BMBAD BJESR Mühendislik Bilimleri ve Araştırmaları Dergisi
Journal of Engineering Sciences and Researches

Patoloji Görüntülerinde Akut Böbrek Hasarı Evresinin Otomatik Belirlenmesi için Derin Öğrenme Tabanlı Yeni Bir Yaklaşım

A Novel Deep Learning-Based Approach for Automatic Determination of Acute Kidney Injury Stage in Pathology Images

[1]Fatma Tuana DOĞU, [2]Mehmet ÖZTÜRK, [3]Hülya DOĞAN

[1]Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik – Elektronik Mühendisliği Anabilim Dalı, Trabzon, Türkiye
[2]Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik – Elektronik Mühendisliği Bölümü, Trabzon, Türkiye
[3]Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Yazılım Mühendisliği Bölümü, Trabzon, Türkiye

[1]448511@ogr.ktu.edu.tr, [2]mehmetozturk@ktu.edu.tr, [3]hulya@ktu.edu.tr

Araştırma Makalesi/Research Article

ARTICLE INFO

Article history

Received : 09 March 2026
Accepted : 16 June 2026

Keywords:

Deep Learning, Pathology Images, Acute Kidney Injury.

ABSTRACT

Kidney diseases are a critical health problem that is widespread globally and difficult to diagnose early. The pathological examination process of kidney biopsies is complicated by the fact that it depends on the experience of the specialist and is time-consuming. Various image analysis approaches exist in the literature, but there is a gap in the field regarding the automated determination of acute kidney injury (AKI) stages using deep learning. This study presents a deep learning approach that can automatically determine AKI stages from pathology images. The proposed deep learning-based approach combines multi-scale feature extraction and attention mechanisms, achieving an accuracy of 0.7668 and a macro F1-score of 0.7252 according to experimental results. The results obtained demonstrate that the proposed approach offers high performance compared to studies in the literature and has the potential for standardization in clinical decision support processes.

© 2026 Bandirma Onyedi Eylul University, Faculty of Engineering and Natural Science. Published by Dergi Park. All rights reserved.

MAKALE BİLGİSİ

Makale Tarihleri

Gönderim : 09 Mart 2026
Kabul : 16 Haziran 2026

Anahtar Kelimeler:

Derin Öğrenme, Patoloji Görüntüleri, Akut Böbrek Hasarı

ÖZET

Böbrek hastalıkları, dünya genelinde yaygın görülen ve erken teşhisi zor olan kritik bir sağlık sorunudur. Böbrek biyopsilerinin patolojik inceleme süreci, uzmanın deneyimine bağlı olması ve zaman alması nedeniyle tanı sürecini zorlaştırmaktadır. Literatürde çeşitli görüntü analizi yaklaşımları bulunmaktadır ancak derin öğrenme ile akut böbrek hasarı (ABH) evrelerinin otomatik belirlenmesine yönelik bir boşluk mevcuttur. Bu çalışma, ABH evrelerini patoloji görüntülerinden otomatik olarak belirleyebilen bir derin öğrenme yaklaşımı sunmaktadır. Önerilen derin öğrenme tabanlı yaklaşım, çok ölçekli özellik çıkarımı ve dikkat mekanizmalarını bir araya getirerek deneysel sonuçlara göre 0.7668 doğruluk ve 0.7252 makro F1-skoru elde etmektedir. Elde edilen sonuçlar, önerilen yaklaşımın literatürdeki çalışmalara kıyasla yüksek bir performans sunduğunu ve klinik karar destek süreçlerinde standardizasyon potansiyeli taşıdığını göstermektedir.

© 2026 Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi. Dergi Park tarafından yayınlanmaktadır. Tüm Hakları Saklıdır.

ORCID ID: [1]0009-0008-0916-9394
[2]0000-0002-8535-3088
[3]0000-0003-3695-8539

1. GİRİŞ

Böbrek hastalıkları, dünya genelinde yaygın olarak görülen ve yaşam kalitesini büyük ölçüde etkileyen bir sağlık problemidir. Dünya Sağlık Örgütü (WHO) 2021 verilerine göre, böbrek hastalıkları dünya çapında ölüm nedenleri arasında 9. sırada yer almaktadır [1]. Bu durum böbrek hastalıklarının erken evrede tespitini ve uygun tedavi planlamasını zorunlu hale getirmektedir. Böbrek hastalıklarının erken evrede belirti vermemesi geç tanıya ve böbrekte geri dönülemez hasarların oluşmasına neden olmaktadır. Ek olarak böbrek biyopsisi üzerinde yapılan patolojik inceleme süreci böbreğin sadece belirli bir kısmını temsil ettiği, gözlemcinin deneyimine bağlı olduğu ve uzun zaman aldığı için tanı süreci zorlaşmaktadır. Derin öğrenme alanındaki gelişmeler, patolojik incelemelerin hızlı, objektif ve yüksek doğrulukta analiz edilmesini sağlamaktadır. Literatürde böbrek patoloji görüntüleri üzerinde yapılan birçok otomatik tespit, segmentasyon ve sınıflandırma çalışması bulunmaktadır. Çalışmaların detayları aşağıdaki gibi özetlenebilmektedir:

