

Atıf İçin: Sari, E. F., Apaydin, N. N., Yaman, O. & Karakose, M. (2025). Kavga Tespiti İçin DenseNet121 ve Transfer Öğrenme Yöntemi. *İğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 15(4), 1165-1177.

To Cite: Sari, E. F., Apaydin, N. N., Yaman, O. & Karakose, M. (2025). DenseNet121 and Transfer Learning Method for Fight Detection. *Journal of the Institute of Science and Technology*, 15(4), 1165-1177.

Akan Görüntülerde Kavga tespiti için DenseNet121 ve Transfer Öğrenme Tabanlı Video Analiz Yöntemi

Elif Feyza SARI¹, Nafiye Nur APAYDIN^{2*}, Orhan YAMAN³, Mehmet KARAKÖSE⁴

Öne Çıkanlar:

- Kavga tespiti
- Transfer öğrenme
- Densnet121
- Video işleme

Anahtar Kelimeler:

- Peliculas veri seti
- Hockey veri seti
- Anahtar kare çıkarımı

ÖZET:

Ani gelişen kavga olaylarının erken tespiti ve bu olaylara hızlı ve etkili bir şekilde müdahale edilmesi, kamu güvenliğinin sağlanması ve olası olumsuz durumların önlenmesi açısından büyük bir öneme sahiptir. Günümüz teknolojik imkanları doğrultusunda, video tabanlı otomatik kavga tespit sistemleri, güvenlik güçlerine ve yetkililere zamanında uyarılar sağlayarak proaktif önlemler alınmasını mümkün kılmaktadır. Bu çalışmada, kavga (fight) ve kavga içermeyen (noFight) durumların doğru bir şekilde sınıflandırılabilmesi amacıyla derin öğrenme tabanlı bir yöntem önerilmektedir. Özellikle, görüntü işleme alanında başarılı sonuçlar elde eden DenseNet121 modeli kullanılarak transfer öğrenme yöntemiyle video tabanlı kavga tespiti gerçekleştirilmiştir. Deneysel çalışmalarda, literatürde yaygın olarak kullanılan Peliculas ve Hockey veri setleri değerlendirilmiş olup, önerilen yöntem ile Peliculas veri seti üzerinde %100 doğruluk (Accuracy), Hockey veri seti üzerinde ise %99.32 doğruluk elde edilmiştir. Elde edilen yüksek başarı oranları, önerilen yöntemin kavga tespitinde etkin bir şekilde kullanılabileceğini göstermektedir.

DenseNet121 and Transfer Learning Based Video Analysis Method for Fight Detection in Streaming Images

Highlights:

- Fight detection
- Transfer learning
- DensNet121
- Video processing

Keywords:

- Peliculas dataset
- Hockey dataset
- Keyframe extraction

ABSTRACT:

Early detection of sudden fight incidents and prompt, effective intervention are crucial for ensuring public safety and preventing potential adverse situations. With the advancements in modern technology, video-based automatic fight detection systems enable security forces and authorities to take proactive measures by providing timely alerts. In this study, a deep learning-based method is proposed to accurately classify fight and non-fight situations. Specifically, video-based fight detection is performed using transfer learning with the DenseNet121 model, which has demonstrated successful results in image processing tasks. In the experimental studies, the widely used Peliculas and Hockey datasets from the literature were evaluated. The proposed method achieved 100% accuracy on the Peliculas dataset and 99.32% accuracy on the Hockey dataset. These high success rates indicate that the proposed method can be effectively used for fight detection.

^{1*}Elif Feyza SARI ([Orcid ID: 0009-0003-4806-9242](https://orcid.org/0009-0003-4806-9242)) Fırat Üniversitesi, Teknoloji Fakültesi, Adli Bilişim Mühendisliği Bölümü, Elazığ, Türkiye,

²Nafiye Nur APAYDIN ([Orcid ID: 0009-0006-3438-7401](https://orcid.org/0009-0006-3438-7401)) Fırat Üniversitesi, Teknoloji Fakültesi, Adli Bilişim Mühendisliği Bölümü, Elazığ, Türkiye,

³Orhan YAMAN ([Orcid ID: 0000-0001-9623-2284](https://orcid.org/0000-0001-9623-2284)) Fırat Üniversitesi, Teknoloji Fakültesi, Adli Bilişim Mühendisliği Bölümü, Elazığ, Türkiye,

⁴Mehmet KARAKÖSE ([Orcid ID: 0000-0002-3276-3788](https://orcid.org/0000-0002-3276-3788)) Fırat Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, Elazığ, Türkiye

***Sorumlu Yazar/Corresponding Author:** Nafiye Nur Apaydin, e-mail: 241144110@firat.edu.tr

Günümüzde teknolojinin gelişmesiyle beraber güvenlik kameralarının kullanımı yaygınlaşmaktadır. Güvenlik kameraları sayesinde insanlar kendilerini 7/24 güvende hissedebilmekte ve basit hırsızlık olayları bile hızlı bir şekilde gün yüzüne çıkarak suçlular kolayca tespit edilebilmektedir. Ancak güvenlik kameraları ele alındığında videoların süresinin uzun olması, birbirini takip eden benzer görüntülerden oluşmasından dolayı şüpheli aktiviteler durumunda ani gelişen olayların takibini insan gücüyle yapmak zor, uzun ve zahmetli olabilir (Akash, Moorthy, Esha ve Nathiya, 2022). Şiddet olayları; özellikle okullar, spor etkinlikleri, metro istasyonları gibi kalabalık ortamlarda ve kamuya açık alanlarda ciddi bir güvenlik sorunu haline gelmektedir. Toplum güvenliğini tehdit eden bu olaylara erken müdahale edebilmek için güvenlik kameralarından hızlı bir şekilde kavga tespiti yapılarak gerekli yerlere bildirmek önemlidir (Hassner, Itcher ve Kliper-Gross, 2012). Bu nedenle literatürde kavga tespiti ile ilgili çalışmalar sürekli olarak yapılmakta ve geliştirilmektedir (Nievas, Suarez, Garcia ve Sukthankar, 2011).

