



Makine Öğrenmesi ile Mesane Kanseri Tespiti

Detection of Bladder Cancer Using Machine Learning

Eda Mukadder Siso¹, Mehmet Ortaburun², Ayşegül Gözaydın³

Geliş Tarihi (Received): 28.11.2024

Kabul Tarihi (Accepted): 31.01.2026

Yayın Tarihi (Published): 27.04.2026

Abstract: Bladder cancer is a type of cancer that is fatal, especially in advanced stages, and is difficult to diagnose early. In this study, a hybrid artificial intelligence approach is presented in which text-based patient complaint data and image data are evaluated together. The aim is to develop a tool that will assist physicians in the diagnosis of bladder cancer through a local decision support system developed in Türkiye and to evaluate the potential contribution of the developed software to health economics. In the study, the performance of the hybrid structure created by the integration of natural language processing (NLP) and image processing techniques was measured with metrics such as accuracy, sensitivity, specificity, F1 score and dice loss, and the clinical significance level was discussed by comparing the results with the literature.

Keywords: Bladder cancer, Machine learning, Decision support system, Deep learning, Hybrid model, Health technologies

&

Öz: Mesane kanseri, özellikle ileri evrede ölümcül seyreden ve erken teşhisi zor olan bir kanser türüdür. Bu çalışmada, metin tabanlı hasta şikâyet verileri ile görüntü verilerinin bir arada değerlendirildiği hibrit bir yapay zekâ yaklaşımı sunulmuştur. Amaç, Türkiye'de geliştirilen yerli karar destek sistemi aracılığıyla mesane kanseri tanısında hekimlere yardımcı olacak bir araç ortaya koymak ve geliştirilen yazılımın sağlık ekonomisine katkı potansiyelini değerlendirmektir. Çalışmada, doğal dil işleme (NLP) ve görüntü işleme tekniklerinin entegrasyonu ile oluşturulan hibrit yapının doğruluk, hassasiyet, özgüllük, F1 skoru ve dice kaybı gibi metriklerle performansı ölçülmüş ve sonuçlar literatürle karşılaştırılarak klinik anlamlılık düzeyi tartışılmıştır.

Anahtar Kelimeler: Mesane kanseri, Makine öğrenmesi, Karar destek sistemi, Derin öğrenme, Hibrit model, Sağlık teknolojileri

Atıf/Cite as Siso EM, Ortaburun M, Gözaydın A. Makine Öğrenmesi ile Mesane Kanseri Tespiti. Abant Sağlık Bilimleri ve Teknolojileri Dergisi. 2026;6(1):1-12.

İntihal-Plagiarism/Etik-Ethic: Bu makale, en az iki hakem tarafından incelenmiş ve intihal içermediği, araştırma ve yayın etiğine uyulduğu teyit edilmiştir. / This article has been reviewed by at least two referees and it has been confirmed that it is plagiarism-free and complies with research and publication ethics. <https://dergipark.org.tr/tr/pub/sabited/policy>

Telif Hakkı/Copyright © Published by Bolu Abant İzzet Baysal University, Since 2021 – Bolu

¹Eda Mukadder Siso, Yönetim Kurulu Üyesi, Sisoft Sağlık Bilgi Sistemleri A.Ş. Ankara, Türkiye, eda.siso@sisoft.com.tr, <https://orcid.org/0009-0001-0591-5803> (Sorumlu Yazar/Corresponding Author)

²Mehmet Ortaburun, Yazılım Geliştirici, Sisoft Sağlık Bilgi Sistemleri A.Ş. Anklara, Türkiye, mehmet.ortaburun@sisoft.com.tr, <https://orcid.org/0009-0009-1644-1840>

³Ayşegül Gözaydın, Kalite Sistem Yöneticisi, Sisoft Sağlık Bilgi Sistemleri A.Ş. Ankara, Türkiye, pelin.usta@sisoft.com.tr, <https://orcid.org/0009-0005-2817-0153>

Giriş

Mesane kanseri, dünyada ve Türkiye’de yaygın görülen, özellikle geç evrede tanı konduğunda ölüm oranlarının yükseldiği önemli bir halk sağlığı sorunudur (1,2). Erken teşhis edilemediğinde tedavi süreçleri karmaşık hale gelmekte ve hastaların yaşam kalitesi belirgin biçimde azalmaktadır. Günümüzde tanıda kullanılan sistoskopi ve biyopsi gibi invaziv yöntemler hem hasta konforunu azaltmakta hem de sağlık hizmetlerine ek maliyet getirmektedir (1). Bu durum, daha hızlı, güvenilir ve hasta dostu alternatif teşhis yaklaşımlarına olan ihtiyacı ortaya koymaktadır.

Son yıllarda yapay zekâ ve makine öğrenmesi teknolojileri, tıbbi görüntü analizinden elektronik sağlık kayıtlarının işlenmesine kadar geniş bir alanda uygulanmaktadır (2). Yapılan literatür araştırmalarında; derin öğrenme tabanlı yöntemlerin tıbbi görüntü analizinde giderek daha yaygın kullanıldığını ve özellikle U-Net mimarisinin medikal görüntü segmentasyonunda yüksek performans sağladığını bildirilmektedir. U-Net ve türevi ağlar, hem lokal ayrıntıları hem de küresel bağlamı birlikte değerlendirebilmesi sayesinde lezyon ve tümör segmentasyonunda etkili sonuçlar sunmaktadır (2). Mesane kanseri özelinde ise, bilgisayarlı tomografi (BT) görüntüleri üzerinde gerçekleştirilen çalışmalarda derin öğrenme tabanlı segmentasyon yaklaşımlarının klinik karar destek açısından umut verici olduğu gösterilmiştir. Baressi Şegota ve arkadaşları (2021), transfer öğrenme destekli U-Net mimarileri kullanarak mesane kanseri kitlelerinin semantik segmentasyonunu gerçekleştirmiştir. Çalışmada farklı görüntü düzlemleri için Dice benzerlik katsayısı (DSC) değerleri 0.93–0.96 aralığına kadar ulaşmış, en yüksek performans sagittal düzlemde $DSC \approx 0.966$ olarak raporlanmıştır. Bu bulgular, U-Net tabanlı derin öğrenme modellerinin mesane kanseri lezyonlarının otomatik belirlenmesinde yüksek doğruluk potansiyeline sahip olduğunu göstermektedir (3). Patel (2019), çok katmanlı karar destek sistemlerinde veri ön işleme tekniklerinin etkinliğini vurgulayarak bu tür uygulamaların tanı verimliliğini artırabileceğini göstermiştir (4). Bununla birlikte, görüntü tabanlı ve metin tabanlı verileri bütünleştiren hibrit karar destek sistemlerinin katkısını gerçek yaşam klinik akışında ve hasta sonuçlarıyla ilişkilendirerek değerlendiren çalışmaların sayısı sınırlıdır.

