01611. <https://doi.org/10.1002/slct.202501611>
- Megias-Alguacil, D., Tervoort, E., Cattin, C., & Gauckler, L. J. (2011). Contact angle and adsorption behavior of carboxylic acids on α -Al₂O₃ surfaces. *Journal of Colloid and Interface Science*, 353(2), 512-518. <https://doi.org/10.1016/j.jcis.2010.09.087>
- Rana, D., & Matsuura, T. (2010). Surface Modifications for Antifouling Membranes. *Chemical Reviews*, 110(4), 2448-2471. <https://doi.org/10.1021/cr800208y>
- Suganthi, S., Vignesh, S., Al-Ansari, M. M., Al-Humaid, L. A., Oh, T. H., & Raj, V. (2024). Development of PVA/sodium alginate incorporated with histidine capped silver nanoparticles for food packaging application. *Polymers for Advanced Technologies*, 35(5), e6413. <https://doi.org/10.1002/pat.6413>
- Sukhlaaied, W., & Riyajan, S.-A. (2013). Synthesis and properties of carrageenan grafted copolymer with poly(vinyl alcohol). *Carbohydrate Polymers*, 98(1), 677-685. <https://doi.org/10.1016/j.carbpol.2013.06.047>
- Tripathy, T., & Singh, R. P. (2000). High performance flocculating agent based on partially hydrolysed sodium alginate–g–polyacrylamide. *European Polymer Journal*, 36(7), 1471-1476. [https://doi.org/10.1016/S0014-3057\(99\)00201-3](https://doi.org/10.1016/S0014-3057(99)00201-3)
- Wei, M., Zhang, Y., Wang, Y., Liu, X., Li, X., & Zheng, X. (2024). Employing Atomic Force Microscopy (AFM) for Microscale Investigation of Interfaces and Interactions in Membrane Fouling Processes: New Perspectives and Prospects. *Membranes*, 14(2), 35. <https://doi.org/10.3390/membranes14020035>
- Yoshida, W., & Cohen, Y. (2003). Topological AFM characterization of graft polymerized silica membranes. *Journal of Membrane Science*, 215(1), 249-264. [https://doi.org/10.1016/S0376-7388\(03\)00019-X](https://doi.org/10.1016/S0376-7388(03)00019-X)
- Zhuikova, Y. V., Zhuikov, V. A., Zubareva, A. A., Akhmedova, S. A., Sviridova, I. K., Sergeeva, N. S., & Varlamov, V. P. (2020). Physicochemical and biological characteristics of chitosan/ κ -carrageenan thin layer-by-layer films for surface modification of nitinol. *Micron*, 138, 102922. <https://doi.org/10.1016/j.micron.2020.102922>

Hybrid Deep Learning Strategies Leveraging Cutting-Edge VGG Architectures for Advanced Oral Cancer Diagnosis

Vahtettin Cem BAYDOĞAN^{1*} 

¹Department of Software Engineering, Faculty of Technology, Firat University, Elazığ, Türkiye

Received: 02.12.2025, Accepted: 09.12.2025, Published: 23.12.2025

ABSTRACT

Oral Cancer (OC) has become a critical public health problem, with its increasing prevalence worldwide and high mortality rate when diagnosed late. Tobacco and alcohol use, Human Papilloma Virus (HPV) infections, and various environmental factors play a significant role in the development of the disease. Early detection of the disease significantly improves treatment success and quality of life. However, traditional clinical examinations and manual assessment methods are both time-consuming and can lead to high misclassification rates due to expert dependency. In this study, a deep learning-based hybrid approach for the automatic classification of OC is proposed. The proposed model utilizes different variants of the Visual Geometry Group (VGG) architecture, namely VGG11, VGG13, VGG16, and VGG19, to extract deep features from OC images. The resulting deep features were processed with various classifiers, including Extreme Gradient Boosting (XGBoost), K-Nearest Neighbors (kNN), Support Vector Machines (SVM), Random Forest (RF), and Light Gradient Boosting Machine (LGBM), and a comprehensive experimental analysis was conducted. Experimental findings demonstrate that the VGG19+SVM hybrid model, in particular, demonstrated superior performance, achieving the highest AUC score (0.9144) for inter-class discrimination. Furthermore, the VGG19+LGBM model achieved the highest accuracy rate (0.9158), demonstrating strong classification performance. The results demonstrate that VGG-based deep feature extraction provides high accuracy and strong discrimination in OC classification. These findings demonstrate that the proposed hybrid approach is a reliable diagnostic tool that can be effectively used in clinical decision support systems.

Keywords: Oral cancer detection; VGG; deep learning; transfer learning.

Gelişmiş Ağız Kanseri Tanısı için Son Teknoloji VGG Mimarilerinden Yararlanan Hibrit Derin Öğrenme Stratejileri

ÖZ

Ağız Kanseri (OC), dünya çapında giderek artan yaygınlığı ve geç teşhis edildiğinde yüksek ölüm oranı ile kritik bir halk sağlığı sorunu haline gelmiştir. Tütün ve alkol kullanımı, Human Papilloma Virüsü (HPV) enfeksiyonları ve çeşitli çevresel faktörler hastalığın gelişiminde önemli rol oynar. Hastalığın erken teşhisi, tedavi başarısını ve yaşam kalitesini önemli ölçüde artırmaktadır. Ancak, geleneksel klinik muayeneler ve manuel değerlendirme yöntemleri hem zaman alıcıdır hem de uzman bağımlılığı nedeniyle yüksek yanlış sınıflandırma oranlarına yol açabilir. Bu nedenle, bu çalışmada, OC'nin otomatik sınıflandırılması için derin öğrenme tabanlı bir hibrit yaklaşım önerilmiştir. Önerilen model, OC görüntülerinden derin özellikler çıkarmak için Görsel Geometri Grubu (VGG) mimarisinin VGG11, VGG13, VGG16 ve VGG19 farklı varyantlarını kullanılmıştır. Elde edilen derin özellikler, Extreme

Gradient Boosting (XGBoost), k-En Yakın Komşu (kNN), Destek Vektör Makineleri (SVM), Rastgele Orman (RF) ve Light Gradient Boosting Machine (LGBM) dahil olmak üzere çeşitli sınıflandırıcılarla işlenmiştir ve kapsamlı bir deneysel analiz gerçekleştirilmiştir. Deneysel bulgular, özellikle VGG19+SVM hibrit modelinin üstün performans gösterdiğini ve sınıflar arası ayrımcılık için en yüksek AUC puanını (0,9144) elde ettiğini göstermektedir. Ayrıca, VGG19+LGBM modeli en yüksek doğruluk (0,9158) oranını elde ederek güçlü sınıflandırma performansı göstermiştir. Sonuçlar, VGG tabanlı derin özellik çıkarmanın OC sınıflandırmasında yüksek doğruluk ve güçlü ayrımcılık sağladığını göstermektedir. Bu bulgular, önerilen hibrit yaklaşımın klinik karar destek sistemlerinde etkili bir şekilde kullanılabilecek güvenilir bir tanı aracı olduğunu göstermektedir.

Anahtar Kelimeler: Ağız kanseri tespiti; VGG; Derin öğrenme; Transfer öğrenme.

1. INTRODUCTION

OC often begins as subtle changes in the soft tissues of the mouth, which can be difficult to detect in the early stages (Warin & Suebnukarn, 2024). The disease is often preceded by visible lesions called Oral Potentially Malignant Disorders (OPMDs). The most common symptoms of oral cancer are a non-healing sore or ulcer (usually lasting more than 14 days), swellings or masses that may cause pain or bleeding, or patches inside the mouth that are white, red, or a combination of both (Yaduvanshi et al., 2025). Clinically, oral cancer often presents as a nodular lesion protruding from the mucosal surface or a deep ulcer with a rough surface and ill-defined borders (Warin et al., 2021).

The development of oral cancer is a multifactorial process, influenced by lifestyle, biological, and environmental factors (Prabhakaran et al., 2020). The primary risk factors include tobacco use (including smoking, smokeless tobacco, or betel nut chewing) and excessive alcohol consumption (Welikala et al., 2020). Alcohol acts as a solvent, increasing the absorption of carcinogens, and when combined with tobacco use, it increases the risk. Another significant risk factor is infection with HPV. Increasing age (men over 50 are particularly at increased risk) and a diet low in fruits and vegetables are also risk factors (Chang et al., 2023).

OC diagnosis has traditionally relied on clinical examination (visual examination) and histopathological examination, followed by surgical biopsy if a suspicious lesion is present (Shah et al., 2025). Histological evaluation remains the gold standard for diagnosing OC. However, these traditional methods have limitations, including subjectivity, interpretation variability, invasiveness, time-consuming nature, and the potential for delays in diagnosis (Li et al., 2021). This negatively impacts patient survival rates, particularly in settings with inadequate healthcare services, such as rural areas. Additionally, early-stage cancer lesions can often resemble common oral diseases, making early detection of malignant lesions difficult (Huang et al., 2023, 2024).

Recent advances in artificial intelligence (AI) and machine learning (ML) offer promising and innovative solutions to improve the early detection and diagnosis of oral cancer. Deep learning (DL) techniques, particularly convolutional neural networks (CNNs), can automatically extract and analyze complex patterns from large medical image datasets (photographs, histopathological images, and radiologic images). AI-powered algorithms can achieve high levels of accuracy, sensitivity, and specificity in oral cancer diagnosis. DL models have performed on par with, or even better than, experts on clinical validation datasets. The use of AI has significant clinical implications, particularly in resource-limited settings where specialist physicians are scarce, by making diagnosis faster, non-invasive, cost-effective, and easier. These DL solutions, integrated with smartphone-based imaging systems, have the potential to

increase the effectiveness and accessibility of oral cancer screening programs. The development of automated systems based on DL promises to have a transformative impact by enabling early-stage detection of oral cancer, both improving patient outcomes and reducing the burden on healthcare systems.

In this study, different DL-based feature extraction methods and traditional ML algorithms were evaluated separately for the automatic detection of OC. Pre-trained or transfer learning-based VGG architectures (VGG11, VGG13, VGG16, and VGG19) were used in the feature extraction phase, and the classification success of the deep features obtained from these models was analyzed independently. In the classification phase, the features of each VGG models (VGG11, VGG13, VGG16, and VGG19) were tested and compared with XGBoost, kNN, SVM, RF, and LGBM algorithms. This demonstrates that VGG-based deep features provide high accuracy and robust discrimination in detecting OC when used in conjunction with different ML classifiers. The primary purpose of using multiple methods is to increase the generalization capacity of the model, reduce the risk of overfitting, and develop a reliable decision support system for early OC diagnosis. Furthermore, the study proposes an explainable and optimized OC classification framework that combines DL and ML components.

The study's contribution to the literature can be summarized as follows:

- The study presents a comprehensive and comparative hybrid evaluation for OC detection by systematically comparing deep features obtained from four different VGG architectures with five different ML algorithms.
- Deep features extracted from VGG11-19 models, combined with XGBoost, kNN, SVM, RF, and LGBM, demonstrate high accuracy and robust discrimination capabilities. An optimized OC classification framework that can serve as a reference in the literature is proposed.
- The multi-model approach offers strong generalization by reducing overlearning, while the DL-ML integration provides an explainable, reliable method for early OC detection that can be integrated into clinical decision support systems.

The remaining sections of the article are structured as follows: Section 2 presents a literature review that provides the theoretical basis for the OC detection study and comprehensively examines related literature. Section 3 details the OC image datasets used and describes the model architecture that combines the proposed VGG-based deep feature extraction approach (VGG11, VGG13, VGG16, and VGG19) with ML classifiers (XGBoost, KNN, SVM, RF, LGBM). Section 4 includes the results of the conducted experimental studies and comparative analyses with other methods. Finally, Section 5 presents an overview of the findings, highlighting the conclusions of the OC detection study and the importance of future studies.

2. LITERATURE REVIEW

Rapid advances in DL in recent years have led to significant advances in the early detection of OC and demonstrated the effectiveness of various architectural approaches. CNN and transfer learning-based models, in particular, are among the most successful methods widely used for extracting meaningful visual features from oral lesion images. Some of these studies are listed below.

Kabir et al. proposed a lightweight and accurate hybrid architecture, LightSE-MobileViT, for OC detection (Kabir et al., 2025). The model was trained on the Oral Cancer Classification Dataset using 131 clinical images. The model achieved 98.39% accuracy, 1.00 ROC-AUC, and 0.98 macro F1 score. This superior performance is attributed to the hybrid architecture, which integrates lightweight convolutional layers, Squeeze-and-Excitation (SE) blocks that provide channel-based attention enhancement, and the MobileViT transformer encoder used to model global contextual dependencies. Although the study's evaluation confirmed the model's clinical applicability and architectural advantages, the small size of the dataset used and the lack of generalizability due to its collection from a single center were noted as limitations.

Soni and colleagues reported remarkable results using the EfficientNetB0 architecture for the detection of oral squamous cell carcinoma (OSCC). The study achieved 99.65% accuracy, 98.36% sensitivity, and 97.83% precision, surpassing leading methodologies in the literature. These results demonstrate the significant potential of deep learning methodologies for automated early detection (Soni et al., 2024).

Warin and his team's study is significant because it compares the performance of deep learning models with human experts, such as oral and maxillofacial surgeons (OMFS) and general practitioners (GPs). In this study, conducted on clinical oral images (including Oral Potentially Malignant Disorders—OPMDs), the cross-validation results of the DenseNet-169 model achieved very high values of 99% sensitivity and 99% specificity for the OSCC class. For comparison, oral and maxillofacial surgeons achieved 90% sensitivity and 89% specificity, while general practitioners achieved 77% sensitivity and 87% specificity. These results demonstrate that DL algorithms perform better than or equivalent to human experts in the classification of OSCC and OPMDs (Warin et al., 2021-2024).

Alosaimi W. et al. used various DL models for the classification of histopathological images of OSCC. In experiments conducted on a dataset of 1224 histopathological images, the overall performance of the model was reported as 98.0% accuracy, 99.0% sensitivity, and 98.0% F1-score. The study highlighted that data augmentation techniques support DL models for OSCC histopathological images, providing high performance (Alosaimi et al., 2022).

Shetty SK. et al. combined deep transfer learning capabilities with duck-pack optimization to achieve promising results in the classification of OSCC histopathological images. The model achieved 97.3% accuracy and 96.4% F1-score on 1224 histopathological images. This optimization approach is considered an effective way to achieve high classification accuracy in histopathological images (Shetty et al.).

Myriam et al. used a CNN and Deep Belief Network (DBN) architecture optimized with Hybrid PSO-BER Optimization for OC detection. Tested on the Kaggle biomedical image dataset, the proposed hybrid algorithm demonstrated high classification performance with 97.35% accuracy. The overall evaluation shows that the proposed hybrid optimization significantly improves both accuracy and efficiency in OC detection using CNN and DBN (Myriam et al., 2023).

Das M. et al. utilized a custom-designed 10-layer CNN for the automatic detection of OSCC from histopathological images (1224 images). The model reported 97.0% accuracy, along with 98.0% sensitivity and 97.0% specificity. The high performance and low error rate (0.03) of the custom CNN architecture, particularly for histopathological analyses, demonstrated the effectiveness of the custom model design in this area (Das et al., 2023).

Fu et al.'s study used a DL algorithm (DenseNet/Cascading CNN) for OSCC detection in clinical images. Using 6176 images during development, the model achieved 92.3% overall accuracy, 95.0% sensitivity, and 0.98% AUC on the clinical validation dataset (401 images). This result is of great clinical significance, as it demonstrates that the model performs comparably to or better than the results achieved by OC specialists (sensitivity 91.7%, specificity 93.1%). This is compelling evidence that DL systems can compete with human experts for early diagnosis using clinical images (Fu et al., 2020). The most current techniques and methods used in the detection of oral cancer were examined in the study conducted by Warin (Warin et al. 2024).

A general review of the literature reveals that CNN-based deep feature extraction approaches exhibit high classification accuracy and reliable performance in OC detection. However, it has also been reported that these models still have room for improvement in terms of generalization ability, data diversity representing different clinical conditions, and sensitivity to early-stage lesions. In recent years, studies on the use of hybrid architectures, as well as deeper and optimized CNN variants, have increased. In particular, combining deep features obtained from different VGG architectures (VGG11, VGG13, VGG16, and VGG19) with machine learning-based classification algorithms has been noted to provide improved performance. Therefore, hybrid systems integrating CNN-based deep feature extraction with algorithms such as XGBoost, KNN, SVM, RF, and LGBM are considered an innovative and promising research direction for automatic, accurate, and early-stage OC detection. Such hybrid approaches enable

the development of explainable, scalable, and reliable solutions in clinical decision support systems by combining the high representation power of DL models with the interpretability of ML algorithms.

3. MATERIALS AND METHODS

In this section, detailed information about the OC dataset and methods used in the study is presented comprehensively.

3.1. Oral cancer data set

This study utilized the Oral Cancer Image Dataset, publicly available on the Kaggle platform (ZaidPy, 2022). This dataset, accessible to researchers on the Kaggle platform, consists of a total of 950 intraoral images, including 500 OC images and 450 non-cancer images. The images have been meticulously labeled by experts to facilitate classification, clearly distinguishing lesions with OC symptoms from healthy or benign tissue. The dataset is divided into two main classes:

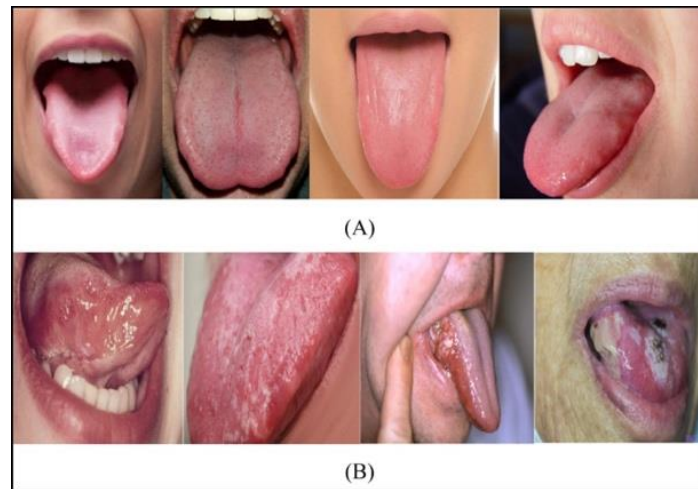


Figure 1: Sample images of OC classes: (A) Non-Cancer (B) Cancer.

Cancer (OC): Contains images with malignant oral tissue findings, including color changes, irregular surface structures, ulcerative formations, and other pathological findings.

Non-Cancer (Healthy/Benign): Contains images representing normal oral tissues or benign lesions without cancer findings.

The balanced distribution of the dataset provides a significant advantage in ML and DL-based classification studies, increasing the model's generalization ability and reducing the risk of overfitting. The collected images aim to contribute to the development of new methods for the early diagnosis of OC and

serve as a valuable resource for researchers in the field of health informatics. Sample images from the dataset are depicted in Fig. 1 (Liu et al., 2025).

3.2. Processed methods

This subsection presents the methodological steps of the proposed hybrid transfer learning and ML-based OC classification method in detail. First, the layer structure configurations of four different CNN-based VGG architectures (VGG11, VGG13, VGG16, and VGG19) designed for deep feature extraction from intraoral images are presented in Table 1. This table demonstrates the theoretical basis for the proposed system's multilayer feature representation by comparing the number of convolution blocks, pooling layers, and fully connected layers for each model. The proposed system architecture has a multistage processing flow, and the general operation scheme is presented in Fig. 2.

Table 1: VGG models configuration for the proposed hybrid system.

Layer Block	VGG-11	VGG-13	VGG-16	VGG-19
Input	224×224×3	224×224×3	224×224×3	224×224×3
Block 1	Conv3-64 MaxPool	Conv3-64, Conv3-64 MaxPool	Conv3-64, Conv3-64 MaxPool	Conv3-64, Conv3-64 MaxPool
Block 2	Conv3-128 MaxPool	Conv3-128, Conv3-128 MaxPool	Conv3-128, Conv3-128 MaxPool	Conv3-128, Conv3-128 MaxPool
Block 3	Conv3-256 MaxPool	Conv3-256, Conv3-256 MaxPool	Conv3-256, Conv3-256, Conv3-256 MaxPool	Conv3-256, Conv3-256, Conv3-256, Conv3-256 MaxPool
Block 4	Conv3-512, Conv3-512 MaxPool	Conv3-512, Conv3-512 MaxPool	Conv3-512, Conv3-512, Conv3-512 MaxPool	Conv3-512, Conv3-512, Conv3-512, Conv3-512 MaxPool
Block 5	Conv3-512, Conv3-512 MaxPool	Conv3-512, Conv3-512 MaxPool	Conv3-512, Conv3-512, Conv3-512 MaxPool	Conv3-512, Conv3-512, Conv3-512, Conv3-512 MaxPool
Resulting Feature Layer	FC-4096 → FC-4096 → FC-1000	FC-4096 → FC-4096 → FC-1000	FC-4096 → FC-4096 → FC-1000	FC-4096 → FC-4096 → FC-1000

In the first stage, a dataset consisting of oral cancer and non-cancer images is input to the system. After undergoing basic preprocessing steps such as resizing and normalization, the images are fed separately into the four different VGG architectures. VGG-based CNN structures learn local spatial patterns such as color, texture, and edge features of lesions through convolution and pooling operations applied across

layers, resulting in high-dimensional deep feature vectors for each model. Following the feature extraction phase, the resulting deep feature vectors are used as input to ML-based classification algorithms. This study evaluated five algorithms offering different decision boundary structures and generalization capabilities: XGBoost, kNN, SVM, RF, and LGBM. The features extracted from each VGG models (VGG11, VGG13, VGG16, and VGG19) were trained separately with these classifiers to predict cancer/non-cancer classes. Thus, the contribution of both the deep feature extraction phase and the classification algorithm to the overall performance was analyzed in detail. This hybrid approach aims to provide a more stable and explainable classification framework for oral cancer detection by combining the high representation power of DL models with the flexible decision structures of ML algorithms (Kaya et al., 2023).

The performance of the proposed system was evaluated using standard metrics commonly used in classification problems. Accuracy, recall, precision, F-score, and Area Under the Receiver Operating Characteristic Curve (AUC) metrics were used for performance analysis. These metrics allow for a multidimensional assessment of the system's ability to accurately distinguish between OC and healthy samples. In particular, the AUC value reveals the model's overall discrimination under different threshold values, while the F-score reflects the balance between sensitivity and precision in unbalanced class distributions.

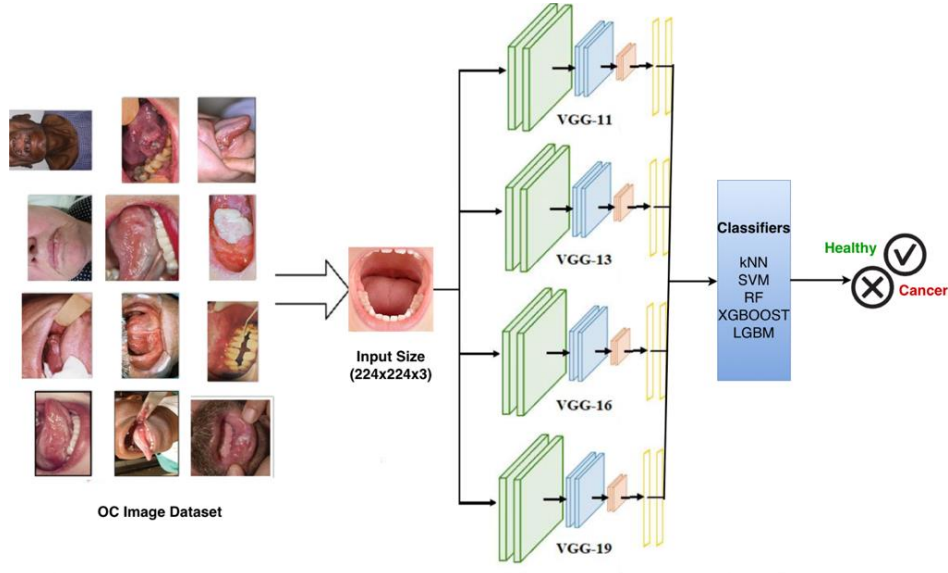


Figure 2: Pipeline of the proposed hybrid system for OC detection.

Summary information on the performance evaluation metrics is presented in Eq. (1-4) and Fig. 3, demonstrating that the proposed multistage hybrid architecture has strong potential to contribute to clinical decision support systems for the automated and highly accurate detection of OC.

		Predict	
		0	1
Actual	0	TN	FP
	1	FN	TP

TP: True Positive
FP: False Positive
FN: False Negative
TN: True Negative

Figure 3: Performance evaluation criteria for the proposed hybrid OC detection system.

$$Accuracy = \left(\frac{(TP + TN)}{(TP + FP + FN + TN)} \right) \quad (1)$$

$$Precision = \left(\frac{(TP)}{(TP + FP)} \right) \quad (2)$$

$$F - Score = 2x \left(\frac{(TP)}{(TP + FP + FN)} \right) \quad (3)$$

$$Recall = \left(\frac{(TP)}{(TP + FN)} \right) \quad (4)$$

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the experimental results of the developed hybrid approach for the automatic classification of OC and non-cancer images are analyzed and presented in Table 2.

The proposed method is based on the separate use of four different CNN architectures (VGG11, VGG13, VGG16, and VGG19) to extract deep features from oral cavity images. Each VGG model learns hierarchical and discriminative local spatial features, enabling the representation of lesion-level texture, color, and structural patterns associated with oral cancer. The extracted deep feature vectors are subsequently classified using five machine learning algorithms, including XGBoost, KNN, SVM, RF, and LGBM.

The performance evaluation results in Table 3 depict that deep features obtained from four different VGG architectures (VGG11, VGG13, VGG16, and VGG19) combined with five ML classifiers (XGBoost, SVM, RF, LGBM, and kNN) provide high accuracy in detecting OCs. When all metrics are examined together, it is clear that the SVM algorithm is the most consistent and provides the highest performance in terms of accuracy, precision, recall, F-score, and AUC. In particular, with VGG13, VGG16, and VGG19, SVM demonstrated balanced superiority in both sensitivity (recall) and precision (precision), directly contributing to the increase in F-score. This demonstrates the model's success in both correctly

recognizing positive classes and reducing the false positive rate. The complexity matrix, ROC curve, and precision-recall curve of the VGG19+SVM hybrid model that provides the highest performance is illustrated in Fig. 4.

Table 2: Performance results for all experiments in the proposed hybrid system for OC detection

	Algorithms	Performance Evaluation Metrics				
		Accuracy	Precision	Recall	F-Score	AUC (ROC)
VGG11	XGBoost	0.8789	0.8785	0.8789	0.8787	0.8789
	SVM	0.9053	0.9077	0.9033	0.9046	0.9033
	RF	0.9000	0.8996	0.9000	0.8998	0.9000
	LGBM	0.8895	0.8900	0.8883	0.8890	0.8883
	kNN	0.8947	0.8944	0.8944	0.8944	0.8944
VGG13	XGBoost	0.9105	0.9123	0.9089	0.9100	0.9089
	SVM	0.9158	0.9201	0.9133	0.9150	0.9133
	RF	0.8526	0.8543	0.8506	0.8516	0.8506
	LGBM	0.8737	0.8746	0.8722	0.8730	0.8722
	kNN	0.8632	0.8656	0.8656	0.8632	0.8656
VGG16	XGBoost	0.8737	0.8746	0.8722	0.8730	0.8722
	SVM	0.9105	0.9123	0.9089	0.9100	0.9089
	RF	0.8526	0.8543	0.8506	0.8516	0.8506
	LGBM	0.8737	0.8757	0.8717	0.8728	0.8717
	kNN	0.8632	0.8639	0.8617	0.8624	0.8617
VGG19	XGBoost	0.9105	0.9101	0.9106	0.9103	0.9106
	SVM	0.9158	0.9170	0.9144	0.9153	0.9144
	RF	0.8895	0.8910	0.8878	0.8888	0.8878
	LGBM	0.9158	0.9170	0.9144	0.9153	0.9144
	kNN	0.9053	0.9050	0.9061	0.9052	0.9061

When the three graphs in Fig. 4 of the VGG19+SVM model are evaluated together, it is seen that the model exhibits strong and balanced performance in OC detection. The confusion matrix demonstrates that the model successfully distinguishes two classes with high correct classification rates. The ROC curve AUC value of 0.9153 confirms the model's overall high discrimination capacity, while the Precision–Recall curve's PR AUC value of 0.9144 demonstrates both high precision and strong sensitivity in detecting the positive class. These results demonstrate that the VGG19+SVM combination provides reliable classification performance in OC detection.

Comparing the VGG architectures, the VGG19 architecture produces the strongest results across all metrics. VGG19's deeper convolutional layers allowed it to produce more detailed and discriminative representation vectors of lesions' color variations, edge structure, and texture characteristics; this was particularly evident in the precision and AUC metrics. The combinations of VGG13-19+SVM and VGG19+LGBM achieved the highest performance with an accuracy value of 0.9158, demonstrating a strong synergy between deep feature extraction and gradient boosting-based and margin-maximization-based classifiers.

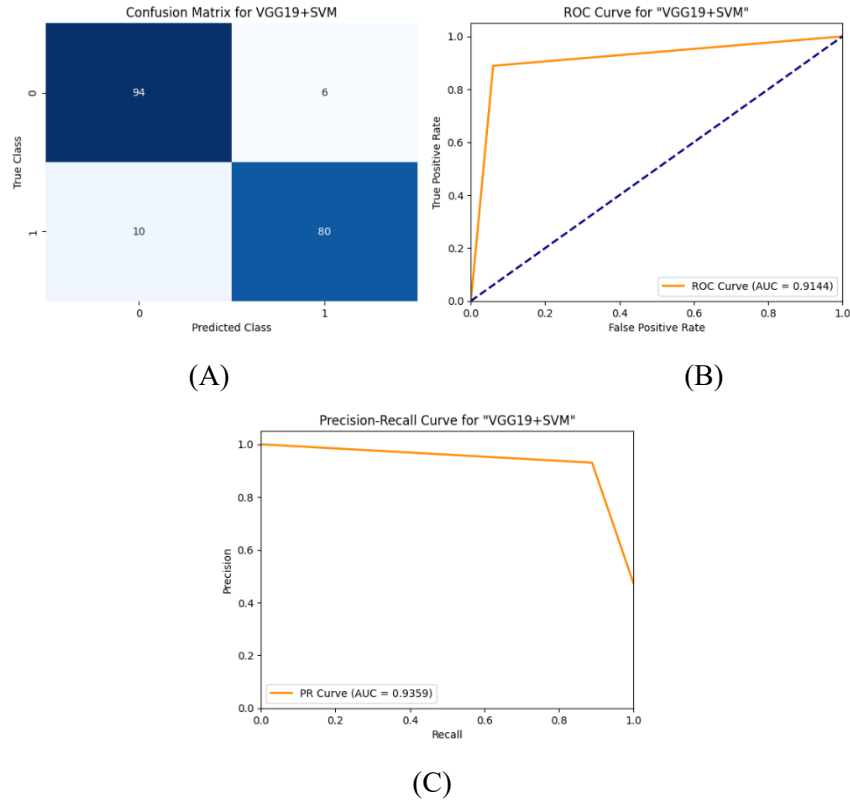


Figure 4: (A) Confusion Matrix, (B) ROC Curve, and (C) Precision-Recall Curve of the VGG19+SVM hybrid model applied for OC detection

In an inter-algorithm comparison, XGBoost demonstrated competitive performance in terms of accuracy and AUC, while kNN produced more stable but moderate results. RF, on the other hand, performed lower than the other algorithms, particularly in precision and recall. This may be attributed to the occasional overgeneralization of tree-based splits in high-dimensional feature vectors. In contrast, the high F-score and AUC values of SVM and LGBM demonstrate that they can further optimize both classification accuracy and discrimination between positive and negative classes. Overall, evaluation of accuracy, precision, recall, F-score, and AUC metrics demonstrates that the proposed hybrid approach offers

balanced and consistent performance in OC detection, with both high sensitivity and low misclassification rates. This comprehensive success profile strongly supports the system's potential as a reliable and applicable solution for clinical decision support mechanisms for early diagnosis.

The results obtained in this study are compared with previously implemented experimental results. A comparison of the methods used and the resulting performance metrics is presented in Table 3. A review of studies in the literature reveals that the most common approaches for OC detection are deep learning-based CNN models and the processing of features derived from these models with machine learning classifiers. In this regard, the VGG-based deep feature extraction implemented in our study, followed by the use of classical machine learning algorithms such as SVM, LGBM, and XGBoost, fully parallels these approaches in the literature. Furthermore, the 92% accuracy rate achieved is considered an acceptable and robust performance compared to results reported in the existing literature. This result demonstrates that the model provides a compatible and valid approach to current AI-based OC detection studies, both in terms of structure and performance level.

Table 3: Comparison of the results obtained with other studies

Author(s)	Data set	Purpose	Type	Performance Metrics
Warin et al., 2021	980 oral images	Classification and detection of oral squamous cell carcinoma (OSCC) and potential malignant lesions with CNN-based models	Classification/Detection	With DenseNet-169 model, AUC = 1.00 (OSCC), 0.98 (OPMD); With Faster R-CNN, AUC = 0.88.
Gomes et al., 2023	5069 oral lesion image set (6 lesion types)	Automatic classification of 6 basic lesion types in oral mucosa using CNN	Classification	Accuracy: 95.09%
Liu et al., 2025	Lip and tongue images	Early-stage oral cancer detection with 19-layer CNN	Classification	Accuracy: 99.54%, Sensitivity: 95.73%, Specificity: 96.21%
Das et al., 2023	Histopathological images of oral mucosa	Developing a CNN-based model for automatic detection of OSCC	Detection	Accuracy: 97.82%
This study	Oral lesion image set (500 cancer and 450 non-cancer images)	Oral cancer detection with deep feature extraction (VGGs) + ML classifiers (XGBoost, SVM, RF, LGBM, kNN)	Classification	Accuracy: VGG13 + SVM $\approx$ 0.92, VGG19 + SVM $\approx$ 0.92, VGG19 LGBM $\approx$ 0.92

5. CONCLUSION

The integration of AI and DL methods in OC diagnosis stands out as a transformative change in the search for solutions to this critical public health problem. Because OC (especially OSCC) is an aggressive disease with high morbidity and mortality rates, the vital importance of early diagnosis has been

reaffirmed. The inherent limitations of classical diagnostic methods (visual examination and invasive histopathology), such as subjectivity, time-consuming nature, and dependence on experienced specialists, have increased the need for DL-based automated systems.

In this study, a hybrid classification approach is proposed for the automatic detection of OC, using deep features obtained from VGG11, VGG13, VGG16, and VGG19 architectures combined with five different ML algorithms (SVM, XGBoost, LGBM, RF, and kNN). Experimental results show that, although all models demonstrate high classification performance, the VGG19 model, in particular, produces more discriminable features due to its deeper structure and achieves the highest success when combined with SVM. All of the obtained accuracy, precision, recall, F-score, and AUC values demonstrate that the proposed system distinguishes both positive (cancer) and negative (non-cancer) classes in a balanced manner. The high AUC values of the ROC and precision–recall curves confirm the stable performance of the model at different thresholds, while the confusion matrix confirms the low misclassification rates. These comprehensive analyses demonstrate that combining VGG-based deep feature extraction with powerful classifiers such as SVM and LGBM provides an effective and reliable solution for detecting OC, and one that can be integrated into clinical decision support systems. The findings allow the hybrid approach to be considered a powerful AI model that can contribute to early diagnosis.

Critical steps are being taken for AI to achieve clinical utility in OC diagnosis. To ensure the reliability and transparency of model predictions, the integration of methods such as Explainable Artificial Intelligence (XAI) and Case-Based Reasoning (CBR) increases clinicians' confidence in system decisions. For example, OCANet's ability to provide both visual and textual explanations using Large Language Models (LLMs) such as GPT-5 supports the interpretability and reliability of automated predictions. However, despite the successful results achieved, larger and more diverse clinical datasets are needed to fully integrate DL algorithms into clinical routines, and the generalizability and robustness of these models across diverse populations and clinical scenarios must be validated. These studies are expected to fully realize the potential of DL, accelerate early-stage OC detection, and ultimately significantly improve survival rates for OC patients.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest among the authors.

CONTRIBUTIONS OF AUTHORS

V.C.B.: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing— review and editing and original draft preparation.