Örneğin; Akti, Tataroğlu ve Ekenel, güvenlik kamerası görüntülerinden kavga anlarının hızlı ve doğru bir şekilde tespitini hedeflemiş ve dikkat katmanı ile Bi-LSTM ağlarını içeren bir yöntem önermişlerdir (Akti, Tataroğlu ve Ekenel, 2019). Carneiro ve ark. videolarda kavga tespiti için çok akışlı derin öğrenme mimarisi önererek mekânsal, zamansal, ritmik ve derinlik bilgilerini entegre etmeyi amaçlamışlardır (Carneiro ve ark., 2019). Venkatesh ve ark. gözetim tabanlı suç tespiti konusunda uç birim üzerinde gerçek zamanlı çalışan bir sistem geliştirmiştir (Venkatesh ve ark., 2020). Nadeem ve Kurugollu, sentetik veriler kullanarak şiddet tespiti amacıyla VGG-16 ve VGG-16+LSTM ağlarını eğitmiş ve iki farklı transfer öğrenme stratejisi ile gerçek hayat verilerinde bu modellerin başarımını değerlendirmiştir (Nadeem ve Kurugollu, 2021). Akti ve ark. sosyal medya platformlarında, özellikle Twitter'daki kavga sahnelerinin tespitine odaklanarak ResNet-50, ViT ve VGG-16 yöntemlerini karşılaştırmışlardır (Akti ve ark., 2022). Iqar ve ark. videolardaki potansiyel tehlikeli davranışları tespit etmek amacıyla CNN ve LSTM yöntemlerini birleştiren bir sistem geliştirmiştir (Iqar ve ark., 2023). Jaiswal ve ark. Hockey ve Peliculas veri setleri üzerinde HOG tabanlı Ensemble Boosting yaklaşımını kullanarak en az sayıda özellik ile etkin şiddet tespiti gerçekleştirmişlerdir (Jaiswal ve ark., 2023). Son olarak, Stanley ve ark. Spatio-Temporal Shift Module (STSM) ve Multi-Scale Attention Module (MSAM) yöntemlerini kullanarak kamuya açık alanlarda meydana gelen şiddet davranışlarını tespit etmeye yönelik etkili bir sistem sunmuşlardır (Stanley ve ark., 2023).

Makine öğrenmesi, son yıllarda teknolojik ilerlemelere önemli katkılar sağlamış ve pek çok alanda yaygın olarak kullanılmaya başlanmıştır. Web arama motorları, e-ticaret öneri sistemleri ve sosyal medya içerik filtreleme gibi uygulamalar makine öğrenmesi yöntemlerinin günlük yaşamdaki yansımalarına örnek teşkil etmektedir. Bu yöntemler, sesli komutların yazıya dökülmesi, görüntülerdeki nesnelerin tanınması veya kullanıcı ilgisine yönelik içeriklerin belirlenmesi gibi işlevleri de kapsamaktadır. Makine öğrenmesinin alt dalı olan derin öğrenme, çok katmanlı hesaplamalı modeller aracılığıyla verilerden otomatik olarak anlamlı temsiller öğrenmeyi amaçlar. Özellikle büyük veri kümelerinde, geri yayılım algoritması sayesinde karmaşık yapıları keşfedebilme kapasitesi ile öne çıkar. Derin öğrenme, konuşma ve görüntü tanıma, nesne algılama, otomatik çeviri ve metin sınıflandırma gibi çok sayıda alanda yüksek başarı göstermektedir. Bu başarı, mimari yapısının karmaşıklığına rağmen öğrenme gücünün yüksek olmasından kaynaklanmaktadır. 2006 yılında Hinton ve arkadaşları tarafından ortaya konan derin öğrenme kavramı, 2012 yılında ImageNet yarışmasında elde edilen çarpıcı sonuçlarla geniş çapta dikkat çekmiştir. Özellikle evrimsel sinir ağları (CNN), görüntü işleme ve tanıma gibi görevlerde en çok tercih edilen modellerden biri haline gelmiştir. CNN mimarisi, veriler üzerinde yerel örüntüleri tanımlayabilen evrimsel katmanlar, boyut küçültme işlemleri yapan havuzlama katmanları ve nihai karar sürecinde görev alan tam bağlı katmanlardan oluşur. CNN'de karşılık gelen filtreler

sayesinde bir görüntünün mekansal ve zamansal bağımlılıkları yakalanabilmektedir (Karaduman ve Akin, 2020). Bu nedenle CNN, modelin öğrenme performansını artırırken işlem verimliliğini de sağlar. Derin öğrenmenin kullanım alanları yalnızca görüntü tanıma ile sınırlı kalmamaktadır. Tıp, finans, savunma sanayi, sinyal işleme, doğal dil işleme, ilaç geliştirme ve hatta sözlük oluşturma gibi çok çeşitli disiplinlerde başarılı şekilde uygulanmaktadır (Kunduracioglu ve Pacal 2024). Örneğin; Ali ve Yaman çevresel seslerin otomatik tanımlanabilmesi için sekiz farklı ortamdan alınan sesleri kullanarak yeni bir ses veri seti elde etmişlerdir. Elde edilen ses veri setine Uzun Kısa Süreli Bellek (LSTM), Çift Yönlü Uzun Kısa Süreli Bellek (BiLSTM) ve Evrişimli Sinir Ağı (CNN) modelleri uygulanarak sırasıyla %99.4, %99.23 ve %93.35 doğruluk elde edilmiştir. Test sonuçlarında ise bu üç model için %97 doğruluk elde edilmiştir (Ali ve Yaman, 2024). Bir başka çalışmada, Ozel ve ark. insan eylem sınıflandırması için KTH (İnsan Eylemi Tanıma) veri kümesine geliştirdikleri CNN modelini uygulayarak %96 doğruluk elde etmişlerdir. Çalışma sonucunda elde edilen doğruluk KTH veri setini kullanan önceki çalışmalarla karşılaştırıldığında, önerilen yöntem literatürdeki diğer yaklaşımlara kıyasla üstün performans göstermiştir. Bu sayede, KTH veri setinde insan eylem tanıma için iyi yapılandırılmış bir derin öğrenme modelinin kullanılmasının avantajlı olduğu görülmektedir (Ozel, Apaydin, Yaman ve Karakose, 2025). Pacal ve ark. ISIC 2019 ve HAM10000 veri setlerini kullanarak cilt kanseri için optimize edilmiş, MetaFormer mimarisine dayalı yenilikçi bir derin öğrenme modeli geliştirmiştir (Pacal ve ark., 2025). Bir diğer çalışmada Pacal ve Attallah kolorektal kanserin doğru tanısı için InceptionNeXt blokları, geliştirilmiş Swin Transformer blokları ve Artık Çok Katmanlı Algılayıcı (ResMLP) yapılarının entegre edildiği hibrit bir derin öğrenme modeli önermiştir (Pacal ve Attallah 2025). Yaygın bir nörolojik rahatsızlık olan ve önemli bir ölüm oranına sahip iskemik inme, hasta sonuçlarının uygun şekilde değerlendirilmesini sağlamak için etkilenen bölgelerin doğru bir şekilde belirlenmesini gerektirir. Bu nedenle bir başka çalışmada, değişken lezyon boyutları, gürültü ve doku yapılarının benzer yoğunluk özellikleri gibi zorluklara karşı, ISLES 2022 veri kümesindeki Difüzyon Ağırlıklı Görüntüler (DWI) üzerinde iskemik inme bölgesinin segmentasyonu amaçlanmış; U-Net, U-Net++ ve Attention U-Net modelleri karşılaştırmalı olarak değerlendirilmiştir (İnce ve ark., 2025). Bir başka çalışmada, serebral damar tıkanıklığı segmentasyonundaki bu zorlukları ele almak için ConvNeXtV2 blokları ve GRN tabanlı Çok Katmanlı Algılayıcılar (MLP) ile geliştirilmiş yeni bir UNet mimarisi önerilmiştir (İnce ve ark., 2025).