Bu çalışma, TÜBİTAK tarafından desteklenen ve kurum bünyesinde yürütülen önceki Ar-Ge projelerinde geliştirilen Sisodess ve Sisowiser platformlarının çıktıları üzerine kurgulanmıştır. Sisodess, klinik verileri analiz ederek hekimlerin tanı ve tedavi kararlarında destek sağlayan, mobil tabanlı bir klinik karar destek sistemidir ve amacı sağlık hizmetlerinde kalite ve güvenliği artırmaktır. Sisowiser ise mesane kanseri verisine özgü olarak optimize edilmiş bir karar destek sistemidir. Bu iki platformdan edinilen deneyimler doğrultusunda, literatürdeki mevcut boşluğu doldurmak, hibrit veri kullanımının tanı süreçlerindeki potansiyel katkısını göstermek ve klinik uygulamaya dönük bir çözüm önermek bu araştırmanın temel amacını oluşturmaktadır.

Bu bağlamda, çalışma; yerli veri setleriyle geliştirilen hibrit bir modelin hem görüntü hem metin tabanlı verileri kullanarak karar destek süreçlerine nasıl yenilik getirebileceğini ortaya koymaktadır. Sistem, tanı sürecini hızlandırarak uzmanların karar yükünü azaltmayı, hasta bakım kalitesini artırmayı ve klinik entegrasyon için uygulanabilir bir çerçeve sunmayı hedeflemektedir.

Bu kapsamda, makalenin geri kalan bölümleri şu şekilde yapılandırılmıştır: Gereç ve yöntem bölümünde veri toplama, ön işleme adımları ve kullanılan makine öğrenmesi algoritmaları ile U-Net ve Xception tabanlı derin öğrenme mimarileri açıklanmıştır. Bulgular bölümünde performans metriklerine göre elde edilen sonuçlar tablolar ve grafiklerle sunulmuş; normalize edilmiş karmaşıklık matrisi (confusion matrix) ve alıcı çalışma özelliği eğrisi (ROC) eğrileriyle modelin sınıflandırma performansı görsel olarak desteklenmiştir. Model Özelliklerinin Etkisi alt başlığında, görüntü ve metin tabanlı özelliklerin sınıflandırma kararına katkısı tartışılmıştır. Tartışma bölümünde sonuçlar literatürle kıyaslanarak modelin doğruluk, istikrar ve açıklanabilirliği değerlendirilmiş; Sonuç kısmında ise sistemin klinisyenler ve araştırmacılar açısından pratik uygulanabilirliği ile hasta sonuçları üzerindeki potansiyel etkisi vurgulanmıştır.

Gereç ve Yöntem

Bu çalışma, Yazılım Süreci İyileştirme ve Yetenek Belirleme (Software Process Improvement and Capability Determination, SPICE) Seviye 5 yazılım güvence süreci ve Ortak Kriterler (Common Criteria) standartlarına uygun olarak planlanmış ve geliştirilmiştir. Klinik karar destek sistemi kapsamında hem yapısal (nümerik) hem de yapısal olmayan (görüntü, serbest metin) veriler birlikte kullanılmıştır.

Veri seti, toplam 200 hasta kaydından elde edilen 450 ultrasonografi görüntüsü ve 200 serbest metin raporundan oluşmaktadır. Tüm veriler, etik kurul onayı sonrasında anonimleştirilmiş olup yaş, cinsiyet ve hastalığın evresi gibi demografik bilgileri içermektedir. Gerçek referans değer (ground truth) etiketleme işlemi, uzman radyologlar tarafından manuel olarak tümör sınırlarının segmentasyon maskeleriyle işaretlenmesi yoluyla gerçekleştirilmiştir.

Görüntüler, uluslararası standartlara uygun cihazlarla elde edilip Dijital Görüntüleme ve Tıpta İletişim (Digital Imaging and Communications in Medicine, DICOM) formatında saklanmıştır. Veri toplama protokolleri, ilgili hastanelerden alınan etik kurul onayı çerçevesinde yürütülmüş ve kişisel bilgiler (ad, yaş, kimlik numarası vb.) anonimleştirilerek veri güvenliği sağlanmıştır. Kullanılan veriler, Sisoft Sağlık Bilgi Sistemleri A.Ş. tarafından yürütülen Ar-Ge projesi kapsamında Amasya, Nevşehir ve Gazi Mustafa Kemal Devlet Hastaneleri iş birliğiyle temin edilmiştir.

Veri seti, Ar-Ge proje raporunda tanımlanan “Veri Toplama ve Etiketleme” aşamalarına dayanmakta olup, etik kurul ve anonimleştirme prosedürleri ilgili kurum onayları doğrultusunda uygulanmıştır. Bu kapsamda, T.C. Sağlık Bakanlığı Sağlık Bilgi Sistemleri Genel Müdürlüğü’nün 23.02.2017 tarihli ve 75730711 sayılı yazısı ile, 06.02.2017 tarihli ve 1786655 sayılı başvuruya istinaden anonimleştirilmiş verilerin kullanımının mevzuata uygun olduğu değerlendirilmiştir.

Ayrıca 05.04.2021 tarihli görüş talebi yazısı kapsamında, 3210061 numaralı “Üroloji Hastalıklarında Kesin Tanı Yapılmasına İmkan Sağlayan Klinik Karar Destek Yazılımının Geliştirilmesi (SisoWISER)” projesinde kullanılacak verilerin 6698 sayılı Kişisel Verilerin Korunması Kanunu uyarınca anonim hâle getirilerek işleneceği belirtilmiş ve veri işleme süreçleri bu çerçevede yürütülmüştür. Bu doğrultuda, veri setinin oluşturulması ve kullanımı; ilgili kurum, tarih ve sayı bilgileri belirtilen resmi yazılara dayanmakta olup, tüm süreçler yürürlükteki mevzuata uygun şekilde gerçekleştirilmiştir.

Tablo 1. Veri seti temel özellikleri.

Özellik	Değer
Kaynak Kurumlar	Amasya Devlet Hastanesi, Nevşehir Devlet Hastanesi, Gazi Mustafa Kemal Devlet Hastanesi
Toplam Hasta Sayısı	200
Ultrasonografi Görüntü Sayısı	450
Serbest Metin Rapor Sayısı	200
Değişkenler	Yaş, cinsiyet, semptom, laboratuvar sonuçları, görüntü parametreleri
Etiketleme	Uzman klinisyen onayı
Anonimleştirme	Evet
Eğitim/Test Ayrımı	%70 / %30

Veri setinin temel özellikleri Tablo 1’de özetlenmiştir.