Hastalık ve lezyon sınıflandırması üzerine yapılan çalışmalar, böbrek hastalıklarının ve lezyonlarının türünün otomatik olarak belirlenmesinde yüksek başarı göstermektedir. Hacking ve Bijol [2], elektron mikroskobu görüntülerinde sekiz farklı böbrek hastalığını ve sağlıklı böbrekleri transfer öğrenme ile %84.1 kesinlikte sınıflandırmaktadır. Chagas vd. [3], evrişimli sinir ağı ve destek vektör makinesi kullanarak önerdikleri yaklaşımla glomerüler hiperhücresellik türünün belirlenmesinde %82 doğruluk elde etmektedir. Zheng vd. [4], SSL-CLAM adındaki zayıf denetimli derin öğrenme modeli ve uzman patolog ile insan-makine füzyonu yöntemini kullanarak berrak hücreli böbrek hücreli karsinomun Fuhrman derecelendirmesini %85.6 doğruluk ile gerçekleştirmektedir. Bu çalışmalar, farklı veri türleri ve modeller kullanılmasına rağmen patoloji görüntülerinde otomatik sınıflandırmanın yüksek doğrulukla gerçekleştirilebileceğini göstermektedir.

Patolojik yapıların tam ve nicel olarak belirlenmesi, derin öğrenmenin başarılı olduğu alanlardandır. Shubham vd. [5], U-Net ve EfficientNetB4 ile segmentasyon yaparak insan böbrek dokusu görüntülerindeki glomerülleri %99.68 doğrulukta otomatik olarak tanımlamaktadır. Bouteldja vd. [6], farklı böbrek hastalığı modellerinden elde edilen görüntülerin çok sınıflı histolojik analizi için evrişimli sinir ağlarını kullanmakta ve farklı kategorilerde tespit aşamasında %67.6 ile %100 aralığında ortalama duyarlılık (AP), segmentasyon aşamasında ise %73.5 ile %98.8 aralığında ortalama nesne düzeyinde Dice katsayısı elde etmektedir. Nguyen vd. [7], akut renal tübüler hasarın histopatolojisini değerlendirmek için DeepLabV3 derin öğrenme modelini kullanmakta ve dört farklı yapının segmentasyonunda %91.78 Dice katsayısı elde etmektedir. Bu sonuçlar, derin öğrenmenin histopatolojik yapıların yüksek doğrulukta segmentasyonunda etkili olduğunu göstermektedir.

Literatürde böbrek biyopsisindeki belirli yapısal hasarları ve lezyonları hedefleyen çalışmalar bulunmaktadır. Zheng vd. [8], böbrek biyopsisinde yerel ve global özellikleri birleştiren özellik piramit ağı yaklaşımıyla interstisyel fibrozis şiddetini %71.8 doğrulukta tahmin etmektedir. Marsh vd. [9], donör böbrek biyopsilerindeki yüzde global glomerüloskleroz miktarını otomatik olarak sayısallaştırmak için derin bir sinir ağı önermekte ve gereksiz organ atılma olasılığını %37 oranında azaltmaktadır. Juang vd. [10], glomerül türünü otomatik olarak tespit ve sınıflandırmak için evrişimli sinir ağı tabanlı bir yaklaşım kullanmakta ve glomerül tespitinde %95.24 F1-skoru ve sınıflandırmada %76.23 ortalama duyarlılık elde etmektedir.

Literatür ABH tespiti konusunda daha çok klinik veriler ve biyobelirteçler üzerinde yoğunlaşmaktadır. Tomašev vd. [11], elektronik sağlık kayıtlarındaki verilerini kullanarak ABH'nin kötüleşmesini tahmin eden yinelemeli sinir ağları tabanlı bir yaklaşım geliştirmekte ve %92.1 AUROC değeriyle tahmin etmektedir. Yue vd. [12], sepsis hastalarında ABH'nin erken tahmini için XGBoost yaklaşımıyla %83.2 doğruluk değerini elde etmektedir. He vd. [13], sepsis ile ilişkili ABH hastalarında ABH'nin 7-90 gün sürmesi olarak tanımlanan akut böbrek hastalığı gelişimini öngörmek için RNN-LSTM modelini kullanmakta ve AUROC 1.00 değeri elde etmektedir.

ABH klinik olarak Kidney Disease Improving Global Outcomes (KDIGO) kriterlerine göre evrelendirilmekte olup bu kriterler biyokimyasal parametrelere dayanmaktadır. Yukarıda özetlenen biyobelirteç ve klinik veri tabanlı çalışmalar ABH riskini ve kötüleşmesini öngörmede değerli bilgiler sunsa da böbrek dokusunda meydana gelen hücresel hasarın düzeyini doğrudan yansıtmamaktadır. Mevcut çalışmaların çoğu klinik verilere, laboratuvar bulgularına veya elektronik sağlık kayıtlarına odaklanmaktadır ve böbrek biyopsi görüntülerine dayalı otomatik ABH evrelemesini doğrudan ele alan bir çalışma literatürde mevcut değildir. Bu çalışma, böbrek biyopsi görüntülerinden ABH evresini otomatik olarak belirleyebilen yeni bir derin öğrenme yaklaşımı önermektedir. Önerilen yöntem, tanı sürecini hızlandırmayı, değerlendirmeyi standardize etmeyi ve klinik karar destek sistemlerine katkı sağlamayı hedeflemektedir.