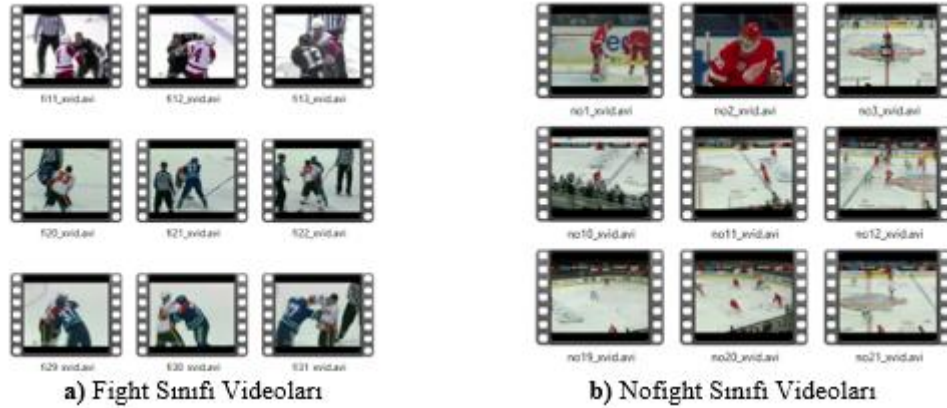
Görüntü işleme alanında derin öğrenme modelleri, son yıllarda önemli bir atılım gerçekleştirmiş ve geleneksel yöntemlere kıyasla üstün başarılar sergilemiştir. Derin öğrenmenin bu alandaki gücü, özellikle karmaşık örüntüleri otomatik olarak öğrenebilme kapasitesinden kaynaklanmaktadır. Çünkü derin öğrenme mimarisi, çoğu öğrenmeye tabi olan ve doğrusal olmayan girdi-çıkı işlemlerini hesaplayan basit modüllerden oluşan çok katmanlı bir yapıdır (LeCun, Bengio ve Hinton, 2015). Bu kapsamda, bir review çalışmasında, 2020'den 2024'e kadar 160 araştırma makalesi analiz edilerek bitki hastalığı tespitinde derin öğrenme tekniklerinin uygulanması sistematik olarak incelenmiştir (Pacal ve ark., 2024). Başka bir review çalışmasında, serebral damar tıkanıklığıyla ilişkili durumların teşhisi için derin öğrenmenin kullanımıyla ilgili 2020 ile 2024 yılları arasında yayınlanmış 61 MRI tabanlı çalışmalar incelenmiştir (Bayram ve ark., 2025). Bu derleme çalışmalarına dayanarak, derin öğrenmenin görüntü işlemeye entegre edilmesi sadece doğruluk açısından değil, aynı zamanda zaman ve iş gücü açısından da verimlilik sağladığı için tercih edildiği görülmektedir. Bu nedenle bu çalışmada, Densenet121 tabanlı Transfer Learning yöntemi önerilmiştir. literatürde bulunan Peliculas ve Hockey veri setlerine önerilen yöntem ayrı ayrı uygulanarak kavga tespiti yapılmıştır. Bu çalışma ile can ile mal kaybının önlenmesi ve toplum düzeninin sağlanmasına katkıda bulunmak amaçlanmıştır.

MATERYAL VE METOT

Veri Seti

Hockey veri seti

Bu veri seti, buz hokeyi maçlarından kavga ve kavga olmayan görüntüler içermektedir. Toplamda 1000 videodan oluşmaktadır. 500'ü “fight” diğer 500'ü ise “noFight” sınıfına ait videolardır. Videoların süresi iki saniyedir. Şekil 1’de örnek Hockey veri seti görüntüleri verilmiştir (Nievas ve ark., 2011).



Şekil 1. Hockey veri seti görüntüleri

Películas veri seti

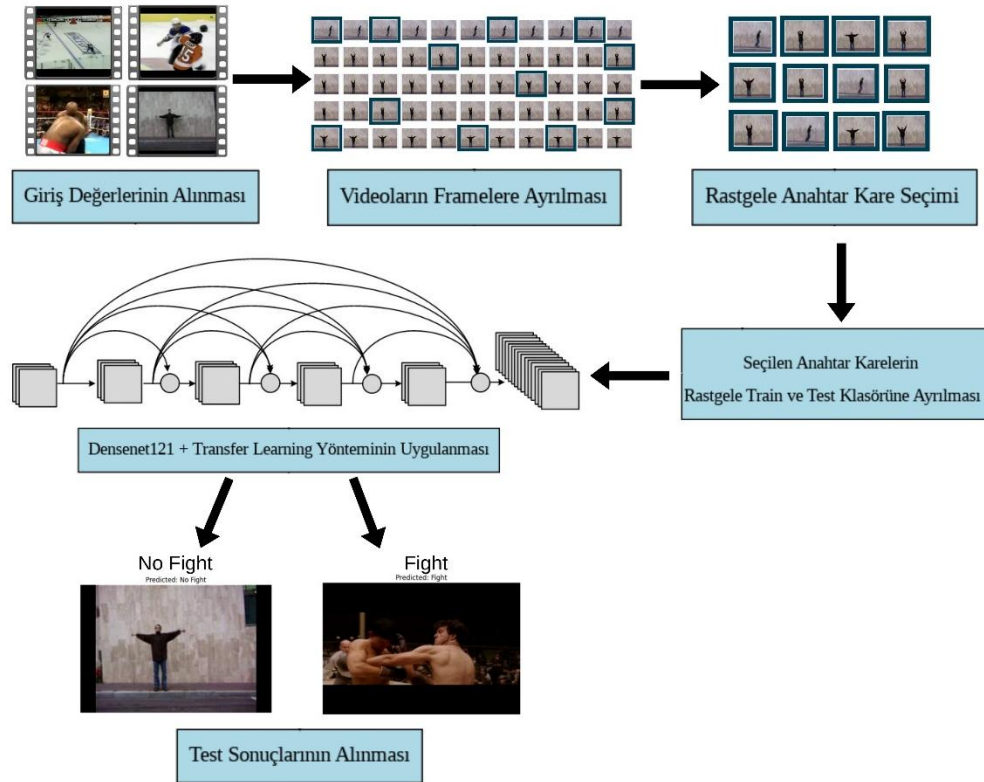
Bu veri seti, Hollywood filmlerinden, futbol maçları gibi olaylardan kavga ve kavga olmayan sahneler içermektedir. Toplamda 200 video bulunmaktadır; bunlardan 100'ü kavga videoları ve 100'ü kavga olmayan günlük yaşamdan alınan yürüyüş, el sallama gibi basit ve sıradan olayları içermektedir. Videolar genellikle 1-2 saniyedir ve kare boyutları değişiklik göstermektedir. Şekil 2’de örnek Películas veri seti görüntüleri verilmiştir (NaveenK903, t.y.).