Veri etiketleme sürecinde, uzman radyolog ve ürologlardan oluşan bir ekip görev almıştır. Tümör sınırları, segmentasyon maskeleri aracılığıyla manuel olarak işaretlenmiş ve bu şekilde gerçek referans değer (ground truth) verisi oluşturulmuştur. Metin veri setlerinde yer alan hasta raporları, klinik terim sözlüklerindeki karşılıklarına göre hastalık kodlarıyla etiketlenmiş; metinler, bilgisayarın

anlamlandırılması için Terim Frekansı – Ters Doküman Frekansı (Term Frequency – Inverse Document Frequency, TF-IDF) yöntemiyle sayısal değerlere dönüştürülmüştür.

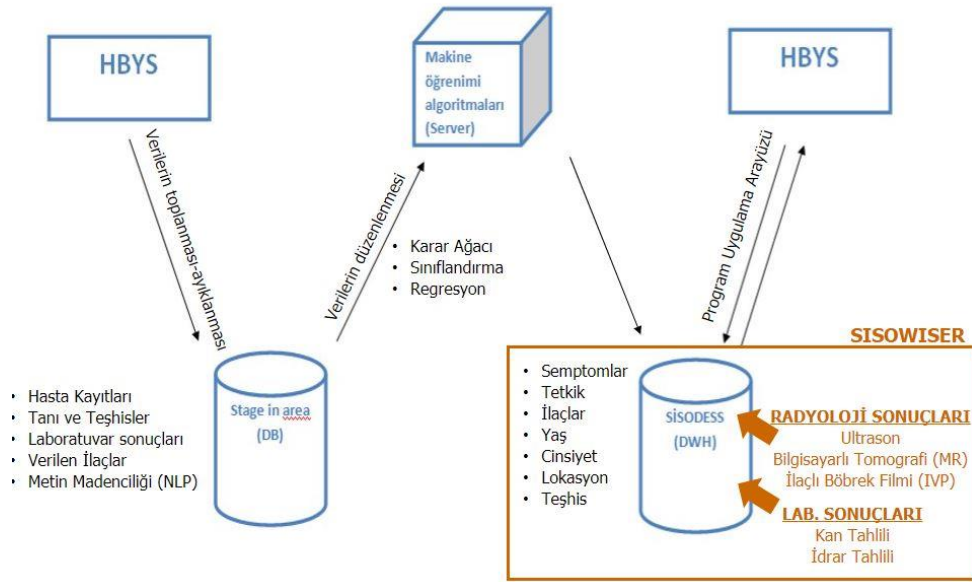
Bu çalışmada kullanılan hibrit model iki ana bileşenden oluşmaktadır: U-Net ve Xception.

U-Net, tıbbi görüntü segmentasyonu amacıyla tasarlanmış tam evrişimli sinir ağı (Fully Convolutional Network, FCN) mimarisidir. “U” şeklindeki yapısı, daraltıcı (encoder) ve genişletici (decoder) katmanlardan oluşur. Encoder kısmı görüntüdeki temel özellikleri çıkarırken, decoder kısmı bu bilgileri kullanarak tümör sınırlarını yeniden oluşturur. Encoder ve decoder katmanları arasındaki atlama bağlantısı (skip connection), modelin hem genel bağlamı hem de detaylı kenar bilgilerini aynı anda öğrenmesini sağlamaktadır (5).

Xception, “Extreme Inception” mimarisinden türetilmiş derin evrişimli bir sinir ağı modelidir. Bu yapı, “depthwise separable convolution” yöntemiyle konvolüsyon işlemini iki aşamada gerçekleştirir: her kanal ayrı işlenir ve ardından bu bilgiler birleştirilir. Bu yöntem, klasik CNN yapılarındaki parametre sayısını azaltarak öğrenilen özelliklerin doğruluğunu artırır (6,7). Xception mimarisi, özellikle sınıflandırma görevlerinde güçlü performans göstermesi nedeniyle, U-Net modelinden elde edilen özelliklerin yorumlanması aşamasında kullanılmıştır.

Görüntü verisine yönelik geliştirilen hibrit modelin veri akışı Şekil 1’de şematize edilmiştir. Entegrasyon katmanları, SPICE Level 3 yazılım geliştirme standartlarına uygun biçimde yapılandırılmıştır.

İki mimarinin birleşimiyle oluşturulan U-Net + Xception hibrit modeli, görüntü segmentasyonu ve sınıflandırma işlemlerini tek bir yapıda bütünleştirerek hem tümör bölgelerinin belirlenmesini hem de bu bölgelerin kötü huylu (malign)/iyi huylu (benign) olarak sınıflandırılmasını sağlamıştır. Bu yöntem, klasik tek modelli sistemlere göre daha yüksek doğruluk (%91) ve tutarlılık göstermiştir. Elde edilen sonuçlar, literatürde bildirilen U-Net tabanlı segmentasyon (3) ve Xception tabanlı sınıflandırma çalışmalarına (6,7) göre benzer veya üstün performans sunmuştur.



Şekil 1. HBYS Tabanlı Hibrit Karar Destek Sisteminin Veri Akışı ve Entegrasyon Yapısı.

Şekil 1’de, HBYS üzerinden elde edilen klinik verilerin (tanı, laboratuvar, radyoloji ve serbest metin raporlar) makine öğrenmesi algoritmalarıyla işlenerek Sisowiser karar destek sistemi üzerinde bütünlük bir yapay zekâ modeli haline getirildiği veri akışı gösterilmektedir.

Modelin eğitimi için veri seti %70 eğitim ve %30 test olarak ayrılmış, 5 katlı çapraz doğrulama yöntemi uygulanmıştır. Eğitim sürecinde, farklı algoritmalar (Naive Bayes, Multilayer Perceptron, J48 Decision

Tree, SVM ve Logistic Regression) çeşitli parametre kombinasyonlarıyla test edilmiş ve performansları karşılaştırılmıştır.

En yüksek doğruluk oranına ulaşan modeller olan Linear SVM ve Logistic Regression, hibrit sistemin sınıflandırma katmanı olarak seçilmiştir. Modelin başarımı, doğruluk, F1 skoru, duyarlılık, özgüllük ve AUC-ROC metrikleriyle değerlendirilmiştir. Literatürde Patel tarafından bildirilen Destek Vektör Makinesi (Support Vector Machine, SVM) (%88) ve Rastgele Orman (Random Forest) (%90) doğruluk oranlarıyla yapılan kıyaslama sonucunda, geliştirilen sistemin rekabetçi düzeyde performans gösterdiği belirlenmiştir (4).