Çalışmanın akışı şu şekildedir: Bölüm 2'de akut böbrek hasarı evresinin patoloji görüntülerinde otomatik belirlenmesi için kullanılan veri seti ve önerilen yaklaşım bulunmaktadır. Bölüm 3'te elde edilen bulgular ve tartışma yer almaktadır. Son bölüm olan Bölüm 4'te sonuçlar ve gelecekteki çalışmalar için öneriler yer almaktadır.

2. MATERYAL VE METOT

Bu çalışma etiketli patoloji görüntüleri ile veri setinin oluşturulması ve derin öğrenme tabanlı evre sınıflandırma yaklaşımının geliştirilmesi olmak üzere iki kısımdan oluşmaktadır.

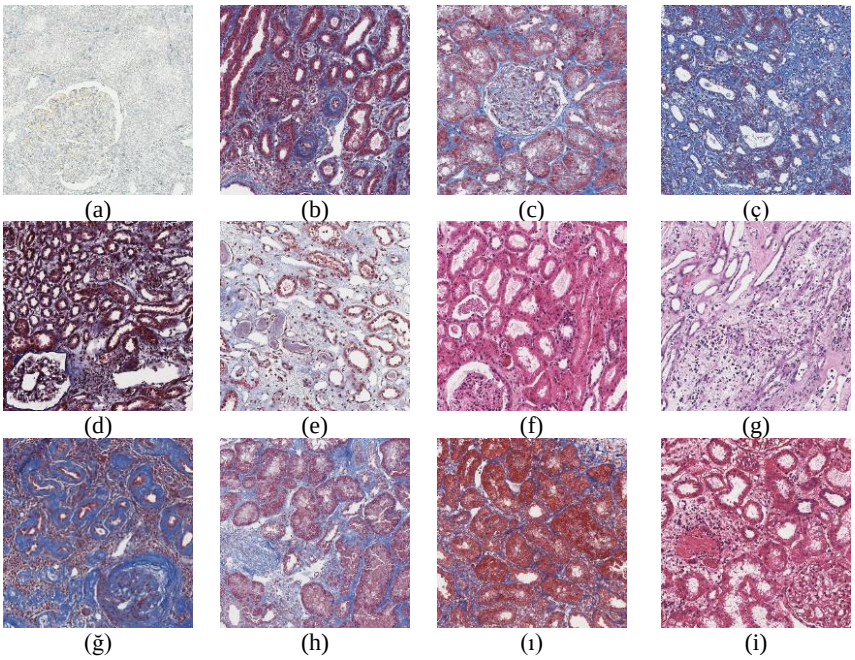
2.1. Etiketli Patoloji Görüntüleri ile Veri Setinin Oluşturulması

Bu çalışmada, ABH evresinin otomatik belirlenmesi için kullanılacak derin öğrenme modellerinin eğitimi, doğrulanması ve test edilmesi amacıyla patolojik görüntü verileri Kidney Precision Medicine Project (KPMP) veri tabanından elde edilmektedir [14]. Görüntülere ek olarak her Tüm Slayt Görüntüsü (WSI) için ABH Evre 1, 2 ve 3 olan etiketleri de bu veri tabanından elde edilmektedir.

Elde edilen veri seti ışık mikroskobu kullanılarak oluşturulmuş patoloji görüntülerinden oluşmaktadır. Veriler yüksek çözünürlüklü WSI formatı olan SVS dosya uzantısındadır. Veri seti Hematoksilen ve Eozin, Gümüş, Periyodik Asit-Schiff, Trikrom ve Toluidin Mavisi olmak üzere farklı boyama türlerini içermektedir. Veri setindeki boyama türü dağılımı, hasta bazında değişkenlik göstermektedir.

WSI'lerin aşırı büyük çözünürlükleri ve büyük ölçüde doku dışı alanlar içermesi, doğrudan derin öğrenme modelleri için verimsiz olmaktadır. Modelin yalnızca histopatolojik olarak anlamlı bölgelere odaklanmasını sağlamak amacıyla WSI içerisindeki tüm doku içeren bölgeler manuel olarak arka plandan ayrılmaktadır. Tespit edilen doku bölgeleri, bir uzman patoloğun mikroskop altında incelediği temsili alanı korumak adına ve patologların rutin tanı sürecinde tek seferde analiz etmeye odaklandığı görüntü alanını modellemek amacıyla 2000×2000 piksel boyutlarında görüntüler halinde kırpılmaktadır. Kırpılan bu görüntüler model eğitiminde kullanılmak üzere JPEG formatında kaydedilerek veri setini oluşturmaktadır.

KPMP veri tabanından elde edilen her bir WSI, uzman patologlar tarafından belirlenmiş bir ABH evresi etiketi taşımaktadır. Bu çalışmada, elde edilen kırpılmış her yeni görüntü ait olduğu WSI'nin ana etiketi ile etiketlenmektedir. Klinik etiketlere göre standardize edilmiş görüntü veri setinin dağılımına bakıldığında, Evre 1 için 836, Evre 2 için 751 ve Evre 3 için 1905 görüntü bulunmaktadır. Her evreye ait örnek görüntüler Şekil 1'de verilmektedir.