Şekil 2. Películas veri seti görüntüleri

Yöntem

Şekil 3’te önerilen yöntemin akış diyagramı verilmiştir.



Şekil 3. Önerilen densenet121 ve transfer learning yönteminin akış diyagramı

Şekil 3’te verilen akış diyagramında görüldüğü üzere ilk olarak videolar alınıp framelerine ayrılmıştır. Daha sonra rastgele bir şekilde bu framelerden anahtar kareler seçilmiştir. Anahtar kareler seçildikten sonra Densenet121 + Transfer Learning yöntemi uygulanarak kavga tespiti için yeni bir model oluşturulmuştur. En son da sonuçlar alınarak modelin örnek görseller üzerinde testinin sonuçları elde edilmiştir.

Videoların framelere ayrılması

Películas veri setinde “fight” sınıfında 4791 tane, “nofight” sınıfında 5050 tane olmak üzere toplamda 9841 frame elde edilmiştir. Hockey veri setinde “fight” sınıfında 9420 tane, “nofight” sınıfında 6371 tane olmak üzere toplamda 15791 frame elde edilmiştir.

Rastgele anahtar kare seçimi

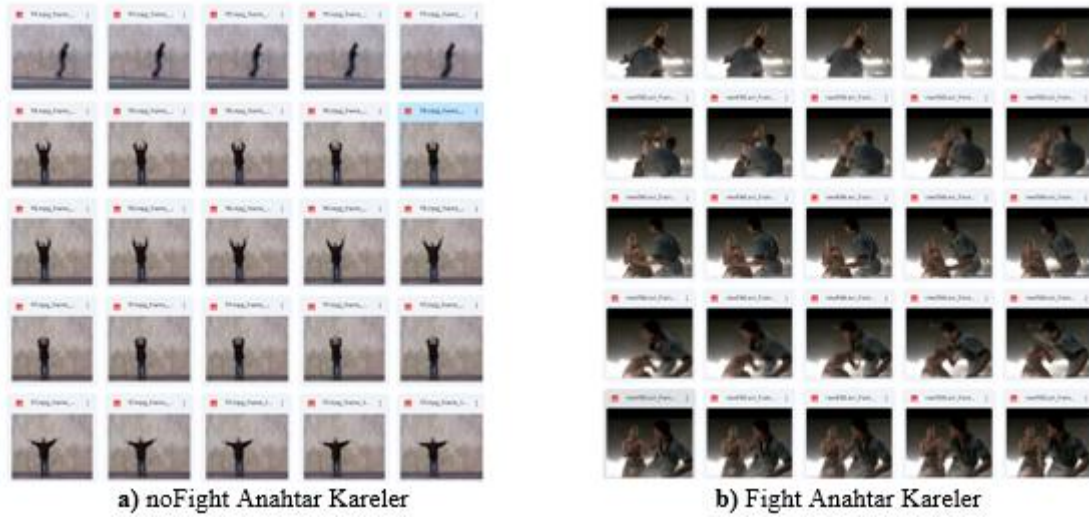
İki veri seti için de videolardan sınıf başına elde edilen framelerin sayısının eşit dağılmaması ve çalışmada kullanılan donanım özelliklerinin yetersizliğinden kaynaklı olarak önerilen yöntemin kullanılıp veri setlerinin eğitilmesinin uzun sürmesinden dolayı rastgele bir şekilde anahtar kare seçimi yapılmıştır. Eğitim için iki veri setinden rastgele anahtar kare seçimi sonucu elde edilen görsel sayıları sınıf bazlı olarak Çizelge 1’de verilmiştir.

Çizelge 1. İki veri seti için rastgele anahtar kare seçme sonucunda eğitim için alınan fight ve nofight sınıfına ait veri sayıları

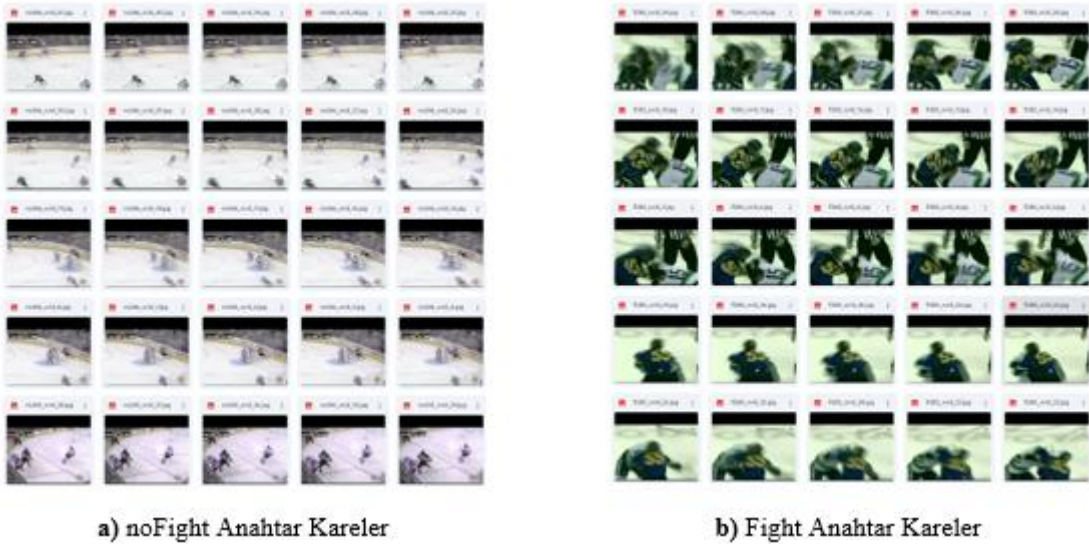
Veri Seti	fight	nofight	Toplam
Hockey	4294	4107	8401
Películas	4500	4500	9000