REFERENCES

- Alosaimi, W., & Uddin, M. I. (2022). Efficient data augmentation techniques for improved classification in limited dataset of oral squamous cell carcinoma. *Computer Modeling in Engineering & Sciences*, 131, 1387–1401.
- Chang, X., Yu, M., Liu, R., Jing, R., Ding, J., Xia, J., Zhu, Z., Li, X., Yao, Q., Zhu, L., & Zhang, T. (2023). Deep learning methods for oral cancer detection using Raman spectroscopy. *Vibrational Spectroscopy*, 126, 103522. <https://doi.org/10.1016/j.vibspec.2023.103522>
- Das, M., Dash, R., & Mishra, S. K. (2023). Automatic detection of oral squamous cell carcinoma from histopathological images of oral mucosa using deep convolutional neural network. *International Journal of Environmental Research and Public Health*, 20(3), 2131. <https://doi.org/10.3390/ijerph20032131>
- Fu, Q., Chen, Y., Li, Z., Jing, Q., Hu, C., Liu, H., ... & Xiong, X. (2020). A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study. *EClinicalMedicine*, 27, 100558. <https://doi.org/10.1016/j.eclinm.2020.100558>
- Gomes, R. F. T., Schmith, J., Figueiredo, R. M. D., Freitas, S. A., Machado, G. N., Romanini, J., & Carrard, V. C. (2023). Use of artificial intelligence in the classification of elementary oral lesions from clinical images. *International journal of environmental research and public health*, 20(5), 3894.
- Huang, Q., Ding, H., & Razmjooy, N. (2023). Optimal deep learning neural network using ISSA for diagnosing the oral cancer. *Biomedical Signal Processing and Control*, 84, 104749. <https://doi.org/10.1016/j.bspc.2023.104749>
- Huang, Q., Ding, H., & Razmjooy, N. (2024). Oral cancer detection using convolutional neural network optimized by combined seagull optimization algorithm. *Biomedical Signal Processing and Control*, 87, 105546. <https://doi.org/10.1016/j.bspc.2023.105546>
- Kabir, M. F., Ahmad, M. Y., Uddin, R., Cordero, M., & Kant, S. (2025). Accurate and lightweight oral cancer detection using SE-MobileViT on clinically validated image dataset. *Discover Artificial Intelligence*, 5, 173. <https://doi.org/10.1007/s44163-025-00442-2>
- Kaya, V., & Akgül, İ. (2023). Classification of skin cancer using VGGNet model structures. *Gümüşhane Üniversitesi Fen Bilimleri Dergisi*, 13(1), 190-198. <https://doi.org/10.17714/gumusfenbil.1069894>
- Li, L., Pu, C., Tao, J., Zhu, L., Hu, S., Qiao, B., Xing, L., Wei, B., Shi, C., Chen, P., & Zhang, H. (2024). Development of an oral cancer detection system through deep learning. *BMC Oral Health*, 24, 1468. <https://doi.org/10.1186/s12903-024-05195-5>
- Liu, P., & Bagi, K. (2025). A tailored deep learning approach for early detection of oral cancer using a 19-layer CNN on clinical lip and tongue images. *Scientific Reports*, 15, 23851. <https://doi.org/10.1038/s41598-025-07957-9>
- Myriam, H., Abdelhamid, A. A., El-Kenawy, E. S. M., Ibrahim, A., & others. (2023). Advanced meta-heuristic algorithm based on particle swarm and Al-Biruni Earth radius optimization methods for oral cancer detection. *IEEE Access*, 11, 23681–23700. <https://doi.org/10.1109/ACCESS.2023.3253430>
- Prabhakaran, R., & Mohana, J. (2020). Detection of oral cancer using machine learning classification methods. *International Journal of Electrical Engineering and Technology*, 11(3), 384–393. <https://ssrn.com/abstract=3638829>
- Shetty, S., & Patil, A. (2023). Duck pack optimization with deep transfer learning-enabled oral squamous cell carcinoma classification on histopathological images. *International Journal of Grid and High Performance Computing*, 15, 1–21.

- Soni, A., Sethy, P. K., Dewangan, A. K., Nanthaamornphong, A., Behera, S. K., & Devi, B. (2024). Enhancing oral squamous cell carcinoma detection: A novel approach using improved EfficientNet architecture. *BMC Oral Health*, 24, 601
- Shah, S. J. H., Albishri, A., Wang, R., & Lee, Y. (2025). Integrating local and global attention mechanisms for enhanced oral cancer detection and explainability. *Computers in Biology and Medicine*, 189, 109841. <https://doi.org/10.1016/j.combiomed.2025.109841>
- Warin, K., & Suebnukarn, S. (2024). Deep learning in oral cancer: A systematic review. *BMC Oral Health*, 24, 212. <https://doi.org/10.1186/s12903-024-03993-5>
- Warin, K., Limprasert, W., Suebnukarn, S., Jinaporntham, S., & Jantana, P. (2021). Automatic classification and detection of oral cancer in photographic images using deep learning algorithms. *Journal of Oral Pathology & Medicine*, 50(9), 911–918. <https://doi.org/10.1111/jop.13227>
- Welikala, R. A., Remagnino, P., Lim, J. H., Chan, C. S., Rajendran, S., Kallarakkal, T. G., Zain, R. B., Jayasinghe, R. D., Rimal, J., Kerr, A. R., Amtha, R., Patil, K., Tilakaratne, W. M., Gibson, J., Cheong, S. C., & Barman, S. A. (2020). Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. *IEEE Access*, 8, 132677–132693. <https://doi.org/10.1109/ACCESS.2020.3010180>
- Yaduvanshi, V., Murugan, R., & Goel, T. (2025). Automatic oral cancer detection and classification using modified local texture descriptor and machine learning algorithms. *Multimedia Tools and Applications*, 84(2), 1031–1055. <https://doi.org/10.1007/s11042-024-19040-y>
- ZaidPy. (2022). Oral Cancer Dataset [Data set]. Kaggle. <https://www.kaggle.com/datasets/zaidpy/oral-cancer-dataset>

Endüstriyel Dişli Kutularının Mekanik Veriminin Operasyonel Parametrelere Göre Deneysel İncelemesi

Umut Murat GÖK^{1*}, Osman Taha ŞEN¹, İbrahim Mehmet PALABIYIK¹

¹Makina Mühendisliği Bölümü, Makina Fakültesi, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

Geliş: 15.09.2025, Kabul: 05.12.2025, Yayınlanma: 23.12.2025

ÖZ

Dişli kutuları, otomotiv, savunma, havacılık ve uzay sanayii başta olmak üzere birçok alanda yaygın biçimde kullanılan bir alt sistem niteliği taşımaktadır. Güç aktarımı ve kontrol amaçlı uygulamalarda kritik rol oynayan dişli sistemlerden kaynaklanan hata ve hasarların önceden kestirimi mühendislik yapıtlarının güvenilirliği açısından kritik önem arz etmektedir. Dişli kutuları, yatak, dişli mekanizması, mil, sızdırmazlık elemanı, gövde gibi farklı bileşenlerden oluşmaktadır ve montaj halinde söz konusu alt bileşenler ile ilgili hataların tekil ve kombinasyon halindeki formları ile karşılaşılabilir. Mühendislik tasarımlarının hafif yapılara dönüştüğü ve güç/kütle parametresinin arttığı teknolojik gelişim sürecinde, geliştirme sürecindeki deneysel çalışmaların önemi giderek artmaktadır. Çalışmada, endüstriyel dişli kutularının dinamik davranışını inceleme kabiliyetine sahip özgün bir deney platformu geliştirilmiş ve mekanik verimin operasyonel parametrelere göre incelemesi gerçekleştirilmiştir. Yük, açısız hız ve yağlayıcı sıcaklığının mekanik verim üzerindeki etkilerinin irdelendiği deneysel çalışmanın gelecek aşamalarında mekanik verimin hatalar ile ilişkisi üzerine araştırmalar yapılacaktır.

Anahtar Kelimeler: Dişli kutusu; deneysel mekanik; mekanik verim; parametrik inceleme

Experimental Investigation of the Mechanical Efficiency on Industrial Gearboxes According to Operational Parameters

ABSTRACT

Gearboxes are widely used as subsystems in many sectors, particularly automotive, defense, aerospace, and other industries. Error detecting and damage predicting caused by geared systems, which play important role in power transmission and control applications, have critical importance for the reliability of engineering systems. Gearboxes consist of different components such as bearing, gear mechanism, shaft, sealing and housing. Singular and/or combined forms of these error can be observed in assembled gearboxes. In the technological development process where engineering designs are transformed into lightweight structures, experimental studies become much more important. In this study, a novel experimental platform, capable of examining the dynamic behavior of industrial gearboxes, was developed and the investigation of the effects of operating parameters on mechanical efficiency was carried out. In the future phase of the experimental study the correlation between mechanical efficiency and errors will be investigated.

Keywords: Gearbox; experimental mechanics; mechanical efficiency; parametric investigation

1. GİRİŞ

Dişli kutuları, otomotiv, havacılık, savunma sanayi ve endüstriyel uygulamalar olmak üzere birçok alanda hız-moment dönüşümü amacıyla kullanılmaktadır. Tarihsel dönemdeki kullanımı, doğanın gözlemi ve zamanın belirlenmesi amacıyla Mısır uygarlığına kadar dayanan dişli sistemler (Matschoß 1940), günümüzde başta havacılık olmak üzere birçok teknoloji alanında kritik alt sistem olma niteliğini korumakta ve halen güncel araştırma konuları arasında bulunmaktadır.

A.B.D. Ulusal Ulaşım Güvenliği Kurulu (NTSB) tarafından yürütülen kapsamlı bir çalışmada, 1963 yılı ortasından 1997'ye kadar 8436 döner kanatlı hava aracı sivil kaza-kırım raporlarına dayanarak kapsamlı bir bilgi veri tabanı geliştirilmiştir. Kaza sayısı ve dağılımı incelendiğinde, farklı konfigürasyonlar için güç aktarma sistemlerinden kaynaklanan arızaların yaklaşık %35 olduğunu görülmektedir. Güç aktarma organı temelli kaza-kırım olayları, tek pistonlu için %37,79, tek türbinli için %36,4 ve çift türbinli motor konfigürasyonları için %35,9 mertebelerine denk gelmektedir (Harris, Kasper ve Iseler 2000). Bununla birlikte, dişli kutularındaki hasarın çoğunlukla (%65) dişli çiftlerinden kaynaklandığı bildirilmektedir (Handbook of Loss Prevention, Allianz Versicherungs-AG 1978).

Havacılık ve uzay sanayii motor tasarımının geleceğinde, geniş itki kuvveti aralığında optimum güç sağlayan, büyük çapta fan boyutları ve yüksek verime sahip sistemler ön plana çıkmaktadır ve bu sistemlerin önemli bir parçası dişli sistemlerdir (Rolls-Royce revs up world's most powerful aerospace gearbox. Designfax 2016).

Dişli kutuları kullanımı, klasik yaklaşım olan mekanik güç iletiminin yanı sıra teknolojinin gelişimi ile birlikte, konum kontrolü amacıyla tıbbi ve elektronik cihazlarda da yaygınlaşmaktadır.

Özet olarak, tarihsel geçmişi oldukça uzun yıllara dayalı olan dişli sistemler, günümüz teknolojik şartlarında kritik rol oynamaya devam etmektedir. Temel olarak, yataklar, dişli çiftleri ve miller gibi temel makina elemanlarını ihtiva eden dişli kutuları için matematik modelleme çalışmaları önemli zorlukları beraberinde getirmektedir. İdeal koşullarda, sistemin matematiksel modeli oluşturulurken dişlilerin, millerin ve yatakların tüm eğilme ve burulma rijitlikleri dikkate alınmalıdır (Özgüven ve Houser 1988). Bununla birlikte, dişli çiftlerinin eş çalışması, matematiksel modellerin giderek daha karmaşık hale gelmesine; hareket denklemlerinin ve analitik çözümlerinin elde edilmesinin zorlaşmasına neden olmaktadır. Bu gerekçeler ışığında, deneysel araştırmalar, dişli sistemlerinin temel dinamiğini, titreşim-akustik karakteristiklerini araştırmak ve matematiksel modelleri doğrulamak için faydalı veriler üretmektedir.

Bu çalışmanın ana amacı çeşitli çalışma parametrelerinin dişli kutularının mekanik verimi üzerindeki etkilerini deneysel olarak incelemektir. Bu amaçla ilk olarak laboratuvar ortamında kontrollü bir test sistemi kurulmuştur ve bu test sistemi ilgili sensörler ile donatılmıştır. Daha sonra, etkisi incelenmesi istenen

çalışma parametreleri belirlenmiştir, testler belirlenen çalışma parametrelerinin belirli değerlerinde yürütülmüştür ve test sistemi üzerindeki sensörlerden veriler toplanmıştır. Sensörlerden toplanan veriler işlenerek, her test senaryosu için üzerinde testler yapılan dişli kutusunun mekanik verimi hesaplanmıştır. Makalenin kapsamı tek kademeli dişli kutuları ile sınırlı tutulmuştur. Ancak test sistemi, yeni güncellemeler ile çok kademeli vites kutularının da test edilebileceği bir yapıda tasarlanmıştır. Ayrıca, çalışma kapsamında belirlenen çalışma parametreleri dişli kutusu giriş açısal hızı, yük ve yağ sıcaklığı olarak belirlenmiştir.

2. TEST SİSTEMİ TASARIMI

Bu çalışmada, dişli kutusunu gerçek çalışma şartlarında (açısal hız, yük) test edebilme yeteneğine sahip özgün bir deney platformu geliştirilmiş ve deneyler esnasında ölçülen dijital sinyaller, zaman ve frekans tabanlarında kaydedilmiştir. Geliştirilen deney platformunda, farklı çevrim oranlarına sahip dişli kutuları, yük altında tahrik edilerek veriler toplanmıştır. Test sistemi üzerinden toplanan sinyaller aşağıda listelenmiştir:

1. Dişli kutusu giriş yatağı üzerinden 3 eksen de ivme
2. Dişli kutusu çıkış yatağı üzerinden 1 eksen de ivme
3. Dişli kutusuna belirli bir mesafeden ses basıncı
4. Dişli kutusu yağ sıcaklığı
5. Dişli kutusu giriş mili üzerinden moment
6. Dişli kutusu giriş mili üzerinden açısal deplasman
7. Dişli kutusu çıkış mili üzerinden moment
8. Dişli kutusu çıkış mili üzerinden açısal deplasman

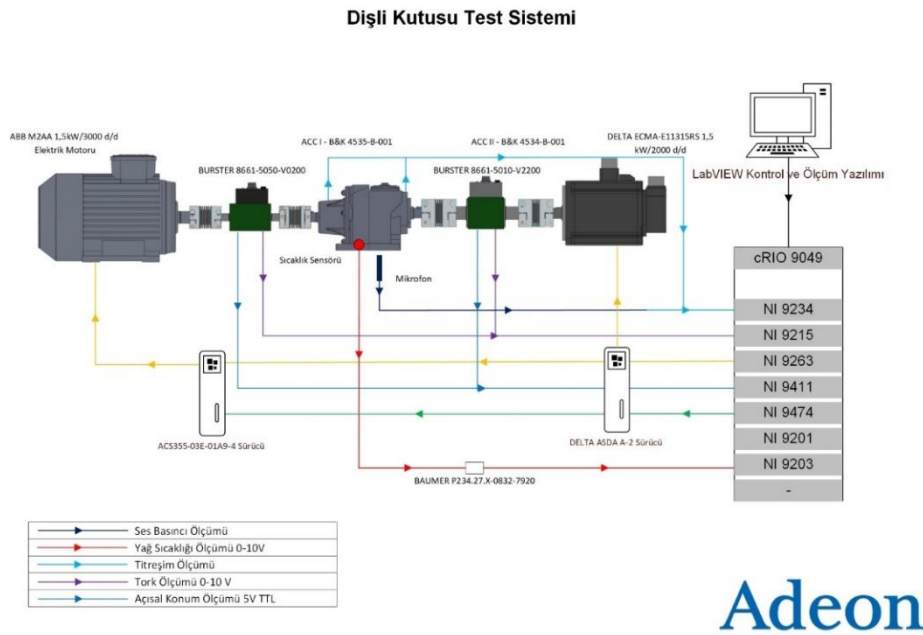
Sonuç olarak veriler toplam 10 kanal üzerinden eş zamanlı olarak kaydedilmiştir ve kaydedilen veriler geliştirilen MATLAB kodu ile işlenmiştir. Deneyler, %25, %50, %75 ve %100 olmak üzere, 4 farklı yük kademesinde, 0-2800 d/d dönme hızı bandında 100 d/d aralıklar ile gerçekleştirilmiştir. Yağ sıcaklığının dinamik davranışa etkilerinin görülmesi ve verilerin sınıflandırılabilmesi için deney öncesi yağ sıcaklığı otomatik olarak koşullandırılmıştır ve veriler 15-30 °C bandında 5 °C aralıklı olarak toplanmıştır. Sonuç olarak 3 farklı çalışma parametresinin (yük, dişli kutusu giriş hızı ve dişli kutusu yağ sıcaklığı) farklı değerlerinde toplam 320 farklı çalışma parametresi setinde testler gerçekleştirilmiştir. Yük motoru sürücü kısıtlamalarından dolayı bütün hızlarda tam yük altında çalıştırılamamıştır. Ayrıca deneyler farklı çevrim oranlarına sahip 3 farklı dişli kutusu üzerinde tekrar edilmiştir.

İlk aşama deneylerinde her çevrim oranında 3 adet olmak üzere, 3 farklı çevrim oranında toplam 9 dişli kutusu ile kalite aralığı belirlenmiştir. Deneyler sırasında yapılan öznel değerlendirmelerde dahi, iyi durumda bilgisi ile imalatçı kuruluş tarafından sağlanan kullanılmamış dişli kutularında dahi belirgin

titreşim ve gürültü farklılıkları saptanmıştır. Bununla birlikte, mekanik verimdeki ayrışmanın da imalat kalitesi ile ilişkili olduğu görülmüştür. Makale kapsamında, mekanik verimin değişimi detaylı olarak incelenmesinin temel motivasyonu da görülen bu anlamlı ilişkidir.

2.1. Test platformu

Literatür araştırması ve kavramsal tasarım çalışmaları ışığında deney platformu açık sistem konfigürasyonunda tasarlanmıştır. Tasarlanan test sistemi, sensörler haricinde temel olarak iki elektrik motoru ve bir adet dişli kutusundan oluşmaktadır. Testler esnasında kullanılan dişli kutularından bazılarında 2 bazılarında ise 3 indirgeme mevcuttur. Dişli kutusu giriş miline bağlı olan elektrik motoru sistemi sürerken, dişli kutusu çıkış miline bağlı olan elektrik motoru ise sistemi yüklemektedir. Tahrik, yük motorları ve diğer enstrümantasyon öğeleri test sisteminin sabit bileşenlerini, dişli kutusu ise sistemdeki test objesini oluşturmaktadır. Sistemin, ölçüm ve kontrol bileşenleri arasındaki veri akışı Şekil 1’de verilen şematik diyagramda görülmektedir.



Şekil 1: Deney platformu şematik diyagramı

2.2. Elektromekanik tasarım

Test sistemi, dişli kutusu girişinde hız çıkışında ise moment kontrolü olacak şekilde tasarlanmıştır. Bu sayede dinamik olarak belirlenen test senaryolarında test edilebilmiştir.

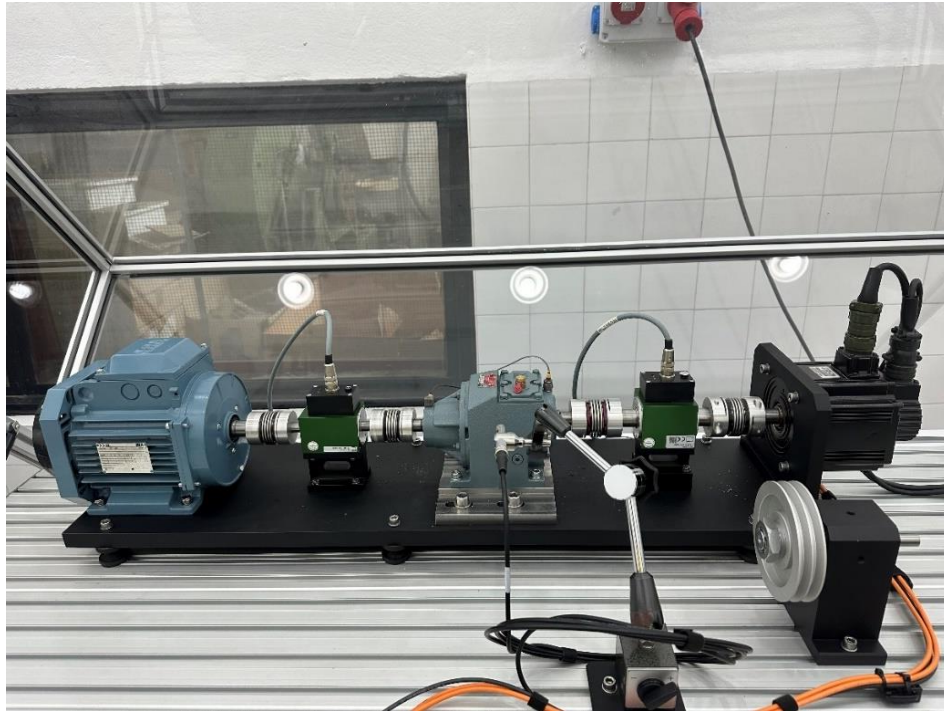
Mekanik verimin ölçüme dayalı olarak hesaplanabilmesi için dişli kutularının giriş ve çıkışında moment ve açısal konum ölçümü yapabilen sensörler yerleştirilmiştir. Toplanan moment ve açısal hız verileri, mekanik

gücün ve dolayısı ile mekanik verimin dinamik şartlar altında hesaplanması için kullanılmakla birlikte sistemin kontrolü için anlık yük ve açısal hız bilgilerini de sağlamaktadır.

Dişli sistemlerde yağ sıcaklığına bağlı olarak viskozitede önemli değişimler gerçekleşmektedir (Doleschel, Höhn ve Michaelis 2000). Yağ viskozitesinin değişimine bağlı olarak, ivme, ses basıncı, moment ve mekanik verimde ciddi farklılaşmalar görülmektedir. Bu nedenle, karşılaştırmalı analizin sağlıklı yapılabilmesi için yağ sıcaklığının koşullandırılması kritik önem arz etmektedir. Geliştirilen sistem arayüzünde, testin gerçekleştirileceği yağ sıcaklığı parametresi girilerek yağın otomatik olarak koşullandırılması yapılmaktadır.

Sistem sensör ve ölçme enstrümanasyonu açısından esnek ve geliştirilebilir bir yapıda geliştirilmiştir. NI cRIO 9049 kontrolcü üzerinden farklı veri toplama ve sinyal modülleri ile kontrol edilen sistem, gerçek zamanlı açısal hız/moment kontrolü yapabilmekte ve dinamik verileri senkronize olarak kaydedebilmektedir.

Sistem, mekanik yapı, kontrol sistemi ve test odası olmak üzere üç ana bileşenden oluşmaktadır. Test odası dış ortamdan tamamen izole edilerek bir kabin içerisinde konumlandırılmıştır. Böylece arka plan gürültüsü ve bozucu etkenlerin ölçüm sonuçlarını etkilemesinin önüne geçilmiştir. (Şekil 2)



Şekil 2: Geliştirilen deney platformunun genel görünümü

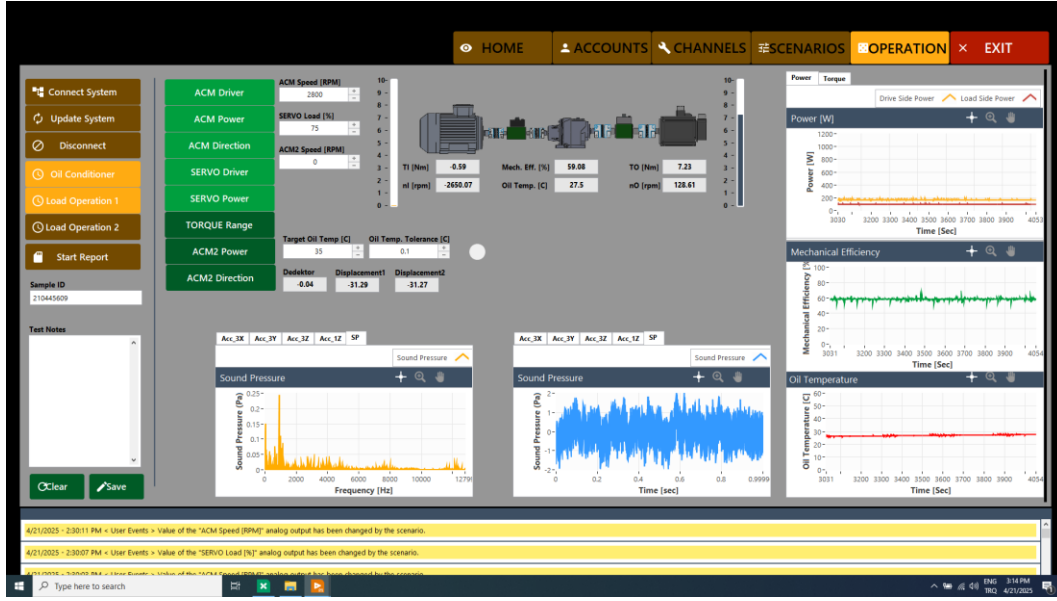
Daha önce de belirtildiği gibi, dişli kutularının mekanik veriminin deneysel verilere dayalı olarak hesaplanabilmesi ve sistemin açısal hız ve moment kontrolü için geri besleme sağlanabilmesi için dişli

kutusu girişinde ve çıkışında moment/açısal deplasman sensörlerinin kullanılması gerekmektedir. Bu nedenle, dişli kutusu girişinde BURSTER 8661-5010-V2200 ve çıkışında BURSTER 8661-5050-V0200 marka sensörler kullanılmıştır. Dişli kutusu girişinde ve çıkışında farklı sensör modellerinin kullanılmasının nedeni, dişli kutusunun mekanik indirgemesinden dolayı girişteki ve çıkıştaki moment ve açısal hızların farklı (girişte düşük moment yüksek açısal hız ve çıkışta yüksek moment ve düşük açısal hız) olmasıdır. Bu nedenle, kısmi yüklerde ölçüm belirsizliğini düşürmek amacı ile çift ölçekli (2 Nm ve 10 Nm), üst sınırı 10 Nm olan BURSTER 8661-5010-V2200 moment sensörü seçilmiştir. Ayrıca, her ne kadar mevcut sistemin fiziksel yük sınırı 10 Nm olsa da gelecekte yapılacak çalışmalar için bir altyapı oluşturması amacı ile çıkış mili üzerinde üst sınırı 50 Nm olan bir moment sensörü tercih edilmiştir. Bu sensörlerin açısal hız bakımından üst sınırı 15.000 d/d ve hassasiyeti %0,2'dir. Her iki sensörün de açısal konum ölçme kabiliyetine sahip 4096ppr çözünürlüğünde dahili enkoderi bulunmaktadır. Enkoder için PPR (Pulse Per Revolution) bir açısal çevrim için enkoderin üretebileceği kare dalga formundaki sinyali ifade etmektedir. Bir başka ifade ile 4096 ppr çözünürlüğündeki enkoder ile ölçülebilecek açısal deplasman hassasiyeti $360^\circ/4096 = 0,089^\circ$ olarak hesaplanabilir. Ölçülen açısal deplasman verisi üzerinden elde edilen açısal hız bilgisi ile iletim hatası, mekanik güç ve dolayısı ile mekanik verim deneysel veriye dayalı olarak hesaplanabilmektedir.

2.3. Yazılım mimarisi

Yazılım mimarisi, LabVIEW ortamında DAQmx modunda geliştirilmiştir. Farklı test çevrimlerini uygulayabilmekte olan yazılım ve arayüz sayesinde, dinamik veriler anlık olarak operatör tarafından gözlemlenebilmekte ve öznel değerlendirmeler kaydedilebilmektedir. Yazılımda, test edilen her dişli kutusu için seri numarası girilerek verilerin sistematik bir biçimde depolanması sağlanmıştır.

Yazılım, NATIONAL INSTRUMENTS cRIO 9049 kontrolcüsüne gömülü olarak gerçek zamanlı çalışan ölçüm/analiz yazılımı ve sistem arayüzü olmak üzere iki temel bileşenden oluşmaktadır. Sistem arayüzü, operasyonel parametrelerinin ayarlanabildiği parametre bölümü, test çevrimlerinin kurgulandığı test senaryosu bölümü ve operatör arayüzünden oluşmaktadır. (Şekil 3)



Şekil 3: Geliştirilen sistem arayüzü

Ölçümler, yağ koşullandırma işlemi tamamlandıktan sonra, istenilen yük kademesi ve açısal hız basamaklarında otomatik olarak gerçekleştirilmektedir. Dosya kayıt formatındaki tüm alfanümerik karakterler MATLAB’de geliştirilen analiz kodunda işlenmekte ve veriler bu bilgilere göre sınıflandırılmaktadır. Ayrıca, 0,1 s periyodundaki ortalama veriler operatör ekranına yansıtılarak test sırasındaki öznel değerlendirmeler de kayıt altına alınmıştır.

Test sisteminden toplanan ivme ve ses basıncı sinyalleri 25600 Hz, moment sinyalleri 2048 Hz ve son olarak da açısal konum sinyalleri 1024 Hz örnekleme frekanslarında toplanmıştır.

3. DENEYSEL ÇALIŞMALAR

Makale çalışması kapsamında yapılan deneyler, her çevrim oranında 3 adet örnek olmak üzere 3 farklı çevrim oranında toplam 9 örnek dişli kutusu ile gerçekleştirilmiştir. İncelenen dişli kutularının seri numarası, her kademe için diş sayıları ve çevrim oranları, modül (m) ve helis açısı (β), yağlayıcı tipi ve test sıcaklıkları Çizelge 1’de verilmiştir.

Çizelge 1: Ayrıntılı deney planı.

Seri numarası	Konstrüktif parametreler				Test Şartları		
		I. Kademe	II. kademe	III. Kademe	Yağlayıcı Tipi ve Hacmi	Test Sıcaklığı °C	
210445604-210445606	z1/z2	21/38	15/30	-	VG220	400 ml	15-30
	m	1,0	1,75	-			
	β	30°	15°	-			
210445610-210445612	z1/z2	18/41	9/44	-	VG220	400 ml	15-30
	m	1	1,5	-			
	β	30°	15°	-			
210445607-210445609	z1/z2	20/39	18/44	12/52	VG220	350 ml	15-35
	m	1	1,25	1,25			
	β	30°	20°	15°			

Mekanik verim, dişli kutusunun çıkış milinden alınan mekanik gücün, giriş miline uygulanan mekanik güce oranıdır. Her mühendislik yapıtında mevcut kayıplar nedeniyle, sistemin çıkış enerjisi, giriş enerjisinden küçüktür ve sistemde kaybedilen enerji, ağırlıklı olarak ısı enerjisi formunda olmak üzere, akustik ve mekanik (titreşim) enerji olarak da çevre ortama transfer olmaktadır.

Mekanik sistemlerde gücün matematiksel ifadesi:

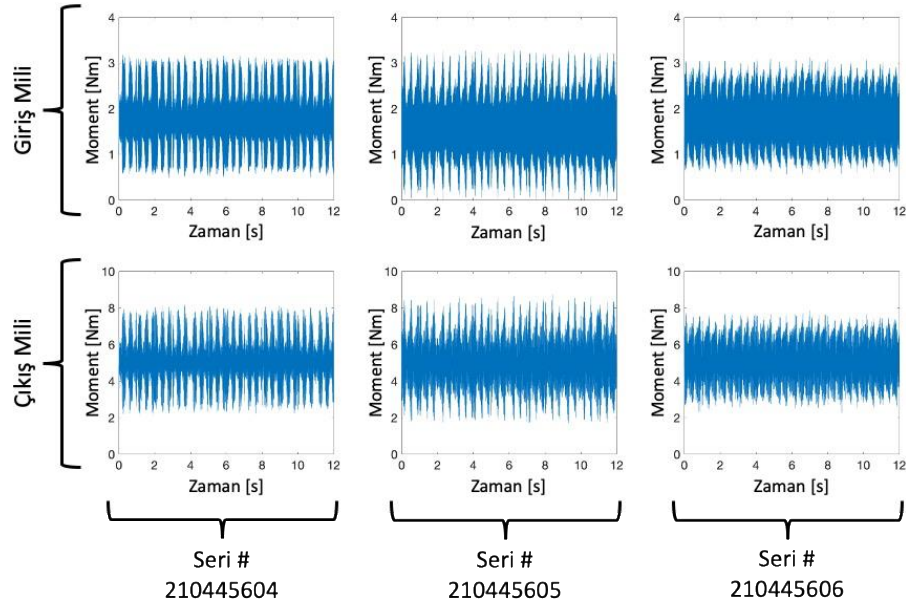
$$P = M\omega \quad (1)$$

şeklindedir. Denklem (1)'de M , momenti, ω ise rad/s cinsinden açısal hızı ifade etmektedir. Dolayısı ile mekanik güç, ölçüm parametreleri olan moment ve açısal konum verileri kullanılarak ölçüme dayalı olarak hesaplanabilmektedir.

Mekanik verim, giriş ve çıkış gücü üzerinden Denklem (2)'de verildiği şekilde hesaplanmaktadır.

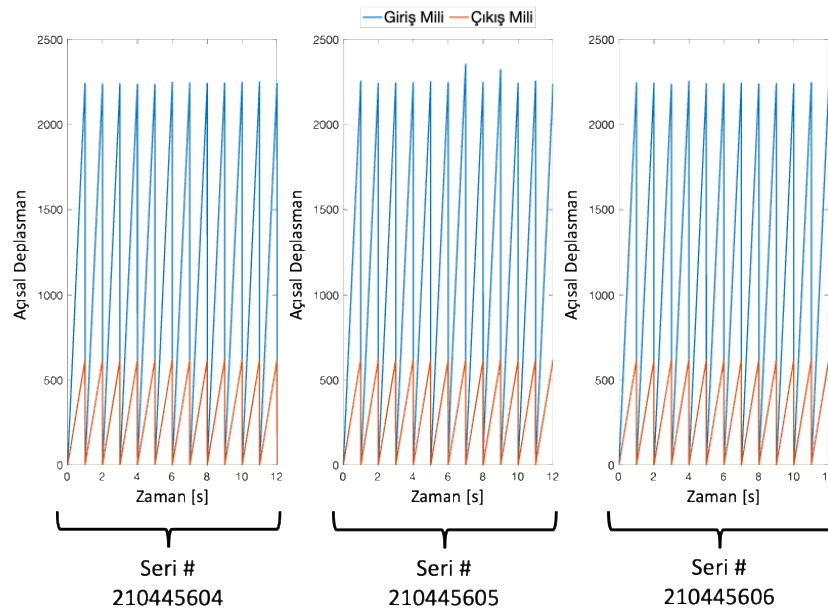
$$\eta = \frac{P_{\text{çıkış}}}{P_{\text{giriş}}} = \frac{M_{\text{çıkış}}\omega_{\text{çıkış}}}{M_{\text{giriş}}\omega_{\text{giriş}}} \quad (2)$$

Yürütülen deneysel çalışmaya bir örnek olması açısından 1400 d/d, %50 yük ve 20 °C sıcaklık şartları altında dişli kutusu giriş ve çıkış milleri üzerinden ölçülen moment verileri Şekil 4'te sunulmuştur. Bu veriler 12 sn'lik bir süre için, 2048 Hz örnekleme frekansı ile kaydedilmiştir. Grafikten de görüleceği üzere sinyalin periyodik karakteristiği dişli sistemlerin temel dinamiğini yansıtmaktadır. Zaman tabanındaki veriler incelendiğinde, moment verisi için giriş ve çıkış millerindeki ortalama genlik ve tepeden tepeye fark esas alınarak hesaplanan salınım değerleri sırasıyla $M_{\text{giriş}} = 1,7665 \pm 0,5169 \text{ Nm}$ ve $M_{\text{çıkış}} = 5,0518 \pm 0,9762 \text{ Nm}$ olarak hesaplanmıştır. Ortalama verim hesaplanırken zaman tabanındaki sinyalden elde edilen bu değerler referans alınmıştır.



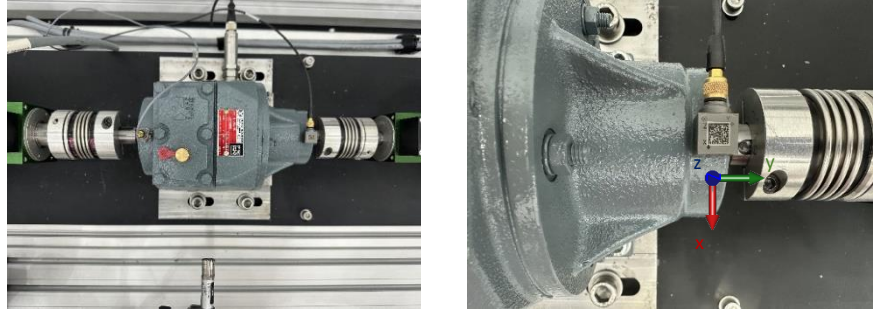
Şekil 4: 210445604-210445606 seri numaralı dişli kutuları için %50 yük, 1200 d/d, 20 °C deney noktasında alınan zaman tabanındaki moment verileri.

Aynı deney şartları için dişli kutusunun giriş ve çıkış milleri üzerinden ölçülen açısal deplasman verileri Şekil 5’te verilmektedir.