Şekil 1. ABH evrelerine ait örnek patoloji görüntüleri ((a-ç) Evre 1, (d-g) Evre 2 ve (ğ-i) Evre 3 olarak etiketlenmiş).

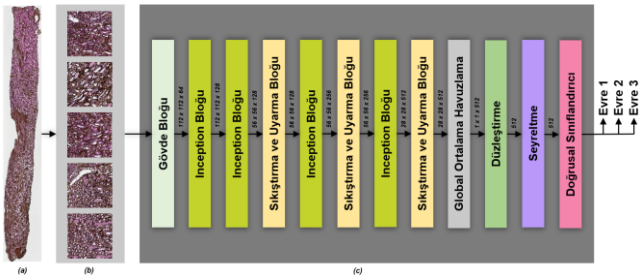
2.2. Otomatik Evre Belirlenmesi için Derin Öğrenme Tabanlı Yaklaşımın Geliştirilmesi

Bu çalışmada ABH evresinin patoloji görüntülerinden otomatik belirlenmesi için yeni bir derin öğrenme tabanlı yaklaşım önerilmektedir.

Şekil 2'de patoloji görüntülerinde ABH evresinin otomatik belirlenmesi için önerilen derin öğrenme tabanlı yaklaşımın genel akışı verilmektedir. Akışın (a) kısmında ham veriler, (b) kısmında ham verilerden elde edilen yeni görüntüler, (c) kısmında ise evrenin otomatik belirlenmesi için oluşturulan yaklaşımın mimari detayları yer almaktadır.

Önerilen yaklaşımın ilk aşaması olan gövde bloğu, giriş görüntüsünden temel görsel özellikleri çıkarmak üzere tasarlanmıştır. Bu aşamada, 224x224x3 boyutundaki giriş görüntüsü, ardışık iki adet 3x3 evrişim katmanından geçirilmektedir. Her iki katmanda da modelin öğrenme kararlılığını artırmak için batch normalization ve ReLU aktivasyon fonksiyonu kullanılarak düşük seviyeli kenar, doku ve renk geçişi özellikleri öğrenilmektedir.

Gövde bloğunun sonunda uygulanan 3x3 maksimum havuzlama işlemi ile uzamsal boyut kontrollü bir şekilde azaltılarak özellik haritası 112x112x64 boyutuna dönüştürülmektedir. Gövde bloğundan elde edilen özellik haritaları, farklı ölçeklerdeki yapıları yakalayabilmek için Inception bloklarına aktarılmaktadır. İlk Inception bloğu



Şekil 2. Patoloji görüntülerinde ABH evresinin otomatik belirlenmesi için önerilen yaklaşımın genel akışı.

çıkış kanal sayısını 128 olarak üretmekte ve uzamsal boyutunu 112x112x128 olarak korumaktadır. İkinci Inception bloğunda uzamsal çözünürlüğü azaltmak amacıyla 56x56x128'e düşürülmektedir. Devamındaki sıkıştırma ve uyarma bloğu (SE), boyut değişmeden kanal bazlı ağırlıklandırma yaparak modelin ayırt edici özellikleri öğrenmesini güçlendirmektedir. Sonrasındaki Inception bloğu çıkış kanal sayısını 256'ya yükseltmekte ve uzamsal boyutu 56x56 olarak korumaktadır. Bunu izleyen SE bloğu, yine boyutu değiştirmeden kanal ağırlıklandırmasını gerçekleştirerek önemli özelliklerin ön plana çıkmasını sağlamaktadır. Son Inception bloğu uzamsal çözünürlüğü düşürmekte ve çıkış boyutu 28x28x512'ye ulaşmaktadır. Bu bloğu takip eden SE bloğu kanal bazlı dikkat mekanizmasını uygulayarak modelin ayırt edici özellikleri öğrenmesini desteklemektedir. Özellik çıkarım aşamasının sonunda elde edilen 28x28x512 boyutlu haritalar global ortalama havuzlama ile 1x1x512 boyutuna dönüştürülmektedir. Daha sonra düzleştirme katmanı ile bu haritalar tek boyutlu bir vektöre çevrilmektedir. Bu vektör, aşırı öğrenmeyi önlemek ve genelleme kabiliyetini artırmak amacıyla 0.5 seyreltme işlemine tabi tutulmaktadır. Son aşamada yer alan doğrusal sınıflandırıcı katman, her bir sınıfa ait logit değerlerini üretmektedir. Bu logit değerleri Softmax fonksiyonu ile olasılığa dönüştürülmekte ve en yüksek olasılık değerine sahip sınıf seçilerek nihai karar verilmektedir. Bu süreç sonucunda üç sınıflı ABH evre sınıflandırması gerçekleştirilmektedir.