Çizelge 1’e göre Hockey veri seti için 4294 tanesi “fight” sınıfından, 4107 tanesi “nofight” sınıfından olmak üzere toplamda 8401 tane anahtar kare alınmıştır. Películas veri seti için ise, 4500 tanesi “fight” sınıfından, 4500 tanesi “nofight” sınıfından olmak üzere toplamda 9000 tane anahtar kare alınmıştır.

Şekil 4'te Peliculas veri setinin “fight” ve “nofight” sınıfları için anahtar kare görüntüleri, Şekil 5'te de Hockey veri setinin “fight” ve “nofight” sınıfları için anahtar kare görüntüleri verilmiştir.



Şekil 4. Peliculas veri setinin anahtar kare görüntüleri

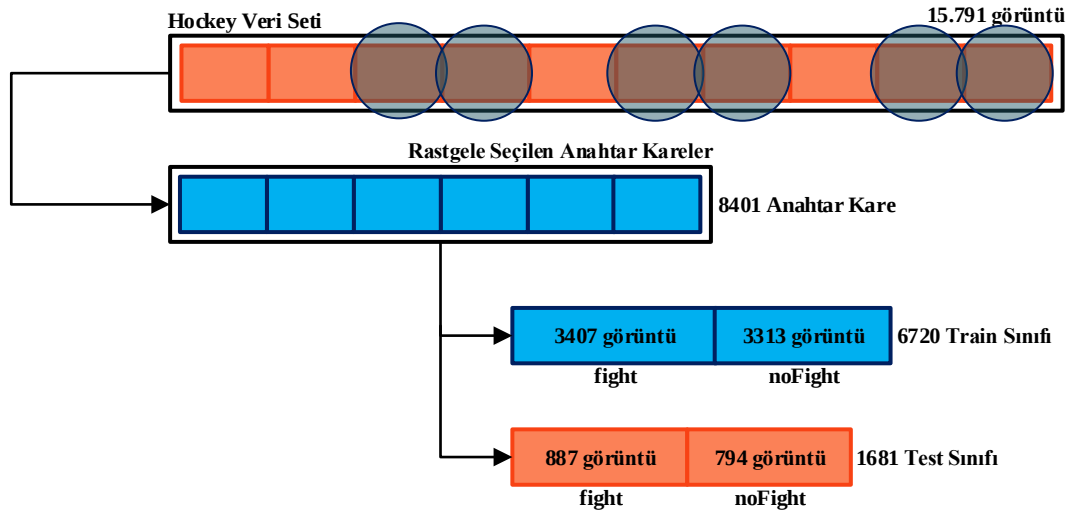


Şekil 5. Hockey veri setinin anahtar kare görüntüleri

Verinin rastgele olarak train ve test sınıflarına ayrılması

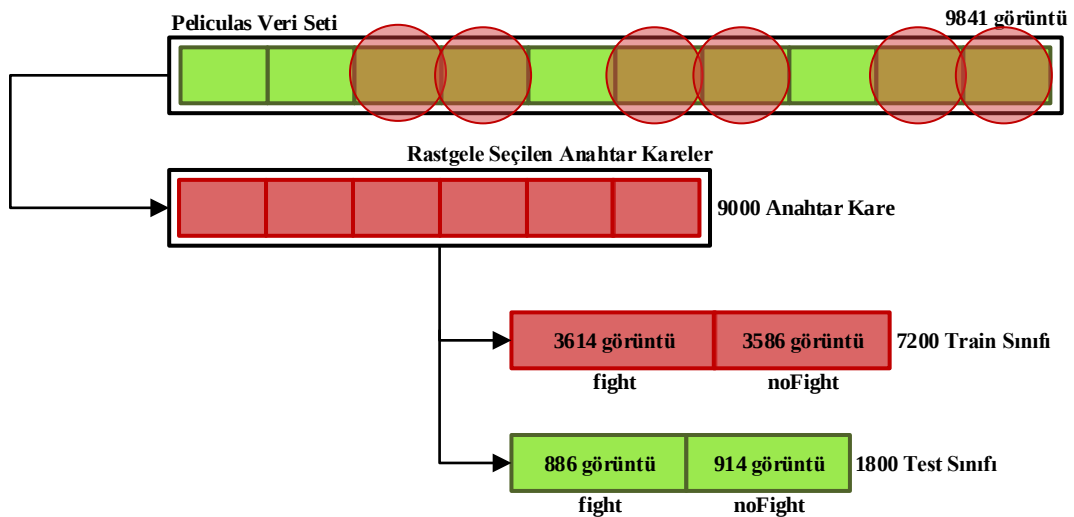
Seçilen anahtar karelerin tamamı rastgele olarak %20'si “Test”, %80'i “Train” olarak seçilmiştir. Bu nedenle Test ve Train de fight ve nofight sınıfları eşit sayıda görsel sahip değildir ancak birbirine yakın sayıda görsel içermektedir.

Hockey veri seti için toplamda 8401 tane anahtar kare sayısı vardır. Bu anahtar karelerin rastgele %20'si yaklaşık olarak 1681 görsel; 887'si “fight”, 794'ü “nofight” sınıfı olarak Test'e, geriye kalan %80'lik kısım yani 6720 görsel 3407'si “fight”, 3313'ü “nofight” sınıfı olarak Train'e seçilmiştir. Şekil 6'da Hockey veri seti için toplam verinin nasıl ele alındığı görülmektedir.