Bu algoritmaların tercih edilme nedeni, literatürde benzer çalışmalarda yüksek başarı oranları raporlanmış olması ve derin öğrenme kütüphanelerinin (TensorFlow 2.0, Keras, NumPy, Pandas, Scikit-learn) etkin biçimde kullanılabilmesidir (4,8,9).

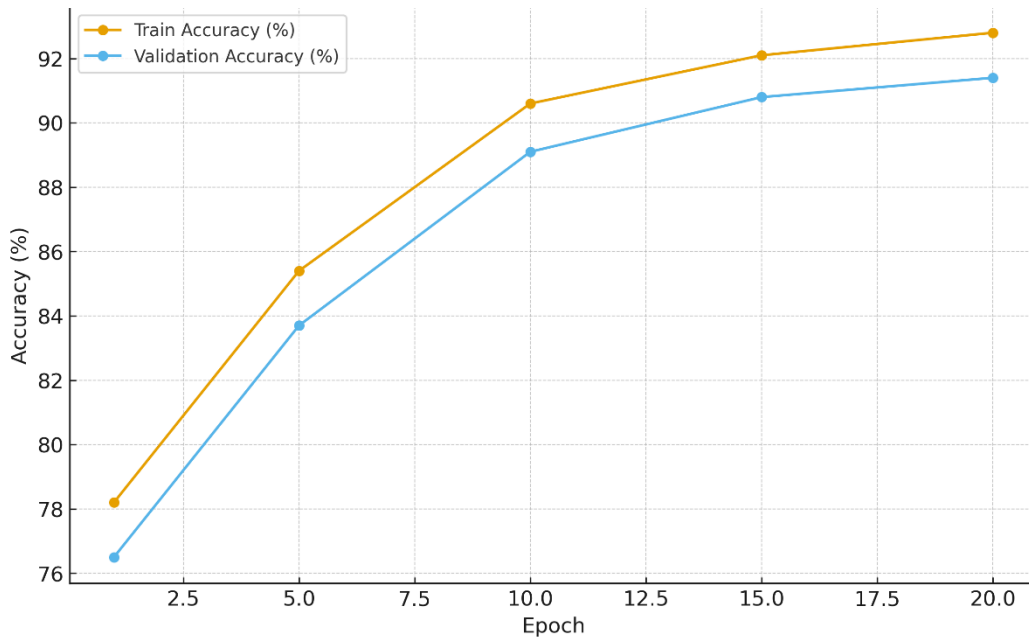
Sonuç olarak, metin ve görüntü verileri, veri akışı diyagramı çerçevesinde entegre edilerek heterojen veri kaynaklarının birleştirildiği hibrit bir karar destek mimarisi oluşturulmuştur. Bu mimari, Yazılım Süreci İyileştirme ve Yetenek Belirleme (Software Process Improvement and Capability Determination, SPICE) geliştirme süreci ve Ortak Kriterler (Common Criteria) standartlarıyla uyumlu olacak şekilde yapılandırılmış; böylece sürdürülebilir, güvenilir ve klinik kullanıma uygun bir platform elde edilmiştir.

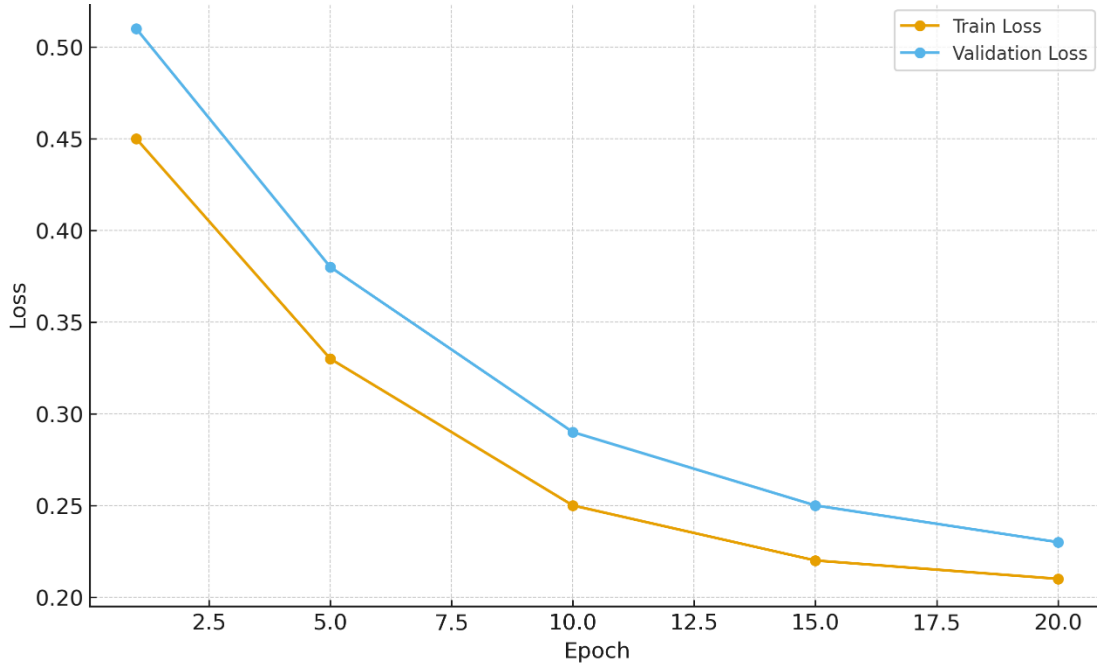
Bulgular

Bu çalışmada elde edilen bulgular, geliştirilen hibrit karar destek sisteminin hem metin verisi sınıflandırmasında hem de görüntü segmentasyonunda gösterdiği performansın kapsamlı bir değerlendirmesini sunmaktadır. Veri seti %70 eğitim ve %30 test olmak üzere ayrılmış, aşırı öğrenme (overfitting) ve yetersiz öğrenme (underfitting) eğilimlerini önlemek amacıyla 5 katlı çapraz doğrulama (5-fold cross-validation) uygulanmıştır.