3. BULGULAR VE TARTIŞMA

Bu çalışmada ABH evresinin otomatik belirlenmesi için gerekli deneyler, Intel Core i7-13620H 2.40 GHz işlemci, 32 GB RAM ve NVIDIA GeForce RTX 4070 GPU'ya sahip bir bilgisayarda Python programlama dili 3.12 sürümünde PyTorch kütüphanesi ile gerçekleştirilmektedir. KPMP veri seti üzerinde daha önce böbrek biyopsisi görüntülerinden ABH evresini otomatik olarak belirleyen bir çalışma bulunmamaktadır. Bu nedenle, literatürdeki derin öğrenme modelleri, önerilen yaklaşımın performansının değerlendirilmesi amacıyla kullanılmıştır. Patoloji görüntülerinde ABH evresinin otomatik belirlenmesi için önerilen derin öğrenme tabanlı yeni yaklaşımı değerlendirmek ve literatürdeki diğer yaklaşımlarla karşılaştırmak için doğruluk, F1-skoru, kesinlik, duyarlılık ve özgüllük metrikleri kullanılmaktadır. Ek olarak, tüm metrikler veri setindeki dengesizlik sebebiyle hem ağırlıklı hem de makro olarak hesaplanmaktadır. Modelin eğitimi ve değerlendirilmesi aşamasında, 5 katlı tabakalı çapraz doğrulama uygulanmıştır. Her katta veri setinin %80'i eğitim, %20'si test için kullanılmış ve raporlanan metrikler beş katın ortalaması olarak verilmiştir. Bölme işlemi WSI düzeyinde yapılmış olup aynı WSI'den elde edilen görüntüler birden fazla katta yer almamaktadır ve aynı WSI'den görüntüler aynı anda hem eğitim hem de test kümesinde bulunmamaktadır. Metrik detayları aşağıdaki gibidir:

1) Doğruluk (Accuracy - ACC): Doğruluk model tarafından doğru sınıflandırılan veri sayısının, toplam veri sayısına oranını ifade etmektedir. TP (Doğru Pozitif) pozitif sınıfların doğru bir şekilde tahmin edildiğini, TN (Doğru Negatif) negatif sınıfların doğru bir şekilde tahmin edildiğini, FP (Yanlış Pozitif) hatalı bir şekilde pozitif tahmin edildiğini ve FN (Yanlış Negatif) hatalı bir şekilde negatif tahmin edildiğini göstermek üzere doğruluk metriğinin formülü aşağıdaki gibidir:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

2) Kesinlik (Precision - PREC): Kesinlik modelin pozitif olarak tahmin ettiği verilerde gerçekte pozitif olan durumların oranını ifade etmektedir.

$$PREC = \frac{TP}{TP + FP} \tag{2}$$

3) Duyarlılık (Recall - REC): Duyarlılık gerçekte pozitif olan tüm veriler arasında modelin doğru bir şekilde pozitif olarak tahmin ettiği verilerin oranını ifade etmektedir.

$$REC = \frac{TP}{TP + FN} \tag{3}$$

4) F1-Skoru: F1-skoru kesinlik ve duyarlılık metriklerinin harmonik ortalaması olmakta ve özellikle dengesiz veri setlerinde modelin performansını göstermektedir.

$$F1 - Skoru = 2\frac{PREC.REC}{PREC + REC} \quad (4)$$

5) Özgüllük (Specificity - SPEC): Özgüllük gerçekte negatif olan tüm verilerde modelin doğru bir şekilde negatif olarak tahmin ettiği verilerin oranını ifade etmektedir.

$$SPEC = \frac{TN}{TN + FP} \quad (5)$$

Çalışmada ABH evresinin patoloji görüntülerinden otomatik belirlenmesi için önerilen yeni yaklaşım, literatürde yer alan 7 derin öğrenme yaklaşımıyla karşılaştırılmaktadır. Tüm derin öğrenme modellerinde giriş görüntüleri eğitime başlamadan önce giriş boyutuna uygun olacak şekilde 2000x2000 boyutundan 224x224 piksel boyutuna yeniden boyutlandırılmaktadır. Ek olarak ImageNet [15] ortalama ve standart sapma değerleri kullanılarak normalize edilmektedir. Veri bölme stratejisi olarak sınıf dağılımını koruyan 5 katlı tabakalı çapraz doğrulama uygulanmıştır. Önerilen yaklaşımın 5 katlı tabakalı çapraz doğrulama süreci yaklaşık 46 dakika sürmekte ve görüntü başına işlem süresi 0.038 ms olarak kaydedilmektedir. Başlangıç öğrenme hızı $3x10^{-5}$ ve ağırlık azalması $1x10^{-4}$ olarak belirlenmektedir. Optimizasyon için AdamW algoritması kullanılmaktadır. Sınıflandırma hatası, sınıf dengesizliğinin etkisini azaltmak amacıyla her sınıf için örnek sayısının tersine orantılı olarak hesaplanan ağırlıkların kullanıldığı ağırlıklı çapraz entropi kaybı fonksiyonu ile hesaplanmaktadır. Eğitim süreci yığın büyüklüğü 32 ile gerçekleştirilmektedir. Eğitim süreci 50 epoch olarak belirlenmektedir. Hiperparametre detayları Tablo 1'de verilmiştir:

Tablo 1. Kullanılan hiperparametreler

Hiperparametre	Değer
Veri Bölme Stratejisi	5 katlı tabakalı
Yığın Büyüklüğü	32
Epoch	50
Optimizasyon	AdamW
Öğrenme Oranı	$3x10^{-5}$
Ağırlık Azalması	$1x10^{-4}$
Kayıp Fonksiyonu	Ağırlıklı çapraz entropi