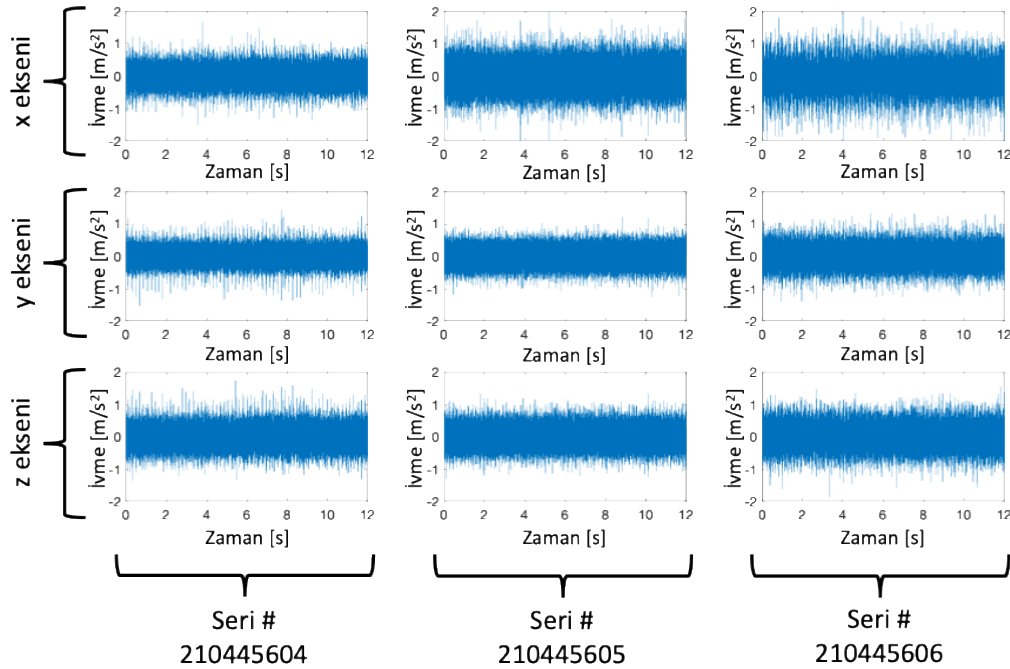
Şekil 5: 210445604-210445606 seri numaralı dişli kutuları için %50 yük, 1200 d/d, 20 °C deney noktasında alınan zaman tabanındaki açısal deplasman verileri.

Dişli kutusu üzerine uygulanan 3-eksenli ivmeölçer, tek eksenli ivmeölçer, mikrofon ve yağ sıcaklık sensörünün yerleşimi ve referans eksen takımı Şekil 6’da verilmiştir.

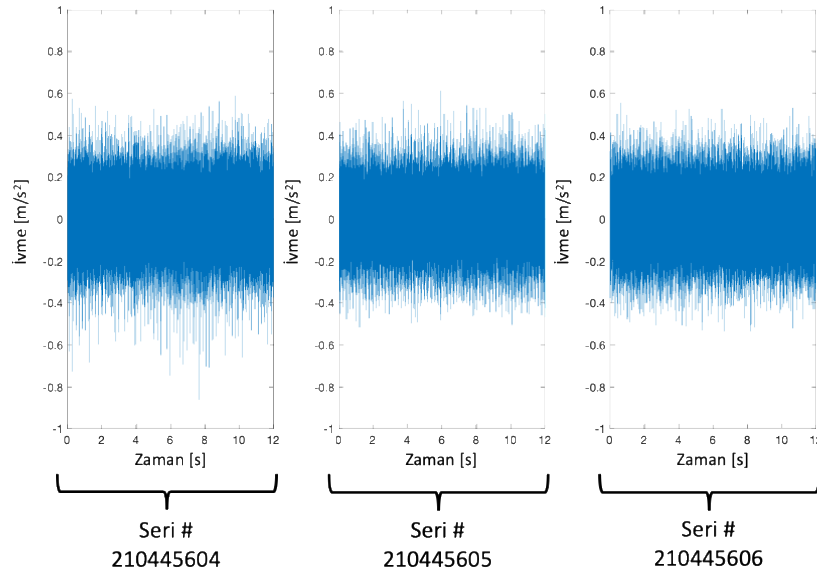


Şekil 6: (a) Sensör yerleşimi. (b) Referans eksen takımı.

Son olarak, yine aynı çalışma şartları için giriş milinin yatak bölgesine yakın olarak konumlandırılan 3-eksenli ivmeölçerden ve çıkış milinin yatak bölgesine yakın olarak konumlandırılan tek eksenli ivmeölçerden alınan veriler Şekil 7 ve Şekil 8’de verilmiştir. Giriş ve çıkış miline yakın olarak konumlandırılan ivmeölçerlerden z-ekseninde alınan veriler karşılaştırmalı olarak incelendiğinde, genliklerin açısal hızın yüksek olduğu giriş bölgesinde daha yüksek olduğu görülmektedir. Bu bağlamda, titreşimin yükten daha çok açısal hız ile doğru orantılı olarak arttığı söylenebilir.



Şekil 7: 210445604-210445606 seri numaralı dişli kutuları için %50 yük, 1200 d/d, 20 °C deney noktasında giriş mili yatak bölgesinden alınan 3-eksenli ivme verileri.



Şekil 8: 210445604-210445606 seri numaralı dişli kutuları için %50 yük, 1200 d/d, 20 °C deney noktasında çıkış mili yatak bölgesinden alınan tek eksenli ivme verileri

Test edilen tüm dişli kutuları için 1200 d/d, %50 yük, 20 °C yağ sıcaklığı noktasındaki zaman tabanında elde edilen sinyallerin RMS değerleri Çizelge 2’de verilmiştir. Sayısal değerler incelendiğinde dişli kutusunun çıkış milinde genliğin azalması, hızın azalmasının doğal bir sonucudur nitekim çevrim oranının yüksek olduğu üçüncü grupta titreşim genlikleri minimum düzeydedir. RMS parametresi, x- ekseninde en geniş dağılımı göstermektedir. Referans eksen takımında da görüleceği üzere x- eksenli dişli kutusu mil eksenini temsil etmektedir. Gözlemlenen bu değişim mil montajı ile ilgili eksen kaçıklığı sorununun varlığına işaret etmektedir.

Çizelge 2: Tüm dişli kutuları için 1200 d/d, %50 yük, 20 °C yağ sıcaklığı noktasındaki zaman tabanında elde edilen sinyallerin RMS değerleri.

Seri numarası	Sensör				
	3-eksenli ivmeölçer			1-eksenli ivmeölçer	Mikrofon
	x-	y-	z-	z-	-
210445604	0,228750	0,197244	0,251046	0,113609	0,178739
210445605	0,329650	0,220687	0,256003	0,108217	0,177146
210445606	0,264883	0,241102	0,259862	0,107724	0,250685
210445610	0,203470	0,115004	0,169520	0,070232	0,155542
210445611	0,371533	0,140019	0,164498	0,073582	0,165854
210445612	0,274123	0,127851	0,169687	0,061544	0,130248
210445607	0,178096	0,135543	0,146872	0,040056	0,148173
210445608	0,237822	0,146502	0,202254	0,079503	0,167159
210445909	0,246557	0,121380	0,176280	0,064713	0,144481

Dişli kutularında güç aktarımı sırasında oluşan kayıplar,

1. Yüke bağlı olan dişli teması ve yatak kaynaklı kayıplar,
2. Yüke bağlı olmayan, yağlayıcı tipi, viskozitesi ve dalma derinliğine bağlı kayıplar,

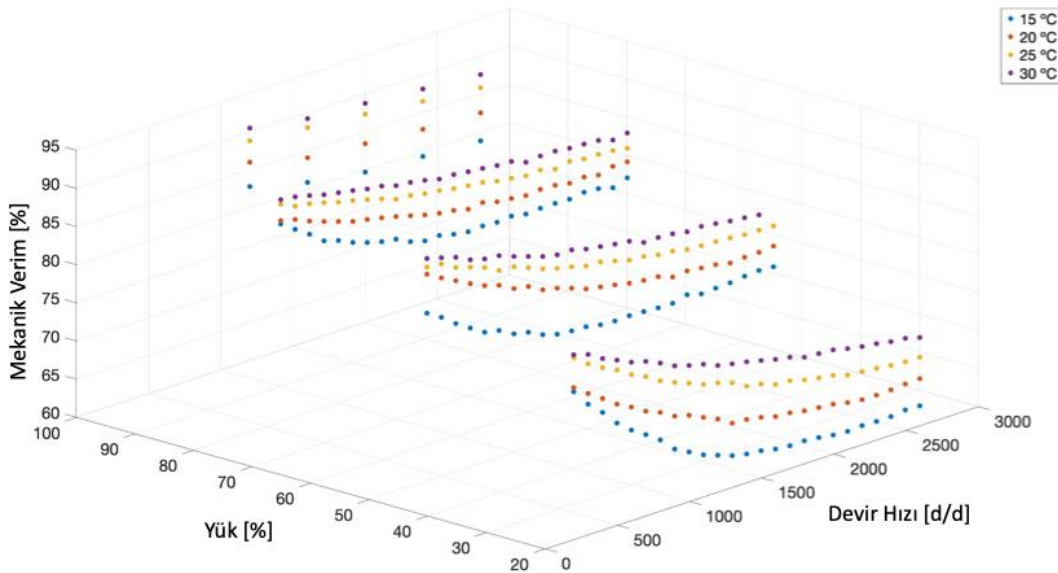
olarak ikiye ayrılabilir (Changenet ve Velez 2007). Yağ sıcaklığının viskozite üzerinde etkili olduğunu bilindiğine göre yağ sıcaklığının da yüke bağlı olmayan bir parametre olduğu ifade edilebilir. Yüke bağlı olan yatak kayıpları, yatak tipine ve boyutuna, yataklama düzenine, yağlayıcı viskozitesine bağlı olarak değişmektedir (Wimmer, Salzgeber ve Haslinger 2003).

Şekil 9’da 210445604 seri numaralı dişli kutusu için toplam 320 deney noktasında ortalama mekanik verim değerlerinden oluşan mekanik verim haritası verilmiştir.

Deney noktaları;

1. 4 farklı yük [25 50 75 100] [%];
2. 25 farklı devir hızı [400 500 ... 2800] [d/d];
3. 4 farklı yağ sıcaklığı [15 20 25 30] [°C];

parametreleri için oluşturulmuştur ancak tam yükte veriler, zaman kısıtlaması olduğundan sadece 5 farklı devir hızı [1200 1600 2000 2400 2800] [d/d] için toplanmıştır.



Şekil 9: 210445604 seri numaralı dişli kutusu için operasyonel parametrelere göre mekanik verim haritası.

Grafikten görüleceği üzere her hız ve yük durumu için artan sıcaklık mekanik verimde artışa neden olmaktadır. Yükselen sıcaklığın etkisi ile birlikte azalan viskozite, sürtünme kaynaklı kayıpları azaltmaktadır.

Test sisteminin fiziksel sınırları gereği, dişli kutuları anma moment değerlerine göre kısmi yükler altında çalıştırılmaktadır, dolayısı ile deney prosedürleri daha yüksek nominal yük altında uygulanırsa, daha yüksek mekanik verim değerlerinin elde edileceği söylenebilir nitekim aynı sıcaklık ve devir hızında, yükün değişiminin mekanik verimi yüke bağlı olarak artan kayıplara rağmen yükseltmesi bu durumun en temel kanıtıdır.

Güç aktarmadaki enerji kaybının hız ile orantılı biçimde artış gösterdiği görülmektedir. 400-1500 d/d bandında güç kaybının hıza duyarlılığının daha yüksek olduğu, 1500-2800 d/d bandında ise asimptotik karakter gösterdiği görülmektedir.

Test edilen tüm dişli kutuları için %50 yük, 20 °C yağ sıcaklığı noktasında, farklı dönme hızlarındaki ortalama mekanik verim Çizelge 3'te verilmiştir. Sayısal değerler incelendiğinde, aynı model dişli kutularının mekanik verim değerlerinde %6 mertebesine varan sapmanın bulunduğu ve bu sapmanın imalat kalitesi ile ilişkili olduğu görülmektedir. Mekanik verimin, aynı grup dişli kutuları içerisinde hıza bağlı düşüşü incelendiğinde, nominal ve oransal olarak önemli farklılıkların bulunduğu saptanmıştır. Hıza bağlı kayıpların çalkalanma etkisi ile oluşması ve dalma derinliği dolayısı ile yağlayıcı hacmi ile ilişkili olması nedeniyle, imalat hattında yağ dolum işleminin hassas olarak kontrol edilemediği söylenebilir.

Çizelge 3: Tüm dişli kutuları için %50 yük, 20 °C deney noktasında mekanik verimin hıza bağlı değişimi.

Seri numarası	Konstrüktif parametreler		Dönme hızı [d/d]				
	Kademe	Çevrim Oranı	1200	1600	2000	2400	2800
210445604	2	3,6190	80,0308	77,9340	76,8131	75,8462	75,7998
210445605			86,4384	86,2525	85,3384	85,0636	84,9043
210445606			79,8012	78,6880	77,2731	76,4781	76,5346
210445610	2	11,1358	66,2977	62,1490	60,7468	59,0601	58,4357
210445611			68,4275	65,1720	63,0755	61,7496	60,8325
210445612			65,6924	61,7134	59,3831	58,2237	57,6978
210445607	3	20,6556	52,9892	49,1787	46,0494	45,3066	44,3474
210445608			49,1989	47,4998	48,4623	43,8205	42,6266
210445909			50,3907	46,3878	44,3812	43,1525	42,5652

Test edilen tüm dişli kutuları için 1200 d/d, 20 °C yağ sıcaklığı noktasında, farklı yüklerde ortalama mekanik verim Çizelge 4'te verilmiştir. Ortalama mekanik verimin tüm dişli kutuları için yüke bağlı olarak artış gösterdiği sayısal değerler üzerinden görülmektedir.

Çizelge 4: Tüm dişli kutuları için 1200 d/d, 20 °C deney noktasında mekanik verimin yüke bağlı değişimi.

Seri numarası	Konstrüktif parametreler		Yük [%]			
	Kademe Sayısı	Çevrim Oranı	25	50	75	100
210445604	2	3,6190	68,9545	80,0308	84,2306	86,0126
210445605			78,8255	86,4384	89,9409	92,0437
210445606			69,1351	79,8012	84,2173	88,4805
210445610	2	11,1358	51,3607	66,2977	72,2112	77,6153
210445611			55,1927	68,4275	75,2995	78,3133
210445612			51,3494	65,6924	72,8514	76,4922
210445607	3	20,6556	37,1468	52,9892	60,4591	66,2549
210445608			34,6011	49,1989	58,0829	61,3469
210445909			35,9465	50,3907	54,6889	60,5214

Hız ve yüke bağlı değişim birlikte incelendiğinde, yağlayıcı hacminin mekanik verim üzerindeki etkisi doğrulanmaktadır. 21445605 seri numaralı dişli kutusu, grup içerisinde tüm noktalarda en yüksek mekanik verim oranına sahiptir. Bununla birlikte, aynı dişli kutusu hıza bağlı kayıplarda da en iyi davranışı göstermektedir. Bu doğrultuda, mekanik verimin oransal değişiminin yağlayıcı hacmi ile ilişkili olduğu kesin olarak söylenebilir nitekim imalatçı kuruluş dişli kutusu iç hacminin konstrüktif nedenlerle daha düşük olduğu 3 kademeli grupta yağlayıcı hacmini 350 ml seviyesine düşürmüştür ancak mekanik verimin önem arz ettiği kritik uygulamalarda yağlayıcı hacminin çevrim oranına göre de optimize edilmesi öneri olarak sunulabilir.

4. SONUÇ

Bu çalışmada dişli kutularının mekanik verimi yük, açılal hız ve sıcaklık gibi temel operasyonel parametrelere göre deneysel olarak incelenmiştir. Yağlayıcı sıcaklığı ve viskozitesinin mekanik verim ile korelasyonu açık olarak görülmüştür. Bununla birlikte, sıcaklık ve viskozitenin açılal hızın neden olduğu kayıplar üzerindeki etkileri saptanmış dolayısı ile ikincil olarak mekanik verim üzerinde etkileri olduğu görülmüştür. Yükün mekanik verim üzerindeki doğru orantılı etkisinin kritik önem arz ettiği belirlenmiştir, nitekim kayıpların geneli hız ile orantılı biçimde değişmektedir ve efektif momentin artışı mekanik verimde artışa neden olmaktadır.

Gelecek çalışmalarda yatak, dişli mekanizması, gövde gibi farklı hataları barındıran dişli kutuları ile yapılacak deneylerde mekanik verimin hatalar ile ilişkisi irdelenecektir.

ÇIKAR ÇATIŞMASI BEYANI

Yazarlar arasında çıkar çatışması bulunmamaktadır.

TEŞEKKÜR

Bu çalışma, Umut Murat GÖK'ün kurucusu olduğu ADEON Makine Ar-Ge Mühendislik Ticaret Limited Şirketi'nin yürütmüş olduğu 7160766 numaralı TÜBİTAK TEYDEB destekli 'Titreşim Akustik Verilere Dayalı Kalite Kontrol Test Sistemi' başlıklı proje kapsamında gerçekleştirilmiştir.

Deneysel çalışmada, İTÜ BAP MGA-21021-43378 numaralı proje kapsamında destek alınmıştır.

YAZARLARIN KATKILARI

U.M.G. : Kavramsallaştırma, yöntem, yazılım, doğrulama, formel analiz, araştırma, kaynaklar, yazı yazma - orijinal taslak hazırlama.

O.T.Ş. : Formel analiz, doğrulama, araştırma, gözden geçirme ve düzenleme.

İ.M.P. : Formel analiz, doğrulama, araştırma, gözden geçirme ve düzenleme.

KAYNAKLAR

Allianz Versicherungs-AG. (1978). Handbook of Loss Prevention, Springer-Verlag, Berlin.

Changenet, C., Velez, P. (2007). Housing influence on churning losses in geared transmissions, ASME.

Doleschel, A., Höhn, B.-R., Michaelis, K. (2000). Frictional Behaviour of Synthetic Gear Lubricants. 27th Leeds-Lyon Symposium on Tribology in Lyon, p. 759 – 768.

F.D. Harris, E.F. Kasper, L.E. Iseler. (2000). US civil rotorcraft accidents, 1963 through 1997, Technical Report NASA/TM-2000-209597, USAAMCOM-TR-00-A-006, NASA and the US Army Aviation and Missile Command.

Matschoß (1940). Geschichte des Zahnrades, VDI, Berlin.

Özgüven, H. Nevzat, Houser, D. R. (1988). Mathematical Models Used In Gear Dynamics- A Review. Journal of Sound and Vibration, 121(3), 383–411.

Rolls-Royce revs up world's most powerful aerospace gearbox. (2016). Designfax, 01636669, 12- 43.

Wimmer, A., Salzgeber, K.; Haslinger, R. (2003). WP1 – Analysis of Minimum Oil Requirements Considering Friction in Gears and Engines. Final Report Oil-free Powertrain, EU Project Contract No: IPS-2001-CT-98006.

Zaman Serisi Tabanlı Makine Öğrenmesi Modelleri ile GitHub Projelerindeki Programlama Dili Popülerliğinin Tahmini (2011–2021)

Bora UĞURLU^{1*}, Bahadır KARASULU¹

¹Bilgisayar Mühendisliği Bölümü, Mühendislik Fakültesi, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, Türkiye

Geliş: 24.09.2025, Kabul: 27.11.2025, Yayınlanma: 24.12.2025

ÖZ

Bu çalışmada GitHub platformunda 2011–2021 dönemine ait farklı programlama dillerinin depo (repository), çekme isteği (PR) ve sorun (issue) verileri kullanılarak, dillerin popülerliği zaman serisi tabanlı makine öğrenmesi yöntemleriyle tahmin edilmiştir. Üç farklı kaynaktan bütünleştirilen veri kümesi, dil–yıl–çeyrek düzeyinde PR, issue ve depo sayılarını içermekte; farklı kaynaklardan elde edilen metrikler tek bir zaman çizelgesinde birleştirilerek her dil için çeyreklik gözlemler üzerinden modelleme yapılmasına olanak vermektedir. Öznitelik mühendisliği sonrasında lojistik regresyon, karar ağaçları, rastgele orman, destek vektör makineleri ve gradyan artırma yöntemleri uygulanmıştır. Bulgular, Lojistik Regresyonun (AUC=0,996), Rastgele Ormanın (AUC=0,994) ve SVM’nin (AUC=0,988) güçlü ayırt edicilik sağladığını; Karar Ağaçları ve Gradyan Artırmanın ise yüksek doğruluk değerlerine rağmen ROC-AUC açısından daha zayıf kaldığını göstermektedir. Bu kapsamda, doğruluk ile ROC-AUC’nin birlikte raporlanması yöntemler arasındaki ayrım gücünü daha görünür kılmaktadır. Ayrıca analizler, Python ve JavaScript gibi dillerin uzun vadeli yükselişini doğrulamış, karar ağaçları ve gradyan artırma nadir dönemlerde öne çıkan dilleri yakalamada daha dengeli sonuçlar sunmuştur.

Anahtar Kelimeler: Programlama Dilleri; Makine Öğrenmesi, Zaman Serisi Analizi

Time Series–Based Machine Learning Models for Forecasting Programming Language Popularity on the GitHub Projects (2011–2021)

ABSTRACT

In this study, popularity trends of programming languages were predicted using time-series–based machine learning methods on GitHub data covering 2011–2021. The integrated dataset, compiled from three different sources, contains counts of repositories, pull requests (PRs), and issues at the language–year–quarter level; by consolidating metrics from multiple sources into a single timeline, it enables quarter-based modeling for each language. Following feature engineering, logistic regression, decision trees, random forests, support vector machines (SVM), and gradient boosting were applied. The findings indicate that Logistic Regression (AUC = 0.996), Random Forest (AUC = 0.994), and SVM (AUC = 0.988) provide strong discriminative performance, whereas Decision Trees and Gradient Boosting remain weaker in terms of ROC-AUC despite achieving high accuracy. In this context, reporting accuracy together with ROC-AUC makes differences in discriminative power across methods more apparent. Moreover, the analyses confirm the long-term rise of languages such as Python and JavaScript; decision trees and gradient boosting yield more balanced results in capturing languages that become prominent during rare periods.

Keywords: Programming Languages, Machine Learning, Time Series Analysis.

1. GİRİŞ

Son on yılda açık kaynak ekosisteminin büyümesi ve çeşitlenmesi, programlama dillerinin kullanım eğilimlerinde belirgin değişimlere yol açmıştır. GitHub gibi büyük ölçekli platformlar, yalnızca depo sayıları açısından değil, aynı zamanda geliştirici etkileşimlerini yansıtan çekme isteği ve sorun gibi ölçütler aracılığıyla da bu dönüşümü somut biçimde belgelemektedir (Wessel, Vargovich, Gerosa ve Treude, 2023). 2024 Octoverse'in 2024 yılındaki raporunda (GitHub Staff, 2024) Python'un ilk kez JavaScript'i geride bırakarak en popüler dil konumuna yükselmesi, diller arası rekabetin ve kullanım alışkanlıklarının ne denli hızlı değişebildiğinin güncel bir göstergesidir. Ancak mevcut raporlar ve çalışmalar çoğunlukla betimleyici istatistiklere odaklanmakta, geleceğe yönelik öngörüler sunmada sınırlı kalmaktadır.

Bu bağlamda, bu çalışma GitHub'ın 2011–2021 yıllarını kapsayan GitHub Programming Languages Data veri kümesi üzerinden, programlama dillerinin popülerliğini tahmin etmeyi amaçlamaktadır. Veri kümesi; dil, yıl, çeyreklik düzeyinde çekme isteği (pull request, PR) ve sorun(issue) sayılarını ve her dil için toplam depo sayısını içermektedir. Çalışmanın özgün katkısı, yalnızca statik büyüklükleri değil topluluk etkileşimini yansıtan ölçütler de dikkate alarak çok boyutlu bir tahmin modeli önermesidir. Zaman serisi tabanlı öznitelik mühendisliği ile klasik makine öğrenmesi algoritmaları (lojistik regresyon (Logistic Regression, LR), karar ağaçları (Decision Trees, DT), rastgele orman (Random Forest, RF), destek vektör makineleri (Support Vector Machine, SVM), gradyan artırma (gradient boosting) bütünleştirilmiş; elde edilen bulgular ise Octoverse eğilimleri ve önceki ampirik çalışmalarla karşılaştırılarak yazılım mühendisliği literatürüne öngörü odaklı yeni bir bakış açısı kazandırılmıştır. Bu çalışma beş bölümden oluşmaktadır. Bu çalışmanın ikinci bölümünde literatürde var olan önceki çalışmalar sunulmaktadır. Üçüncü bölümde veri kümesinin hazır hale getirilmesi ve bu veri kümesi üzerinde uygulanan beş farklı makine öğrenme algoritmaları ele alınmaktadır. Dördüncü bölümde veri kümesi üzerinde denenen makine öğrenmesi algoritmalarının ulaşılmış olduğu başarımlar ölçütleri yer almaktadır. Beşinci ve son bölümde ise elde edilen sonuçların ve bulguların bir değerlendirilmesi yapılmaktadır.

2. ÖNCEKİ ÇALIŞMALAR

Endüstriyel ve akademik çalışmalar incelendiğinde; 2014 ila 2024 yılları arasında açık kaynak ekosisteminde yaşanan hızlı büyüme ve topluluk çeşitliliği, programlama dillerinin kullanım eğilimlerinde önemli kaymalara yol açmıştır. Örneğin 2024 Octoverse raporu, Python'un ilk kez JavaScript'i geride bırakarak GitHub üzerinde en popüler dil olduğunu ve Jupyter Notebook kullanımının %92 artış gösterdiğini belirtmiştir (GitHub Staff, 2024).

Borges, Hora ve Valente'nin (2016) çalışması, GitHub'daki projelerin popüleritesini yıldız sayıları üzerinden öngörmeye odaklanmaktadır. Çoklu doğrusal regresyon modelleriyle yapılan analizlerde, özellikle son altı aylık veriler kullanıldığında tahminlerin oldukça doğru olduğu gösterilmiştir. Ayrıca benzer büyüme eğilimleri sergileyen depolar için özgül modellerin, yavaş büyüyen veya az yıldızlı projelerde daha başarılı sonuçlar verdiği belirtilmektedir (Borges, Hora ve Valente, 2016).

Ray, Posnett, Devanbu ve Filkov'a (2017) çalışmalarında, GitHub üzerinde geliştirilen 728 projeden toplanan geniş ölçekli veriler (63 milyon satır kod, 1,5 milyon işleme (commit)) kullanılarak programlama dillerinin yazılım kalitesi üzerindeki etkilerini incelemektedir. Bulgular, programlama dili seçiminin hata oranları üzerinde istatistiksel olarak anlamlı ancak sınırlı bir etkiye sahip olduğunu göstermektedir. Özellikle statik tipli ve tip karışıklığını engelleyen dillerin daha az hata eğilimli olduğu, işlevsel dillerin ise prosedürel dillere kıyasla daha güvenilir sonuçlar verdiği belirlenmiştir. Bununla birlikte, proje büyüklüğü, ekip boyutu ve işleme yoğunluğu gibi süreçsel faktörlerin yazılım hataları üzerinde programlama dili tercihlerinden çok daha güçlü bir etkiye sahip olduğu vurgulanmaktadır.

Bissyandé, Lo, Jiang, Réveillère, Klein ve Le Traon'un (2013) çalışması, GitHub'daki geniş ölçekli projeler üzerinden sorun takip sistemi (issue tracker) kullanımını incelemiş ve bu araçların özellikle büyük ve aktif ekipler için daha yaygın olduğunu göstermiştir. Araştırma ayrıca böyle bir sistemin kullanımının proje popüleritesiyle güçlü biçimde ilişkili olduğunu, ancak hata raporlarının sayısı ile düzeltilme süresi arasındaki ilişkinin zayıf kaldığını ortaya koymuştur (Bissyandé, Lo, Jiang, Réveillère, Klein ve Le Traon, 2013).

Rahman ve Roy'un (2014) çalışması, GitHub'daki 78 projeden 100 binden fazla çekme isteğini inceleyerek başarılı ve başarısız taleplerin hangi koşullarda gerçekleştiğini araştırmaktadır. Çalışma, özellikle programlama dili, proje alanı, proje yaşı ve geliştirici deneyimi gibi faktörlerin başarı oranları üzerinde belirleyici olduğunu göstermektedir. Bulgulara göre, büyük sayıda geliştiriciye veya aşırı ayrışmaya (fork) sahip projelerde başarısız çekme isteği oranı artarken, belirli alanlardaki (ör. framework, IDE) projeler daha yüksek başarı oranı sergilemektedir (Rahman ve Roy, 2014).

Bu makalede, 2011–2021 dönemini kapsayan GitHub Programming Languages Data veri kümesini kullanarak (dil, yıl, çeyrek düzeyinde PR ve sorun sayıları ile dil başına toplam depo sayıları), programlama dillerinin popülerliğini zaman serisi tabanlı klasik makine öğrenmesi yöntemleriyle tahmin edilmesi amaçlanmıştır. Çalışmanın katkıları üç ekseninde toplanmaktadır. İlk olarak yalnızca statik büyüklükleri değil, PR/sorun gibi topluluk etkileşim göstergelerini de içeren çok boyutlu bir popülerlik modeli önermek, ikinci olarak lojistik regresyon, karar ağaçları, rastgele orman, SVM ve gradyan artırma gibi klasik makine öğrenmesi algoritmalarını zaman gecikmeli (lag) ve değişim oranı gibi öznelik mühendisliği ile

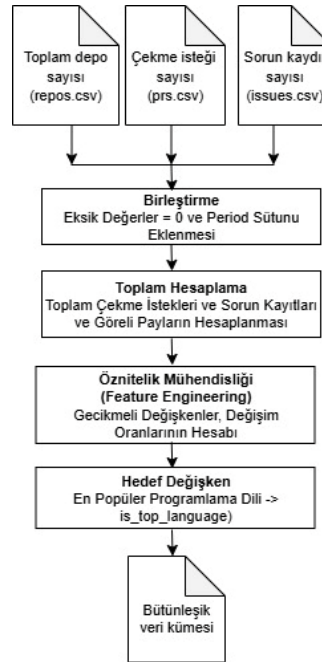
bütünleştirerek tahmin performansını değerlendirmek ve son olarak bulguları Octoverse eğilimleri ve önceki ampirik çalışmalar ışığında tartışarak yazılım mühendisliği yazınına öngörü odaklı bir çerçeve sunmaktır.

3. MATERYAL ve METOT

Bu bölümde çalışmada kullanılan veri kümesinin hazırlanma süreci ayrıntılı olarak açıklanacak, ardından kullanılan makine öğrenmesi yöntemleri tanıtılacaktır. İlk alt başlık altında GitHub'dan elde edilen verilerin nasıl birleştirildiği, dönüştürüldüğü ele alınacaktır. Daha sonraki alt başlık altında ise lojistik regresyon, karar ağaçları, rastgele orman, destek vektör makineleri ve gradyan artırma yöntemlerinin temel prensipleri ve bu çalışmada kullanılma gerekçeleri sunulacaktır.

3.1. Veri Kümesi ve Hazırlık

Bu çalışmada kullanılan veri kümesi, GitHub üzerinde 2011–2021 yılları arasında farklı programlama dillerine ait depo, çekme isteği (PR) ve sorun kaydı sayılarını içeren GitHub Programming Languages Data kümesidir (Wen, t.y.). Bu veri kümesi Şekil 1'de belirtilen adımlardan sonra beş farklı algoritma tarafından kullanılabilir hale getirilmiştir.



Şekil 1: Bütünleşik veri kümesi dosyasının oluşturulması

Şekil 1'de görülen akış diyagramı, çalışmada kullanılan bütünleşik veri kümesinin elde edilme adımlarını özetlemektedir. Süreç üç ayrı kaynaktan (depo sayıları repos.csv, çekme isteği sayıları prs.csv ve sorun kaydı sayıları issues.csv) başlayan ham veriler ile başlamakta ve aşamalı bir işleme sürecinden geçerek nihai veri kümesine ulaşmaktadır. İlk aşamada veriler, programlama dili, yıl ve çeyrek bazında birleştirilmiş,

eksik değerler 0 ile doldurulmuş ve dönemsel analizlerin kolaylaştırılması amacıyla her satır için bir period değişkeni tanımlanmıştır. İkinci aşamada, dönemsel toplam çekme isteği ve sorun kayıtları hesaplanarak dillerin göreceli popülerliğini yansıtan paylar (pr_share, issue_share) elde edilmiştir. Üçüncü aşamada, zaman serisi yapısının modele yansıtılabilmesi amacıyla öznitelik mühendisliği yapılmıştır. Bu kapsamda gecikmeli değişkenler (örneğin lag1, lag2) oluşturulmuş, ayrıca dönemler arası eğilimleri göstermek üzere bir önceki çeyreğe göre değişim oranları (chg1) hesaplanmıştır. Bu adım, dillerin yalnızca mutlak değerlerle değil, aynı zamanda büyüme ve azalış hızlarıyla da temsil edilmesini sağlamıştır. Son aşamada, her çeyrekte diller PR payına göre sıralanmış ve en yüksek değere sahip dil “en popüler dil” olarak işaretlenmiştir. Bu işlem sonucunda, sınıflandırma görevine olanak tanıyan ikili hedef değişken (is_top_lang) oluşturulmuştur. Böylece, nihai bütünsel veri kümesi hem ham verileri hem de türetilmiş değişkenleri içeren çok boyutlu bir yapıya kavuşmuş ve makine öğrenmesi algoritmalarının uygulanmasına hazır hale getirilmiştir. Veri kümesinde hem yaygın diller (Python, Java, JavaScript, C++, R, MATLAB, Go, Swift, TypeScript) hem de daha özel/az kullanılan diller (Eiffel, Nim, Smalltalk, Fortran, Ada, Xtend, ZenScript vb.) olmak üzere toplamda 208 farklı programlama dili yer almaktadır.

Çizelge 1’de, eğitim ve test veri kümelerinin zaman dilimlerine göre dağılımını özetlemektedir. Eğitim verisi 2011’in birinci çeyreğinden 2018’in dördüncü çeyreğine kadar olan yaklaşık sekiz yıllık dönemi kapsamaktadır. Bu veri kümesi, toplam verinin yaklaşık %70’ini oluşturmakta ve modellerin öğrenme süreci için geniş bir tarihsel arka plan sağlamaktadır.

Çizelge 1: Eğitim ve test veri aralıkları

Veri Kümesi	Yıl Aralığı	Kapsadığı Çeyrekler	Oran (%)
Eğitim	2011–2018	2011 Q1 – 2018 Q4	%70
Test	2019–2021	2019 Q1 – 2021 Q4	%30

Test verisi ise 2019’un birinci çeyreğinden 2021’in dördüncü çeyreğine kadar olan üç yıllık dönemi içermekte ve toplam verinin %30’unu temsil etmektedir. Böylelikle modeller, geçmiş verilerle eğitildikten sonra daha sonraki döneme ait gözlemler üzerinde sınanmış, zaman serisi tabanlı öngörü problemi bağlamında ileriye dönük tahminlerin güvenilirliği değerlendirilmiştir.

Bu dağılım, makine öğrenmesi çalışmalarında yaygın biçimde kullanılan %70–%30 eğitim/test oranıyla uyumlu olup, hem modelin öğrenmesi için yeterli veri sağlamış hem de bağımsız bir test kümesi aracılığıyla genelleme yeteneğinin değerlendirilmesine imkân tanımıştır.

Bu çalışmada kullanılan bütünsel veri kümesi, GitHub Programming Languages Data setindeki depo (repos), çekme isteği (PR) ve sorun (issue) kayıtlarının programlama dili–yıl–çeyrek düzeyinde birleştirilmesiyle oluşturulmuş çoklu zaman serisi/panel bir yapıya sahiptir. Veri, 2011’in birinci çeyreği ile

2021'in dördüncü çeyreği arasındaki dönemi kapsamakta ve her satır bir dilin ilgili çeyrekteki etkinliğini temsil etmektedir. Öznitelik mühendisliği adımıyla ham PR/issue/depo sayılarının yanı sıra görelî paylar ve zaman bağımlılığını yansıtan gecikmeli değişkenler (lag1, lag2) ile çeyrekler arası değişim oranları (chg1) türetilmiş; böylece dillerin yalnızca mutlak büyüklükleri değil dönemsel dinamikleri de modele aktarılmıştır. Nihai veri kümesi, yaygın ve nadir dilleri birlikte içeren toplam 208 programlama dilini kapsamaktadır. Zaman serisi doğasına uygun biçimde veriler kronolojik olarak ayrılmış, 2011Q1–2018Q4 dönemi eğitim, 2019Q1–2021Q4 dönemi ise test kümesi olarak kullanılmıştır. Bu kurulum, modellerin geçmiş dönem örüntülerini öğrenip daha ileri döneme ait gözlemler üzerinde sınanmasına imkân tanımaktadır.

Modelleme aşamasında lojistik regresyon, karar ağaçları, rastgele orman, destek vektör makineleri ve gradyan artırma algoritmaları scikit-learn altyapısı ile uygulanmış, karşılaştırılabilirliği korumak amacıyla hiperparametrelerin büyük bölümü varsayılan değerlerde bırakılmıştır. Lojistik regresyon modelinde optimizasyonun yakınsamayı garanti etmesi için iterasyon sınırı yükseltilmiş (max_iter=2000), karar ağacı ve gradyan artırmada tekrarlanabilirlik sağlamak üzere sabit rastgelelik tohumu atanmıştır (random_state=42). Rastgele orman modeli, topluluk öğrenmesinin kararlılığını arttırmak amacıyla 300 ağaçla (n_estimators=300) eğitilmiş ve tüm çekirdekleri kullanacak şekilde paralelleştirilmiştir (n_jobs=-1). SVM modeli doğrusal olmayan ayrımı yakalayabilmek için RBF çekirdeğiyle kurulmuş (kernel="rbf") ve olasılık tabanlı ölçütlerin hesaplanabilmesi için probability=True seçeneği etkinleştirilmiştir. Bu hiperparametre stratejisi, modellerin temel karakteristiklerini öne çıkaran, yanlılığı azaltılmış bir kıyaslama hattı sağlamaktadır.