Şekil 6. Hockey veri seti için anahtar karelerin tamamının rastgele olarak test ve train sınıfına ayrılması

Películas veri seti için toplamda 9000 tane anahtar kare sayısı vardır. Bu anahtar karelerin rastgele %20'si yaklaşık olarak 1800 görsel; 886'sı "fight", 914'ü "nofight" sınıfı olarak Teste, geriye kalan %80'lik kısım yani 7200 görsel 3614'ü "fight", 3586'sı "nofight" sınıfı olarak Train için seçilmiştir. Şekil 7'de Películas veri seti için toplam verinin nasıl ele alındığı görülmektedir.

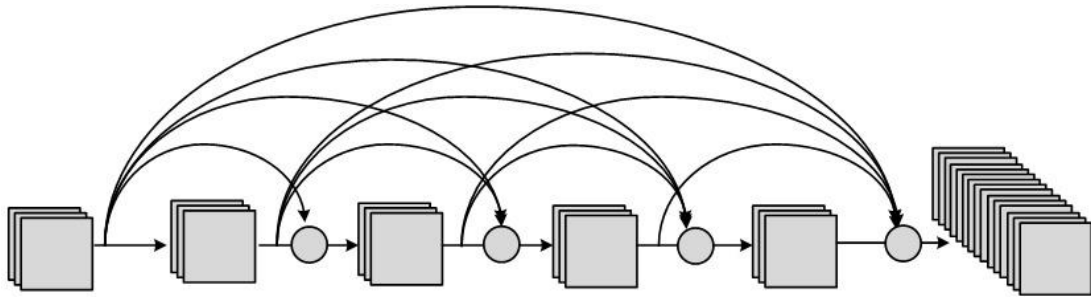


Şekil 7. Películas veri seti için anahtar karelerin tamamının rastgele olarak test ve train sınıfına ayrılması

Rastgele anahtar kare seçimiyle rastgele olarak verilerin Train ve Test'e ayrılması sağlanarak overfitting (aşırı öğrenme) önlenmeye çalışılmıştır.

DenseNet121 modeli

Bu çalışmada önerilen DenseNet121 mimarisi, DenseNet (Densely Connected Convolutional Networks) derin öğrenme mimarisinin 121 katmanlı versiyonudur (Huang, Liu, Van Der Maaten ve Weinberger, 2017). DenseNet121 mimarisinin ismi, "Dense" kelimesinden gelmektedir ve modelin yoğun bağlantılar (densely connections) kullandığını anlatmaktadır (Albawi, Mohammed ve Al-Zawi, 2018). "121" ise modelin toplamda 121 katmandan oluştuğunu belirtir. DenseNet121'de toplam 4 adet yoğun bloklar (Dense Block) bulunur. Yoğun bloklar arasında geçiş katmanları yer alır (Gu ve ark., 2018). Şekil 8'de önerilen DenseNet121 modelinin katman yapısı verilmiştir (Pleiss ve ark., 2017).



Şekil 8. Önerilen DenseNet121 modelinin katman yapısı

Önerilen yöntemin değerlendirilmesi amacıyla Precision (Kesinlik), Recall (Duyarlılık), F-1 Skor, Accuracy (Doğruluk) gibi parametreler kullanılmıştır.

Precision

Precision, modelin pozitif sınıf için yaptığı tahminlerin ne kadarının gerçekten doğru olduğunu gösterir. Precision performans değeri Eşitlik 1. ile hesaplanmaktadır (Kilincer, Ertam ve Sengur, 2021).

$$Precision = TP / (TP + FP) \quad (1)$$

Recall

Recall, pozitif sınıfa ait olan bütün örneklerin kaç tanesinin önerilen model tarafından doğru tespit edildiğini gösterir. Recall performans değeri Eşitlik 2. ile hesaplanmaktadır (Yaman, Yetiş ve Karaköse, 2021).

$$Recall = TP / (TP + FN) \quad (2)$$

F-1 Skor

F1-Skoru, Precision ve Recall'un harmonik ortalamasıdır. Bu metrik, Precision ve Recall arasındaki dengeyi sağlar. F1-Skor performans değeri Eşitlik 3. ile hesaplanmaktadır (Demir ve Yaman, 2024).

$$F1 \text{ Skoru} = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (3)$$

Accuracy

Accuracy, bir modelin yaptığı doğru tahminlerin toplam tahminlere oranını gösteren bir performans ölçütüdür. Eşitlik 4. ile hesaplanmaktadır (Ertam, Kilincer ve Yaman, 2017).

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (4)$$

BULGULAR VE TARTIŞMA

Bu çalışmada sınıflandırma sonuçlarının elde edilmesinde Google Colab kullanılmıştır.

Películas Veri Seti Sonuçları

Şekil 9'da Películas veri setinden elde edilen karmaşıklık matrisi verilmiştir.

		Confusion Matrix	
True Label	fight	886	0
	noFight	0	914
		fight	noFight

Şekil 9. Películas veri seti karmaşıklık matrisi

Şekil 9’da verilen karmaşıklık matrisi incelendiğinde önerilen yöntemin, toplam 1800 örneğin hepsini doğru sınıflandırarak %100 doğruluğa ulaştığı görülmüştür. Çizelge 2’de önerilen yöntemin Peliculas veri seti için 10 iterasyon sonucunda elde edilen performans değerleri verilmiştir.

Çizelge 2. Önerilen yöntemin Peliculas veri seti için 10 iterasyon sonucu performans değerleri

Sınıflar	Accuracy	Precision	Recall	F1-Score
fight	%100	%100	%100	%100
noFight	%100	%100	%100	%100

Çizelge 2’de; önerilen yöntemin, "fights" sınıfı için ortalama %100 precision, %100 recall ve %100 F1-skoru ile "noFights" sınıfı için de %100 precision, %100 recall ve %100 F1-skoru başarı oranlarına ulaştığı görülmektedir. Bu sonuçlar, önerilen yöntemin Peliculas veri seti için her iki sınıfı da hatasız bir şekilde ayırt ettiğini göstermektedir.

Şekil 10’da önerilen yöntem sonucunda elde edilen modelin eğitimde kullanılmayan örnek görseller üzerinde testi verilmiştir.



Şekil 10. Peliculas veri setine ait eğitimde kullanılmayan örnek görsellerin test sonuçları

Hockey Veri Seti Sonuçları

Şekil 11’de Hockey veri setinden elde edilen karmaşıklık matrisi verilmiştir.