Tablo 2. Önerilen yaklaşımın eksiltme çalışması

Model	ACC
4 I	0.7230
I + SE + 3 I	0.7357
I + SE + I + SE + 2 I	0.7443
I + SE + I + SE + I + SE + I	0.7415
I + SE + I + SE + I + SE + I + SE	0.7413
I + I + SE + I + I	0.7365
I + I + SE + I + SE + I	0.7622
I + I + SE + I + SE + I + SE	**0.7668**
I + I + I + SE + I + SE	0.7125
I + I + I + I + SE	0.7194

Önerilen yaklaşımın performans başarısının irdelenmesi amacıyla bir eksiltme çalışması gerçekleştirilmektedir. Bu analiz kapsamında önerilen yaklaşımın Inception (I) blokları sabit tutulmuş ve SE bloklarının sayısı değiştirilerek 10 farklı model varyantı oluşturulmaktadır.

Tablo 2'de SE bloğu içermeyen, 4 Inception bloğundan oluşan temel modelin 0.7230 doğruluk elde ettiği gözlemlenmektedir. SE bloklarının çeşitli sıralama ve sayıda kullanıldığı çoğu modelde ise temel modelin üzerinde doğruluk tespit edilmektedir. Bu durum SE bloklarının içerdiği dikkat mekanizmasının özellik haritalarına odaklanmayı artırarak doğruluğun artmasında etkisi olduğunu göstermektedir. Her Inception bloğunun sonunda SE bloğunun kullanıldığı modelde (ACC: 0.7413) doğruluğun en iyi modelin gerisinde kaldığı görülmektedir. Bu durumun dikkat mekanizmasının ilk katmanlardan elde edilen özelliklere erken odaklanmasından kaynaklandığı düşünülmektedir. Bazı modellerde SE blokları yer almasına rağmen doğruluk bakımından temel modelin de gerisinde kaldığı gözlemlenmektedir (Örn. I + I + I + SE + I + SE ACC: 0.7125). Bu eksiltme çalışmasının sonucunda en iyi blok diziliminin I + I + SE + I + SE + I + SE modeli olduğu tespit edilmektedir (ACC: 0.7668).

Karşılaştırılan yaklaşımlar InceptionV1 [16], ResNet18 [17], ResNeXt50 [18], SqueezeNet v1.0 [19], ShuffleNetV2 [20], MobileNetV1 [21] ve EfficientNetB0 [22] modelleridir. Elde edilen metrik sonuçları ortalama şeklinde Tablo 3 ve Tablo 4'te verilmektedir. Veri setindeki dengesizlik sebebiyle F1-skoru, PREC ve REC metrikleri Tablo 3'te makro, Tablo 4'te ise ağırlıklı olarak verilmektedir. Sonuçlar incelendiğinde önerilen yaklaşımın 0.7668 doğruluk, 0.7252 makro F1-skoru, 0.7334 makro PREC, 0.7298 makro REC, 0.8674 SPEC, 0.7703 ağırlıklı F1-skoru, 0.7823 ağırlıklı PREC metrik sonuçları vererek diğer yaklaşımlardan daha yüksek başarı gösterdiği gözlemlenmektedir. Elde edilen bu metrik değerleri, önerilen yaklaşımın dokularda oluşan farklı şiddetteki hücresel bozulmaları yakalamadaki başarısını kanıtlamaktadır. Özellikle veri setindeki sınıf dengesizliğine rağmen makro F1-skoru, modelin her evre için güvenilir tahminler üretebildiğini ve klinik pratikte karşılaşılan karmaşık doku desenlerini başarıyla genelleyebileceğini göstermektedir. ResNet18 ise doğrulukta 0.7393 ile önerilen yaklaşıma yakın doğruluk göstermektedir ancak makro F1-skoru (0.6821), makro PREC (0.6975) ve makro REC (0.6991) değerleri ile önerilen yaklaşımdan daha düşük bir başarı göstermektedir. En düşük performans ise ShuffleNetV2'de gözlemlenmiştir (ACC: 0.5803, makro F1-skoru: 0.4893). SqueezeNet v1.0 küçük ve hafif bir model olmasına rağmen makro F1-skoru ve duyarlılık değerleri yaklaşık 0.7 olarak gözlemlenmektedir. Parametre verimliliğine sahip MobileNetV1 ise makro F1-skoru 0.58 ve duyarlılık 0.59 ile daha sınırlı bir performans göstermektedir. InceptionV1 ve ShuffleNetV2, daha düşük performans (ACC: 0.6397, 0.5803) göstermekte olup eski veya daha az parametre içeren mimarilerin akut böbrek hasarı evresinin patoloji

görüntülerinde tespiti için sınırlı kaldığını göstermektedir. Önerilen yaklaşım, diğer tüm yaklaşımlara göre bir performans artışı sağlamakta ve veri setinin dengesiz olduğu göz önüne alındığında özellikle makro F1-skoru (0.7252) ve REC (0.7298) açısından daha dengeli performans göstermektedir. Bu sonuçların elde edilmesinde, önerilen yaklaşımın çok ölçekli Inception blokları ve kanal bazlı SE dikkat mekanizmaları sayesinde patoloji görüntülerindeki hem düşük seviyeli dokusal detayları hem de yüksek seviyeli yapıları etkili şekilde öğrendiği düşünülmektedir.