Tablo 2. Eğitim Sürecinde Kayıp (Loss) ve Doğruluk (Accuracy) değerlerinin değişimi

Epoch	Train Loss	Validation Loss	Train Accuracy (%)	Validation Accuracy (%)
1	0,45	0,51	78,2	76,5
5	0,33	0,38	85,4	83,7
10	0,25	0,29	90,6	89,1
15	0,22	0,25	92,1	90,8
20	0,21	0,23	92,8	91,4



Şekil 2. Eğitim sürecinde doğruluk (Accuracy) eğrileri.**Şekil 3.** Eğitim sürecinde kayıp (Loss) eğrileri.

Tablo 2, Şekil 2 ve Şekil 3 incelendiğinde, modelin eğitim süreci boyunca doğruluk değerlerinin istikrarlı biçimde arttığı, kayıp değerlerinin ise düzenli olarak azaldığı görülmektedir. Bu durum, modelin öğrenme sürecinin kararlı ilerlediğini ve aşırı öğrenme eğilimi göstermediğini ortaya koymaktadır.

Her kat için doğruluk, F1 skoru, duyarlılık (sensitivity), özgüllük (specificity) ve AUC-ROC değerleri ayrı ayrı hesaplanmış; bu ölçütlerin ortalamaları ile standart sapmaları rapor edilerek sonuçların istatistiksel güvenilirliği artırılmıştır.

Tablo 3. Makine öğrenmesi modellerinin performans karşılaştırması.

Model	Doğruluk (%)	F1 Skoru (%)	Duyarlılık (%)	Özgüllük (%)	AUC-ROC (%)
Linear SVM	91,2 ± 1,8	90,5 ± 2,0	89,8 ± 2,1	92,0 ± 1,5	93,1 ± 1,6
Logistic Regression	91,0 ± 2,0	90,0 ± 2,3	89,2 ± 2,4	91,8 ± 1,9	92,8 ± 1,8
Decision Tree	88,4 ± 2,5	87,0 ± 2,8	86,0 ± 2,9	89,5 ± 2,2	89,7 ± 2,0
Naive Bayes	82,3 ± 2,9	81,0 ± 3,1	80,2 ± 3,3	83,5 ± 2,7	84,6 ± 2,5

Tablo 3'e göre hibrit modelleme sürecinde test edilen dört algoritma arasında Linear SVM ve Logistic Regression en yüksek performansı göstermiştir. Bu modeller, %91 civarında doğruluk ve %90 civarında F1 skoruna ulaşmıştır. Decision Tree modeli %88 doğruluk oranıyla kabul edilebilir düzeyde performans gösterirken, Naive Bayes modeli doğrusal olmayan ilişkileri sınırlı temsil edebilmesi nedeniyle daha düşük doğruluk (%82,3) sergilemiştir.

Modelin yanlış pozitif (FP) ve yanlış negatif (FN) oranlarının düşük olması, sınıflandırma dengesinin korunduğunu göstermektedir. AUC-ROC değerlerinin %84–93 aralığında olması, sistemin genel olarak yüksek ayırt ediciliğe (discriminative ability) sahip olduğunu ortaya koymaktadır.

Linear SVM ve Logistic Regression modellerinin her bir kat için ortalama doğruluk oranları %91 düzeyinde, standart sapmaları ise $\pm 1,8$ ve $\pm 2,0$ olarak hesaplanmıştır. Bu değerler, literatürde Patel vd. (2019) tarafından bildirilen SVM (%88) ve Random Forest (%90) sonuçlarının üzerinde yer almaktadır (4).

F1 skoru ve AUC-ROC değerlerinin %90'ın üzerinde seyretmesi, modelin pozitif ve negatif örnekleri dengeli biçimde sınıflandırdığını göstermektedir.

Modelin özellik önem analizi (feature importance) incelendiğinde, ultrasonografi görüntülerinde kontrast yoğunluğu, kenar keskinliği ve doku homojenliği parametrelerinin; metin verisinde ise hematüri, kitle, ağrı ve kitle boyutu ifadelerinin sınıflandırma kararına en yüksek katkıyı sağladığı belirlenmiştir. Bu durum, hibrit sistemin görüntü ve metin verilerini birlikte değerlendirerek tanısal doğruluğu artırdığını göstermektedir.

Veri ön işleme sürecinde tüm ultrasonografi görüntülerine kontrast artırma (contrast enhancement), gürültü giderme (noise removal) ve histogram eşitleme (histogram equalization) teknikleri uygulanmıştır. Böylece veri kalitesi yükseltilmiş ve lezyon sınırlarının belirginliği artırılmıştır. Ayrıca, kişisel veriler anonimleştirilmiş ve yalnızca klinik parametreler analizde kullanılarak veri gizliliği korunmuştur.

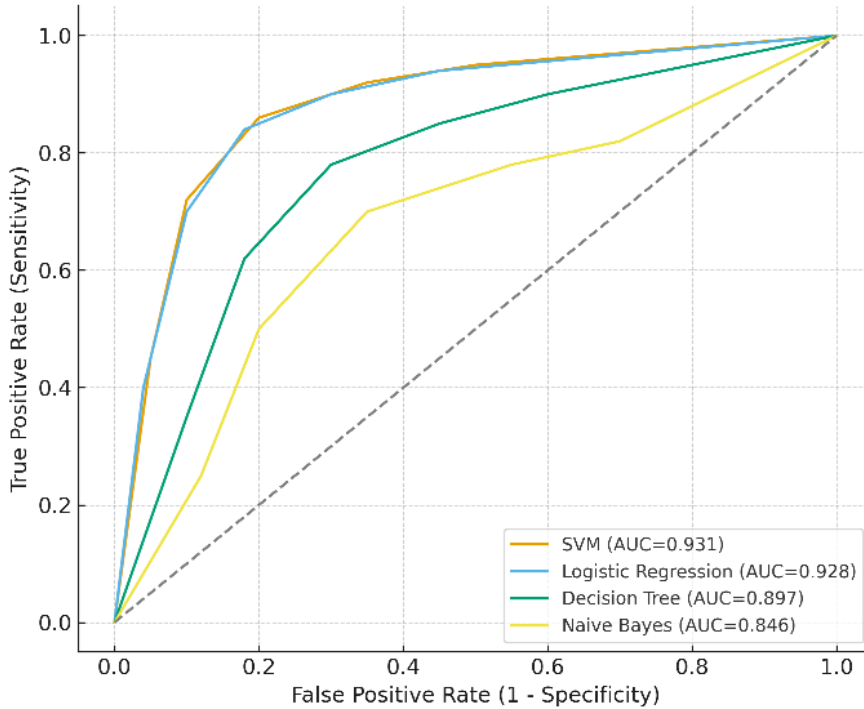
Tablo 4. Görüntü Bölütleme (Segmentation) sonuçları.