Veri ön işleme aşamasında, özellikle lojistik regresyon ve SVM için standartlaştırma (standard scaling) uygulanmıştır. Bu modeller girdilerin ölçeklerine duyarlı olduklarından, büyük ölçek farklılıkları parametre tahminlerini bozabilir. Örneğin, PR sayıları on binler düzeyindeyken sorun oranları [0,1] aralığında kalabilir. Standartlaştırma ile tüm öznitelikler ortalaması 0, standart sapması 1 olacak şekilde dönüştürülmüştür. Böylece lojistik regresyon ve SVM daha istikrarlı hale gelirken, farklı boyutlardaki değişkenler modelin öğrenmesini olumsuz etkilememiştir. Buna karşılık, karar ağaçları, rastgele orman ve gradyan artırma yöntemleri ölçek farklılıklarına duyarlı olmadıkları için bu algoritmalara ayrıca ölçeklendirme yapılmamıştır. Veri kümesinde 2011–2021 yılları arasında en çok tercih edilen programlama dili 2011'de Ruby, 2012–2020 arasında JavaScript, ve 2021'de ise Python olmuştur.

3.2 Modelleme Yaklaşımları

Programlama dillerinin popülerliğini tahmin etmek amacıyla klasik makine öğrenmesi algoritmaları uygulanmıştır. Bu kapsamda sınıflandırma ve regresyon problemleri için beş farklı yöntem kullanılmıştır. Bunlar lojistik regresyon, karar ağaçları, RF, SVM ve gradyan artırma yöntemidir.

İlk yöntem olan lojistik regresyon, bağımlı değişkenin kategorik olduğu durumlarda kullanılan istatistiksel bir modeldir. En yaygın kullanımı ikili sınıflandırma (binary classification) problemlerinde olup, bir gözlemin belirli bir sınıfa ait olma olasılığını tahmin etmeyi amaçlar. Model, olasılık değerlerini doğrudan tahmin etmek yerine, bu değerleri logit dönüşümü ile sınırsız bir aralığa dönüştürür. Lojistik regresyonun temel formülasyonu Denklem 1’de verilmiştir (Cox, 1958; Hosmer, Lemeshow & Sturdivant, 2013).

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (1)$$

Yukarıdaki Denklem 1’de p , bağımlı değişkenin “başarı” ya da “1” olarak kodlanan sınıfa ait olma olasılığıdır. X_i bağımsız değişkenleri temsil etmektedir. β_i , katsayılar olup her bir bağımsız değişkenin etkisini göstermektedir (Menard, 2002).

İkinci yöntem ise karar ağaçlarıdır. Karar ağaçları, sınıflandırma ve regresyon problemlerinde kullanılan, veriyi özniteliklere göre dallara ayırarak tahmin yapan algoritmalarıdır. Her iç düğüm belirli bir öznitelik üzerinde koşul içerirken, yaprak düğümler sınıf veya tahmin edilen değeri temsil eder. Bu yöntem, özellikle yorumlanabilirlik açısından güçlüdür ve görsel olarak karar sürecini anlaşılır kılar (Quinlan, 1986). Karar ağacında en önemli aşama, veriyi bölecek en iyi özelliğin seçilmesidir. Bu seçim genellikle bilgi kazancı (information gain) veya Gini safsızlığı (Gini impurity) ölçütleriyle yapılır (Breiman vd., 1984). Her iki kavram ise Denklem 2’de verilen entropi kavramıyla ilişkilidir (Shannon, 1948).

$$\text{Entropi}(S) = - \sum_{i=1}^c p_i \log_2(p_i) \quad (2)$$

Denklem 2’de verilen p_i , sınıf i ’nin olasılığıdır. Bilgi kazancı ise entropiyi temel alarak sırasıyla Denklem 3’ deki gibi ifade edilir.

$$\text{Bilgi Kazancı}(S, A) = \text{Entropi}(S) - \sum_{v \in \text{Değerler}(A)} \frac{|S_v|}{|S|} \text{Entropi}(S_v) \quad (3)$$

Denklem 3 karar ağaçlarının bölünme kriterini belirtir. Bir öznitelik olan A , veri kümesi S ’yi böldüğünde entropi ne kadar azalırsa bilgi kazancı da o derece yüksek olacaktır. Diğer bir değişle en yüksek kazancı sağlayan öznitelik ağacın düğümünde seçilir. Seçilen düğümün rastgele seçilen iki örneğin farklı sınıflara ait olma olasılığı ise Gini safsızlığı ile ifade edilir (Breiman vd., 1984). Bu safsızlık, matematiksel olarak Denklem 4’teki gibi ifade edilir. Bu denklem incelendiğinde p_i , sınıf i ’nin düğümdeki olasılığı olup c ise toplam sınıf sayısıdır.

$$\text{Gini}(S) = 1 - \sum_{i=1}^c p_i^2 \quad (4)$$

Denklem 3 ve 4’ün her ikisi de “*hangi özneliğe göre bölünme yapılmalı?*” sorusuna yanıt arar ve her ikisinin dayandıkları temel kavram Denklem 2’de matematiksel olarak ifade edilen bilgi teorisidir. Gini safsızlığı matematiksel olarak daha basit ve hızlı bir alternatif olarak karşımıza çıkmaktadır.

Çalışmadaki kullanılan bir diğer yöntem ise rastgele ormandır. Rastgele orman, birden fazla karar ağacının birlikte (ensemble) kullanıldığı topluluk tabanlı güçlü bir makine öğrenmesi yöntemidir (Breiman, 2001). Bu yöntem, tekrar örnekleme yoluyla birleştirme (bootstrap aggregating) yöntemine dayanır. Veri kümesinden tekrar örnekleme yoluyla birçok alt küme oluşturulur. Her alt küme üzerinde bağımsız bir karar ağacı eğitilir. Tahmin aşamasına gelindiğinde sınıflandırma problemleri için ağaçların verdiği sınıf tahminleri arasında çoğunluk oylaması yapılırken regresyon problemlerinde ise ağaçların tahminleri aritmetik ortalama alınarak sonuca ulaşılır. Bu yöntemi matematiksel olarak Denklem 5’teki gibi ifade etmek mümkündür.

$$y = \frac{1}{T} \sum_{t=1}^T h_t(x) \quad (5)$$

Yukarıdaki Denklem 5’te T toplam ağaç sayısını ifade etmektedir. $h_t(x)$, t. ağacın tahmin fonksiyonudur. y ise rastgele ormanın nihai tahmindir. Bu yöntemde ayrıca, her düğümde bölünme yapılırken tüm öznelilikler yerine rastgele seçilen bir alt küme kullanılır. Bu yaklaşım, ağaçlar arasında korelasyonu azaltarak modelin genelleme performansını artırır (Biau ve Scornet, 2016).

Çalışmadaki kullanılan bir diğer yöntem ise SVM’dir. Bu yaklaşım sınıflandırma ve regresyon problemlerinde kullanılan güçlü bir makine öğrenmesi yöntemi olup temel amacı, farklı sınıflara ait verileri en iyi şekilde ayıran hiper düzlemi bulmaktır. Destek Vektör Makineleri, veriyi iki sınıfa ayıran ve bu iki sınıf arasındaki marjini (margin) maksimize eden hiper düzlemi seçer. Hiper düzleme en yakın veriler “destek vektörleri” olarak adlandırılır ve karar sınırının konumunu belirler. Hiper düzlemin ağırlık vektörünün hesaplanma şekli ise Denklem 6’da verilmiştir. Bu denklemde belirtilen w , hiper düzlemin ağırlık vektörü olup y_i ise sınıf etiketini gösterir (Cortes ve Vapnik, 1995).

$$\min_{w,b} = \frac{1}{2} \|w\|^2 \mid y_i(w \cdot x_i + b) \geq 1 \forall i \quad (6)$$

Beşinci ve son yöntem ise gradyan arttırmadır. Gradyan artırma, birden fazla “zayıf öğrenen modelin” (genellikle karar ağaçları) ardışık olarak eğitilip bir araya getirilmesiyle güçlü bir tahmin modeli oluşturan bir topluluk öğrenmesi (ensemble learning) yöntemidir. Temel fikir, her yeni modelin, önceki modellerin yapamadığı hataları gidermeye odaklanmasıdır. Bu yapıda ilk model eğitilir ve hatalar hesaplanır. Yeni

model bu hataları tahmin etmeye çalışır ve en sonunda oluşacak olan model tüm zayıf modellerin ağırlıkları toplamından oluşur. Modelin matematiksel ifadesi Denklem 7’deki gibidir (Friedman, 2001).

$$F_M(x) = \sum_{m=1}^M \gamma_m h_m(x) \quad (7)$$

Yukarıdaki denklemde M, toplam ağaç sayısı olup $h_m(x)$ m. karar ağacıdır. γ_m ise her ağacın ağırlığıdır. Bu yöntemin yaygın uygulamaları ve bu çalışmada kullanılan ise klasik makine öğrenmesi algoritmalarından olan Gradient Boosting Machines (GBM)’dir. Günümüzde kullanılan XGBoost (Extreme Gradient Boosting), ve LightGBM (Light Gradient Boosting), GBM algoritmasının performans, ölçeklenebilirlik ve verimlilik açısından geliştirilmiş, modernleştirilmiş varyantlarıdır (Chen ve Guestrin, 2016; Ke vd., 2017).

4. DENEYSEL BULGULAR

Bu çalışmada önerilen beş farklı makine öğrenmesi yöntemi (Lojistik Regresyon, Karar Ağaçları, Rastgele Orman, Destek Vektör Makineleri ve Gradyan Artırma) 2011–2018 döneminde eğitilmiş ve 2019–2021 döneminde test edilmiştir. Sınıflandırma görevi, her çeyrekte en popüler programlama dilinin doğru tahmin edilmesi üzerine kurgulanmıştır. Deneysel sonuçlar Çizelge 2’de sunulmuştur.

Çizelge 2: Beş farklı sınıflandırma yönteminin test dönemi performansları

Model	Doğruluk	Kesinlik (Ağırlıklı)	Duyarlılık (Ağırlıklı)	Makro-F1	MCC
Lojistik Regresyon	0,9891	0,9784	0,9891	0,4972	0
Karar Ağaçları	0,99	0,9901	0,99	0,5689	0,2759
Rastgele Orman	0,9891	0,9784	0,9891	0,4972	0
Destek Vektör Makineleri	0,9883	0,9828	0,9883	0,5595	0.1559
Gradyan Arttırma	0,99	0,99	0,99	0,5689	0,2759

Çizelge 2’de sunulan bulgular, beş farklı makine öğrenmesi yönteminin test dönemindeki performanslarını ortaya koymaktadır. Çizelge 2’deki koyu harflerle gösterilen değerler en iyi sonuçları ifade etmektedir. Genel olarak tüm modellerin doğruluk oranlarının %98,8 ile %99 arasında yoğunlaşması, modellerin popüler programlama dillerini tahmin etmede yüksek başarı sağladığını göstermektedir. Ancak bu yüksek doğruluk değerleri, çoğunluk sınıfın doğru tahmin edilmesinden kaynaklanabileceği için tek başına güvenilir bir performans göstergesi değildir. Çizelge 2’ye göre bütün modeller yüksek doğruluk ($\approx$ %99) sergilese de Karar Ağaçları ve Gradyan Arttırma yöntemleri öne çıkmaktadır. Bu iki model hem doğruluk (%99) hem de Kesinlik (%99 civarı) ve Duyarlılık (%99) değerlerinde en yüksek performansı göstermiştir. Ayrıca sınıf dengesizliğini dikkate alan Makro-F1 skorunda da (0,5689) en iyi sonuçlara ulaşmışlardır.

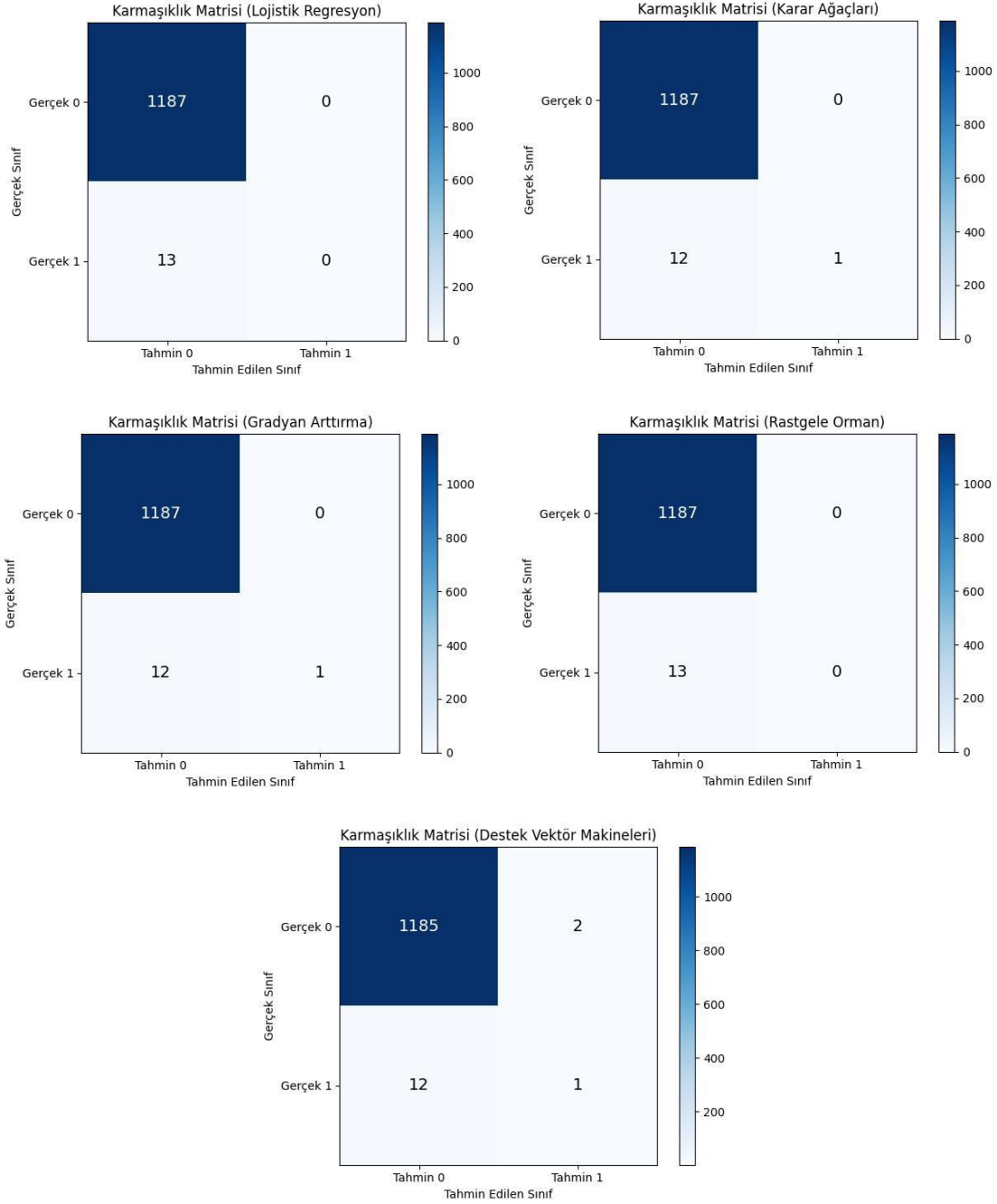
Karar Ağaçları ve Gradyan Artırma yöntemlerinin diğer algoritmalarından daha iyi sonuç vermesi, öznelilikler arasındaki doğrusal olmayan ilişkileri yakalayabilme ve sınıflar arası ayrımı daha etkin gerçekleştirebilme güçlerinden kaynaklanmaktadır. Veri kümesinde popüler dillerin uzun süre baskın kalması nedeniyle sınıf dengesizliği mevcuttur; bu durum sınıflandırma doğruluğu değerlerinin tüm modellerde yüksek çıkmasına yol açarken, Makro-F1 skorları azınlık sınıfları ayırt etme becerisini daha doğru yansıtmaktadır. Karar Ağaçları ve Gradyan Artırma, bu dengesiz yapıya rağmen az temsil edilen sınıfları da görece başarılı şekilde tahmin ederek dengeli performans sergilemiş; bu başarıları, gecikme ve değişim oranı gibi türetilmiş öznelilikleri daha etkin kullanabilmelerine bağlanabilir. Sonuçlar, modellerin test kümesinde de tutarlı performans göstermesi nedeniyle aşırı öğrenme (overfitting) riskinin düşük olduğunu ortaya koymaktadır.

Kesinlik ve duyarlık ölçütlerinin %97,8 ile %99 arasında seyretmesi, yanlış pozitif ve yanlış negatif tahminler açısından modellerin dengeli bir performans sunduğunu ortaya koymaktadır. Özellikle Karar Ağaçları ve Gradyan Artırma algoritmalarının %99 seviyesine ulaşması, bu yöntemlerin sınıflar arasında ayırım yapmada daha başarılı olduğunu göstermektedir. Bu durum, söz konusu algoritmaların sadece baskın sınıfı tahmin etmekle kalmayıp, daha az temsil edilen sınıfları da görece iyi ayırt edebildiğine işaret etmektedir.

Buna karşın sınıf dengesizliğini dikkate alan Makro-F1 metriği daha ayırt edici bir tablo sunmaktadır. Karar Ağaçları ve Gradyan Artırma yöntemleri 0,5689 değeriyle en yüksek performansı sergilerken, onları SVM 0,5595 ile takip etmiştir. Lojistik Regresyon ve Rastgele Orman ise 0,4972 değeriyle görece daha düşük bir performans göstermiştir. Bu sonuçlar, özellikle ağaç tabanlı yöntemlerin dengesiz veri koşullarında avantajlı olduğunu, doğrusal modellerin ise bu bağlamda sınırlı kaldığını göstermektedir.

Matthews Correlation Coefficient (MCC) değerleri, sınıf dengesizliğinin belirgin olduğu bu ikili sınıflandırma probleminde modellerin gerçek ayırt etme gücünü daha güvenilir biçimde yansıtmaktadır. MCC, doğruluk (accuracy) gibi ölçütlerin aksine dört hücreyi (TP, TN, FP, FN) birlikte dikkate alarak “her iki sınıfta da tutarlı doğru tahmin” yapıp yapılmadığını ölçer. Test sonuçlarında Karar Ağacı ve Gradyan Artırma modellerinin MCC değerlerinin ~0,276 ile diğerlerine göre daha yüksek olması, bu iki modelin azınlık sınıfını (top language) en azından sınırlı da olsa yakalayabildiğini ve rastgele tahmine kıyasla pozitif bir korelasyon ürettiğini gösterir. SVM’in MCC’sinin ~0,156 seviyesinde kalması ise modelin çoğunluk sınıfı lehine ağırlıklandığını, azınlık sınıfındaki iyileşmenin daha kısıtlı olduğunu düşündürmektedir. Buna karşılık Lojistik Regresyon ve Rastgele Ağaç (Random Forest) için MCC’nin 0 çıkması, bu modellerin test aşamasında pratikte “tek sınıflı” bir karar kuralına kaydığını diğer bir deyişle neredeyse tüm örnekleri çoğunluk sınıfı olarak etiketlemektedir. Böyle bir durumda TP ve FP üretilmediği (veya son derece ihmal edilebilir kaldığı) için MCC doğal olarak sıfıra çöker. Bu davranışın temel nedeni, veri setindeki ciddi sınıf

dengeşizliği altında her iki modelin de varsayılan eşik (0,5) ve sınıf ağırlıksız öğrenme ayarlarıyla azınlık sınıfını cezalandırmaktansa çoğunluk sınıfını doğru tahmin etmeyi optimize etmesidir. Dolayısıyla yüksek doğruluk değerlerine rağmen, sınıflar arası dengeli bir karar üretmedikleri MCC tarafından net biçimde ortaya konmaktadır.

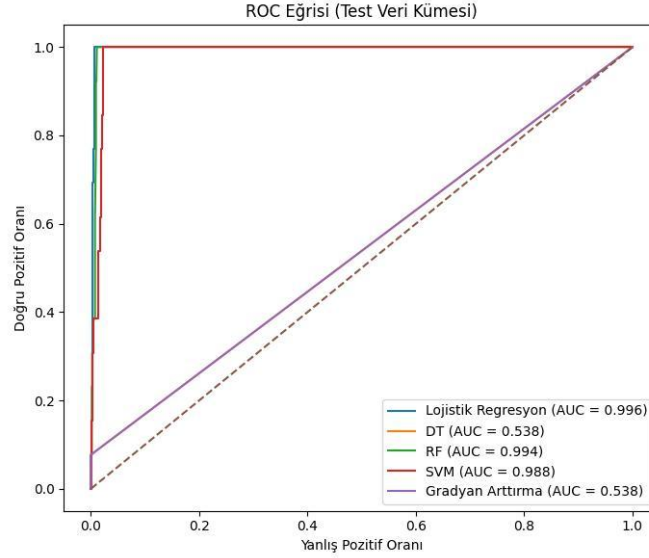


Şekil 2: Algoritmaların test veri kümesi üzerindeki karmaşıklık matrisleri

Şekil 2’de verilmiş olan karmaşıklık matrisleri birlikte değerlendirildiğinde, test kümesinde sınıf dağılımının belirgin biçimde dengesiz olduğu görülmektedir: “Gerçek 0” sınıfı 1187 örnekle baskın iken “Gerçek 1” sınıfı yalnızca 13 örnekten oluşmaktadır. Karar Ağaçları ve Gradyan Arttırma modelleri bu dengesizlik altında çoğunluk sınıfını neredeyse hatasız yakalamış (1187 doğru negatif, yalancı pozitif yok), ancak azınlık sınıfında duyarlılık oldukça sınırlı kalmıştır; her iki model de 13 pozitif örneğin yalnızca 1’ini doğru sınıflandırabilmiş, 12’sini kaçırmıştır. Lojistik Regresyon ve Rastgele Orman matrisleri ise modelin karar eşiği ve öğrenme dinamikleri nedeniyle pratikte tek-sınıf tahminine kaydığını göstermektedir: tüm örnekler “0” olarak etiketlenmiş, böylece pozitif sınıfa ait doğru tahmin (TP) hiç üretilenmemiş ve 13 pozitif örneğin tamamı yanlış negatifte kalmıştır. Destek Vektör Makineleri sonuçları, azınlık sınıfını yakalama bakımından sınırlı bir iyileşme sağlasa da (1 doğru pozitif), bunun bedeli olarak az sayıda da olsa yalancı pozitif üretmiş (2 örnek); buna karşın 12 pozitif örnek yine kaçırılmıştır. Tüm modellerde yüksek doğruluğun temel olarak çoğunluk sınıfını doğru tahmin etmekten kaynaklandığı, asıl ayırt edici başarının azınlık sınıfında düşük kaldığı anlaşılmaktadır.

Şekil 3’te verilmiş olan Alıcı İşletim Karakteristik (Receiver Operating Characteristic, ROC) eğrileri incelendiğinde, çalışmada en güvenilir sonuçların Lojistik Regresyon, Rastgele Orman ve SVM tarafından elde edildiği anlaşılmaktadır. Bu modellerin ROC eğrisi altında kalan alan (Area Under Curve, AUC) değerleri %98’in üzerinde olup, sınıfları ayırt etme kapasitesi oldukça yüksektir. Buna karşın Karar Ağaçları ve Gradyan Arttırma yöntemlerinin düşük AUC değerleri, model yapılandırmalarının yeniden gözden geçirilmesi gerektiğini göstermektedir.

Şekil 3’teki ROC eğrileri ile Çizelge 2’nin birlikte ele alındığında, modellerin performansları arasında hem benzerlikler hem de ayrışmalar olduğunu göstermektedir. Lojistik Regresyon, Rastgele Orman ve Destek Vektör Makineleri, tabloda yüksek doğruluk, kesinlik ve duyarlılık değerleri sergilerken ROC eğrilerinde de AUC değerlerinin 0,98’in üzerinde olmasıyla tutarlı biçimde güçlü ayırt edicilik göstermektedir. Buna karşın Karar Ağaçları ve Gradyan Arttırma yöntemlerinde bir farklılaşma dikkat çekmektedir: Çizelge 2’de doğruluk ve Makro-F1 değerleri oldukça yüksek çıkmasına rağmen, ROC eğrilerinde AUC değerlerinin 0,538 seviyesinde kalması bu modellerin eşik bağımlı ayırt edicilikte başarısız olduklarını göstermektedir. Bu durum, tek ağaç yapısının aşırı uyuma (overfitting) eğiliminden ve boosting algoritmasının olasılık kalibrasyonundaki zayıflıklarından kaynaklanabilir. Dolayısıyla, bazı modeller için farklı ölçütlerin birbirini doğruladığı görülürken, diğerlerinde ölçütler arasında ayrışma gözlenmiş ve bu da model değerlendirmesinde birden fazla ölçütün dikkate alınmasının gerekliliğini ortaya koymuştur. 2011–2021 döneminde GitHub verilerine göre en çok tercih edilen programlama dili 2011’de Ruby, 2012–2020 arasında JavaScript, 2021’de ise Python olmuştur. Deneysel bulgular da bu sıralamanın doğru olduğunu kanıtlamıştır.



Şekil 3: Algoritmaların test veri kümesi üzerindeki ROC eğrileri

Sonuç olarak, yalnızca doğruluk değerlerine bakıldığında tüm modeller benzer seviyede başarılı görünse de Makro-F1 sonuçları daha kapsamlı bir değerlendirme sunmaktadır. Bu bağlamda Karar Ağaçları ve Gradyan Artırma yöntemleri, programlama dili popülerliğini öngörmeye daha güvenilir bir performans sergileyerek öne çıkmıştır. Destek Vektör Makineleri güçlü bir alternatif olarak dikkat çekerken, Lojistik Regresyon ve Rastgele Orman'ın daha sınırlı bir performans sunduğu anlaşılmaktadır.

5. SONUÇ

Bu çalışmada GitHub üzerinde 2011–2021 yılları arasındaki programlama dili etkinliklerini (repository, pull request ve issue sayıları) içeren bir veri kümesi kullanılarak programlama dillerinin popülerliğini öngörmeye yönelik bir yaklaşım geliştirilmiştir. Veri kümesi üzerinde gerçekleştirilen öznelik mühendisliği adımlarıyla lag, değişim oranı ve pay (share) değişkenleri türetilmiş; böylece yalnızca statik değerler değil, zaman serisi dinamikleri de modele yansıtılmıştır. Çalışmada beş farklı klasik makine öğrenmesi algoritması (Lojistik Regresyon, Karar Ağaçları, Rastgele Orman, Destek Vektör Makineleri ve Gradyan Artırma) uygulanarak performansları karşılaştırılmıştır.

Deneysel sonuçlar, tüm yöntemlerin yüksek doğruluk oranlarına ulaştığını, ancak özellikle sınıf dengesizliğinin etkisini dengeleyen Makro-F1 ölçütünde farklılıkların ortaya çıktığını göstermiştir. Karar Ağaçları ve Gradyan Artırma yöntemleri, nadir dönemlerde öne çıkan dilleri daha iyi yakalayarak diğer yöntemlere kıyasla daha dengeli bir performans sergilemiştir. Bu bulgu, öngörü problemlerinde tek başına doğruluk oranının yeterli olmadığını, farklı ölçütlerle değerlendirme yapmanın gerekli olduğunu ortaya koymaktadır.

Elde edilen sonuçlar, yazılım ekosisteminde programlama dili tercihlerinin evriminin öngörülmesine katkı sunmaktadır. Bu yaklaşım, yazılım geliştiricileri, teknoloji şirketleri ve eğitim kurumları açısından stratejik kararlar alınmasına yardımcı olabilir. Örneğin, popülerliği hızla artan dillerin erken tespit edilmesi, müfredatların güncellenmesi veya endüstriyel yatırımların yönlendirilmesi açısından önem taşımaktadır. Yapay zekâ altyapısının programlama dillerinin kullanımındaki etkisi de bir çeşit varyasyon veya bias değeri oluşturmaktadır.

Gelecekte yapılacak çalışmalar, daha uzun dönemleri kapsayan verilerin kullanılması, derin öğrenme tabanlı yöntemlerin denenmesi ve sosyal kodlama platformlarından elde edilebilecek ek göstergelerin (örn. yıldız sayıları, fork oranları) modele entegre edilmesi ile genişletilebilir. Böylece programlama dillerinin popülerlik dinamiklerini daha bütüncül bir biçimde modellemek ve yazılım mühendisliği alanına öngörü odaklı katkılar sağlamak mümkün olacaktır.

ÇIKAR ÇATIŞMASI BEYANI

Yazarlar arasında çıkar çatışması bulunmamaktadır.

TEŞEKKÜR

Bu çalışma herhangi bir kurum ve/veya kuruluş tarafından desteklenmemiştir.

YAZARLARIN KATKILARI

B.U.: Yazılımsal yapının geliştirilmesi, veri kümesi üzerinde uygulanması, sonuçların elde edilmesi ve yazım.

B.K.: Uygun çözümün belirlenmesi için literatürün araştırılarak yöntemlerin belirlenmesi, sonuçların analiz edilmesi, yazım ve gözden geçirme.

KAYNAKLAR

- Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Bissyandé, T. F., Lo, D., Jiang, L., Réveillère, L., Klein, J., & Le Traon, Y. (2013). Got issues? Who cares about it? A large scale investigation of issue trackers from GitHub. 2013 IEEE 24th International Symposium on Software Reliability Engineering (ISSRE), 188–197. IEEE. <https://doi.org/10.1109/ISSRE.2013.6698917>
- Borges, H., Hora, A., & Valente, M. T. (2016). Predicting the popularity of GitHub repositories. *Proceedings of the 12th International Conference on Predictive Models and Data Analytics in Software Engineering (PROMISE 2016)*, 1–10. ACM. <https://doi.org/10.1145/2972958.2972966>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth International Group.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>

- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
<https://doi.org/10.1145/2939672.2939785>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
<https://doi.org/10.1007/BF00994018>
- Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), 215–242.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232
- GitHub Staff. (2024, October 29). Octoverse: AI leads Python to top language as the number of global developers surges. *GitHub Blog*. <https://github.blog/news-insights/octoverse/octoverse-2024/>
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 3146–3154.
- Menard, S. (2002). *Applied Logistic Regression Analysis* (2nd ed.). Sage.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106.
<https://doi.org/10.1007/BF00116251>
- Rahman, M. M., & Roy, C. K. (2014). An insight into the pull requests of GitHub. *Proceedings of the 11th Working Conference on Mining Software Repositories (MSR 2014)*, 364–367. ACM.
<https://doi.org/10.1145/2597073.2597076>
- Ray, B., Posnett, D., Devanbu, P., & Filkov, V. (2017). A large-scale study of programming languages and code quality in GitHub. *Communications of the ACM*, 60(10), 91–100. <https://doi.org/10.1145/3126905>
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.
<https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Wen, I. (2021). *GitHub Programming Languages Data (2011–2021)* [Dataset]. Kaggle.
<https://www.kaggle.com/datasets/isaacwen/github-programming-languages-data>
- Wessel, M., Vargovich, J., Gerosa, M. A., & Treude, C. (2023). GitHub Actions: The impact on the pull request process. *Empirical Software Engineering*, 28(131). <https://doi.org/10.1007/s10664-023-10369-w>

İyileştirici Tasarım: Mimarlık, Sağlık Ortamları ve İyileşme Arasındaki İlişkinin Bibliyometrik Analizi

Gülcan MİNSOLMAZ^{*1}, Ayyüce AVŞAR²

¹Mimarlık Bölümü, Mimarlık Fakültesi, Kırklareli Üniversitesi, Kırklareli, Türkiye

²Mimarlık Anabilim Dalı, Fen Bilimleri Enstitüsü, Kırklareli Üniversitesi, Kırklareli, Türkiye

Geliş: 14.10.2025, Kabul: 25.11.2025, Yayınlanma: 24.12.2025

ÖZ

Sağlık ortamlarında tasarımın, yalnızca tedavi süreçlerini desteklemekle kalmayıp bireylerin fiziksel, psikolojik ve sosyal iyilik hâlini de güçlendirdiği yönündeki araştırmalar, mimarlık ve sağlık arasındaki ilişkinin önemini ortaya koymaktadır. Bu doğrultuda yapılan çalışma, mimarlık ve iyileştirici tasarım arasındaki kavramsal çerçeveyi genişletmek ve derinleştirmek amacıyla alan yazında yer alan akademik üretimi bibliyometrik yöntemlerle analiz etmektedir. Araştırma kapsamında; sağlık yapıları, mimarlık ve iyileşme kavramlarının kesişiminde yer alan literatür, web of science veri tabanı üzerinden sistematik biçimde taranmıştır. Çalışmada iki temel bibliyometrik yöntem kullanılmıştır: Birlikte bulunma (co-occurrence) ve ortak-yazar (co-authorship) analizleri. Birlikte bulunma analizi, yazarların kullandığı anahtar kelimeler üzerinden alandaki kavramsal kümeleri görünür kılarken; ortak-yazar analizleri ülkeler, kurumlar ve araştırmacılar arasındaki işbirliği ağlarını ortaya koymaktadır. Seçilen anahtar kelimeler ile yapılan analizler sonucunda 2010–2025 yılları arasında yayımlanmış 8 belge tespit edilmiştir. Bu belgelerde biyofilik tasarım, salutojenik yaklaşım, çocuk dostu tasarım, empatik mimari ve sürdürülebilirlik gibi temaların öne çıktığı belirlenmiştir. Bulgular, sağlık yapılarında iyileştirici tasarımın, kullanıcı deneyimi ve çevresel sürdürülebilirliği önceleyen insan merkezli yaklaşımlarla gelişim gösterdiğini ortaya koymaktadır. Çalışma literatüre; mevcut eğilimleri belirlemek, araştırma boşluklarını görünür kılmak ve gelecekteki çalışmalar için yol gösterici bir zemin sunmaktadır.

Anahtar Kelimeler: İyileştirici Tasarım; Sağlık Ortamları; Mimarlık; Bibliyometrik Analiz; Kullanıcı Odaklı Yaklaşım

Healing Design: A Bibliometric Analysis of the Relationship Between Architecture, Healthcare Environments, and Healing

ABSTRACT

Research in healthcare environments shows that design not only supports treatment processes but also enhances individuals' physical, psychological, and social well-being, emphasizing the significance of the relationship between architecture and health. In this context, this study analyzes academic production in the literature using bibliometric methods to expand and deepen the conceptual framework between architecture and healing design. The literature at the intersection of healthcare facilities, architecture, and healing was systematically examined through the Web of Science database. Two main bibliometric techniques were employed: co-occurrence and co-authorship analyses. The co-occurrence analysis visualizes conceptual clusters in the field based on keywords used by authors, while the co-authorship analysis reveals collaboration networks among countries, institutions, and researchers. As a result of the

analyses conducted with the selected keywords, eight documents published between 2010 and 2025 were identified. These publications highlight themes such as biophilic design, salutogenic approach, child-friendly design, empathetic architecture, and sustainability. The findings indicate that healing design in healthcare facilities has evolved through human-centered approaches prioritizing user experience and environmental sustainability. The study contributes to the literature by identifying current trends, making research gaps visible, and providing a guiding framework for future studies.

Keywords: Healing Design; Healthcare Environments; Architecture; Bibliometric Analysis; User centered Approach

1. GİRİŞ

Son yıllarda sağlık alanında hastalıkların tedavi edilmesinin yanında bireylerin fiziksel, ruhsal ve sosyal boyutlarıyla tam bir iyilik hâline ulaşmaları da önemli bir hedef haline gelmiştir. Dünya Sağlık Örgütü'nün tanımı da bu yaklaşımı desteklemektedir. Dünya Sağlık Örgütü'ne (1948) göre sağlık, yalnızca bedensel bozuklukların yokluğu değil, çok yönlü bir esenlik durumudur. Bu bakış açısı, sağlık yapılarının tasarımında da etkisini göstermekte; mekânların yalnızca tedavi sürecini desteklemesini değil, aynı zamanda kullanıcıların iyilik hâlini de güçlendirmesi gerektiğini ortaya koymaktadır (Dilani, 2006). Mekânın tedavi süreçlerinde belirleyici bir rol oynadığına ilişkin bulgular özellikle “kanıta dayalı tasarım” literatüründe vurgulanmaktadır. Ulrich'in (1984) çalışması, hastaların odalarından doğayı görebilmelerinin iyileşme sürelerini kısalttığını göstererek çevresel faktörlerin sağlık üzerindeki etkisini açık biçimde ortaya koymuştur. Bu bulgu, sonraki araştırmaların da yönünü belirlemiş; ışık, akustik, renk, hava kalitesi ve doğayla kurulan bağ gibi unsurların tedaviye katkısı farklı bağlamlarda incelenmiştir (Ulrich et al., 2008; Lawson & Phiri, 2003). Ayrıca, mekânsal kalitenin yalnızca hastaların değil, sağlık çalışanları ve ziyaretçilerin de deneyimlerini doğrudan etkilediği; kullanıcı odaklı tasarımın stresin azalmasına, hasta memnuniyetinin artmasına ve çalışanların verimliliğine katkı sağladığı çalışmalarla (Harris et al., 2002; Verderber & Fine, 2000) ortaya konmuştur.