Tablo 3 ve Tablo 4'teki metrik değerleri 5 katlı tabakalı çapraz doğrulamanın kat ortalaması olarak verilirken Şekil 3'teki karışıklık matrisleri tüm katların tahminleri birleştirilerek oluşturulan global karışıklık matrislerini yansıtmaktadır. Şekil 3'te yer alan karışıklık matrisleri incelendiğinde önerilen yaklaşımın Evre 1 olarak etiketlenen 836 görüntüden 613'ünü (REC: %73.3), Evre 2 olarak etiketlenen 751 görüntüden 494'ünü (REC: %65.8), Evre 3 olarak etiketlenen 1905 görüntüden 1487'sini (REC: %78.1) doğru sınıflandırdığı gözlemlenmektedir. Önerilen yaklaşıma en yakın performansı ise Evre 1 için %71.8, Evre 2 için %56.5 ve Evre 3 için %77.9 duyarlılık değeri elde eden ResNet18 göstermektedir. Tüm karışıklık matrislerinin hata desenleri incelendiğinde en çok hatanın gerçek etiketi Evre 3 olan görüntülerin Evre 2 olarak (Örn. InceptionV1: 515 görüntü, ResNeXt50: 484 görüntü, ShuffleNetV2: 484 görüntü, MobileNetV1: 489 görüntü, EfficientNetB0: 514 görüntü) ve gerçek etiketi Evre 2 olan görüntülerin Evre 3 olarak (Örn. InceptionV1: 274 görüntü, ResNeXt50: 234 görüntü, ShuffleNetV2: 420 görüntü, MobileNetV1: 269 görüntü, EfficientNetB0: 305 görüntü) tahmin edilmesinde görüldüğü tespit edilmektedir. Bu durumun evreler arası geçişlerin patolojik düzeyde birbirine çok yakın desenler sergilemesinden kaynaklandığı düşünülmektedir. Bazı görüntülerde Evre 2'nin ilerlemiş aşamaları ile Evre 3'ün erken aşamaları arasındaki dokusal benzerlik, otomatik sınıflandırma için kullanılan modellerin evreler arası sınırları kesin olarak ayırt etmesini zorlaştırmaktadır.

Tablo 3. Derin öğrenme modellerinin performans karşılaştırması 1

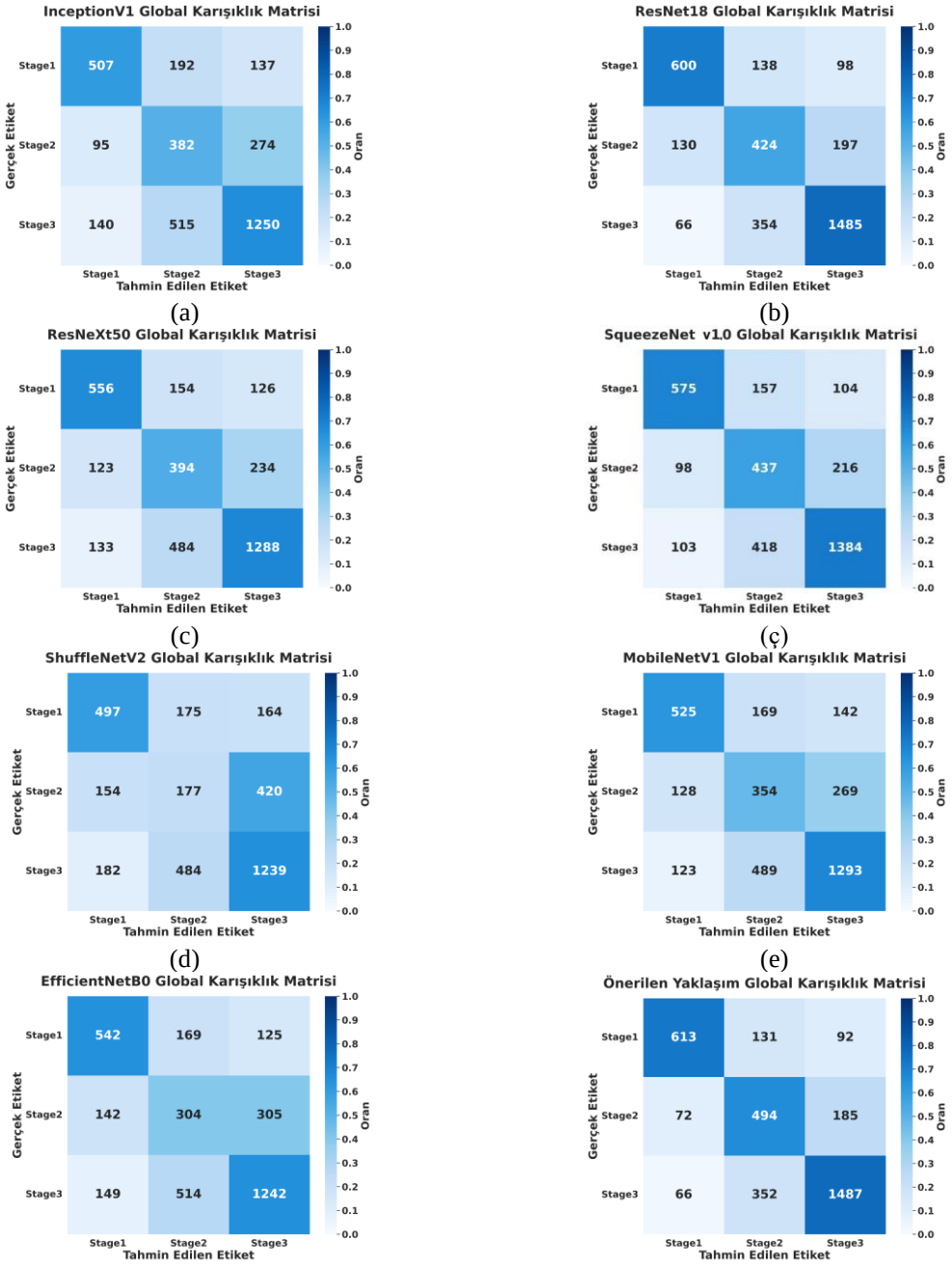
Model	ACC	F1-Skoru$_{makro}$	PREC$_{makro}$	REC$_{makro}$
InceptionV1 [16]	0.6397	0.6054	0.6142	0.6297
ResNet18 [17]	0.7393	0.6821	0.6975	0.6991
ResNeXt50 [18]	0.7030	0.5872	0.6296	0.6038
SqueezeNet v1.0 [19]	0.7110	0.6629	0.6652	0.6849
ShuffleNetV2 [20]	0.5803	0.4893	0.5075	0.5129
MobileNetV1 [21]	0.6973	0.5795	0.6248	0.5948
EfficientNetB0 [22]	0.6412	0.5566	0.5743	0.5740
Önerilen Yaklaşım	0.7668	0.7252	0.7334	0.7298