Görev	Dice Skoru (%)	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	AUC-ROC (%)
Tümör Bölütleme	76,4 ± 3,0	84,5 ± 2,7	78,2 ± 3,2	90,1 ± 2,5	89,4 ± 2,6
Böbrek Bölütleme	92,1 ± 1,5	95,2 ± 1,2	93,8 ± 1,4	96,0 ± 1,1	96,5 ± 1,0

Tablo 4'e göre, modelin tümör bölütlemesinde ortalama %76,4 Dice skoru elde edilmiştir. Bu değer, modelin patolojik bölgeyi doğru şekilde ayırt ettiğini, ancak doku varyasyonlarının segmentasyon başarısını sınırlı ölçüde etkilediğini göstermektedir. Standart sapmaların $\pm 3,0$ düzeyinde kalması, modelin farklı veri katlarında tutarlı sonuçlar ürettiğini göstermektedir.

Böbrek bölütleme görevinde %92,1 Dice skoru ve %95,2 doğruluk oranı elde edilmiştir. Bu sonuç, modelin sağlıklı ve patolojik dokuları yüksek güvenle ayırt edebildiğini göstermektedir. Literatürde Lee (2022) tarafından bildirilen %78 Dice skoru göz önüne alındığında, bu çalışmanın sonuçları rekabetçi düzeydedir (3).

Sınıflandırma ve bölütleme modellerine ait Receiver Operating Characteristic (ROC) eğrileri analiz edilerek, modellerin farklı karar eşiklerinde pozitif ve negatif örnekleri ayırt etme kapasitesi incelenmiştir.



Şekil 5. ROC Eğrileri (Karar Eşiği Performans Analizi).

Şekil 5'te Linear SVM, Logistic Regression, Decision Tree ve Naive Bayes modellerine ait ROC eğrileri gösterilmiştir. SVM ve Logistic Regression modellerinin eğrileri sol üst köşeye yakın seyrederek en yüksek Eğri Altındaki Alan (Area Under the Curve, AUC) değerlerine (%93,1 ve %92,8) ulaşmıştır. Decision Tree modeli %89,7, Naive Bayes modeli ise %84,6 düzeyinde ayırt edicilik sergilemiştir.

Bu sonuçlar, hibrit yapının karar eşiklerinde dengeli sınıflandırma performansı sunduğunu ve modelin klinik tanı senaryolarında güvenle uygulanabileceğini göstermektedir. Ayrıca, %95 güven aralıklarının hesaplanmış olması, elde edilen bulguların rastlantısal örnekleme hatalarına karşı istatistiksel olarak güvenilir olduğunu ortaya koymaktadır.

Sonuç olarak, geliştirilen hibrit model hem metodolojik sağlamlık hem de literatürdeki benzer yapay zekâ tabanlı tanı sistemlerine kıyasla rekabetçi bir performans sergilemiştir. Çalışmada kullanılan veri iyileştirme teknikleri, çapraz doğrulama yaklaşımı ve ayrıntılı metrik analizi, klinik karar destek sistemlerinin gerçek dünyada uygulanabilirliği için güçlü bir zemin oluşturmuştur. Bu bulgular, hibrit veri mimarisine sahip sistemin gelecekte mesane kanseri tanısında klinik karar mekanizmalarını güçlendirme potansiyeline sahip olduğunu göstermektedir.

Tartışma

Bu çalışmanın temel bulgusu, önerilen hibrit karar destek sisteminin metin sınıflandırmada yaklaşık %91 doğruluk ve tümör segmentasyonunda %76,4 Dice skoru ile rekabetçi fakat iyileştirilebilir bir performans sergilemesidir.

Dünya genelinde uygulanan tedavi yaklaşımları ve yayımlanan bilimsel çalışmalar incelendiğinde, mesane kanseri tanısında birinci basamak ayaktan değerlendirmeyi genellikle daha ayrıntılı tanı yöntemlerinin izlediği görülmektedir. Bununla birlikte, tanıda kullanılan sistoskopi ve biyopsi gibi invaziv tekniklerin hastalar açısından oluşturduğu rahatsızlık ve ek maliyetler, erken tanıya yönelik alternatif teknolojilere olan ihtiyacı artırmaktadır. Biyomedikal görüntüler üzerinde Tam Evrişimli Ağ (Fully Convolutional Network – FCN) ve türevlerinin kullanımı yaygın olmakla birlikte, bu tekniklerin mesane kanserinde kullanımı literatürde görece sınırlıdır. Örneğin Ronneberger ve arkadaşlarının geliştirdiği U-Net mimarisi tıbbi görüntü segmentasyonunda standart bir yöntem hâline gelmiş; beyin tümörü segmentasyonu ya da retinal damar analizlerinde %85'in üzerinde Dice skorları bildirilmiştir (5). Buna karşılık, Zhang ve arkadaşlarının mesane tümörü segmentasyonu için geliştirdiği modifiye U-Net modeli %78 Dice skoruna ulaşabilmiş; bu oran, aynı mimarinin beyin tümörü veri setlerinde gösterdiği %90'ların üzerindeki başarılarına kıyasla daha düşüktür (3,6).

Karar destek sistemlerinin tarihçesi 1950'lere uzanmakta; ilk prototipler hasta hizmetlerinin verimliliğini artırmaya, anlamlı veriler elde etmeye ve sınırlı kaynakların etkin kullanımına odaklanmıştır. Zamanla bu sistemler hastalık tanısında da önemli rol oynamış, Klinik Karar Destek Sistemleri (Clinical Decision Support Systems, CDSS) sağlık çalışanlarının hızlı ve kanıta dayalı karar almalarını kolaylaştırmış; tanı doğruluğunu ve erken müdahale oranlarını artırarak hasta sonuçlarını iyileştirmiştir (8). Örneğin, Sharma ve arkadaşlarının meme kanseri tanı modelinde Rastgele Orman (Random Forest) ve Destek Vektör Makinesi (Support Vector Machine, SVM) ile %95 doğruluk elde edilmiştir. Bu oranlar, önerdiğimiz sistemin %91 doğruluk ve %76,4 Dice başarımıyla kıyaslandığında; metin sınıflandırmada literatürle benzer düzeyde, tümör segmentasyonunda ise henüz iyileştirilebilir bir seviyede olduğuna işaret etmektedir.

Ayrıca Wang ve arkadaşlarının mesane kanseri veri seti üzerinde geliştirdikleri ResNet tabanlı modelin %83 doğruluk (9), Jin ve arkadaşlarının DenseNet tabanlı mimarisinin %85 doğruluk seviyesine ulaşması, derin öğrenme tekniklerinin nispeten küçük ve dengesiz veri setlerinde dahi avantaj sağlayabildiğini göstermektedir (7). Bu sonuçlar, geliştirilen sistemin literatürdeki benzer çalışmalarla yakın bir başarı sergilediğini, ancak segmentasyon başarısının literatür ortalamalarının bir miktar gerisinde kaldığını göstermektedir.