Günümüzde sağlık yapıları, işlevsellik sınırlarını aşarak “iyileştirici mekân” anlayışı çerçevesinde değerlendirilmektedir. Mimarlık yalnızca teknik bir çözüm aracı değil, aynı zamanda iyileşme sürecini destekleyen ve kullanıcıya ilham veren bir unsur olarak kabul edilmektedir. Purves (2002), iyi tasarlanmış sağlık yapılarının hem hastalara hem de sağlık çalışanlarına olumlu katkılar sunduğunu vurgulamaktadır. Son yıllarda ön plana çıkan biyofilik ve salutojenik tasarım yaklaşımlarının, doğayla bütünleşen peyzaj unsurları, ışık, renk ve mekânsal düzenlemelerin tedavi sürecine katkı sağladığını ortaya koyan bulgular artmaktadır (Abdelaal & Soebarto, 2019; Sánchez-Jáuregui et al., 2024). Bunun yanında çocuk dostu tasarımlar (Bahrami et al., 2025) ve empatiyi merkeze alan mimari yaklaşımlar (Banerji, 2016), kullanıcı odaklı tasarım anlayışının sağlık yapılarında giderek belirginleştiğini göstermektedir. Bu çerçevede mimarlık, sağlık ortamları ve iyileşme arasındaki ilişkilerin anlaşılabilmesi için bibliyometrik çalışmalar önemli bir yöntem olarak öne çıkmaktadır. Bibliyometri, sadece yayın sayılarını ortaya koymakla kalmamakta; kavramların zaman içindeki gelişimini, yazarlar arasındaki işbirliklerini ve araştırma alanlarının yönelimlerini de görünür kılmaktadır (Zupic & Čater, 2015; Cobo et al., 2011). Bu çalışma kapsamında da iki bibliyometrik yöntem kullanılmıştır: birlikte bulunma (co-occurrence) analizi ve ortak-yazar (co-authorship) analizleri. Birlikte bulunma analizi, yazarların kullandığı anahtar kelimelerin makalelerde hangi ölçüde birlikte geçtiğini ortaya koyarken; ortak-yazar analizleri ise yazarlar, kurumlar ve ülkeler arasındaki işbirliklerini inceleyerek bilimsel haritalama teknikleri aracılığıyla bu ilişkileri görünür kılmaktadır. Dolayısıyla bu makale, sistematik inceleme ve bibliyometrik analiz yöntemlerini bir

araya getirerek mimarlık ve sağlık temalı literatürün mevcut durumunu, eksik yönlerini ve gelecekteki potansiyel araştırma alanlarını tartışmayı amaçlamaktadır.

2.YÖNTEM

Bu araştırma, sağlık disiplininde mimarlık konulu çalışmaların bibliyometrik bir perspektiften değerlendirilmesine dayanmaktadır. Bibliyometri, mevcut akademik üretimi sistematik biçimde analiz eden ve literatürün çeşitli boyutlarını ortaya koyan bir yöntemdir (Ruhanen ve ark., 2015: 6; Meena ve ark., 2021: 3). Bu yaklaşım, yayınların objektif ölçütler doğrultusunda sınıflandırılmasına olanak tanıyarak, belirli bir araştırma alanındaki bilimsel üretime bütüncül bir bakış sunmaktadır (Cavalcante ve ark., 2021: 5). Bibliyometrik analiz kapsamında, çok yazarlı çalışmaların incelenmesi, atıf ve ortak atıf ilişkilerinin değerlendirilmesi, anahtar kelime analizi ve bibliyometrik eşleştirme gibi yöntemler sıklıkla kullanılmaktadır (Demir & Erigüç, 2018: 97; Nova-Reyes ve ark., 2020: 5). Süreç içerisinde, literatürde önerilen beş aşamalı iş akışı benimsenmiştir (Börner ve ark., 2003: 9; Cobo ve ark., 2011: 1382; Zupic & Cater, 2015: 433; Aria & Cuccurullo, 2017: 960; Kammerer ve ark., 2021: 28). Bu iş akışı sırasıyla araştırma tasarımı, veri toplama, analiz, görselleştirme ve yorumlama aşamalarını içermektedir. Çalışmaya ilişkin ayrıntılı iş akışı diyagramı ise Çizelge 1’de sunulmuştur.

Çizelge 1: Araştırma sürecini özetleyen iş akış çizelgesi.

Araştırma Tasarımı	Araştırma amacının, sorularının ve yönteminin belirlenmesi
Bibliyometrik Verilerin İncelenmesi	Bibliyometrik veri tabanı: Web of Science Bibliyometrik verilerin filtrelenmesi ve veri setinin indirilmesi
Analiz	Bibliyometrik yazılım: Microsoft Excel
Görselleştirme yöntemi	Ağ Analizi :VOSviewer
Yorumlama	Bulguların tanımlanması ve yorumlanması

2.1. Araştırma Sorusu

Araştırmanın odağı, aşağıdaki dört araştırma sorusu etrafında şekillenmektedir:

1. Mimarlık ve sağlık ortamları arasındaki ilişkide, iyileştirici tasarımı etkileyen temel yaklaşımlar ve kavramlar nelerdir?
2. Bu konuda öne çıkan yazarlar, ülkeler, anahtar kelimeler, kaynaklar, belgeler, kurumlar ve referanslar hangileridir?
3. İyileştirici tasarım üzerine yapılan araştırmalar, zaman içinde hangi kavramlar üzerinde gelişmiş ve derinleşmiştir?
4. Hangi kavramlar son yıllarda ortaya çıkmış, görece daha yenidir ve araştırmaya daha açıktır?

2.2. Bibliyometrik Yöntem

Bu çalışmada, araştırma amacı doğrultusunda literatürden anlamlı veriler elde edebilmek amacıyla bibliyometrik bir yaklaşım benimsenmekte, (Budler, Župič ve Trkman 2021) bibliyometrik yöntemler ve bilimsel haritalama teknikleri kullanılmaktadır (Cabo ve ark., 2011). Bibliyometrik araştırmanın temel amacı ve bu araştırmalardan beklenen sonuçlar, bir araştırma alanındaki üretkenliği tanımlamak, bilimsel unsurların performansını analiz etmek ve araştırma alanının entelektüel, sosyal ve kavramsal yapısını ortaya koyarak alanın evrimsel gelişimini ortaya çıkarmaktır (Öztürk ve Gürler, 2021). Sistematik literatür taramalarının aksine, bibliyometrik yöntemler, literatür taraması yapılırken konuyu daha geniş bir perspektiften ele alarak araştırma alanının hem daha nesnel hem de görsel olarak tanımlanmasını sağlar (Zupic ve Čater, 2015). Bu çalışma kapsamında birlikte bulunma (co-occurrence) analizi ve ortak-yazar (co-authorship) analizleri olmak üzere iki bibliyometrik yöntem kullanılmıştır. Birlikte bulunma analizi (co-occurrence) / author keywords analiz birimi başlığı altında incelenir ve yazarlar tarafından kullanılan anahtar kelimelerin makalelerde birlikte bulunma durumunu gösterir. Ortak-yazar (co-authorship) analizleri ise yazarlar (authors), kuruluşlar (organizations) ve ülkeler (countries) analiz birimleri olmak üzere üç farklı başlık altında incelenir ve bir araştırma alanındaki ülkeler, kurumlar, yazarlar arası iş birliklerini ortaya koyarak bilimsel haritalama tekniği ile bu ilişkileri gösterir.

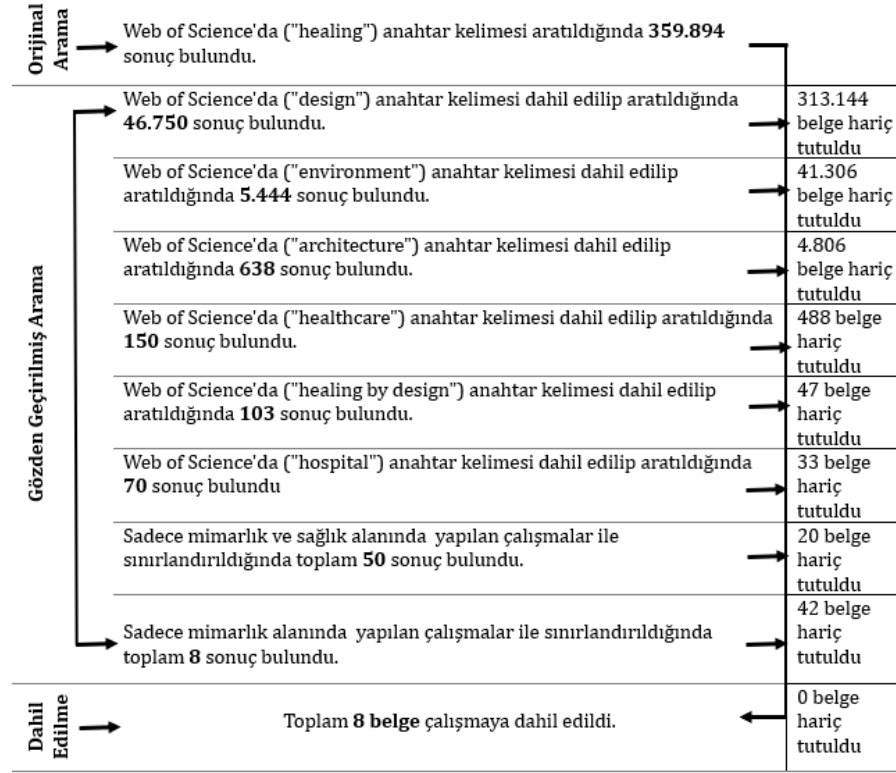
2.3. Veri Tabanı ve Anahtar Kelimelerin Seçimi

Bibliyometrik yaklaşımı benimseyen çalışmalar sıklıkla Web of Science, Scopus ve PubMed gibi veri tabanlarını kullanmaktadır. “Mimarlık” ve “iyileştirici tasarım”ın kesişiminde yer alan bu çalışmada, bilimsel literatürü detaylı bir şekilde incelemek, bilimsel ve sistematik veriler elde etmek amacıyla veri tabanı olarak Web of Science tercih edilmiştir. Web of Science veri tabanı, fiziksel bilimler ve mimarlık gibi sosyal bilimler arasındaki disiplinlerarası alanlar için kapsamlı ve güvenilir bir veri tabanıdır ve olağanüstü kalitedeki bilimsel yayınları içermektedir. Veri görselleştirme ve analiz aracı olarak ise VOSviewer kullanılmıştır. Çalışmada, araştırma stratejisi olarak tündengelim yaklaşımı benimsenmiştir. Bu bağlamda, “iyileştirici” ve “tasarım” (healing and design) kavramları araştırmanın temel noktasını teşkil etmekte olup ilk sorgu bu anahtar kelimeler üzerinden gerçekleştirilmiş, sonrasında çalışmanın diğer odak konuları olan “mimarlık”, “çevre”, “sağlık hizmeti”, “tasarımla iyileşme” ve “hastane yapıları” anahtar kelimeleri sorguya dahil edilmiştir. Araştırma soruları çerçevesinde konuyla ilgili çeşitli ve benzer anahtar kelimeler kullanılarak yapılan birçok deneme sorgusu sonucunda anahtar kelimeler belirlenmiştir. Kapsamlı bir araştırma yapmak için arama sürecinden hiçbir anahtar kelime dışlanmamıştır. Bu doğrultuda ilk sorgu şu şekilde gerçekleştirilmiştir: ALL= (“healing”) AND ALL= (“design”). Bu orijinal arama sonucunda, Web of Science veri tabanında 46.750 döküman bulunmuştur. Daha sonra yapılan sorguya ALL= (“environment”) AND ALL= (“architecture”) AND ALL= (“healthcare”) AND ALL=

(“healing by design”) AND ALL= (“hospital”) anahtar kelimeleri eklenmiştir. Bu orijinal arama sonucunda da, Web of Science veri tabanında 70 döküman bulunmuştur.

2.4. Veritabanı Araması ve Belgelerin Öncelik Seçimi

Web of Science veri tabanında, gelişmiş arama sekmesi üzerinden anahtar kelimeler ile yapılan arama sonuçları ve belge seçim süreci Şekil 1'de ifade edilmiştir.



Şekil 1: Belge seçim süreci şeması.

Şekil 1’de görüldüğü gibi, “healing” anahtar kelimesi ile yapılan ilk arama sonucunda 359.894 belge bulunmuştur. Sonrasında, tüm alanlar tekrar seçilerek ve "design" anahtar kelimesi ile 46.750; "environment" anahtar kelimesi ile 5.444; "architecture" anahtar kelimesi ile 638; "healthcare" anahtar kelimesi ile 150; "healing by design" anahtar kelimesi ile 103; "hospital" anahtar kelimesi ile 70 belge bulunmuştur. Bu 70 belge alanlarına göre mimarlık ve sağlık alanları ile kısıtlandırıldığında 50 belge bulunmuştur. Söz konusu 50 belge içinden sadece mimarlık alanında yapılmış çalışmalar seçildiğinde toplam 8 belge bulunmuş ve çalışmaya bu belgeler dahil edilmiştir. Araştırmada bibliyometrik verilerine ulaşılabilen en eski belge 2010 tarihli belgedir. Bu nedenle, sadece 2010 yılından 2025 yılına kadar olan belgeler çalışmaya dahil edilmiş ve dil sınırı getirilmemiştir.

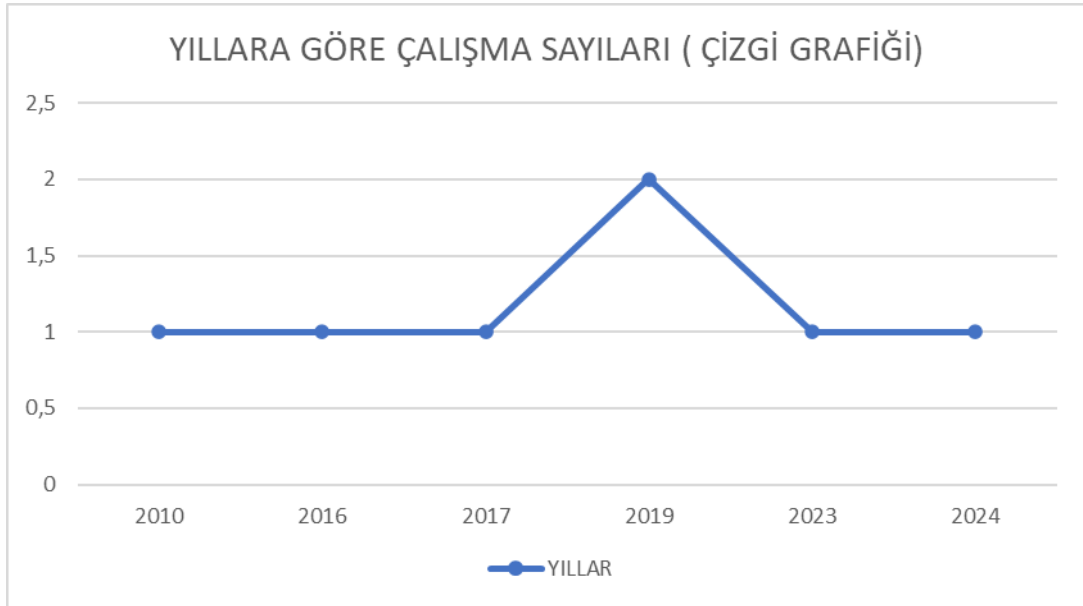
3. BULGULAR

Web of Science veritabanından sorgu sonucunda elde edilen çalışmalar, belge türlerine göre analiz edilmiştir ve sonuçlar Çizelge 2'de gösterilmiştir. Bu bağlamda, 8 yayından, 4'ü makaleden, 2'si bildiri metninden, geriye kalan 2 'si derleme makalesinden oluşmaktadır.

Çizelge 2: Yayınların belge türüne göre dağılımı.

Belge Tipi	Belge Sayısı	% Değeri
Makale	4	50
Bildiri	2	25
Derleme Makalesi	2	25

Sorgu sonuçları, en eski tarihlenmiş belgenin 2010 yılına ait olduğunu ve en güncel belgenin 2025 yılına ait olduğunu ortaya koymaktadır. Analiz doğrultusunda, ilgili alanda 2025, 2024, 2023, 2017, 2016 ve 2010 yıllarında birer çalışma yapıldığı 2019 yılında 2 çalışma yapıldığı Şekil 2' de gösterilmiştir.



Şekil 2: Yıllara göre belge sayısının dağılımı (Web of Science).

Çizelge 3, iyileştirme ve tasarımın kesişimindeki çalışmanın odak konularının dahil edildiği arama sonucu ulaşılan çalışmaya dahil edilen 8 belgeye ilişkin bilgileri sunar.

Çizelge 3: Çalışmaya dahil edilen 8 belgenin içerik çizelgesi.

Yazar Adı	Yayın Yılı	Yayın Başlığı	Konunun İçeriği	Makalede Geçen Anahtar Kelimeler	Yayın Türü	Atılan Referans Sayısı
Farshin Salehi, Sima Vaez Eslami	2010	Green strategies of healthcare design: Case studies of medical centers	Sağlık tasarımı için yeşil stratejiler. Vaka çalışması örnekleri.	healthcare, sustainable, case study, green building, hospitals	Konferans Bildirisi	9
Haimanti Banerji	2016	An attempt to explore components of empathic architecture in hospitals – a study of Indian hospitals	Hint hastanelerinde empatik mimarinin bileşenlerinin keşfedilmesi.	empathic architecture, hospitals, India	Derleme	44
Anna Avrutch, Agnieszka Gebczynska-Janowicz	2017	Art and healthcare – healing potential of artistic interventions in medical settings	Tıbbi ortamlarda sanatsal müdahalelerin iyileştirici potansiyeli.	art, healthcare, artistic interventions, medical settings	Konferans Bildirisi	36
Mohamed S. Abdelad, Veronica Soebarto	2019	Biophilia and salutogenesis as restorative design approaches in healthcare architecture	Sağlık mimarisinde onarıcı tasarım yaklaşımları olarak biyofili ve salutogenezin incelenmesi.	healing environment, biophilic, salutogenic, restorative healthcare, environmental design, attention restoration theory	Makale	98
Yasushi Nagasawa	2019	Global Hospitals in 2050 – A review of the historical development of hospital building studies from a global perspective	Hastane binalarının tarihsel gelişiminin küresel bir bakış açısıyla incelenmesi.	hospital, healthcare, architectural planning, design, historical review, global perspectives	Derleme	75
Traitwat Viryasari, Vikrom Lavorrstichai, Kullathida Sangnin, Kavin Dhanakoses, Palkwan Roopkaew, Pundabere Viryasiri	2023	Hospital design principles implementation: Reflections from practitioners in Thailand	Gelişmekte olan ülkelerde hastane tasarım ilkelerinin uygulanması.	hospital design principles, healthcare architecture, developing countries, healing environment, sustainability, construction	Makale	53
Teresa Sánchez-Jáuregui, Nicolás Gutiérrez-Pérez, Tomás Abad Balboa, Pilar Chías	2024	The relationship between healthcare architecture and nature: interaction with the landscape as therapy	Sağlık mimarisi ve doğa arasındaki ilişkiyi, peyzajla etkileşimi bir terapi olarak incelemek.	healing gardens; health centers; analysis; universal accessibility; environment	Makale	45
Bakhtiar Bahrami, Sara Kariminejad, Goran Erfau	2025	Child-friendly design in paediatric healthcare environments: exploring children's preferences for nature, space, and play	Pediyatrik sağlık ortamlarında çocuk dostu tasarım, çocukların doğa, mekân ve oyun tercihlerinin araştırılması.	child-friendly design; paediatric environment; spatial preferences; socio-spatial characteristics; playful environment	Makale (Erken Erişim)	55

Çizelge 4. Belgelerin araştırma alanları.

Web of Science Kategorileri	Belge Sayısı	% Değeri
Halk Sağlığı, Çevre ve İş Sağlığı	19	38
İnşaat ve Yapı Teknolojisi	16	32
İnşaat Mühendisliği	13	26
Mimarlık	8	16
Çevre Mühendisliği	4	8
Kentsel Çalışmalar	4	8
Çevre Çalışmaları	3	6
Yeşil Sürdürülebilir Bilim ve Teknoloji	2	4
Sağlık Hizmetleri (Genel)	2	4
Sağlık Politikaları Hizmetleri	2	4
Sanat	1	2
Yoğun Bakım Tıbbı	1	2
Etik	1	2
Ormancılık	1	2
Yönetim	1	2
Malzeme Bilimi (Multidisipliner)	1	2
Tıbbi Etik	1	2
Dahiliye (Genel Tıp)	1	2
Hemşirelik	1	2
Bitki Bilimleri	1	2
Psikiyatri	1	2
Bölgesel Kentsel Planlama	1	2
Sosyal Konular	1	2
Sosyal Bilimler (Biyomedikal)	1	2

Çizelge 3’deki yayınlar, sağlık mimarisinde öne çıkan eğilimleri göstermektedir. Çalışmaların çoğu makale olup biyofili, salutogenez, doğayla bütünleşme, empati temelli yaklaşımlar ve çocuk dostu tasarım gibi insan odaklı konulara odaklanmaktadır. Ayrıca sürdürülebilirlik, çevreci stratejiler ve gelişmekte olan ülkelerde hastane tasarımı gibi bağlama özgü başlıklar da incelenmiştir. Alıntı sayılarındaki farklılık, bazı araştırmaların (örneğin Abdelaal & Soebarto, 2019) kuramsal düzeyde önemli etki yarattığını, bazılarının ise uygulamaya dönük katkılar sunduğunu ortaya koymaktadır. Genel olarak bulgular, sağlık mimarisi literatürünün hem iyileştirici hem de sürdürülebilir ve kullanıcı merkezli yaklaşımlarla şekillendiğini göstermektedir.

Çizelge 4, taramalar sonucunda elde edilen 50 belgenin araştırma alanlarını ifade etmektedir. Bu alanlar içerisinde mimarlıkla ilgili olan 8 belge çalışma kapsamında detaylıca incelenmiştir. Not: Belgeler birden fazla alana dahil edilmiş olabilir. Çalışmaların büyük ölçüde **halk sağlığı, çevre ve iş sağlığı (%38), inşaat ve yapı teknolojisi (%32) ve inşaat mühendisliği (%26)** alanlarında yoğunlaştığı görülmektedir. Bununla birlikte **mimarlık (%16) ve çevre mühendisliği (%8)** kategorilerinin de önem kazandığı görülmektedir.

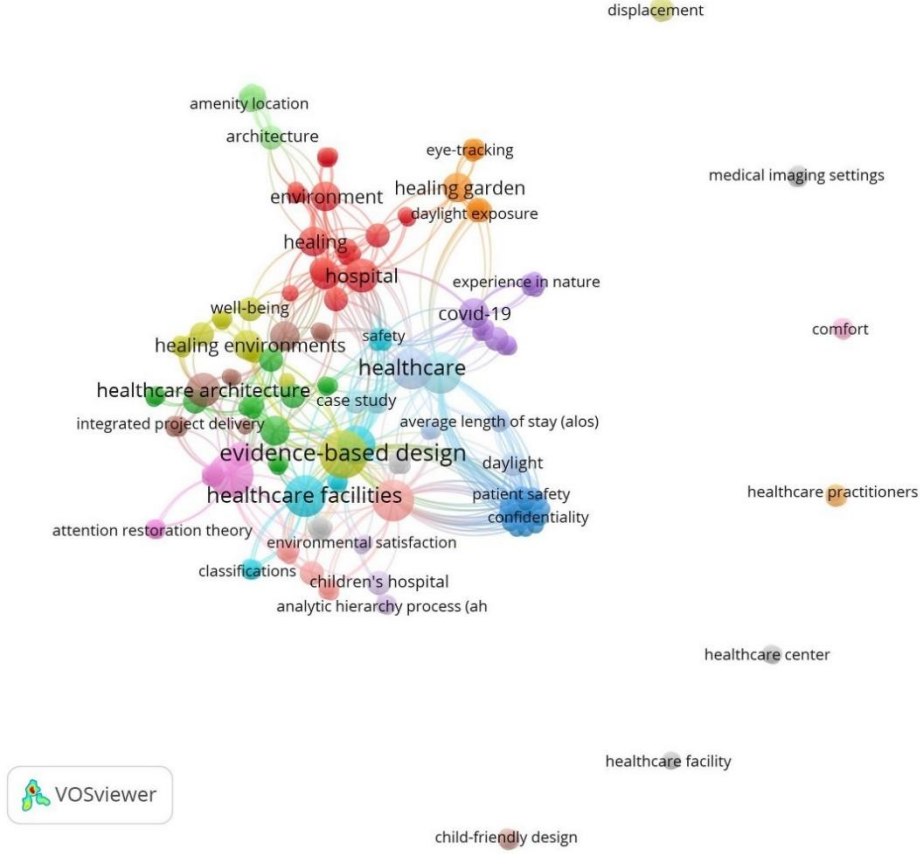
4. ANALİZ VE SONUÇLAR

Araştırma iki aşamalı olarak yürütülmüştür. İlk aşamada, Web of Science veritabanından elde edilen mimarlık ve sağlık alanlarındaki 50 belge; en çok kullanılan terimler, araştırma sıklığı, yayın tarihi, gibi faktörlerin kapsamlı analizi incelenmiş ve VOSviewer yazılımı aracılığıyla ağ analizleri gerçekleştirilmiştir. Bu aşamada amaç, konuyu genel bir çerçevede ele alarak mimarlık ve sağlık alanlarının birbiriyle ilişkisini ortaya koymak ve yorumlamaktır. İkinci aşamada ise, yalnızca mimarlık alanına ait seçilen 8 belgenin analizi yapılmış; böylece araştırma daha spesifik bir çerçevede sınırlandırılmıştır. Her iki aşamada da bibliyometrik verilerin değerlendirilmesi, VOSviewer’da dört farklı bilimsel haritalama tekniği kullanılarak yapılmıştır.

4.1. Anahtar Kelimelerinin Birlikte Görünüm Analizi

Bibliyometrik analizler, mimarlık disiplininde “iyileştirme” kavramının oldukça yeni olduğunu ve dolayısıyla bu konu üzerine yapılan çalışmaların sayısının oldukça az olduğunu ortaya koymaktadır. Bu nedenle “iyileştirici tasarım” kavramı da incelemeye dahil edilmiş, sonrasında “sağlık çevreleri”, “mimarlık” ve “hastane yapıları” kavramları da konuya dahil edilip inceleme detaylandırılmıştır. Araştırma alanına bütünsel bir bakış açısı kazanmak amacıyla, belgelerde sıkça kullanılan anahtar kelimeleri analiz etmek için en az 1 görünüm, eşik olarak belirlenmiştir. Buna göre arama yapılan Web of Science veri tabanından VOSviewer yazılımına aktarılan 50 belgenin verileri sonucunda toplamda 220 anahtar kelime elde edilmiştir. Ayrıca, her bir anahtar kelime için en yüksek toplam bağlantı gücüne sahip 220 anahtar kelime seçilerek toplam bağlantı gücü hesaplanmıştır. Tüm anahtar kelimeler Şekil 3 ‘te anahtar kelimelerin ağ analizi incelendiğinde haritanın merkezi kısmı mimarlık ile sağlık kavramlarının sıkı biçimde iç içe geçtiğini göstermektedir. Özellikle “evidence-based design”, “healthcare facilities”, “hospital”, “healthcare”, “healing / healing environments”, “healthcare architecture” gibi düğümler büyük ve birbirlerine yoğun bağlantılarla bağlıdır. Bu durum tasarım kararlarının sağlık çıktılarıyla (iyileşme, hasta güvenliği, çevresel memnuniyet vb.) sıkça ilişkilendirildiğini göstermektedir. Haritada “comfort”, “medical imaging settings” ve “healthcare practitioners” gibi kavramların ağın merkezinden uzak konumlandığı görülmektedir. Bu durum, söz konusu temaların mimarlık ve sağlık ilişkisi bağlamında

literatürde daha az ele alındığını göstermektedir. Benzer şekilde, child-friendly design (çocuk dostu tasarım) gibi belirli alt konuların da diğer kavramlardan daha izole olduğu fark edilmektedir. Bu bulgular, bu temaların gelecekte daha ayrıntılı biçimde incelenebileceğini ve disiplinler arası çalışmalarda yeni araştırma alanları için önemli bir potansiyel taşıdığını ortaya koymaktadır.



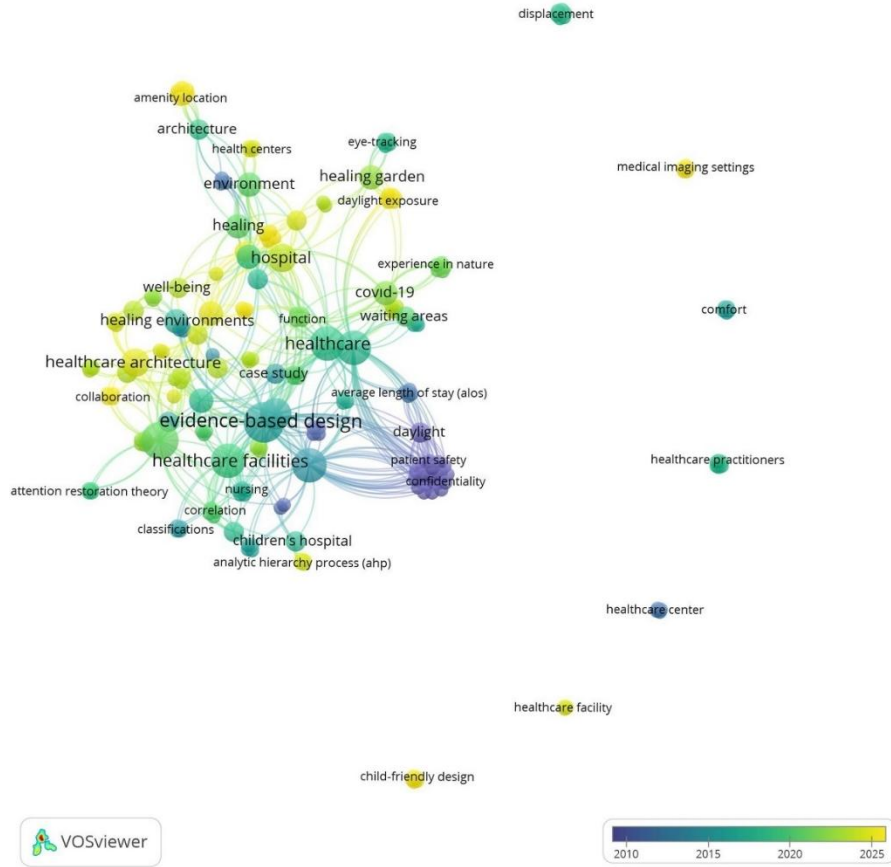
Şekil 3: Anahtar kelimelerinin birlikte görünüm analizinin ağ görselleştirmesi (VOSviewer, 2025).

Bir anahtar kelimenin kümesi, bu kelimenin rengine atıfta bulunmaktadır. Şekil 3'teki, her renk bir noktanın farklı kümelerin renklerinin karıştırılmasıyla belirlendiği Küme Yoğunluğu Görselleştirmesini göstermektedir (Van Eck ve Waltman, 2010).

- **Kırmızı/turuncu kümeler:** hospital, healing, healing garden, daylight exposure gibi sağlık mekânı ve iyileştirici çevre öğelerine odaklanmaktadır — klinik mekanların fiziksel özellikleri (ışık, bahçe, çevre) ile iyileşme ilişkisi.
- **Yeşil/koyu yeşil kümeler:** healthcare architecture, healing environments, well-being gibi mimarlık perspektifli, tasarım odaklı kavramları gruplamaktadır.
- **Mavi/açık mavi kümeler:** healthcare, patient safety, confidentiality, average length of stay gibi sağlık hizmeti performans ölçütleri ve hasta güvenliğiyle ilişkili terimlerdir.

- **Mor ve benzeri kümeler:** covid-19, experience in nature gibi daha spesifik veya dönemsel temaları (COVID etkisi, doğa deneyiminin rolü) göstermektedir.
- **Uzak/dışta kalan düğümler:** comfort, medical imaging settings, displacement, healthcare practitioners, child-friendly design vb. — bunlar çekirdek kümeye göre daha az bağlantılı, dolayısıyla literatürde daha izole veya az çalışılmış alt-temaları işaret etmektedir.

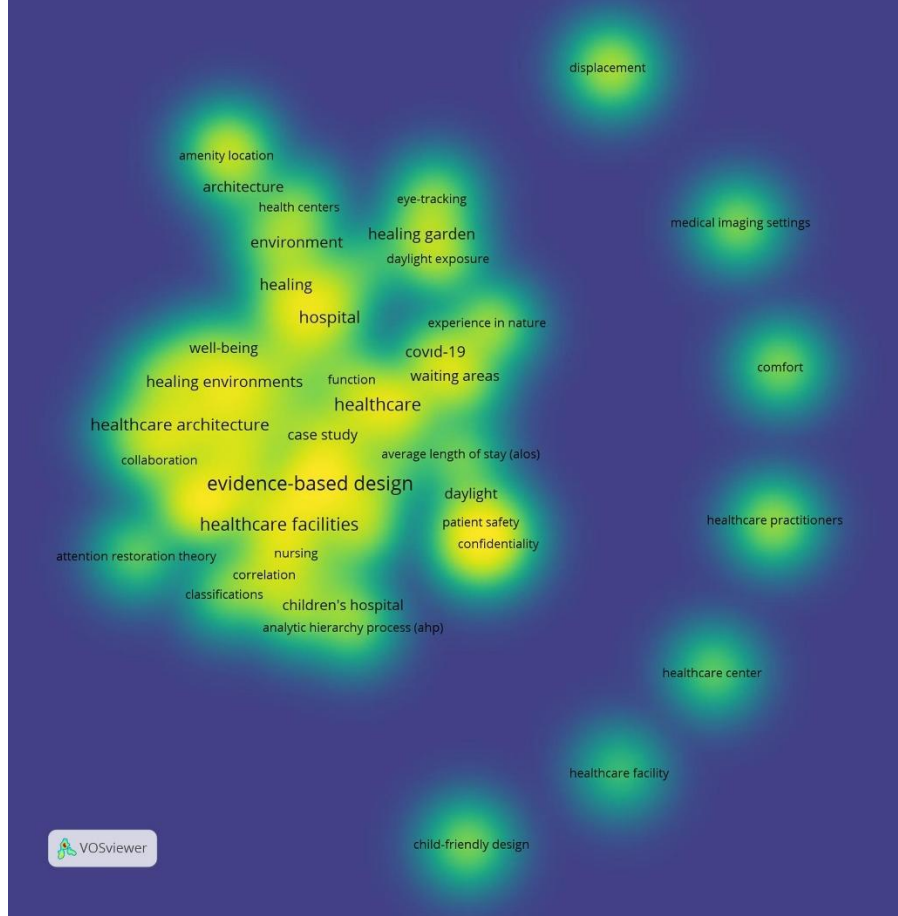
Şekil 4, anahtar kelimelerin ağırlık ve puan analizini göstermektedir. Kelimelerin ağırlığı, kelimelerin ortaya çıkma sıklığına göre seçilmiştir ve puanlar, yayın yılı ortalamasına göre seçilmiştir. Resmin sağ alt köşesinde, anahtar kelimeler ve yayın yılları arasındaki ilişkiyi gösteren renk kodlu bir ölçek bulunmaktadır. Çağdaş araştırma alanları sarı ve açık yeşil ile temsil edilmektedir.



Şekil 4: Anahtar kelimelerin yıllara göre birlikte görünüm analizi için katmanlı görselleştirme (VOSviewer, 2025).

Bu ağ görselleştirmesi (Şekil 4), mimarlık ve sağlık alanlarında yapılan araştırmalarda anahtar kelimelerin yıllara göre dağılımını ve bu kavramlar arasındaki ilişkileri göstermektedir. Haritada yer alan renkler, anahtar kelimelerin araştırmalarda öne çıktığı dönemleri temsil eder; mavi tonlar daha erken yılları (2010–2015), yeşil tonlar orta dönemi (2016–2020), sarı tonlar ise daha yeni çalışmaları (2021–2025) göstermektedir. Haritanın merkezinde yer alan “evidence-based design”, “healthcare facilities”, “hospital”

ve “healthcare” kavramlarının mavi-yeşil tonlarda olması, bu konuların özellikle 2010’lu yıllardan itibaren sürekli olarak araştırmalarda yer aldığını ve disiplinin temelini oluşturduğunu göstermektedir. Bu kavramlar, mimarlık ve sağlık ilişkisini açıklayan ana eksen oluşturmakta ve uzun süredir literatürde istikrarlı bir şekilde ele alınmaktadır.



Şekil 5: Anahtar kelimelerin yoğunluk görselleştirmesi (VOSviewer, 2025).