		Confusion Matrix	
True Label	fights	881	6
	noFights	2	792
		fights	noFights

Şekil 11. Hockey veri seti karmaşıklık matrisi

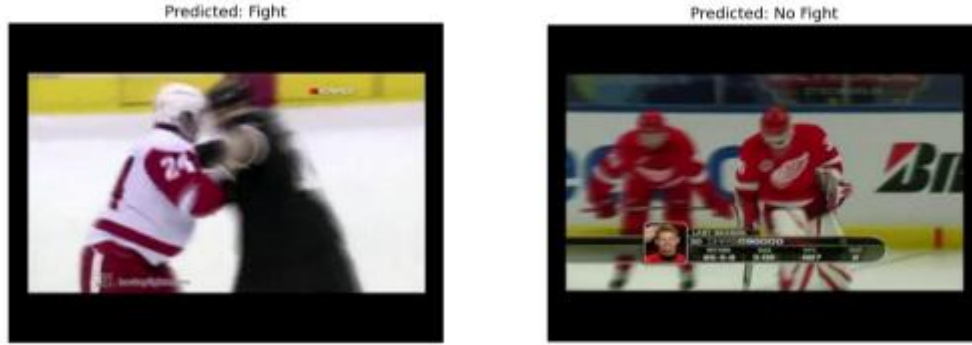
Çizelge 3. Önerilen yöntemin Hockey veri seti için 10 iterasyon sonucu performans değerleri

Sınıflar	Accuracy	Precision	Recall	F1-Score
fight	%99.32	%99.32	%99.77	%99.54
noFight	%99.75	%99.75	%99.30	%99.52

Şekil 11’de verilen sonuçlar incelendiğinde bu confusion matraste rastgele seçilen toplamda 1681 görselin büyük çoğunluğunun doğru sınıflandırıldığı görülmektedir. Matriste, model "fights" sınıfında 881 doğru sınıflama yapmıştır ve sadece 6 örneği yanlış bir şekilde "noFights" olarak tahmin etmiştir.

Önerilen yöntem "noFights" sınıfı için ise 792 doğru tahmin yapmıştır ve sadece 2 örneği yanlış "fights" sınıfı olarak sınıflandırmıştır. Çizelge 3'te önerilen yöntemin 10 iterasyon için Accuracy, Precision, Recall ve F1-Score değerleri verilmiştir.

Şekil 12'de önerilen yöntem sonucunda elde edilen modelin eğitimde kullanılmayan örnek görseller üzerinde testi verilmiştir.



Şekil 12. Hockey veri setine ait eğitimde kullanılmayan örnek görsellerin test sonuçları

Çizelge 4'te literatürde var olan bazı çalışmaların Peliculas ve Hockey veri seti için elde ettiği performans sonuçları ile yapılan çalışmanın sonucunda elde edilen sonuçların karşılaştırılması verilmiştir.

Çizelge 4. Literatürde bulunan bazı çalışmaların Peliculas ve Hockey veri seti için elde ettiği performans değerleri

Çalışma	Yılı	Performans Sonuçları		Yöntem
		Peliculas	Hockey	
Carneiro ve ark.	2019	%89.10	%100	Multi-Stream Convolutional Neural Networks Xception, Bi-LSTM, Attention VGG-16 ve VGG-16 ile birlikte LSTM kullanan versiyonu
Akti, Tataroğlu ve Ekenel	2019	%100	%98	
Nadeem ve Kurugollu	2021	%100	%97	
Akti ve ark.	2022	%100	%98	ResNet-50, ViT, VGG-16, STSM, MSAM
Stanley ve ark.	2023	%99.3	%99.8	
Jaiswal ve ark.	2023	%91.66	%90	Histogram of Oriented Gradients (HOG) CNN, LSTM
Iqrar ve ark.	2023	%98	%90	
Önerilen Yöntem	2025	%100	%99.32	DenseNet121 Transfer Öğrenme

Çizelge 4 incelendiğinde önerilen yöntemin Peliculas ve Hockey veri seti için diğer yöntemlerden daha iyi sonuçlar aldığı görülmektedir.

SONUÇ

Kavga tespiti bireysel ve kamusal güvenliğin sağlanmasında büyük öneme sahiptir. Bu çalışmanın amacı; kavga tespiti için DenseNet121 ve Transfer Öğrenmeyi kullanan yeni bir yöntem önererek kavga durumlarında yetkililerin hızlıca haberdar edilip müdahale etmesini sağlayıp can ve mal kaybının önüne geçmektir. Bu sayede kamusal alanlarda (hastane, toplu taşıma, okul, avm vb.), kalabalık etkinliklerde (konser, maç vb) olası şiddet eylemleri daha hızlı tespit edilerek can, mal kaybı önlenecektir. Çalışmada kavga tespiti yapabilmek için; önce Peliculas ve Hockey veri setleri bütün framelelerine ayrıldıktan sonra rastgele anahtar kare seçme işlemi gerçekleştirilmiştir. Daha sonra seçilen bu anahtar karelerin %80'i train, %20'si ise test sınıfına rastgele olarak ayrılmıştır. DenseNet121 modeli ve Transfer Öğrenme yöntemi kullanılarak train sınıfındaki görüntülerle önerilen yöntem eğitilmiştir ve test sınıfı ile kavga ve kavga olmayan durumlar başarılı bir şekilde tespit edilmiştir. Çalışmada elde edilen sonuçlara göre: Peliculas veri seti için %100 Accuracy, Hockey veri seti için %99,32 Accuracy elde edilmiştir.

TEŞEKKÜR

Bu çalışma TÜBİTAK (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu) tarafından 5220154 numaralı proje kapsamında desteklenmiştir.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

Yazar Katkısı

Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan eder.