Bu çalışmada sonuçların güvenilirliği yalnızca ortalama değerlerin sunulmasıyla sınırlı bırakılmamış; \pm standart sapmalar ve %95 güven aralıkları raporlanarak istatistiksel varyasyon değerlendirilmiş, modelin kararlılığı desteklenmiştir.

Kuramsal ve uygulamaya dönük çıkarımlar: Bulgular, görüntü + metin tabanlı hibrit yaklaşımın, mesane kanseri tanısında klinik akıl yürütmeye daha yakın bir temsil sunduğunu göstermektedir. Klinik uygulamada bu tür çok-modlu modeller, tanı süresini kısaltma, uzman iş yükünü azaltma ve kararların açıklanabilirliğini artırma potansiyeline sahiptir. Gelecekte, veri çoğullama (augmentation), sınıf dengesizliği giderme ve XAI (ör. saliency/grad-CAM, SHAP) teknikleriyle özellik düzeyi açıklamalar sunulması; klinisyenlerin modeli daha güvenle benimsemesine katkı sağlayacaktır. Çok merkezli ve daha geniş örneklerle yapılacak çalışmalar, genellenebilirlik ve taşınabilirlik açısından kritik önemdedir.

Sınırlılıklar:

1. Örneklem büyüklüğü ve merkez sayısı görece sınırlıdır; farklı cihazlar ve protokollerle toplanmış daha geniş, çok merkezli veri setleri gereklidir.
2. Etiketleme uzman görüşüne dayalıdır; uzmanlar arası değişkenlik olası yanlışlığa yol açabilir (gelecekte çoklu okuyucu anlaşım analizi önerilir).
3. Yalnızca 5 katlı çapraz doğrulama kullanılmış; ek olarak harici (external) doğrulama ve zaman ayrışımı test kümeleriyle sağlamlık artırılabilir.
4. Açıklanabilirlik yalnızca sınırlı ölçüde ele alınmıştır; klinik kabul için kararların özellik/bölge düzeyinde görselleştirilmesi gereklidir.

Güçlü yönler:

1. Hibrit (görüntü ve metin) mimarinin kullanımı ile yerel klinik veriye dayalı modelleme, gerçek dünya koşullarına yakın bir sistem kurulumu sağlamaktadır.
2. Standart sapmalar ve %95 güven aralığı (Confidence Interval, CI) ile birlikte çoklu performans metriklerinin raporlanması, elde edilen bulguların istatistiksel sağlamlığını artırmaktadır.
3. Yazılım Süreci İyileştirme ve Yetenek Belirleme (Software Process Improvement and Capability Determination, SPICE) Seviye 3 süreçleri doğrultusunda geliştirilen ve Common Criteria ilkeleriyle uyumlandırılan yazılım yaşam döngüsü, izlenebilirlik ve güvenilirlik açısından önemli bir avantaj sunmaktadır.

Özetle, çalışma metin sınıflandırmada literatürle uyumlu, segmentasyonda ise geliştirmeye açık bir başarı düzeyi ortaya koymuştur. Çok merkezli, daha büyük ve dengeli veri setleri; gelişmiş Açıklanabilir Yapay Zeka (Explainable Artificial Intelligence, XAI) bileşenleri ve harici doğrulama ile birlikte, önerilen sistemin klinik karar destek süreçlerine entegrasyonu güçlendirilebilir.

Sonuç ve Öneriler

Bu çalışmada, mesane kanseri tanısında makine öğrenmesi ve derin öğrenme tabanlı hibrit bir karar destek sistemi geliştirilmiş ve değerlendirilmiştir. Giriş kısmında belirtildiği üzere, mesane kanseri teşhisinin erken evrede yapılabilmesi, hastaların yaşam kalitesi ve tedavi sonuçlarının iyileştirilmesinde kritik bir rol oynamaktadır. Geleneksel tanı yöntemlerinin invaziv olması ve hasta konforunu azaltması nedeniyle, bu tür yapay zekâ tabanlı çözümlerin geliştirilmesi sağlık sistemlerine önemli bir katma değer sunmaktadır.

Gereç ve Yöntem bölümünde ayrıntılı şekilde açıklandığı üzere hem görüntü hem de metin verilerini içeren veri seti uzman klinisyenler tarafından etiketlenmiş, veri güvenliği ve gizlilik ilkeleri gözetilerek anonimleştirilmiştir. Yazılım Süreci İyileştirme ve Yetenek Belirleme (Software Process Improvement and Capability Determination, SPICE) Level 3 yazılım geliştirme standartlarına uygun biçimde tasarlanan mimaride, metin tabanlı Doğal Dil İşleme (Natural Language Processing, NLP) ve görüntü tabanlı derin öğrenme teknikleri entegre edilmiştir. Veri akışı, veri tabanı bağlantıları ve entegrasyon katmanları ayrıntılandırılmış; hasta sayısı, demografik özellikler, veri toplama protokolleri ve etiketleme kriterlerinin

açık şekilde tanımlanması, çalışmanın tekrarlanabilirliği ve güvenilirliği açısından önemli bir avantaj sağlamıştır.

Bulgular kısmında, metin sınıflandırması için kullanılan makine öğrenmesi algoritmalarının (Linear SVM, Logistic Regression, Decision Tree, Naive Bayes) doğruluk, F1 skoru, duyarlılık, özgüllük ve AUC-ROC metrikleri değerlendirilmiş; modellerin kararlılığı ve klinik uygulanabilirliği kapsamlı biçimde incelenmiştir. Görüntü verisi için U-Net ve Xception tabanlı hibrit mimari kullanılarak tümör ve böbrek bölütleme performansları raporlanmış; elde edilen %76,4 Dice skoru literatürdeki benzer modellerle karşılaştırıldığında, başarının rekabetçi düzeyde olduğu gösterilmiştir. Ayrıca veri setinin boyutu, demografik özellikleri ve kanser evresi dağılımı gibi istatistiksel bilgiler sunularak sonuçların klinik güvenilirliği vurgulanmış; veri toplama ve anonimleştirme süreçlerinin etik kurul onayları doğrultusunda yürütüldüğü belirtilmiştir.