Son yıllarda ise (sarı renkle temsil edilen alanlarda) “healing garden”, “daylight exposure”, “well-being” ve “child-friendly design” gibi konuların daha fazla ilgi görmeye başladığı görülmektedir. Bu durum, özellikle kullanıcı deneyimi, doğa ile etkileşim ve iyileştirici çevrelerin tasarım sürecinde önem kazandığını ve yeni araştırma eğilimlerinin daha insan merkezli bir yaklaşıma yöneldiğini göstermektedir. Şekil 5’te, anahtar kelimelerin ortaya çıkma sıklığı, Eleman Yoğunluk Görselleştirmesi’nde boyutlar ve sıcak-soğuk renkler ile temsil edilmektedir. Bu görselleştirmedeki her nokta, anahtar kelimenin yoğunluğunu temsil eden bir renge sahiptir. Varsayılan renkler mavi, yeşil ve sarıdır ve noktanın yanındaki daha fazla nesne ve bu nesnelerin daha büyük ağırlığı ile renk daha sarı hale gelir. “Evidence-based design” hem merkezde hem de yüksek yoğunluktadır, bu da son yıllarda sağlık mimarisinde bilimsel kanıt temelli yaklaşımların baskın hale geldiğini göstermektedir. “Healing environments”, “daylight”, “well-being” gibi kelimelerin yakın konumlanması, kullanıcı deneyimi ve psikolojik iyileşme

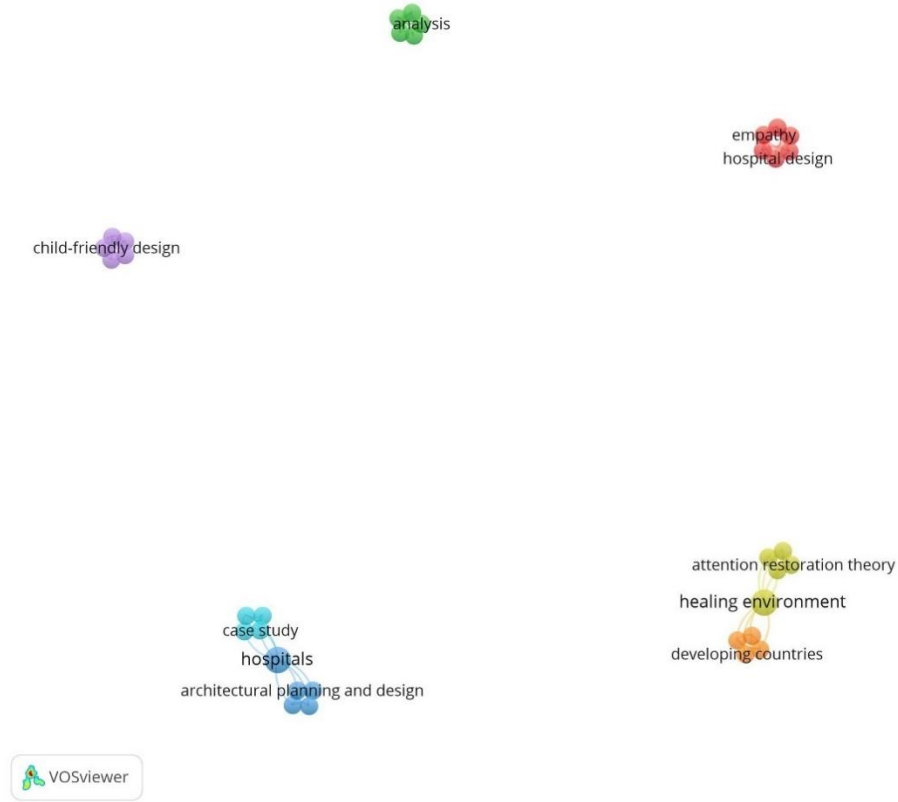
odaklı tasarımların önem kazandığını göstermektedir. “Covid-19” kelimesinin merkezde yer alması, pandeminin sağlık mekânları üzerine etkisini artıran güncel bir araştırma teması olduğunu göstermektedir. Dış halkalarda yer alan kelimeler (ör. “displacement”, “comfort”, “healthcare practitioners”) daha spesifik ve az bağlantılı alt konular olarak değerlendirilebilir.

Çizelge 5: Anahtar kelimelerin oluşumu ve toplam bağlantı gücü (Web of Science, 2025).

Anahtar Kelime	Geçiş Sayısı (Occurrences)	Toplam Bağlantı Gücü (Total Link Strength)
1. healing environment	2	8
2. hospitals	2	8
3. empathy	1	5
4. healing architecture	1	5
5. hospital design	1	5
6. psychological comfort	1	5
7. stressor	1	5
8. therapeutic environment	1	5
9. analysis	1	4
10. architectural planning and design	1	4
11. attention restoration theory	1	4
12. biophilic	1	4
13. case study	1	4
14. child-friendly design	1	4
15. developing countries	1	4
16. environment	1	4
17. global perspectives	1	4
18. green building	1	4
19. healing gardens	1	4
20. health centers	1	4
21. healthcare	1	4
22. healthcare architecture	1	4
23. healthcare facilities	1	4
24. historical review	1	4
25. hospital design principles	1	4
26. paediatric healthcare environment	1	4
27. playful environment	1	4
28. restorative healthcare environmental design	1	4
29. salutogenic	1	4
30. socio-spatial characteristics	1	4
31. spatial preferences	1	4
32. sustainability	1	4
33. sustainable	1	4
34. universal accessibility	1	4

Araştırmanın ikinci aşamasında arama yapılan Web of Science veri tabanından VOSviewer yazılımına aktarılan 8 belgenin verileri sonucunda toplamda 34 anahtar kelime elde edilmiştir. Ayrıca, her bir anahtar kelime için en yüksek toplam bağlantı gücüne sahip 34 anahtar kelime seçilerek toplam bağlantı gücü hesaplanmıştır. Tüm anahtar kelimeler Çizelge 5’te listelenmiş ve görünüm sayısına göre sıralanmıştır. Tablodaki anahtar kelime dağılımı incelendiğinde, **“healing environment”** ve **“hospitals”** kavramlarının

en yüksek geçiş sayısı (2) ve toplam bağlantı gücü (8) ile öne çıktığı görülmektedir. Bu durum, sağlık mimarisi literatüründe iyileştirici ortamların ve hastane yapılarının merkezî bir araştırma odağı olduğunu göstermektedir. Bunun yanında “empathy”, “healing architecture”, “hospital design” ve “psychological comfort” gibi kullanıcı deneyimine ve kullanıcı odaklı tasarıma vurgu yapan kavramların da dikkate değer bağlantı gücüyle öne çıktığı anlaşılmaktadır. Ayrıca “child-friendly design”, “biophilic”, “sustainability” ve “green building” gibi kavramlar, sağlık yapılarında sürdürülebilirlik ve doğa ile bütünleşmenin önemini ortaya koymaktadır. Genel olarak tablo, sağlık mimarisi çalışmalarının yalnızca yapısal ve işlevsel değil, aynı zamanda psikolojik, sosyal ve çevresel boyutları kapsayan çok yönlü bir araştırma alanına dönüştüğünü işaret etmektedir.



Şekil 6: Anahtar kelimelerinin birlikte görünüm analizinin ağ görselleştirmesi (VOSviewer, 2025).

Şekil 6, anahtar kelimelerin kullanım sıklığına göre ağ görselleştirmesini göstermektedir. Haritada yer alan düğümler, ilgili anahtar kelimeleri temsil etmekte; düğümler arasındaki çizgiler ise bu kavramların akademik çalışmalarda birlikte kullanım sıklığını göstermektedir. Ayrıca renkler, farklı kümeleri yani birbirine daha sıkı bağlanmış araştırma alanlarını ortaya koymaktadır. 34 anahtar kelime toplanmış ve bu kelimeler özelliklerine, uygulama alanlarına ve VOSviewer bağlantılarına bağlı olarak 7 kümede gruplandırılmıştır. Küme içlerindeki anahtar kelimelerin hepsi ağ görselleştirmesinde gözükmemektedir. Bunun sebebi birbirleriyle bağlantılı kümeler içindeki en yüksek bağlantı gücüne sahip anahtar

kelimelerin ön plana çıkmasıdır, diğer bağlantılı kelimeler VOSviewer programı içerisinde kümeye yaklaştıkça belli olmaktadır. Ağdaki her anahtar kelime yalnızca bir kümenin parçasıdır.

Bir anahtar kelimenin kümesi, bu kelimenin rengine atıfta bulunmaktadır. Bu analiz, sağlık yapıları ve iyileştirici çevreler temaları etrafında şekillenen bir literatür taraması veya bibliyometrik çalışmanın sonuçlarını göstermektedir. Harita (Şekil 6), kümelerin net bir şekilde ayrıldığını ve aralarındaki bağlantıların zayıf veya seyrek olduğunu gösteriyor; bu da farklı temaların **izole** kaldığını veya aralarında çok fazla ortak yayın olmadığını düşündürmektedir. Ana temaların her biri farklı bir alana odaklanmış görünmektedir:

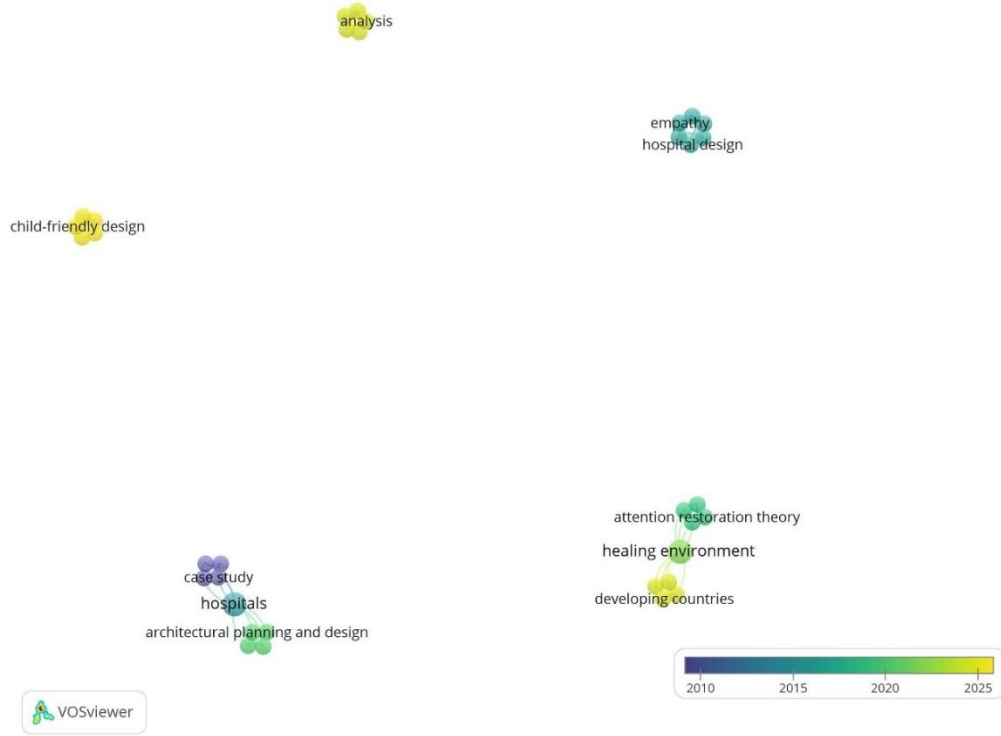
- **Kırmızı Küme :** “empathy” ve “hospital design” merkezli (Küme 1’e karşılık geliyor), tasarımın insancıl yönüne odaklanıyor.
- **Mavi/Turkuaz Küme :** “case study”, “hospitals”, “architectural planning” and “design” merkezli (Küme 3 ve 6’ya karşılık geliyor), mimari uygulama ve hastane yapılarına odaklanıyor.
- **Yeşil Küme:** Sadece “analysis” kelimesini gösteriyor (Küme 2’nin tek görünür bileşeni).
- **Mor Küme** Sadece “child-friendly design” kelimesini gösteriyor (Küme 5’in tek görünür bileşeni).
- **Sarı/Turuncu Küme:** “attention restoration theory”, “healing environment”, “developing countries” merkezli (Küme 4 ve 7’den unsurlar içeriyor). Çevre, dikkat yenilenme kuramına ve küresel bağlama odaklanıyor.

Çizelge 6, anahtar kelime ve küme kombinasyonlarının listesini göstermektedir. Küme içindeki en sık meydana gelen anahtar kelimeler kalın harfle vurgulanmıştır. Analiz, 34 anahtar kelimeyi 7 farklı tematik gruba ayırmıştır.

Çizelge 6: Anahtar kelime ve küme kombinasyonlarının listesi.

Kümeler	Anahtar Kelimeler
Küme 1 (6 kelime)	empathy, healing architecture, hospital design, psychological comfort, stressor, therapeutic environment
Küme 2 (5 kelime)	analysis , environment, healing gardens, health centers, uni-versal accessibility
Küme 3 (5 kelime)	architectural planning and design , global perspectives, healthcare facilities, historical review, hospitals
Küme 4 (5 kelime)	attention restoration theory , biophilic, healing environment, restorative healthcare environment, salutogenic
Küme 5 (5 kelime)	child-friendly design , paediatric healthcare environment, playful environment, socio-spatial characteristics, spatial preferences
Küme 6 (4 kelime)	case study, green building, healthcare, sustainable
Küme 7 (4 kelime)	developing countries, healthcare architecture, hospital design principles, sustainability

Bu kümeler haritada belirgin bir şekilde ayrılmış görünerek, literatürde nispeten izole ve farklı alt alanlara odaklanıldığını gösterir. Temel kümeler arasında Duygusal ve terapötik tasarım (Empathy, Hospital design), Salutogenez ve Restoratif İyileşme Yaklaşımları (Attention restoration theory, Salutogenic, Healing environment) ve Sürdürülebilirlik ve Kapsayıcı Tasarım (Child-friendly design, Developing countries, Green building) bulunmaktadır. Hastane tasarımı (Hospital design), Empati (Empathy) ile güçlü bir bağ (5) kurarak, konunun insan odaklı ve psikolojik yönünün önemini vurgulamaktadır. Bu kümelenme ve bağlantı güçleri, sağlık mimarisi araştırmalarının, teoriden pratiğe uzanarak “çocuk dostu tasarım” ve “sürdürülebilirlik” gibi özel alanlarda da geliştiğini göstermektedir.

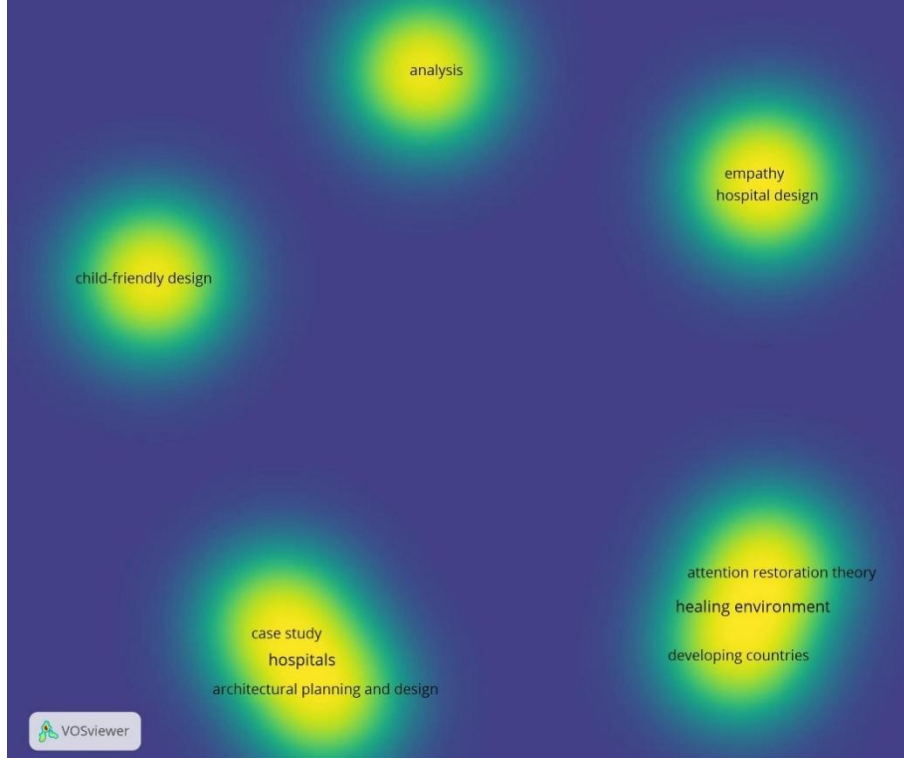


Şekil 7: Anahtar kelimelerin yıllara göre birlikte görünüm analizi için katmanlı görselleştirme (VOSviewer, 2025).

Çizelge 7: Anahtar kelimelerin yayın yılına göre analizi (VOSviewer, 2025).

Yayın Dönemi (Yaklaşık)	Anahtar Kelimeler
2010–2015 (erken dönem)	hospitals, case study, architectural planning and design, historical review, healthcare facilities, global perspectives
2015–2020 (orta dönem)	healing architecture, hospital design, empathy, therapeutic environment, psychological comfort, health centers, environment, universal accessibility, stressor, healing gardens
2020–2023 (güncel dönem)	healing environment, attention restoration theory, biophilic, restorative healthcare environmental design, salutogenic, green building, healthcare
2023–2025 (en güncel dönem)	child-friendly design, paediatric healthcare environment, playful environment, socio-spatial characteristics, spatial preferences, developing countries, healthcare architecture, hospital design principles, sustainability, sustainable

Şekil 7, anahtar kelimelerin ağırlık ve puan analizini göstermektedir. Kelimelerin ağırlığı, kelimelerin ortaya çıkma sıklığına göre seçilmiştir ve puanlar, yayın yılı ortalamasına göre seçildi. Resmin sağ alt köşesinde, anahtar kelimeler ve yayın yılları arasındaki ilişkiyi gösteren renk kodlu bir ölçek bulunmaktadır. Çağdaş araştırma alanları sarı ve açık yeşil ile temsil edilmektedir. Çizelge 7, anahtar kelimelerin yayın yılına göre analizini göstermektedir.



Şekil 8: Öğelerin yoğunluk görselleştirmesi (VOSviewer, 2025).

Şekil 8’de, anahtar kelimelerin ortaya çıkma sıklığı, Eleman Yoğunluk Görselleştirmesi’nde boyutlar ve sıcak-soğuk renkler ile temsil edilmektedir. Bu görselleştirmedeki her nokta, anahtar kelimenin yoğunluğunu temsil eden bir renge sahiptir. Varsayılan renkler mavi, yeşil ve sarıdır ve noktanın yanındaki daha fazla nesne ve bu nesnelerin daha büyük ağırlığı ile renk daha sarı hale gelir. Bu açıdan, “healing environment”, “hospital design” , ve “architectural planning and design” ve “child-friendly design” , “analysis” anahtar kelimelerinin yoğun olduğu görülmektedir. Bu analize göre mimarlık ve hastane tasarımında bilimsel araştırmaların, teknik detayların yanı sıra sosyal, psikolojik ve insani boyutları da içeren kapsamlı bir yaklaşımla ele alındığı söylenebilir.

4.2. Yazarların Atıf Analizi

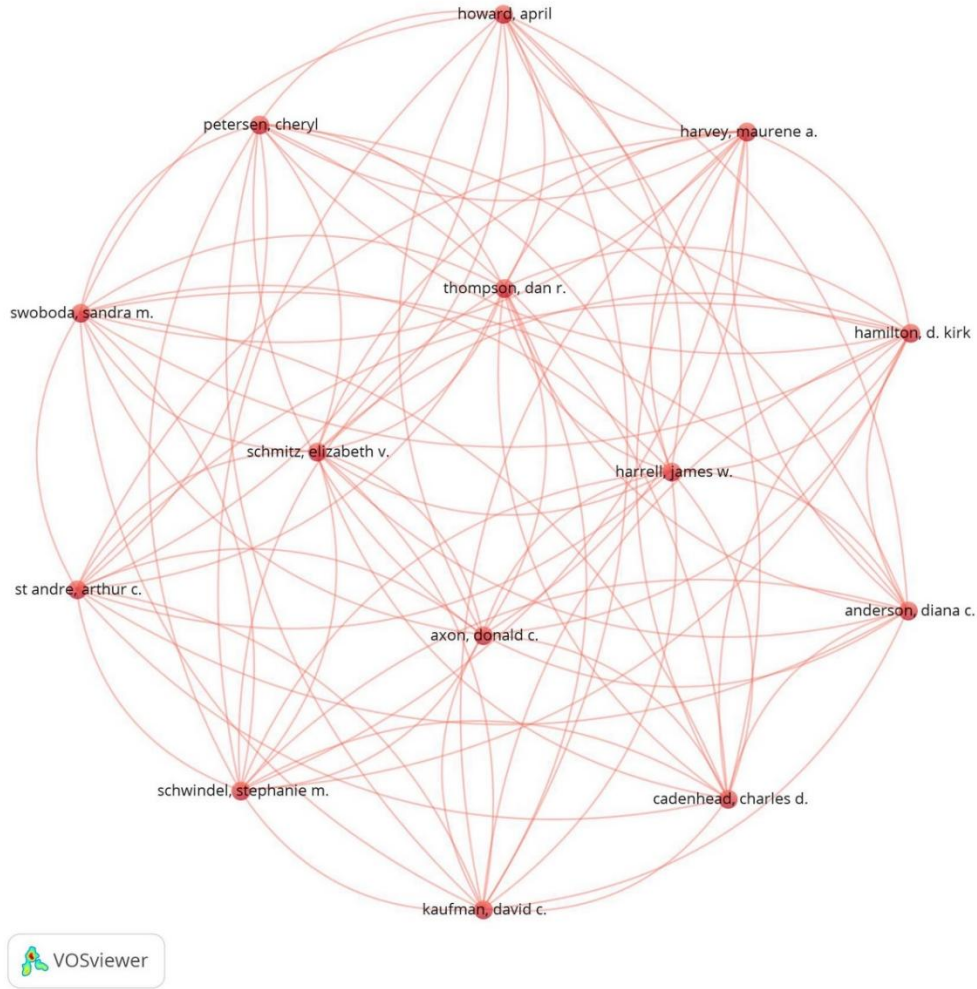
Yazarların atıf analizinde, VOSviewer’a “en az 1 belgeden bir yazar” ve “en az 1 atıf yapan bir yazar” gibi eşik değerleri girilmiştir. Analiz sonucunda, 173 yazar bu eşik değerlerini karşılamıştır. 173 yazardan 14 tanesi birbirleriyle bağlantı gücüne sahip yazarlardır. Çizelge 8, bu bağlantı gücüne ve atıf sayısına sahip 14 yazarı listelemektedir. Bu analize göre, ilk dört yazar (Coelho, Granato, Marin, Suzuki) en yüksek atıfa (104) sahiptir. Sonraki beş yazar 38 atıf ile eşittir, 10. Sırada ise Beutel, Bryan G. 32 atıf ile yer almaktadır. Diğer yazarlara yapılan alıntı bağlantılarının toplam gücü, 173 yazarın her biri için hesaplanmıştır. VOSviewer, ağdaki 173 yazardan bazılarının hiçbir yazara bağlı olmadığını göstermiştir ve bağlı yazarların en kalabalık kümesi 14 yazardan oluşmaktadır. Vosviewer’ dan elde edilen verilere göre, 173 yazardan en yüksek atıfa sahip yazarlar: Guan Xiaobo ve Sea Hyun-Bo (8741 atıf). Onları Dubose, Jennifer, Choi, Young-Seon, Ulrich, Roger S., Zhu, Xuemei ve Zimring, Craig takip etmektedir (874 atıf).

Çizelge 8: Yayın yıllarına göre birbirlerine atıf yapan ve bağlantı gücüne sahip 14 yazar çizelgesi (VOSviewer, 2025).

Yazar	Yayın Yılı (yaklaşık)	Atıf Sayısı	Bağlantı Gücü
1. Howard, April	2015 civarı	13	13
2. Petersen, Cheryl	2016–2017	13	13
3. Harvey, Maurene A.	2016	13	13
4. Thompson, Dan R.	2017–2018	13	13
5. Hamilton, D. Kirk	2018	13	13
6. Harrell, James W.	2018–2019	13	13
7. Anderson, Diana C.	2020 civarı	179	1
8. Cadenhead, Charles D.	2020	179	1
9. Kaufman, David C.	2017	13	13
10. Schwindel, Stephanie M.	2016–2017	13	13
11. St Andre, Arthur C.	2020	179	1
12. Schmitz, Elizabeth V.	2017 civarı	13	13
13. Swoboda (Sandra M.)	2019–2020	179	1
14. Axon, Donald C.	2020 civarı	179	1

Bu grup da çok güçlü bir çekirdek oluşturmaktadır. Geri kalan (örn. Schwindel, Axon, Thompson) daha düşük atıf seviyesindedir (179 civarı).Bağlantı gücü 13 olan yazarlar (örn. Anderson, Schmitz, Cadenhead, Capolongo, Gala, Hamilton, St Andre, Schwindel, Axon) ağı merkezinde olan araştırmacılarıdır. Bu yazarların atıf sayıları birbirinden farklı olsa da (175, 179, 105 vs.), hepsi ağ yapısında “köprü” görevi görmektedir. Bağlantı gücü 7 olan yüksek atıf grubu (Guan, Sea, Ulrich, Zhu, Zimring, Choi, Dubose). Howard, Petersen, Harvey, Thompson, Hamilton, Harrell, Kaufman, Schwindel, Schmitz hepsi 13 atıf ve güçlü bağlantı (13) ile aynı kümede yer almaktadır. Anderson, Cadenhead, St Andre, Swoboda, Axon hepsi 179 atıf ama bağlantı gücü düşüktür (1). Bu da gösteriyor ki ilk grup küçük çaplı ama sıkı bağlı bir ağda, ikinci grup ise daha büyük atıf almış ama diğerleriyle çok az bağlantıdadır.Vosviewer programından elde edilen yazarlar ve küme kombinasyonları verilerine göre; 173 yazar toplam 43 kümede toplanmıştır.

Bu analiz, en kalabalık işbirliği yapılarından bireysel çalışmalara kadar bir dağılım sergilemektedir. En geniş katılımlı grup, 14 yazar ile Küme 1’dir. Bu kümenin ana üyeleri arasında Anderson, Diana C., Axon, Donald C. Ve Thompson, Dan R. Gibi isimler bulunmaktadır. Şekil 9’da bu kümede bulunan yazarların alıntı bağlantılarının ağ görselleştirmesi gösterilmektedir. 173 yazarın bulunduğu kümelerin hepsi ağ görselleştirmesinde gözükmemektedir. Sadece birbirleri ile bağlantısı olan 14 yazarın bulunduğu küme ağ görselleştirmesinde gözükmemektedir. Diğer kümelerdeki yazarların birbirleri ile bağlantı gücü bulunmadığı için ağ görselleştirmesinde isimleri yer almamaktadır.



Şekil 9: 14 yazarın alıntı bağlantılarının ağ görselleştirmesi (VOSviewer, 2025).

VOSviewer programından alınan verilere göre diğer kümeler şu şekilde listelenmektedir: Küme 2; 11 üye ile (Botta, Monica, Capolongo, Stefano ve Vittori, Gail dâhil) güçlü bir ortak çalışma alanını işaret etmektedir. Her biri 10 yazar içeren Küme 3 (Annemans, Margo, Heylighen, Ann ve Wekenborg, Magdalena dâhil) ve Küme 4 (Elf, Marie, Lindahl, Goran ve Wijk, Helle dâhil) de önemli büyüklüktedir.

Bir sonraki kademede, Küme 5; 9 yazara (Dhanakoses, Kawin ve Viryasiri, Traiwat dâhil) ve Küme 6; 8 yazara (Choi, Young-Seon, Ulrich, Roger S. Ve Zimring, Craig dâhil) sahiptir. Küme 7; 7 üyeyle (Demiris, Nikos, Economou, Marina ve Theleritis, Christos dâhil) orta yoğunluklu bir yapı sunmaktadır. Küme 8'den Küme 13'e kadar olan grupların tamamı beşer yazardan oluşmaktadır. Bu 5 kişilik gruplarda yer alan isimler Arasında Bosman, Caryl , Chi, Ziyuan , Grey, Tom , Shetty, Sham Ranjan , Varni, James W. Ve Zheng, Haohua gibi yazarlar bulunmaktadır. Daha küçük ve odaklanmış işbirlikleri, dörder yazarlı (Küme 14-18, Örn: Balboa, Tomas Abad , Raynolds, Mary , Ford, Katrina , Van Galen, Lana , Hu, Dan dahil) ve üçer yazarlı (Küme 19-26, Örn: Long, Hao , Ahmad, Safeer , Fuller, Robert , Cervinka, Renate , Kim, Hway-Suh , Brown, Steven D. , Shepley, Mardelle Mccu dahil) kümelerde yoğunlaşmaktadır. Küme 27'den küme 39'a kadar olan 13 küme ikişer yazar içermektedir (Örn: Hussein, Mohammed Fa , Mustafa, Faris Ali , Dincyurek, Ozgur dahil) . Küme 40'tan küme 43'e kadar olan dört küme ise tek bir yazarın bireysel katkısını (Banerji, Haimanti , Eckenwiler, Lisa , Mccann, Sarah , Nagasawa, Yasushi) temsil etmektedir.

4.3. Kuruluşların Alıntı Analizi

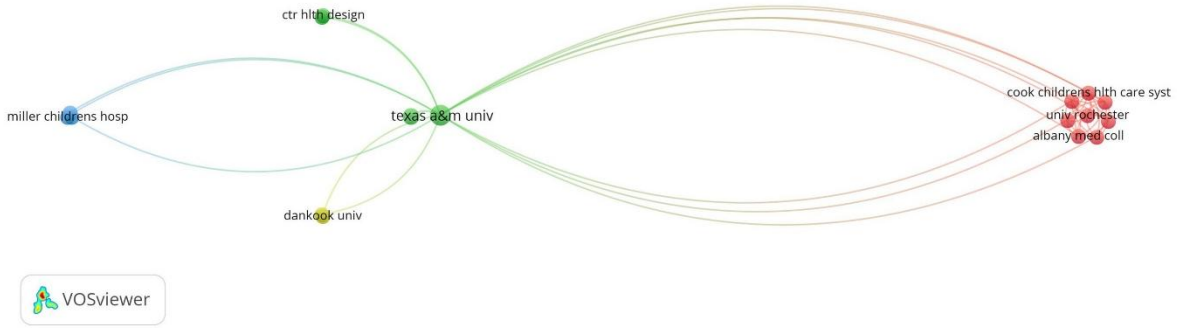
Toplam 97 kuruluştan 97'si “bir kuruluş için minimum belge sayısı = 1” eşliğini karşılamıştır. Bu 97 kuruluş için diğer kuruluşlarla olan alıntı bağlantılarının gücü hesaplanmıştır. VOSviewer, ağdaki 97 kuruluştan bazılarının hiçbir kuruma bağlı olmadığını göstermiştir ve çizelge 9'da, birbirlerine atıf yapan ve bağlantı gücüne sahip 7 kuruluş listelenmektedir. Kuruluşlar incelendiğinde, en çok öne çıkan kurumun Texas A&M Üniversitesi olduğu görülmektedir; hem 29 atıf hem de 179 bağlantı gücüyle hem üretkenlikte hem de işbirliği ağında lider konumdadır. Albany Medical College ve University of Rochester ise görece az sayıda atıf almış olsalar da (sırasıyla 4 ve 3), her ikisinin de bağlantı gücü 179 olduğu için ağın merkezinde yer almaktadır. Bu durum, az yayın yapmalarına rağmen güçlü işbirlikleri sayesinde etkili bir konum kazandıklarını göstermektedir. Cook Children's Health Care System 4 atıf ve 92 bağlantı gücüyle orta düzeyde bir etkiye sahipken, Dankook Üniversitesi yalnızca 1 atıfa sahip olmasına rağmen 100 bağlantı gücüyle dikkat çekmektedir; tek bir çalışmayla bile geniş bir işbirliği ağına ulaşmış olduğu görülmektedir. Öte yandan Miller Children's Hospital ve Center for Health Design düşük atıf ve sınırlı bağlantı gücüyle (sırasıyla 29 ve 13) ağda daha kenarda kalan kurumlar arasında yer almaktadır. Genel olarak tabloya bakıldığında, bazı kurumların yayın sayısından çok işbirliği gücüyle öne çıktığı, bazılarının ise hem üretim hem de bağlantılar açısından liderlik gösterdiği söylenebilir.

Şekil 10, kuruluşların ağ görselleştirme analizini göstermektedir. Toplam 97 kuruluştan 17'si birbirine bağlıdır. Bu görselleştirmede, bir kuruluşun etiketinin ve etrafının ne kadar büyük olması, o kuruluşun ağırlığının yüksek olduğunu göstermektedir. Ağ görselleştirmesinde (Şekil 10) en yüksek ağırlığa (bağlantı gücüne) sahip kurumlar başta “Texas A&M University”, “University of Rochester “, ve “Albany

Medical College “olmak üzere, Dankook University, Cook Children’s Health Care System, Miller Children’s Hospital, Center for Health Design (Ctr Hlth Des)’dir.

Çizelge 9: Birbirlerine atıf yapan ve bağlantı gücüne sahip 7 kuruluşun listesi (VOSviewer, 2025).

Kurum Adı	Atıf Sayısı	Bağlantı Gücü
1. Miller Children’s Hospital	1	29
2. Center for Health Design (Ctr Hlth Des)	1	13
3. Texas A&M University	29	179
4. Dankook University	1	100
5. Cook Children’s Health Care System	4	92
6. University of Rochester	3	179
7. Albany Medical College	4	179



Şekil 10: Kuruluşların bağlantı gücü analizine ilişkin ağ görselleştirmesi (VOSviewer, 2025).

Vosviewer programından elde edilen kuruluşlar ve küme kombinasyonları verilere göre, 97 kurum 38 kümede listelenmektedir. 97 kurumun bulunduğu kümelerin hepsi ağ görselleştirmesinde gözükmemektedir. Sadece birbirleri ile bağlantı gücü olan 7 kuruluşun ismi ağ görselleştirmesinde yer almaktadır. VOSviewer programından alınan verilere göre diğer kümeler şu şekilde listelenmektedir: En büyük kümelene olan Küme 1, 8 kurumla (Albany Med Coll, Cook Childrens Hith Care, Healthcare Gbbn Archite, Johns Hopkins Univ, Univ Rochester, Wake Forest Univ, Washington Hosp Ctr, Whr Architects Inc) en kalabalık grubu oluşturmaktadır. Bunu, her biri 7 kurum içeren iki grup izlemektedir: Küme 2 (Chalmers Univ Technol, Dalarna Univ, Gothenburg Univ, Karolinska İnst, Sahlgrens Univ Hosp, Univ Surrey, Zurich Univ Appl Sci) ve Küme 3 (Chengdu Univ Technol, Ctr Maximum Potential, Hub4mind, İst Super Sanita, Politecn Milan, Studio Monica Botta, Univ Engn & Technol Lah). Dördüncü en büyük grup olan Küme 4 ise 6 kurum (Cornell Univ, Ctr Hlth Design, Dankook Univ, Georgia İnst Technol, Missouri Univ Sci & Tech, Texas A&M Univ) içerirken; Küme 5 (Agr Univ Athens, Athens Univ Econ & Bus, Eugenides Fdn, Umhri, Univ Athens) ve Küme 6 (Cardiff Univ, Guangzhou Acad Fine Art, Guangzhou Xinhua Univ, Loughborough Univ, South China Univ Techn) beşer kurumdan oluşmaktadır. Küme 7’den Küme 14’e kadar olan grupların her biri 3 kurum içermekteyken, Küme 15’ten

Küme 24’e kadar olan gruplar ikişer kurumdan oluşur. Listede yer alan son 14 küme (Küme 25’ten Küme 38’e kadar) ise yalnızca tek bir kurum içermektedir (Örn: Curtin Univ Technol, Med Univ Vienna, Sharif Univ Technol, Islamic Azad Univ vb.).