KAYNAKLAR

- Akti, S., Ofli, F., Imran, M., & Ekenel, H. K. (2022, January). Fight detection from still images in the wild. In 2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW) (pp. 550–559). IEEE. <https://doi.org/10.1109/WACVW54805.2022.00061>
- Akti, S., Tataroglu, G. A., & Ekenel, H. K. (2019, November). Vision-based fight detection from surveillance cameras. In 2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1–6). IEEE. <https://doi.org/10.1109/IPTA.2019.8936070>
- Akash, S. A., Moorthy, R. S. S., Esha, K., & Nathiya, N. (2022). Human violence detection using deep learning techniques. *Journal of Physics: Conference Series*, 2318(1), 012003. <https://doi.org/10.1088/1742-6596/2318/1/012003>
- Ali, Y. Y., & Yaman, O. (2024). Analysis of LSTM, BiLSTM, and CNN methods for environmental sound identification in smart cities. In *Proceedings of the 2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1–6). IEEE. <https://doi.org/10.1109/IDAP64064.2024.10711138>
- Albawi, S., Mohammed, T. A. S., & Al-Zawi, S. (2018). Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET) (pp. 1–6). IEEE. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- Bayram, B., Kunduracioglu, I., Ince, S., & Pacal, I. (2025). A systematic review of deep learning in MRI-based cerebral vascular occlusion-based brain diseases. *Neuroscience*, 568, 1–22. <https://doi.org/10.1016/j.neuroscience.2025.01.020>
- Carneiro, S. A., Da Silva, G. P., Guimaraes, S. J. F., & Pedrini, H. (2019, October). Fight detection in video sequences based on multi-stream convolutional neural networks. In 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI) (pp. 8–15). IEEE. <https://doi.org/10.1109/SIBGRAPI.2019.00010>
- Demir, K., & Yaman, O. (2024). Projector deep feature extraction-based garbage image classification model using underwater images. *Multimedia Tools and Applications*, 83, 79437–79451. <https://doi.org/10.1007/s11042-024-18731-w>
- Ertam, F., Kilincer, L. F., & Yaman, O. (2017). Intrusion detection in computer networks via machine learning algorithms. In *Proceedings of the 2017 International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1–4). IEEE. <https://doi.org/10.1109/IDAP.2017.8090165>
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77, 354–377. <https://doi.org/10.1016/j.patcog.2017.10.013>
- Hassner, T., Itcher, Y., & Kliper-Gross, O. (2012). Violent flows: Real-time detection of violent crowd behavior. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 1–6). <https://doi.org/10.1109/CVPRW.2012.6239348>

- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708). <https://doi.org/10.1109/CVPR.2017.243>
- Ince, S., Kunduracioglu, I., Algarnic, A., Bayram, B., & Pacal, I. (2025). Deep learning for cerebral vascular occlusion segmentation: A novel ConvNeXtV2 and GRN-integrated U-Net framework for diffusion-weighted imaging. *Neuroscience*, 541, 113–129. <https://doi.org/10.1016/j.neuroscience.2025.04.010>
- Iqrar, W., Shahzad, A., Hameed, W., & Abidien, M. Z. (2023, January). A real-time sequence based human activity detection system. In *2023 International Multi-disciplinary Conference in Emerging Research Trends (IMCERT)* (pp. 1–5). IEEE. <https://doi.org/10.1109/IMCERT57083.2023.10075257>
- İnce, S., Kunduracioglu, I., Bayram, B., Pacal, I. (2025). U-Net-Based Models for Precise Brain Stroke Segmentation. *Chaos Theory and Applications*, 7(1), 50-60. <https://doi.org/10.51537/chaos.1605529>
- Jaiswal, S. G., Mohod, S. W., & Sharma, D. (2023, September). HOG ensembled boosting machine learning approach for violent video classification. *Indian Journal of Science and Technology*, 16(34), 2709–2718. <https://doi.org/10.17485/IJST/v16i34.1777>
- Karaduman, G., & Akin, E. (2020). A deep learning based method for detecting of wear on the current collector strips' surfaces of the pantograph in railways. *IEEE Access*, 8, 183799–183812. <https://doi.org/10.1109/ACCESS.2020.3029555>
- Kilincer, I. F., Ertam, F., & Sengur, A. (2021). Machine learning methods for cyber security intrusion detection: Datasets and comparative study. *Computer Networks*, 188, Article 107840. <https://doi.org/10.1016/j.comnet.2021.107840>
- Kunduracioglu, I., & Pacal, I. (2024). Advancements in deep learning for accurate classification of grape leaves and diagnosis of grape diseases. *Journal of Plant Diseases and Protection*, 131, 1061–1080. <https://doi.org/10.1007/s41348-024-00896-z>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Nadeem, M. S., & Kurugollu, F. (2021, October 25). Effectiveness of synthetic images in violence detection. *TechRxiv*. <https://doi.org/10.36227/techrxiv.16822123.v1>
- Nievas, E. B., Suarez, O. D., Garcia, G. B., & Sukthankar, R. (2011). Violence detection in video using computer vision techniques. In *Computer Analysis of Images and Patterns (CAIP 2011)*, Lecture Notes in Computer Science, vol 6854 (pp. 332–339). Springer. https://doi.org/10.1007/978-3-642-23678-5_39
- Ozel, M. E., Apaydin, N. N., Yaman, O., & Karakose, M. (2025). A CNN-based method using optimized parameters for dynamic human action recognition. In *Proceedings of the 2025 29th International Conference on Information Technology (IT)* (pp. 1–4). IEEE. <https://doi.org/10.1109/IT64745.2025.10930265>
- Pacal, I., Kunduracioglu, I., Alma, M. H., Deveci, M., Kadry, S., Nedoma, J., Slany, V., & Martinek, R. (2024). A systematic review of deep learning techniques for plant diseases. *Artificial Intelligence Review*, 57, 304. <https://doi.org/10.1007/s10462-024-10944-7>
- Pacal, I., Ozdemir, B., Zeynalov, J., Gasimov, H., & Pacal, N. (2025). A novel CNN-ViT-based deep learning model for early skin cancer diagnosis. *Biomedical Signal Processing and Control*, 104, 107627. <https://doi.org/10.1016/j.bspc.2025.107627>

- Pacal, I., & Attallah, O. (2025). Hybrid deep learning model for automated colorectal cancer detection using local and global feature extraction. *Knowledge-Based Systems*, 319, 113625. <https://doi.org/10.1016/j.knosys.2025.113625>
- Pleiss, G., Chen, D., Huang, G., Li, T., van der Maaten, L., & Weinberger, K. Q. (2017). Memory-efficient implementation of DenseNets (arXiv:1707.06990) [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.1707.06990>
- Stanley, B. F., R. K. R., J., & P. P., B. (2023, December). Efficient violence recognition system using spatio-temporal shift multi-scale attention model. In 2023 IEEE 20th India Council International Conference (INDICON) (pp. 509–514). IEEE. <https://doi.org/10.1109/INDICON59947.2023.10440925>
- Venkatesh, S., Anand, A., G. S., Ramakrishnan, A., & Vijayaraghavan, V. (2020). Real-time surveillance based crime detection for edge devices. In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (pp. 801–809). SCITEPRESS. <https://doi.org/10.5220/0008990108010809>
- Yaman, O., Yetiş, H., & Karaköse, M. (2021). Image processing and machine learning-based classification method for hyperspectral images. *The Journal of Engineering*, 2021, 85–96. <https://doi.org/10.1049/tje2.12012>