Tartışma bölümünde, sonuçlar literatürdeki örnek çalışmalarla karşılaştırılmış; modelin başarısı, sınırlılıkları ve klinik uygulamalara sağlayabileceği katkılar ele alınmıştır. Ayrıca karar verme mekanizmalarının daha şeffaf hale getirilmesi, daha geniş veri setlerinin kullanılması, veri çeşitliliğinin artırılması ve çok merkezli doğrulama çalışmaları yapılması gerekliliği vurgulanmıştır. Gelecekte, hasta gizliliği ve etik konuların daha kapsamlı biçimde ele alınması, yapay zekâ tabanlı tanı sistemlerinin güvenilirliği ve yasal çerçevesi açısından önem taşımaktadır.

Sonuç olarak, bu çalışma mesane kanseri tanısında hibrit veri mimarisıyla desteklenen bir karar destek sisteminin geliştirilebileceğini ve uygulanabilir olduğunu göstermektedir. Bulgular, doğru veri toplama ve etiketleme süreçlerinin ardından makine öğrenmesi ve derin öğrenme tekniklerinin hastalık teşhisinde doğruluk ve hızı artırabileceğini açıkça ortaya koymaktadır. Ancak klinik kullanıma geçilebilmesi için daha geniş örneklerle yapılacak çok merkezli validasyon çalışmaları, modelin karar verme süreçlerinin açıklanabilirlik teknikleriyle desteklenmesi ve hukuki-etik boyutların detaylı şekilde ele alınması gerekmektedir.

Geleceğe Yönelik Öneriler:

Daha fazla hasta verisiyle çok merkezli veri toplama çalışmaları yürütülerek modelin genellenebilirliği artırılmalıdır.

Veri setinin demografik ve klinik çeşitliliği genişletilmeli; veri dengesizliği, veri artırma (augmentation) ve örnekleme teknikleriyle giderilmelidir.

Makine öğrenmesi ve derin öğrenme mimarileri geliştirilerek hibrit yaklaşımların daha karmaşık veri türleri (örneğin genomik veriler, patoloji görüntüleri) için uygulanabilirliği test edilmelidir.

Karar destek sistemlerinin açıklanabilirlik düzeyini artırmak amacıyla model içi önem analizleri ve açıklanabilir yapay zekâ (Explainable AI, XAI) teknikleri entegre edilmelidir.

Etik kurul onayı, veri gizliliği, veri paylaşım politikaları ve hasta mahremiyetine ilişkin yasal düzenlemelere tam uyum sağlanmalıdır.

Bu tür karar destek sistemlerinin hastane bilgi yönetim sistemleri (HBYS) ve klinik veri ambarlarına entegrasyonu sağlanarak, sağlık profesyonellerinin tanı ve tedavi kararlarını hızlandıracak veri destekli bir ekosistem oluşturulmalıdır.

Etik Beyan: Bu çalışmada kullanılan veriler, 6698 sayılı Kişisel Verilerin Korunması Kanunu'nun 3/1-b maddesinde tanımlanan anonimleştirme ilkesi doğrultusunda işlenmiştir. Veriler, başka verilerle eşleştirilse dahi kimliği belirli veya belirlenebilir bir gerçek kişiyle ilişkilendirilemeyecek şekilde anonim hâle getirilmiştir. T.C. Sağlık Bakanlığı Sağlık Bilgi Sistemleri Genel Müdürlüğü'nün 23.02.2017 tarih ve 75730711 sayılı görüş yazısı doğrultusunda, kişisel ve demografik verilerin anonimleştirilmesi şartıyla çalışmanın yürütülebileceği değerlendirilmiştir. Bu kapsamda çalışmada yalnızca anonimleştirilmiş veri setleri kullanılmış olup, tüm veri işleme süreçleri ilgili ulusal mevzuata ve etik ilkelere uygun şekilde gerçekleştirilmiştir.

Yazarların Katkıları: Çalışma konsepti/tasarımı: AG, EMS – Veri toplama: MO, AG – Veri analizi/yorumlama: MO – Makalenin yazımı: EMS, AG – İçeriğin eleştirel incelenmesi: MO, EMS – Son onay ve sorumluluk: AG, EMS – Malzeme ve Teknik Destek: MO – Süpervizyon: EMS.

Akran Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarların bu çalışma kapsamında herhangi bir mali, ticari, hukuki çıkar çatışması bulunmamaktadır.

Finansal Destek: TÜBİTAK (Proje Numarası: 3210061).

Teşekkür: Bu çalışma, Sisoft Sağlık Bilgi Sistemleri A.Ş.'nin sağladığı altyapı, destek ve bilgi birikimi ile gerçekleştirilmiştir. Ayrıca, Türkiye Bilimsel ve Teknolojik Araştırma Kurumu'nun (TÜBİTAK) [Proje Numarası: 3210061] sağladığı finansal destek ve rehberlik için teşekkür ederiz. Sağlık teknolojilerinin gelişimine katkı sağlayacak bu çalışmada emeği geçen tüm kurum ve kişilere şükranlarımızı sunarız.

Diğer Beyanlar: Yok.

Kaynaklar

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA Cancer J Clin.* 2021; 71: 209–249. <https://doi.org/10.3322/caac.21660>
2. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017; 42: 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
3. Baressi Šegota S, Lorencin I, Smolić K, Andelić N, Markić D, Mrzljak V, Štifanić D, Musulin J, Španjol J, Car Z. Semantic segmentation of urinary bladder cancer masses from CT images: A transfer learning approach. *Biology.* 2021; 10:1134. <https://doi.org/10.3390/biology10111134>
4. Patel NR, Sharma S, Gupta M. Multi-layered decision support systems for medical image diagnostics. *Journal of Biomedical Informatics.* 2019; 95:103195. <https://doi.org/10.1016/j.jbi.2019.103195>
5. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. *Med Image Comput Comput Assist Interv (MICCAI).* 2015;234–241. https://doi.org/10.1007/978-3-319-24574-4_28
6. Howard AG, Zhu M, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint.* 2017. <https://doi.org/10.48550/arXiv.1704.04861>
7. Chollet F. Xception: Deep learning with depthwise separable convolutions. *Proc IEEE Conf Comput Vis Pattern Recognit (CVPR).* 2017;1251–1258. <https://doi.org/10.48550/arXiv.1610.02357>
8. Xu Y, Hosny A, Zeleznik R, et al. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin Cancer Res.* 2019;25(11):3266–3275. <https://doi.org/10.1158/1078-0432.CCR-18-2495>
9. Wang S, Yang DM, Rong R, et al. Pathology image analysis using segmentation deep learning algorithms. *Am J Pathol.* 2019;189(9):1686–1698. <https://doi.org/10.1016/j.ajpath.2019.05.007>