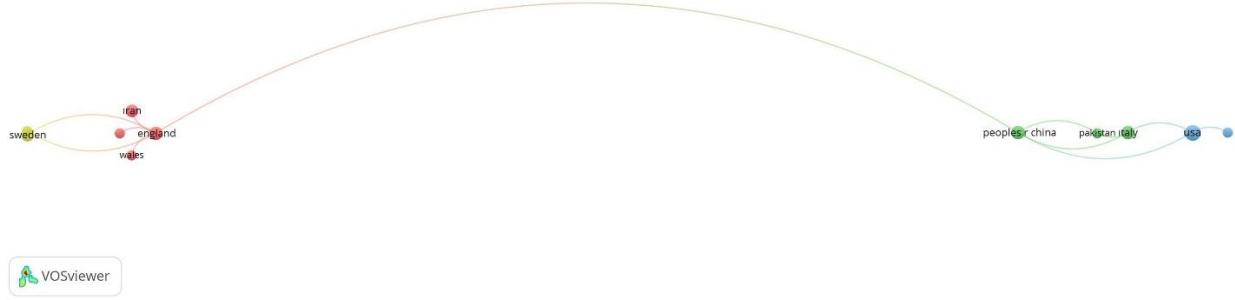
4.4.Ülkelerin Alıntı Analizi

Toplam 29 ülkeden 29’u “bir kuruluş için minimum belge sayısı = 1” eşiğini karşılamıştır. Bu 29 ülke için diğer ülkelerle olan alıntı bağlantılarının gücü hesaplanmıştır. VOSviewer, ağdaki 29 ülkeden bazılarının hiçbir ülkeye bağlı olmadığını göstermiştir ve çizelge 10’da, birbirlerine atıf yapan ve bağlantı gücüne sahip 11 ülke listelenmiştir. Çizelgedeki veriler incelendiğinde, ülkeler arasındaki akademik etki ve işbirliği ağının farklılaştığı görülmektedir. En dikkat çekici ülke ABD; 1326 gibi yüksek bir atıf sayısına sahip olmasına rağmen toplam bağlantı gücünün yalnızca 3 olması, yayınlarının çok etkili olduğunu fakat işbirliği ağında merkezî bir rol üstlenmediğini göstermektedir. Buna karşılık İngiltere 47 atıfla orta düzeyde görünse de 6 bağlantı gücüyle en fazla işbirliği yapan, ağın “bağlantı düğümü” olan ülke konumundadır. İtalya (112), İsveç (92), Avustralya (67), Mısır (64) ve Suudi Arabistan (50) görece yüksek atıf alan ülkeler olarak öne çıkarken, İsveç’in yalnızca 2 bağlantı gücüne sahip olması, çok sayıda atıf alan çalışmalarının sınırlı bir işbirliği ağı içinde üretildiğini göstermektedir. Çin (18) ve Galler (18) düşük atıf sayılarıyla daha sınırlı bir etki alanına sahiptir. Pakistan ve İran ise hiç atıf alamayarak ağın en dış halkasında yer almaktadır. Toplam 29 ülkeden 11’i birbirine bağlantı gücü ile bağlıdır. Çizelge 10 ‘da birbirlerine atıf yapan ve bağlantı gücüne sahip 11 ülke listelenmiştir. Şekil 11 ‘deki görselleştirmede, bir ülkenin etiketinin ve etrafının ne kadar büyük olması, o ülkenin ağırlığının yüksek olduğunu göstermektedir. En yüksek ağırlığa (bağlantı gücüne) sahip ülkelerin başında ilk sırada England (İngiltere), People’s Republic of China (Çin Halk Cumhuriyeti) ve Egypt (Mısır) yer almaktadır. Bunları USA (Amerika Birleşik Devletleri), Saudi Arabia (Suudi Arabistan) ve Australia (Avustralya) izlemektedir.

Çizelge 10: Birbirlerine atıf yapan ve bağlantı gücüne sahip 11 ülkenin listesi.

Ülke	Atıf Sayısı	Bağlantı Gücü
1. Sweden	92	2
2. Iran	0	0
3. England	47	6
4. Wales	18	1
5. People’s China	18	4
6. Pakistan	0	2
7. Italy	112	2
8. USA	1326	3
9. Egypt	64	4
10. Saudi Arabia	50	3
11. Australia	67	3

Vosviewer programından elde edilen ülkeler ve küme kombinasyonları verilerine göre, 29 ülke 17 kümede listelenmektedir. 29 ülkenin her birinin yer aldığı küme Şekil 11'deki ağ görselleştirmesinde gözükmemektedir.



Şekil 11: Ülkelerin bağlantı gücü analizine ilişkin ağ görselleştirmesi (VOSviewer, 2025).

Sadece birbirleriyle bağlantılı olan 11 ülkenin kümesi gözükmemektedir. VOSviewer programından elde edilen 17 küme şu şekilde listelenmektedir: Küme 1, 5 ülkeyle (İtalya, Pakistan, Çin Halk Cumhuriyeti, Güney Kore, ABD) en çok ülkeyi barındıran gruptur. Bunu, 4 ülkeyle (Danimarka, İngiltere, İran, Galler) küme 2 takip ederken; küme 3'te 3 ülke (Avustralya, Mısır, Suudi Arabistan) bulunmaktadır. Küme 4, küme 5 ve küme 6'da ise ikişer ülke yer alır: Küme 4'te Belçika ve Almanya; küme 5'te İsveç ve İsviçre; küme 6'da Singapur ve Tayland. Geriye kalan küme 7'den küme 17'ye kadar olan 11 kümenin her biri yalnızca tek bir ülke içermektedir. Bu tekli kümeler Avusturya (küme 7), Yunanistan (küme 8), Japonya (küme 9), Malezya (küme 10), Yeni Zelanda (küme 11), Polonya (küme 12), İspanya (küme 13), Türkiye (küme 14), Hindistan (küme 15), Irak (küme 16) ve İrlanda (küme 17)'dan oluşmaktadır.

4.5. Analiz Sonuçları

Web of Science'ta seçilen anahtar kelimelerle yapılan sorgulama sonucunda, ilk adımda sadece mimarlık ve sağlık alanlarında yapılmış olan çalışmalar sonucunda listelenen 50 belge, bibliyometrik verilere dayalı bilim haritalama teknikleri kullanılarak analiz edilmiştir. İkinci adımda çalışma daha da sınırlandırılmış ve sadece mimarlık alanındaki 2010 yılından itibaren yayımlanmış makaleler, bildiriler ve derleme makalelerinden oluşan 8 belge, bibliyometrik verilere dayalı bilim haritalama teknikleri kullanılarak analiz edilmiştir. Bu çalışmanın ilk adımında 220 anahtar kelimenin analizi sonucunda özellikle “evidence-based design”, “healthcare facilities”, “hospital”, “healthcare”, “healing / healing environments”, “healthcare architecture” gibi anahtar kelimelerin ön planda olduğu görülmüştür. Bu durum da sağlık yapılarında verilen tasarım kararlarının, mekânın estetik veya işlevsel niteliklerinin ötesine geçerek; iyileşme, hasta güvenliği ve çevresel memnuniyet gibi temel sağlık göstergeleriyle yakından ilişkili olduğunu göstermektedir. Çalışmanın ikinci adımında yapılan analizler ise 34 anahtar kelime, 173 yazar, 97 kuruluş, 29 ülkeyi kapsamaktadır. Analizlerden elde edilen bulgular aşağıdaki gibidir: En sık karşılaşılan anahtar kelimeler empathy (empati), hospital design (hastane tasarımı), attention restoration theory (dikkat

yenileme teorisi), salutogenic (salutojenik), healing environment (iyileşme ortamı)’dır. 2010 'da sıkça karşılaştığımız mimari anahtar kelimesi ise tasarım ve sağlık ortamı terimleri ile yakından ilişkilidir. 2010-2020 arasında sıkça karşılaşılan hastane tasarımı ve psikolojik rahatlık kelimeleri zamanla kişi-çevre uyumu, mimari planlama, vaka çalışması, sistematik inceleme, özellikle küresel perspektifler kelimeleri ile derinleşmiş ve farklı nitelikler kazanmıştır. 2010-2015 arasında ön plana çıkan anahtar kelime hastane ortamıdır. 2011, 2012,2013, 2014 ve 2015 yıllarında ilgili konu üzerine herhangi bir çalışma yapılmamıştır. 2015 ve 2020 arası öne çıkan anahtar kelimesi “iyileştirici mimari”dir. 2020 ile 2023 arası önem kazanan anahtar kelimeleri “biyofilik” ve iyileştirici sağlık ortamı”; bunu “salutojenik”, “yeşil bina” ve “hastane” takip etmektedir. 2023 ile 2025 arasında kullanılmaya başlanan ve giderek önem kazanarak kullanım sıklığı artan kelimeler ise “çocuk dostu tasarım”, “pediatrik sağlık hizmeti ortamı”, “eğlenceli ortam”, “sosyo-mekansal özellikler”, “mekânsal tercihler”, “sağlık hizmeti mimarisi”, “hastane tasarım ilkeleri” ve “sürdürülebilirlik”tir.

5. SONUÇ

Bu çalışma, mimarlık, sağlık çevreleri ve iyileştirici tasarım arasındaki ilişkiyi sistematik inceleme ve bibliyometrik analiz yöntemleriyle ele alarak, literatürde öne çıkan kavramları, yazarları, kurumları ve eğilimleri ortaya koymuştur. Çalışmanın amacı, sağlık yapılarında “iyileştirici tasarım” anlayışının gelişimini görünür kılmak ve alanın kavramsal çerçevesini bütüncül bir bakış açısıyla değerlendirmektir. Bu yönüyle makale, hem mimarlık hem de sağlık disiplinleri için insan odaklı tasarım yaklaşımlarının önemini vurgulamaktadır. Bulgular, 2010 yılından itibaren artan bir ivmeyle gelişen literatürde, biyofilik tasarım, salutojenik yaklaşım, empatik mimari ve çocuk dostu mekânların giderek öne çıktığını göstermektedir. Anahtar kelime analizleri, sağlık yapılarının yalnızca işlevsel ve teknik yönleriyle değil, aynı zamanda psikolojik konfor, sürdürülebilirlik, doğayla bütünleşme ve kullanıcı deneyimi gibi çok boyutlu parametrelerle ele alındığını ortaya koymuştur. Ayrıca, ABD, İngiltere ve İtalya gibi ülkelerin bu alandaki üretimlerde öncü olduğu, fakat gelişmekte olan ülkelere konunun henüz yeterince derinlemesine incelenmediği görülmüştür. Çalışmanın bir diğer önemli sonucu, sağlık yapılarında mekânın yalnızca tedavi sürecine eşlik eden bir unsur değil, aynı zamanda bireylerin ruhsal ve sosyal iyiliğini destekleyen aktif bir etken olduğunun anlaşılmasıdır. Bu bulgu, mimarlıkta insan merkezli yaklaşımın giderek güçlendiğini ve sağlık yapılarının tasarımında estetik, duyuşsal ve çevresel faktörlerin temel birer bileşen hâline geldiğini göstermektedir. Gelecek araştırmalar için birkaç öneri sunulabilir. Öncelikle, gelişmekte olan ülkelere sağlık yapıları bağlamında iyileştirici tasarım ilkelerinin nasıl uygulandığına dair ampirik çalışmalar yapılmalıdır. Ayrıca kültürel farklılıkların, bireylerin mekân algısı ve iyileşme deneyimi üzerindeki etkisini araştırmak, literatüre yeni bir boyut kazandıracaktır. Sürdürülebilirlik ve çocuk dostu tasarım gibi son dönemde öne çıkan yaklaşımların, farklı ölçeklerde (örneğin kliniklerden büyük hastanelere kadar) nasıl hayata geçirildiğinin incelenmesi de önemlidir. Son olarak, mimarlık, psikoloji,

sağlık bilimleri ve sosyoloji gibi farklı disiplinlerin ortak katkısıyla yürütülecek çalışmaların, alandaki bilgi birikimini daha da zenginleştireceği düşünülmektedir. Genel olarak, bu makale mimarlık-sağlık ilişkisi çerçevesinde iyileştirici tasarımın önemini vurgulamakta, mevcut literatürü bütüncül bir bakışla değerlendirmekte ve geleceğe yönelik araştırmalara yol gösterici bir temel sunmaktadır.

ÇIKAR ÇATIŞMASI BEYANI

Yazarlar arasında çıkar çatışması bulunmamaktadır.

YAZARLARIN KATKILARI

G.M: Kavramsallaştırma, yöntem, araştırma, kaynaklar, yazı yazma-gözden geçirme ve düzenleme.

A.A: Kavramsallaştırma, yöntem, araştırma, kaynaklar, yazı yazma- gözden geçirme ve düzenleme.

DESTEKLEYEN KURUM

Bu çalışma, Kırklareli Üniversitesi BAP Koordinatörlüğü bünyesinde gerçekleştirilen KLUBAP-303 Nolu ve “İyileştirici Tasarım Yaklaşımlarının Hastane Kullanıcıları Üzerindeki Etkilerinin Araştırılması: Kırklareli Eğitim ve Araştırma Hastanesi Giriş Mekânı Örneği” başlıklı proje kapsamında üretilmiştir.

KAYNAKLAR

- Abdelaal, M. S., & Soebarto, V. (2019). *Biophilia and salutogenesis as restorative design approaches in healthcare architecture*. Architectural Science Review, 62(3), 195–205. <https://doi.org/10.1080/00038628.2019.1604313>
- Aria, M., & Cuccurullo, C. (2017). *Bibliometrix: An R-tool for comprehensive science mapping analysis*. Journal of Informetrics, 11(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Awtuch, A., & Gebczynska-Janowicz, A. (2017). *Art and health – The healing potential of artistic interventions in medical environments*. IOP Conference Series: Materials Science and Engineering, 245(4), 042037. <https://doi.org/10.1088/1757-899X/245/4/042037>
- Bahrami, B., Kariminejad, S., & Erfani, G. (2025). *Child-friendly design in paediatric healthcare environments: Exploring children's preferences for nature, space, and play*. Health Environments Research & Design Journal, 18(1), 55–73. <https://doi.org/10.1080/13467581.2025.2524532>
- Banerji, H. (2016). *An attempt to explore components of empathic architecture in hospitals: A study of Indian hospitals*. International Journal of Architectural Research, 10(2), 44–63. <https://doi.org/10.3846/20297955.2016.1150220>
- Börner, K., Chen, C., & Boyack, K. W. (2003). *Visualizing knowledge domains*. Annual Review of Information Science and Technology, 37(1), 179–255. <https://doi.org/10.1002/aris.1440370106>
- Cavalcante, L. R., Lucena, R. F., & Almeida, A. C. (2021). *A bibliometric review of healing environments in healthcare architecture*. Journal of Environmental Psychology, 75, 102553. <https://doi.org/10.3390/buildings15173196>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). *Science mapping software tools*:

- Review, analysis, and cooperative study among tools.* Journal of the American Society for Information Science and Technology, 62(7), 1382–1402. <https://doi.org/10.1002/asi.21525>
- Dilani, A. (2001). *Psychosocially supportive design: Scandinavian healthcare design.* World Hospitals and Health Services, 37(4), 11–15.
- Harris, D. D., McBride, G., Ross, C., & Curtis, L. (2002). *A place to heal: Environmental sources of satisfaction among hospital patients.* Journal of Applied Social Psychology, 32(6), 1276–1299. <https://doi.org/10.1111/j.1559-1816.2002.tb01436.x>
- Lawson, B., & Phiri, M. (2003). *The architectural healthcare environment and its effects on patient health outcomes.* London: The Stationery Office.
- Nagasawa, Y. (2019). Global hospitals in 2050: *A historical review of hospital construction from a global perspective.* Japan Architectural Review, 3(2), 123–135. <https://doi.org/10.1002/2475-8876.12131>
- Ruhanen, L., Weiler, B., Moyle, B. D., & McLennan, C. J. (2015). *Trends and patterns in sustainable tourism research: A 25-year bibliometric analysis.* Journal of Sustainable Tourism, 23(4), 517–535. <https://doi.org/10.1080/09669582.2014.978790>
- Sánchez-Jáuregui, T., Gutiérrez-Pérez, N., Abad Balboa, T., & Chias, P. (2024). *The relationship between healthcare architecture and nature: Interaction with the landscape as therapy.* Landscape and Urban Planning, 239, 104879. <https://doi.org/10.20365/disegnarecon.33.2024.19>
- Salehi, F., & Vaez Eslami, S. (2010). *Green strategies of healthcare design: Case studies of medical centers* (Konferans bildirisi). Sustainable Architecture and Urban Development (s. 431–446). Erişim: https://www.irbnet.de/daten/iconda/CIB_DC22689.pdf
- Sarı, M., & Genç, G. (2022). *Sağlık hizmetlerinde performans kavramının bibliyometrik analizi.* Karadeniz Sosyal Bilimler Dergisi, 14(26), 238–254. <https://doi.org/10.38155/ksbd.1059592>
- Ulrich, R. S. (1991). *Effects of interior design on wellness: Theory and recent scientific research.* Journal of Health Care Interior Design, 3(1), 97–109.
- Ulrich, R. S., Zimring, C., Zhu, X., DuBose, J., Seo, H. B., Choi, Y. S., Quan, X., & Joseph, A. (2008). *A review of the research literature on evidence-based healthcare design.* HERD: Health Environments Research & Design Journal, 1(3), 61–125. <https://doi.org/10.1177/193758670800100306>
- Van Eck, N. J., & Waltman, L. (2010). *Software survey: VOSviewer, a computer program for bibliometric mapping.* Scientometrics, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Viryasiri, T., Laovisutthichai, V., Sangnin, K., Dhanakoses, K., Roopkaew, P., & Viryasiri, P. (2023). *Hospital design principles implementation: Reflections from practitioners in Thailand.* Journal of Asian Architecture and Building Engineering, 23(1), 43–56. <https://doi.org/10.1080/13467581.2023.2215842>
- World Health Organization. (1948). *Constitution of the World Health Organization.* Geneva: World Health Organization. <https://www.who.int/about/governance/constitution>
- Yan, S., Azmi, A., Mansor, N., Wang, Z., & Wang, Y. (2024). *Healing spaces as a design approach to optimize emotional regulation for patients with mood disorders.* Buildings, 14(2), 472. <https://doi.org/10.3390/buildings14020472>
- Yılmaz, N., Taş, F., & Uğurluoğlu, D. (2019). *Yöntemler üzerine bibliyometrik bir analiz: Hacettepe Sağlık İdaresi Dergisi Örneği.* Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, (35), 83–106.
- Zupic, I., & Čater, T. (2015). *Bibliometric methods in management and organization.* Organizational Research Methods, 18(3), 429–472. <https://doi.org/10.1177/1094428114562629>

A SURF-Based Rigid Feature Tracking (RiFT) Approach for Cost-Effective Analysis of Fracture Behaviour in Asphalt Pavements

Yavuz ABUT*

Department of Transportation Engineering, Faculty of Engineering, Yalova University, Yalova, Türkiye

Received: 25.11.2025, Accepted: 19.12.2025, Published: 29.12.2025

ABSTRACT

Cracking in asphalt pavements significantly affects durability and service life, making accurate characterization of fracture essential. The Semi-Circular Bending (SCB) test is widely used to evaluate fracture resistance; however, conventional optical techniques such as Digital Image Correlation (DIC) require high-resolution cameras, controlled lighting, and speckle, leading to high costs and complex setups. This study introduces Rigid Feature Tracking (RiFT), a SURF-inspired (Speeded-Up Robust Features), feature-based displacement estimation framework designed for quasi-rigid materials under controlled loading. RiFT is specifically introduced to provide a robust, low-cost optical method for monitoring crack initiation and propagation during SCB test, overcoming the hardware and sample preparation limitations of DIC in this critical application. RiFT departs from correlation-heavy approaches by detecting sparse but robust landmark points and reconstructing dense displacement fields through edge-preserving interpolation, aim to reduce sensitivity to angular and illumination variations. SCB tests were performed on asphalt specimens with an INSTRON 5982, recording load and CMOD at 10 Hz. Images were captured every five seconds using a webcam without additional lighting or surface speckle. This study presents RiFT primarily as a methodological framework validated through controlled SCB experiments, demonstrating its proof-of-concept capability for real-time tracking, with further wide-scale validation planned for future work.

Keywords: Asphalt crack analysis; Digital image correlation; Semi-circular bending test; RiFT; SURF.

Asfalt Kaplamalarda Kırılma Davranışının Maliyet Etkin Analizi için SURF Tabanlı Rijit Özellik Takibi (RiFT) Yaklaşımı

ÖZ

Asfalt kaplamalarda çatlak oluşumu, dayanıklılığı ve hizmet ömrünü ciddi şekilde etkilemektedir. Bu nedenle kırılmanın doğru biçimde analiz edilmesi kritik öneme sahiptir. Yarı Dairesel Eğilme (SCB) testi, kırılma direncini değerlendirmek için yaygın olarak kullanılmaktadır. Ancak, Dijital Görüntü Korelasyonu (DIC) gibi optik yöntemler; yüksek çözünürlüklü kameralar, kontrollü aydınlatma ve benekler gerektirdiğinden maliyet ve karmaşık kurulum sorunları doğurmaktadır. Bu çalışmada, yarı-rijit malzemeler için tasarlanmış, SURF (Speeded-Up Robust Features) algoritmasından esinlenen ve özellik tabanlı yer değiştirme tahmin yöntemi olan Rijit Özellik Takibi (RiFT) sunulmaktadır. RiFT, bu kritik uygulamada DIC'nin donanım ve numune hazırlama kısıtlarını bertaraf ederek, SCB testi sırasında çatlak başlangıcını ve ilerlemesini izlemek için sağlam, düşük maliyetli bir optik yöntem sağlamak amacıyla özel olarak sunulmuştur. RiFT, korelasyon ağırlıklı yaklaşımlardan farklı olarak, seyrek güvenilir nirengi noktaları belirlemekte, kenar koruyucu enterpolasyonla yer değiştirme alanlarını yeniden oluşturarak,

farklı açı ve ışık değişimlerine duyarlılığı azaltmayı hedeflemektedir. SCB testleri, INSTRON 5982 cihazı kullanılarak asfalt numuneleri üzerinde gerçekleştirilmiş; yük ve CMOD 10 Hz'de kaydedilmiştir. Görüntüler, ek aydınlatma veya yüzey benekleri olmadan, her beş saniyede bir web kamerasıyla alınmıştır. Bu çalışma, RiFT'i öncelikle kontrollü SCB deneyleriyle doğrulanmış metodolojik bir çerçeve olarak sunmakta ve RiFT'in gerçek zamanlı izleme yapabildiğinin kavramsal ispatını ortaya koymaktadır. Daha kapsamlı doğrulama ise gelecek çalışmalar için planlanmaktadır.

Anahtar Kelimeler: Asfalt çatlak analizi; Dijital görüntü korelasyonu; Yarı dairesel eğilme testi; RiFT; SURF.

1. INTRODUCTION

Cracking in asphalt concrete is one of the most critical issues in pavement design and maintenance, as it directly affects the durability, service life, and overall performance of road infrastructure. Asphalt pavements are subjected to complex loading and environmental conditions, which lead to progressive damage over time. Among the various forms of distress, cracking is particularly detrimental because it accelerates moisture infiltration, weakens structural integrity, and increases maintenance costs ([Min et al., 2025](#)). Understanding the mechanisms of crack initiation and propagation is therefore essential for designing resilient pavements and optimizing mixture compositions ([Saha & Biligiri, 2016](#)).

Cracking mechanisms in asphalt concrete are generally classified into fracture cracking and fatigue cracking, each requiring distinct evaluation methods. Fatigue cracking typically occurs under repeated traffic loading, where micro-cracks gradually accumulate and coalesce into visible cracks ([Alsheyab et al., 2024](#)). In contrast, fracture cracking is associated with sudden failure under monotonic loading or extreme thermal conditions ([Marasteanu et al., 2012](#)), and reflective cracking occurs when underlying joints or cracks in the base layers propagate upward through the asphalt surface due to differential movement and stress concentration ([Son, 2014](#)). Over the years, researchers have developed various experimental setups to characterize these behaviours. According to [Zhou et al. \(2016\)](#), extensive research has focused on experimental designs for field validation of laboratory tests to evaluate the cracking resistance of asphalt mixtures, and standards for fatigue and fracture characterization have evolved over time. For fatigue assessment, commonly used tests include the four-point bending test ([AASHTO, 2017](#); [ASTM, 2021](#)), the Texas overlay test ([Texas, 2025](#)), and the indirect tensile fatigue test ([AASHTO, 2014](#)). For fracture characterization, widely adopted methods comprise SCB test—typically performed at low (-30°C to 0°C) and intermediate (5 – 25°C) temperatures ([AASHTO, 2020a, 2020b](#); [ASTM, 2016](#); [EN, 2010](#))—along with the indirect tensile test ([ASTM, 2017](#)) and the disk-shaped compact tension test ([ASTM, 2013](#)).

The SCB test has gained prominence for its simplicity, efficiency, and adaptability. Using cylindrical specimens, which are easier to prepare than slabs or beams, SCB enables multiple samples to be extracted from a single core—an advantage for laboratory studies and quality control where material availability and time are critical ([Lu et al., 2021](#)). Under Mode I loading, the SCB test generates dominant tensile stress at the specimen's bottom, effectively simulating pavement cracking under tensile conditions while enabling the determination of critical fracture parameters such as fracture toughness, stress intensity factor, stiffness, compliance, fracture energy, and the J-integral within the frameworks of LEFM (Linear Elastic Fracture Mechanics) and EPFM (Elastic Plastic Fracture Mechanics) ([Saha & Biligiri, 2015](#)).

While primarily used for monotonic loading ([Birgisson et al., 2008](#); [Li & Marasteanu, 2010](#); [Saha & Biligiri, 2015](#)), some studies have extended SCB to cyclic loading ([Hassan, 2013](#); [Huang et al., 2013](#)) for

fatigue analysis, complemented by numerical simulations that model crack propagation and stress distribution. Several standards govern SCB testing, including EN 12697-44:2010 ([EN, 2010](#)), AASHTO TP 105-13 ([AASHTO, 2020a](#)), ASTM D8044-16 ([ASTM, 2016](#)), and AASHTO TP 124-18 ([AASHTO, 2020b](#)). These differ in specimen dimensions, notch lengths, loading rates, and fracture indicators—such as fracture toughness, J-integral, or Flexibility Index—though most recommend a 150 mm diameter. Recent research suggests smaller diameters (e.g., 100 mm) may also yield reliable results, but further validation is needed to ensure repeatability and conversion between sizes ([Lu et al., 2021](#)).

While SCB is a robust and widely accepted fracture test, capturing detailed crack evolution and displacement fields during testing remains challenging. Conventional optical measurement techniques such as Digital Image Correlation (DIC) have been extensively used in experimental mechanics to quantify deformation and strain fields ([Jiang-san et al., 2022](#); [Zhao et al., 2017](#); [Zieliński et al., 2025](#)). These methods rely on correlation-based tracking of intensity patterns between successive frames, which works well for fluid-like or granular motions where displacements are relatively large and directional variability is significant ([Asswad et al., 2024](#); [Liang et al., 2020](#)). In such cases, deformation fields are characterized by small amplitudes, local coherence, and abrupt discontinuities associated with crack initiation and propagation. As a result, traditional DIC algorithms often suffer from unstable correlation peaks, excessive false matches, and sensitivity to minor variations in illumination or surface texture. These limitations are particularly pronounced when displacements fall below the pixel scale, making accurate tracking challenging without costly high-resolution imaging systems and elaborate lighting setups. Moreover, implementing DIC in SCB experiments typically requires high-speed cameras, specialized lighting, and extensive calibration procedures to ensure accuracy ([Birgisson et al., 2008](#); [Wang et al., 2020](#); [Zhao et al., 2017](#)). These requirements increase both the cost and complexity of the experimental setup, limiting the practicality of such methods for routine testing or field applications. Additionally, the large volume of image data generated during high-frequency acquisition imposes significant computational demands, making post-processing time-consuming and resource-intensive. These challenges highlight the need for alternative approaches that are cost-effective, robust to imaging conditions, and capable of capturing fine-scale displacement and crack evolution without relying on expensive hardware or complex calibration.

To address these challenges, this study introduces a feature-based displacement estimation framework, termed RiFT, which builds upon the principles of the SURF algorithm developed by [Bay et al. \(2006\)](#), adapting them to the kinematic characteristics of quasi-rigid solids. Unlike correlation-heavy approaches, RiFT focuses on detecting sparse yet highly reliable landmark points and subsequently reconstructing a dense displacement field through edge-preserving interpolation guided by neighbourhood consistency. This strategy minimizes the need for extensive outlier suppression and instead leverages the spatial

coherence inherent in solid deformation. By doing so, RiFT provides a robust and computationally efficient alternative that is less sensitive to angular variations and lighting inconsistencies, eliminating the dependency on expensive imaging hardware or complex calibration procedures. The proposed framework is validated through controlled SCB experiments, demonstrating its capability to capture sub-pixel displacements and crack evolution patterns that remain elusive to conventional DIC analyses.

Beyond its technical robustness, RiFT offers significant practical advantages. By relying on standard imaging devices such as webcams and avoiding the need for specialized lighting or high-speed acquisition systems, the method drastically reduces experimental costs and setup complexity. [Cabo et al. \(2019\)](#) proposed a hybrid SURF-DIC algorithm that uses image matching and cross-correlation to estimate local displacements with subpixel precision, offering a fast, low-cost, and practical solution suitable for industrial applications. Although the original study by [Cabo et al. \(2019\)](#) focused on estimating local displacements in general structural applications using a hybrid SURF-DIC algorithm, the same concept is highly relevant for asphalt pavement cracking analysis, where accurate displacement tracking around crack tips is critical for evaluating fracture behaviour under SCB and similar tests. Furthermore, its adaptability to varying image quality and minimal calibration requirements make it suitable for field applications where controlled laboratory conditions are not feasible. These attributes position RiFT as a scalable solution for structural health monitoring, pavement performance evaluation, and other engineering contexts where crack detection and displacement tracking are critical yet resource constraints limit the use of traditional optical measurement systems. The SCB test remains a cornerstone in asphalt fracture characterization, but its full potential can only be realized through accurate and efficient measurement of crack evolution. The proposed RiFT framework addresses key limitations of existing optical methods, offering a practical and reliable alternative that aligns with the growing demand for cost-effective and adaptable testing solutions in pavement engineering.

By making full-field deformation analysis accessible and affordable during the standard SCB procedure, RiFT enhances the test's potential as a cornerstone for reliable, routine fracture characterization.

2. DIC-BASED ANALYSIS OF ASPHALT MIXTURES WITH SCB TEST

To provide a clear overview of existing research, Table 1 synthesizes DIC-based studies relevant to asphalt mixtures. It consolidates related work and organizes it across key dimensions such as test and DIC variants, imaging and surface preparation, primary outputs and metrics, as well as limitations and research opportunities, offering a comprehensive reference for understanding methodological trends and gaps in this field.

Table 1: Summary of DIC-based studies on asphalt mixtures and SCB testing.

Reference	Theme	DIC techniques	Camera, Surface Preparation	Key Outputs
Birgisson et al. (2008)	Determination & prediction of crack patterns in HMA using DIC	Plane 2D DIC (least-squares matching of grey-value windows)	Basler AF101, 1300×1030 px, 5 fps; ~5×4 cm ROI; white base + black speckle	Full-field displacement/strain; crack initiation/propagation; agreement with gauges (RMS ≈ 0.034%)
Zhao et al. (2017)	Heterogeneous fracture simulation under SCB (cohesive crack model)	DIC used for displacement tracking; numerical FE focus	Industrial camera 1626×1236, 14 fps; focal 25 mm; speckle not reported	Load–displacement curve; final crack path for comparison with simulation
Zhang et al. (2018)	Crack growth rate of sulfur-extended asphalt (cyclic SCB + DIC)	2D DIC for crack length at selected intervals; CMOD–crack length polynomial	“DIC camera” (unspecified); speckled (white + black dots)	Crack length, tip strain, FPZ; Paris law coefficients via da/dN from CMOD–a correlation
Górszczyk et al. (2019)	Application of DIC for road materials	2D DIC (Istra 4D); NCC/LSM facets	Single camera; images at load steps; multi-LED lighting; white + black speckle	Displacement/strain maps; material moduli; crack visualization; rigid-body motion removal
Wang et al. (2020)	Polymer-modified asphalt before/after aging (SCB + DIC)	Stereo (3D) DIC (VIC-3D); ZNSSD + Newton–Raphson	Dual CCD 1624×1224 @10 Hz; matte black base + white speckle	Exx(t); indices DE and JE; three cracking stages identified
Liang et al. (2020)	“Use of Digital Image Processing” for fracture evaluation	Digital Image Processing (Otsu + median filtering) for crack pixels	HD camera (unspecified); no speckle; offline frame extraction	Crack path/length/area; fracture energy; aggregate effects
Radeef et al. (2021)	Cracking resistance under repeated loading using Digital Image analysis	DIC + ImageJ DIP (fractal dimension, density, tortuosity)	High-res video camera (no fps/res given); speckle applied	Full-field strain; localization zones; geometric crack metrics for modifiers
Pei et al. (2021)	Basalt fiber diameter effects (SCB + Ncorr)	2D DIC (Ncorr/Matlab); Cauchy strain; real-time L via DIC	Sony digital camera (unspecified); natural surface; no extra lighting noted	Three stages; fibers reduce strain; 7 μm fibers delay initiation (~4 s)
Jiang-san et al. (2022)	Warm-mixed rubber powder modified asphalt (fatigue SCB)	Stereo (3D) DIC (VIC-3D 8)	Dual CCD; blue light; 1 Hz; white base + black speckle	Local strain evolution at notch; three damage stages; strain-rate turning point → fatigue life
Cheng et al. (2022)	Long-term service pavement (SCB + stereo DIC)	Stereo (3D) DIC (GOM Correlate)	Two industrial cams (48×2048 px sensor), capable of 80 fps; blue light; acquired at 1 Hz; speckle	Horizontal strain maps; crack-tip strain evolution; strain-based fatigue damage parameter; 3-stage damage
Kong et al. (2023)	Fiber-reinforced emulsified asphalt (cold-recycled)	DIC (GOM Correlate 2019); likely stereo	Two 2 MP CCD (Sony ICX274), 5 fps; LED spots; black-on-white speckle	U,V; Exx,Eyy; dU/dt, dExx/dt; FPZ size (e.g., 19.3 mm vs 16.4 mm)
Ouyang et al. (2023)	Cement bitumen emulsion mixture fracture via DIC	3D (stereo) DIC; mean global Exx as failure indicator	Two cameras; “light system”; surface prep not specified	3D displacement/strain; mean Exx peak as failure; plastic zone area vs cement dosage (optimum ≈2.6%)
Asswad et al. (2024)	Two-Part DIC (2PDIC) for crack propagation	2PDIC splits subset across discontinuity; pseudo-strain thresholds ($\epsilon=0.005/0.05$)	3840×2748 px CCD, 1 fps; thermal chamber (−5 °C, 10 °C); black base + white droplets	COD; true crack length (L_{crack}); cumulative damage length (L_{damage}); tortuosity; energy reduction 22–63% vs ligament length
Ma et al. (2024)	Cyclic SCB protocol at intermediate temperature	Stereo DIC (VIC-3D); virtual extensometer for CMOD; quadratic CMOD–a mapping	Stereo cams 1920×1200, 1 Hz; lighting system; white + black speckle	exx; CMOD; crack length time-history; validation vs clip-on extensometer (≤0.1 mm error); Paris-law n & logA
Zieliński et al. (2025)	SCB + DIC for fracture properties (CivEng Vision)	2D DIC (ZNCC; sub-pixel interpolation)	DSLR Nikon D5300 (24 MP); intervalometer; strong LED (4000–5700 K); no speckle reported	Displacement maps; Exx,Eyy,Exy; crack path from initiation to failure; heterogeneity effects (REF vs RAS; aging)

To critically examine the data summarized in Table 1, several recurring technical and practical constraints in DIC applications for asphalt mixtures become evident. Early works such as [Birgisson et al. \(2008\)](#) and [Górszczyk et al. \(2019\)](#) demonstrated the ability of 2D DIC to provide full-field strain maps and visualize crack initiation, outperforming traditional point sensors. However, these setups relied heavily on painted speckle patterns and controlled lighting, which increases preparation time and cost. Similar requirements appear in [Jiang-san et al. \(2022\)](#), [Ma et al. \(2024\)](#), and [Zhang et al. \(2018\)](#) where white-and-black speckle patterns and LED or blue light sources were essential for correlation accuracy. This dependency limits field applicability and introduces variability when surface preparation is inconsistent. Another major limitation is temporal resolution. Many studies, including [Zhang et al. \(2018\)](#), [Jiang-san et al. \(2022\)](#), ([Cheng et al., 2022](#)), and ([Ma et al., 2024](#)), captured images at 1 Hz or selected intervals, making it difficult to track rapid crack propagation or micro-displacement events. Even when high-speed cameras were available ([Cheng et al., 2022](#); [Wang et al., 2020](#)), acquisition was often down-sampled to low frequencies for data management, sacrificing detail. This constraint is critical because SCB fracture involves sudden transitions that require finer temporal granularity. A third recurring issue is dimensionality and robustness near discontinuities. Most works ([Birgisson et al., 2008](#); [Pei et al., 2021](#); [Zieliński et al., 2025](#)) employed 2D DIC, which struggles with out-of-plane motions and decorrelation at crack tips. Stereo (3D) systems ([Cheng et al., 2022](#); [Jiang-san et al., 2022](#); [Kong et al., 2023](#); [Wang et al., 2020](#)) improve accuracy but add complexity, cost, and calibration overhead, making them impractical for routine or field testing. Furthermore, studies like [Liang et al. \(2020\)](#) and [Radeef et al. \(2021\)](#) combined DIC with image processing (DIP) for crack length estimation, introducing indirect methods and additional post-processing steps that increase computational burden and reduce reproducibility. Finally, reporting gaps and reproducibility concerns persist. Several papers ([Pei et al., 2021](#); [Zhang et al., 2018](#); [Zieliński et al., 2025](#)) omit key imaging parameters such as resolution, frame rate, and calibration details, hindering benchmarking and standardization. [Pei et al. \(2021\)](#) explored natural surface textures to avoid speckle preparation, but robustness under varying lighting and aggregate heterogeneity remains uncertain.

3. EXPERIMENTAL FRAMEWORK

To ensure a coherent experimental workflow, the preparation of specimens and their subsequent SCB testing form the foundation for reliable fracture performance evaluation. Building on these physical tests, an advanced imaging strategy is integrated to capture and analyse crack propagation and deformation patterns in detail. This transition from mechanical testing to feature-based tracking enables a comprehensive understanding of material behaviour under loading, bridging traditional laboratory methods with digital analysis.

3.1. Specimen preparation and SCB testing

Building on the detailed methodology reported by [Abut \(2024\)](#), this study ensured consistency in specimen preparation and SCB test reliability. Cylindrical specimens ($\Phi 150 \text{ mm} \times 50 \text{ mm}$) were fabricated following the Marshall procedure and [EN \(2022\)](#) standards to ensure uniformity in mixture design and compaction. Each specimen was compacted in Proctor moulds ($\Phi 150 \times 116.4 \text{ mm}$) using a vibrating hammer (BOSCH GBH 11 DE, 11 kg, 900 W, 2000 impacts/min) at 135°C . Compaction time was optimized through trial and error, and acceptance criteria required $\geq 98\%$ of the target unit weight and air voids near 4%. Specimens failing these conditions were excluded. After curing for 24 hours, discs were demoulded and cut into semi-circular halves using a precision disc saw. A notch was introduced at the mid-span to create the ligament necessary for fracture testing. In accordance with [AASHTO \(2020a\)](#), specimens were conditioned at -12°C for 2 ± 0.5 hours prior to testing. Monotonic bending tests were performed using an INSTRON 5982 universal testing machine (capacity: 100 kN; sensitivity: 0.16%). Crack Mouth Opening Displacement (CMOD) was monitored with a gauge (range: 0–4 mm; resolution: 0.003 mm), synchronized with load acquisition at 10 Hz. The system is capable of recording CMOD and load data at intervals of one-tenth of a second, ensuring high temporal resolution for correlating mechanical response with crack evolution. The loading rate was maintained at 0.5 mm/min to ensure quasi-static conditions. Fracture parameters—work of fracture, fracture energy, fracture toughness, and stiffness—were calculated according to [AASHTO \(2020a\)](#) to characterize crack initiation and propagation under tensile-dominated conditions.

3.2. Imaging Strategy for the Feature-Based Tracking

To capture crack evolution during SCB tests, a low-cost imaging approach was implemented using a webcam (Logitech HD Pro C910). Images were recorded every 5 seconds at $2304 \times 1536 \text{ px}$ ($\approx 3.5 \text{ MP}$, 96 dpi, bit depth: 8) without additional lighting or artificial speckle patterns (Figure 1). To enhance surface visibility, a thin layer ($\sim 1\text{--}2 \text{ mm}$) was trimmed from the specimen face, exposing aggregate texture for natural feature tracking. After trimming, fine dust was removed using compressed air to ensure a clean surface. This minimalist setup contrasts sharply with conventional DIC systems that require high-speed cameras, controlled illumination, and speckle preparation.

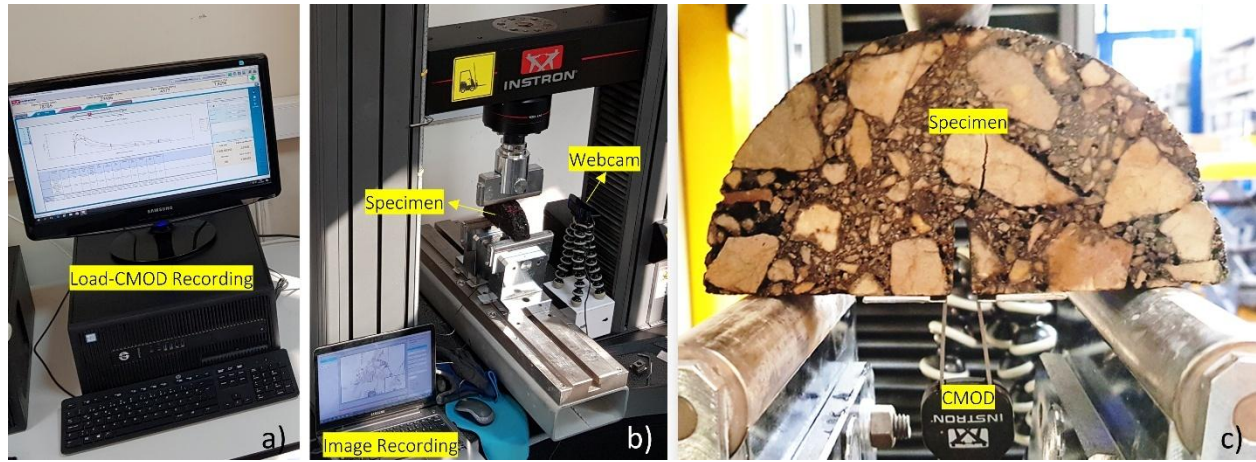


Figure 1: Experimental setup for SCB tests: (a) recording of load and CMOD during fracture test; (b) test frame, specimen, and low-cost imaging system using a webcam; (c) specimen and CMOD

4. EXTRACTION OF ROI AND FEATURE BASED DISPLACEMENT ANALYSIS

The analysis began with a targeted pre-processing stage to isolate the ROI surrounding the crack mouth and suppress background motion. Using manual annotation, a binary mask was generated on the reference image and then applied to the subsequent frame so that both images contained an identical, tightly bounded ROI. The masked images were converted to grayscale to normalize intensity and reduce computational load, yielding consistent inputs for feature detection (Figure 2).

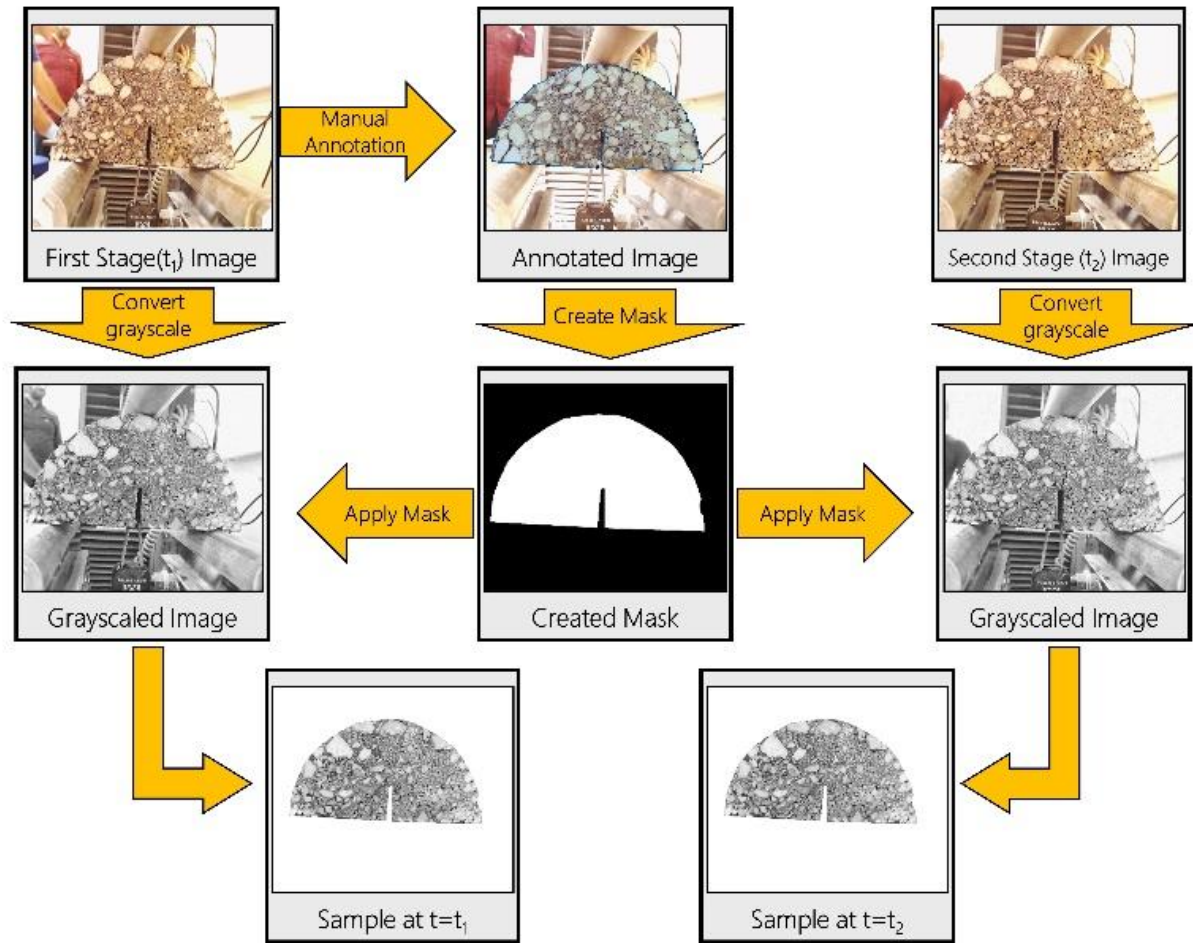


Figure 2: ROI definition workflow.

Within the ROI, landmark points were detected using SURF, which is well suited to the heterogeneous, high contrast textures of asphalt mixtures (aggregate boundaries, binder matrix, and fine surface relief). The detector produced thousands of landmark points in a typical frame (e.g., 5,786 in the example shown), each encoded by a 64-element descriptor that captures the local neighbourhood structure around the feature (Figures 3–4). These descriptors provide invariance to modest geometric and photometric changes—an essential property given our low-cost imaging (webcam) and the absence of controlled illumination or artificial speckle.

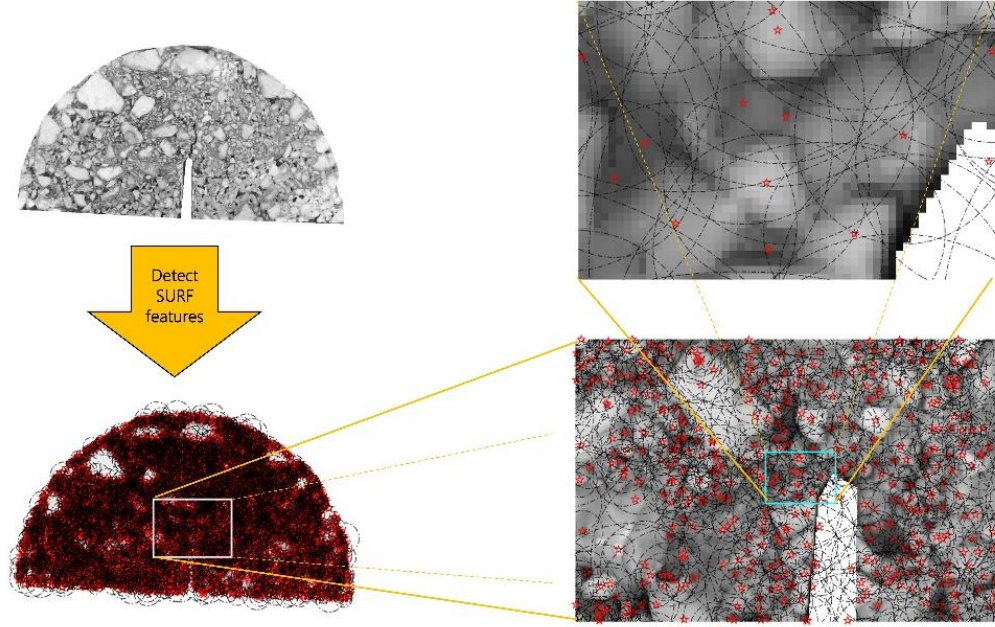


Figure 3: SURF landmark points (single frame).

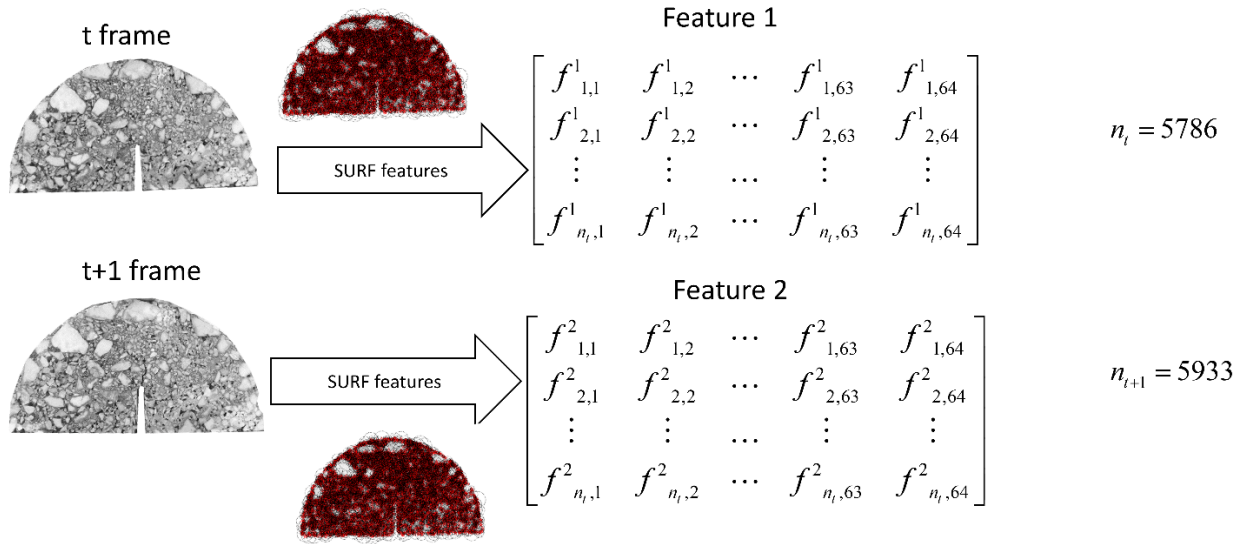


Figure 4: Descriptor inventory

To estimate motion between frames, descriptors were paired using Sum of Absolute Differences (SAD). Matches with $SAD < 0.6$ were retained while others were rejected, providing an initial, sparse displacement graph across the ROI (Figure 5). This thresholding step efficiently removed many ambiguous correspondences without the heavy iterative outlier suppression often required in correlation based DIC, aligning with the lightweight philosophy of our RiFT pipeline.

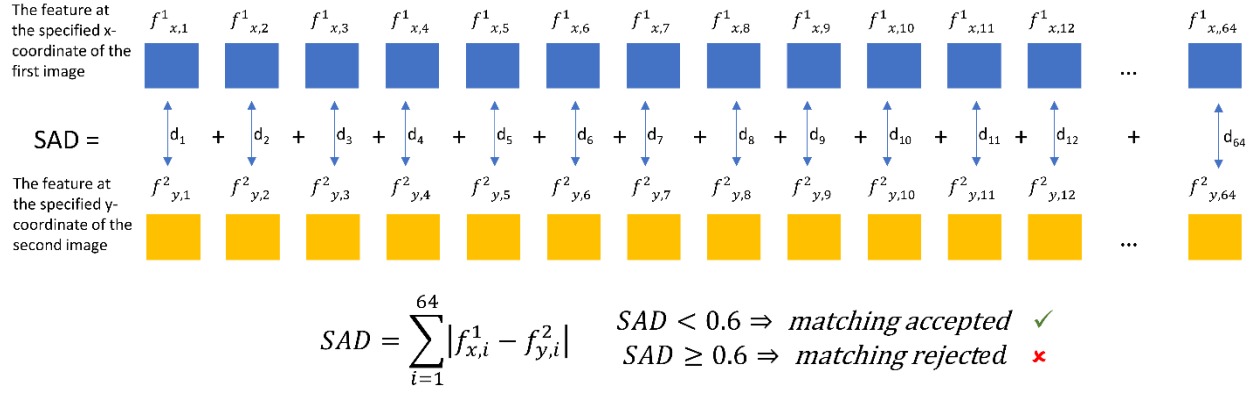


Figure 5: Inter frame matching ($t \rightarrow t + 1$).

Residual spurious vectors were then filtered statistically. The empirical distribution of displacement magnitudes was computed, and vectors beyond the 95th percentile were discarded; these typically correspond to mismatches on weak texture patches or minor background contamination.

The effect is visible in the histogram guided culling and the cleaned vector field: the remaining tracks cluster along physically plausible deformation paths (Figure 6). This percentile-based rule adapts automatically to each image pair and avoids hand tuning for different mixtures or lighting conditions.

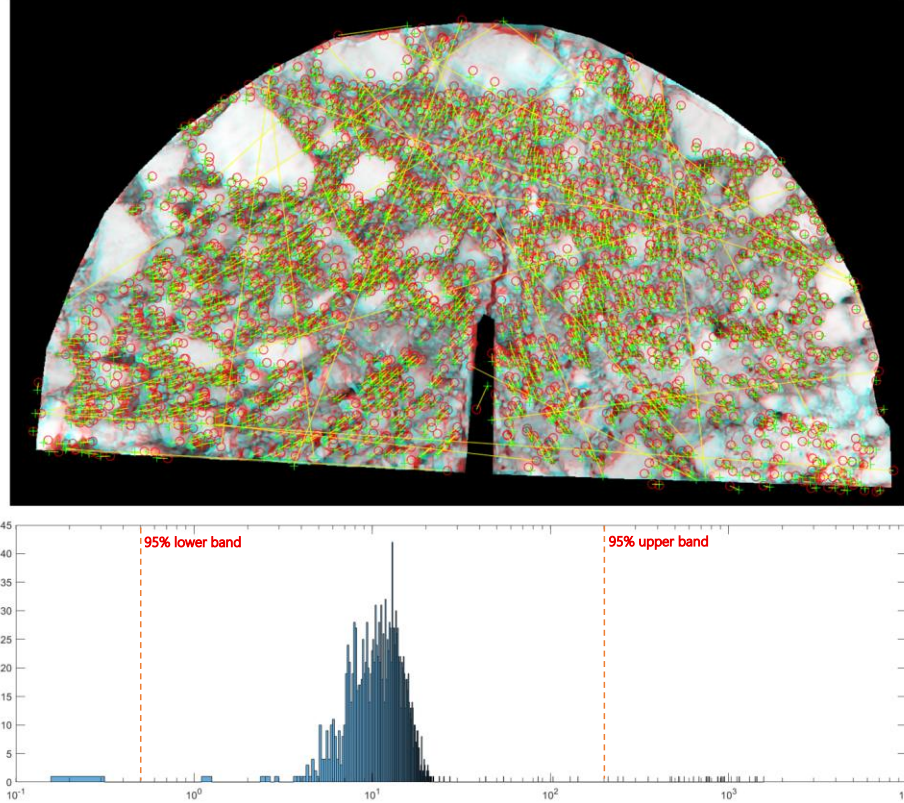


Figure 6: Outlier rejection.

The robust but sparse motion samples were interpolated to a dense displacement field using an edge aware scheme that respects crack driven discontinuities. The resulting component maps horizontal displacement (U_x), vertical displacement (U_y) and the magnitude field reveal a clear localization of deformation in the ligament region adjacent to the notch, with gradients intensifying as the crack mouth opens (Figure 7). All three contour maps utilize two-unit systems: arbitrary position units (U_{pos}) for the coordinates of landmark points and arbitrary displacement units (U_{disp}) (on a 0-10 scale) for the displacements of the landmark points. While all plots confirm that the highest deformation is clearly localized to the ligament region compared to the rest of the specimen, a critical observation is that the maximum values ($\approx 9-10 U_{disp}$) are not concentrated at a single point but are distributed across multiple interconnected centres. These distributed maxima in both the U_x and U_y maps suggest a non-uniform stress distribution within the material (likely due to its heterogeneous nature) and indicate that the Fracture Process Zone (FPZ) develops as an area rather than a single crack line. U_x model the natural crack formation (opening) in the horizontal direction due to crack propagation, while U_y model the compression effect resulting from the loading applied by the support areas at the top and bottom points of the specimen. Consequently, the magnitude map confirms that this dispersed area peaks in total displacement, strongly demonstrating that the deformation within the ligament is localized yet simultaneously heterogeneous, occurring around multiple regions of micro-cracking.

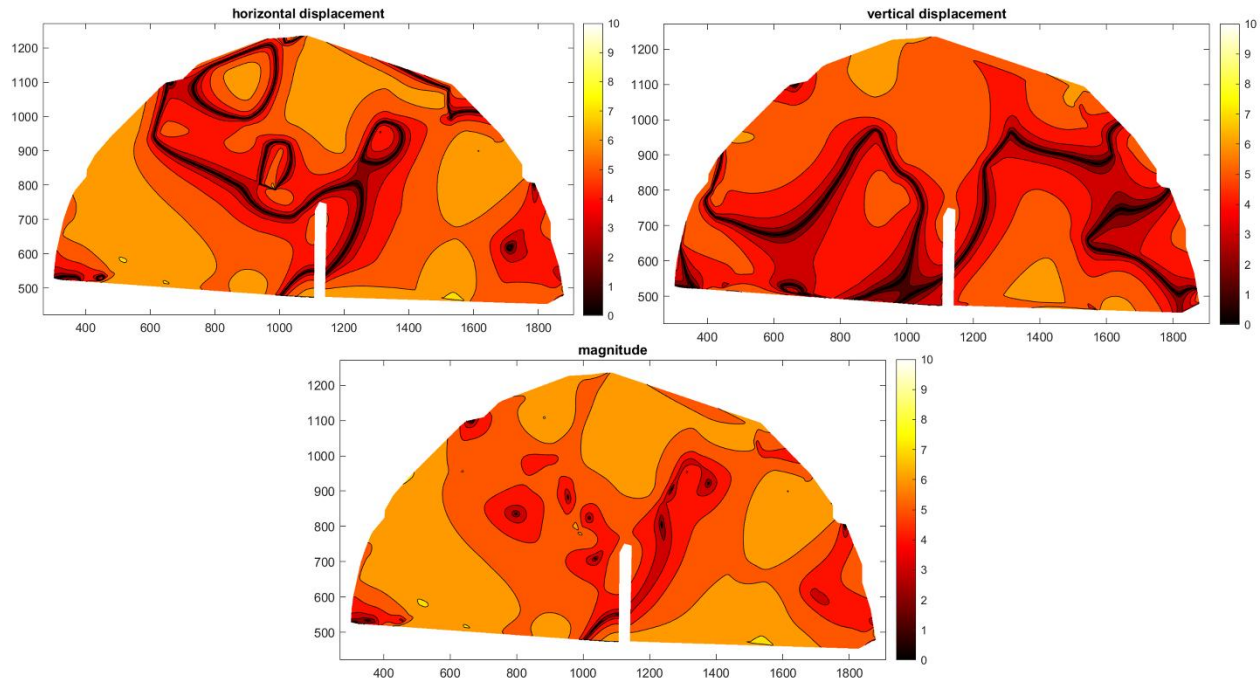


Figure 7: Dense field maps.

Finally, crack evolution was visualized by overlaying the displacement magnitude on the ROI and inspecting successive frames. The concentration bands align with the visually observed fracture path and expand temporally in a manner consistent with the Load–CMOD response captured by the test machine; the final panel shows the mature crack trajectory within the ROI (Figure 2). Taken together, Figures 2–7 demonstrate that a SURF based, threshold and percentile filtered matching pipeline, followed by edge preserving interpolation, can recover sub pixel scale, spatially coherent motions with minimal hardware and pre-processing—capturing the onset and growth of cracking in SCB tests using standard webcam imagery.

5. CONCLUSION AND RECOMMENDATIONS

This study introduced a feature-based displacement tracking framework, RiFT, as an alternative to conventional correlation-heavy methods such as DIC for monitoring crack evolution in asphalt mixtures under SCB loading.

These advantages position RiFT as a scalable solution for fracture analysis in asphalt pavements and other quasi-rigid materials, particularly where resource constraints or field conditions limit the feasibility of conventional DIC systems, thus maximizing the utility of the standard SCB test.

The experimental results demonstrated that RiFT can accurately capture sub-pixel displacements and visualize crack propagation using standard-resolution images acquired from a webcam, without the need for speckle patterns, controlled lighting, or high-speed acquisition systems. By leveraging SURF-based landmark point detection and robust matching combined with percentile-based outlier filtering and edge-preserving interpolation, RiFT achieved reliable displacement fields while significantly reducing computational complexity and hardware costs.

While this study successfully validates the RiFT framework under controlled laboratory conditions using a single asphalt mixture type, future research is necessary to fully benchmark its accuracy across different mixtures and to validate its robustness under diverse field lighting and texture conditions (See Section 6).

The findings confirm that RiFT addresses key limitations of traditional optical methods:

- Eliminates dependency on expensive cameras and elaborate calibration procedures.
- Maintains accuracy under modest angular and illumination variations, making it suitable for field conditions.
- Sparse feature matching and interpolation reduce processing time compared to full-field correlation approaches.

These advantages position RiFT as a scalable solution for fracture analysis in asphalt pavements and other quasi-rigid materials, particularly where resource constraints or field conditions limit the feasibility of conventional DIC systems.

6. RECOMMENDATIONS FOR FUTURE WORK

- Extend RiFT to fatigue SCB tests and cyclic loading scenarios, where crack evolution occurs over thousands of cycles.
- Explore feature selection and displacement prediction using deep learning to further improve robustness under noisy conditions.
- Validate RiFT in real pavement sections under in-service conditions, assessing its adaptability to uncontrolled lighting and environmental variability.
- Benchmark RiFT against DIC-based fracture indicators (e.g., J-integral, flexibility index) to establish conversion relationships for design standards.
- Incorporate automated segmentation techniques to eliminate manual annotation, enabling fully autonomous image processing pipelines.

By addressing these directions, RiFT can evolve into a comprehensive, low-cost, and field-ready tool for structural health monitoring and pavement performance evaluation.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest among the authors.

ACKNOWLEDGEMENT

The author of this article would like to gratefully acknowledge KORDSA Teknik Tekstil A.Ş. for the fracture tests of SCB specimens into the experimental works.

CONTRIBUTIONS OF AUTHORS

Y.A.: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing, review and editing —original draft preparation

REFERENCES

- AASHTO. (2014). *AASHTO T322: Determining the Creep Compliance and Strength of Hot Mix Asphalt (HMA) Using the Indirect Tensile Test Device* AASHTO.
- AASHTO. (2017). *AASHTO T321: Standard Method of Test for Determining the Fatigue Life of Compacted Hot Mix Asphalt (HMA) Subjected to Repeated Flexural Bending*. In. Washington, DC: AASHTO.

- AASHTO. (2020a). *AASHTO TP 105: Standard Method of Test for Determining the Fracture Energy of Asphalt Mixtures Using the Semicircular Bend Geometry (SCB)*. American Association of State Highway and Transportation Officials (AASHTO).
- AASHTO. (2020b). *AASHTO TP 124: Standard Method of Test for Determining the Fracture Potential of Asphalt Mixtures Using the Illinois Flexibility Index Test (I-FIT)*. American Association of State Highway and Transportation Officials (AASHTO).
- Abut, Y. (2024). Experimental Investigation of the Fracture Properties for Crumb Rubber Modified (CRM) Asphalt Mixtures [Granül Kauçuk Modifiyeli (CRM) Asfalt Karışımların Kırılma Özelliklerinin Deneysel Olarak İncelenmesi]. *Politeknik Dergisi*, 27(2), 515-522. <https://doi.org/10.2339/politeknik.1125034>
- Alsheyab, M. A., Khasawneh, M. A., Abualia, A., & Sawalha, A. (2024). A critical review of fatigue cracking in asphalt concrete pavement: a challenge to pavement durability. *Innovative Infrastructure Solutions*, 9(10), 386. <https://doi.org/10.1007/s41062-024-01704-1>
- Asswad, J., Ziade, E., Tehrani, F. F., & Pop, I. O. (2024). Assessment of resistance to crack propagation of asphalt mixtures using two parts digital image correlation. *Theoretical and Applied Fracture Mechanics*, 133, 104578. <https://doi.org/https://doi.org/10.1016/j.tafmec.2024.104578>
- ASTM. (2013). *ASTM D7313: Standard Test Method for Determining Fracture Energy of Asphalt Mixtures Using the Disk-Shaped Compact Tension Geometry*. ASTM International.
- ASTM. (2016). *ASTM D8044: Standard Test Method for Evaluation of Asphalt Mixture Cracking Resistance using the Semi-Circular Bend Test (SCB) at Intermediate Temperatures*. ASTM International.
- ASTM. (2017). *ASTM D6931: Standard Test Method for Indirect Tensile (IDT) Strength of Asphalt Mixtures*. ASTM International.
- ASTM. (2021). *ASTM D8237: Standard Test Method for Determining Fatigue Failure of Asphalt-Aggregate Mixtures with the Four-Point Beam Fatigue Device*. ASTM International. <https://doi.org/10.1520/D8237-21>
- Bay, H., Tuytelaars, T., & Van Gool, L. (2006, 2006/). SURF: Speeded Up Robust Features. *Computer Vision – ECCV 2006*, Berlin, Heidelberg.
- Birgisson, B., Montepara, A., Romeo, E., Roncella, R., Napier, J. A. L., & Tebaldi, G. (2008). Determination and prediction of crack patterns in hot mix asphalt (HMA) mixtures. *Engineering Fracture Mechanics*, 75(3), 664-673. <https://doi.org/https://doi.org/10.1016/j.engfracmech.2007.02.003>
- Cabo, C., Ordóñez, C., Muñoz-Calvente, M., Lozano, M., & Ismael, G. (2019). A hybrid SURF-DIC algorithm to estimate local displacements in structures using low-cost conventional cameras. *Engineering Failure Analysis*, 104, 807-815. <https://doi.org/https://doi.org/10.1016/j.engfailanal.2019.06.083>
- Cheng, L., Zhang, L., Liu, X., Yuan, F., Ma, Y., & Sun, Y. (2022). Evaluation of the fatigue properties for the long-term service asphalt pavement using the semi-circular bending tests and stereo digital image correlation technique. *Construction and Building Materials*, 317, 126119. <https://doi.org/https://doi.org/10.1016/j.conbuildmat.2021.126119>
- EN. (2010). *EN 12697-44: Bituminous mixtures – Test methods – Part 44: Crack propagation by semi-circular bending test*. European Committee for Standardization (CEN). <https://www.en-standard.eu/din-en-12697-44-bituminous-mixtures-test-methods-part-44-crack-propagation-by-semi-circular-bending-test/>
- EN. (2022). *EN 12697-32: Bituminous Mixtures, Test Methods for Hot Mix Asphalt, Laboratory Compaction of Bituminous Mixtures by Vibratory Compactor*. European Committee for Standardization (CEN). <https://www.en-standard.eu/csn-en-12697-32-bituminous-mixtures-test-methods-part-32-specimen-preparation-by-vibratory-compactor/>
- Górszczyk, J., Malicki, K., & Zych, T. (2019). Application of Digital Image Correlation (DIC) Method for Road Material Testing. *Materials*, 12(15). <https://doi.org/10.3390/ma12152349>
- Hassan, M. M. (2013). Effect of secondary aggregates on relationship between creep time dependent index and Paris Law parameters. *Journal of Civil Engineering and Construction Technology*, 4(2), 45-50. <https://doi.org/https://doi.org/10.5897/JCECT12.040>

- Huang, B., Shu, X., & Zuo, G. (2013). Using notched semi circular bending fatigue test to characterize fracture resistance of asphalt mixtures. *Engineering Fracture Mechanics*, 109, 78-88. <https://doi.org/https://doi.org/10.1016/j.engfracmech.2013.07.003>
- Jiang-san, H., Lan, W., & Xin, L. (2022). Anti-fatigue performance of warm-mixed rubber powder modified asphalt mixture based on the DIC technique. *Construction and Building Materials*, 335, 127489. <https://doi.org/https://doi.org/10.1016/j.conbuildmat.2022.127489>
- Kong, L., Ren, D., Zhou, S., He, Z., Ai, C., & Yan, C. (2023). Evaluating the evolution of fiber-reinforced emulsified asphalt cold-recycled mixture damage using digital image correlation. *International Journal of Pavement Engineering*, 24(1), 2176495. <https://doi.org/10.1080/10298436.2023.2176495>
- Li, X. J., & Marasteanu, M. O. (2010). Using Semi Circular Bending Test to Evaluate Low Temperature Fracture Resistance for Asphalt Concrete. *Experimental Mechanics*, 50(7), 867-876. <https://doi.org/10.1007/s11340-009-9303-0>
- Liang, H., Wang, D., Shi, L., Liang, X., & Tang, C. (2020). Use of digital images for fracture performance evaluation of asphalt mixtures. *Construction and Building Materials*, 253, 119152. <https://doi.org/https://doi.org/10.1016/j.conbuildmat.2020.119152>
- Lu, D. X., Bui, H. H., & Saleh, M. (2021). Effects of specimen size and loading conditions on the fracture behaviour of asphalt concretes in the SCB test. *Engineering Fracture Mechanics*, 242, 107452. <https://doi.org/https://doi.org/10.1016/j.engfracmech.2020.107452>
- Ma, Y., Mohammad, L. N., Liu, J., Asghar, M., Elnaml, I., Cooper, S. B., & Cooper, S. B. (2024). Development of a cyclic semi-circular bending test protocol to characterize fatigue cracking of asphalt mixture at intermediate temperature. *Construction and Building Materials*, 443, 137669. <https://doi.org/https://doi.org/10.1016/j.conbuildmat.2024.137669>
- Marasteanu, M., Buttlar, W., Bahia, H., Williams, C., Moon, K. H., Teshale, E. Z., Falchetto, A. C., Turos, M., Dave, E., & Paulino, G. (2012). *Investigation of low temperature cracking in asphalt pavements national pooled fund study-phase II*.
- Min, W., Lu, P., Liu, S., & Wang, H. (2025). A Review of Crack Sealing Technologies for Asphalt Pavement: Materials, Failure Mechanisms, and Detection Methods. *Coatings*, 15(7). <https://doi.org/10.3390/coatings15070836>
- Ouyang, J., Yang, W., Cao, P., & Han, B. (2023). The fracture behaviour of cement bitumen emulsion mixture through the digital image correlation (DIC) method. *International Journal of Pavement Engineering*, 24(1), 2220065. <https://doi.org/10.1080/10298436.2023.2220065>
- Pei, Z., Lou, K., Kong, H., Wu, B., Wu, X., Xiao, P., & Qi, Y. (2021). Effects of Fiber Diameter on Crack Resistance of Asphalt Mixtures Reinforced by Basalt Fibers Based on Digital Image Correlation Technology. *Materials*, 14(23). <https://doi.org/10.3390/ma14237426>
- Radeef, H. R., Abdul Hassan, N., Mahmud, M. Z. H., Zainal Abidin, A. R., Ismail, C. R., Abbas, H. F., & Al-Saffar, Z. H. (2021). Characterisation of cracking resistance in modified hot mix asphalt under repeated loading using digital image analysis. *Theoretical and Applied Fracture Mechanics*, 116, 103130. <https://doi.org/https://doi.org/10.1016/j.tafmec.2021.103130>
- Saha, G., & Biligiri, K. P. (2015). Fracture damage evaluation of asphalt mixtures using Semi-Circular Bending test based on fracture energy approach. *Engineering Fracture Mechanics*, 142, 154-169. <https://doi.org/https://doi.org/10.1016/j.engfracmech.2015.06.009>
- Saha, G., & Biligiri, K. P. (2016). Fracture properties of asphalt mixtures using semi-circular bending test: A state-of-the-art review and future research. *Construction and Building Materials*, 105, 103-112. <https://doi.org/https://doi.org/10.1016/j.conbuildmat.2015.12.046>
- Son, S. (2014). *Development of a Phenomenological Constitutive Model for Fracture Resistance Degradation of Asphalt Concrete with Damage Growth Due to Repeated Loading* [Dissertation, University of Illinois at Urbana-Champaign]. Urbana, Illinois. <https://www.ideals.illinois.edu/items/49876>
- Texas. (2025). *Tex-248-F: Test procedure for the overlay test*. Texas Department of Transportation (TxDOT).

- Wang, L., Shan, M., & Li, C. (2020). The cracking characteristics of the polymer-modified asphalt mixture before and after aging based on the digital image correlation technology. *Construction and Building Materials*, 260, 119802. <https://doi.org/https://doi.org/10.1016/j.conbuildmat.2020.119802>
- Zhang, J., Sakhaeifar, M., Little Dallas, N., Bhasin, A., & Kim, Y.-R. (2018). Characterization of Crack Growth Rate of Sulfur-Extended Asphalt Mixtures Using Cyclic Semicircular Bending Test. *Journal of Materials in Civil Engineering*, 30(12), 04018311. [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0002517](https://doi.org/10.1061/(ASCE)MT.1943-5533.0002517)
- Zhao, Y., Ni, F., Zhou, L., & Jiang, J. (2017). Heterogeneous fracture simulation of asphalt mixture under SCB test with cohesive crack model. *Road Materials and Pavement Design*, 18(6), 1411-1422. <https://doi.org/10.1080/14680629.2016.1230071>
- Zhou, F., Newcomb, D., Gurganus, C., Banihashemrad, S., Park, E. S., Sakhaeifar, M., & Lytton, R. L. (2016). *Experimental design for field validation of laboratory tests to assess cracking resistance of asphalt mixtures* (NCHRP Project, Issue. https://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP09-57_FR.pdf
- Zieliński, P., Klimczak, M., Tekieli, M., & Strzpek, M. (2025). An Evaluation of the Fracture Properties of Asphalt Concrete Mixes Using the Semi-Circular Bending Method and Digital Image Correlation. *Materials*, 18(5). <https://doi.org/10.3390/ma18050967>