Tablo 4. Derin öğrenme modellerinin performans karşılaştırması 2

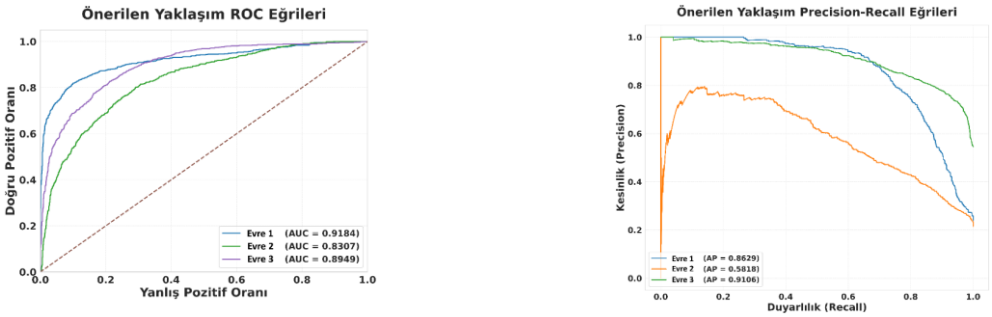
Model	F1-Skoru$_{ağırlıklı}$	PREC$_{ağırlıklı}$	SPEC
InceptionV1 [16]	0.6540	0.6900	0.8102
ResNet18 [17]	0.7376	0.7571	0.8518
ResNeXt50 [18]	0.6686	0.6839	0.8079
SqueezeNet v1.0 [19]	0.7159	0.7354	0.8437
ShuffleNetV2 [20]	0.5666	0.5847	0.7575
MobileNetV1 [21]	0.6695	0.6871	0.8095
EfficientNetB0 [22]	0.6308	0.6477	0.7935
Önerilen Yaklaşım	0.7703	0.7823	0.8674

Şekil 4 (a)'da önerilen yaklaşımın ROC eğrileri incelendiğinde tüm eğrilerin rastgele tahmin eğrisinin üzerinde kaldığı gözlemlenmektedir. AUC değerinin en yüksek olduğu sınıfın Evre 1 (AUC: 0.9184) olduğu görülmektedir. Evre 3 ROC eğrisi (AUC: 0.8949) Evre 1 eğrisine yakın bir seyir izlemekte ve ikinci sırada yer almaktadır. Evre 2 eğrisi ise AUC: 0.8307 ile diğer eğrilerin gerisinde kalmaktadır. Bu durumun Evre 2'nin bir geçiş evresi niteliği taşımasından kaynaklandığı düşünülmektedir. Şekil 4 (b)'de önerilen yaklaşımın PR eğrileri incelendiğinde Evre 3 eğrisinin (AP: 0.9106) genel olarak en üstte yer aldığı ve en yüksek başarıyı gösterdiği tespit edilmektedir. Bu durum modelin Evre 3'ü ayırt etmede yüksek kesinliğe sahip olduğunu göstermektedir. Evre 1 eğrisinin (AP: 0.8629) ise Evre 3 eğrisine yakın bir seyir izlediği gözlemlenmektedir. Evre 2 eğrisinde (AP: 0.5818) ise duyarlılık değeri arttıkça kesinlik değerinin belirgin biçimde azaldığı ve Evre 2 sınıflandırmasının en düşük performansa sahip olduğu gözlemlenmektedir. Bu durum Evre 2 sınıflandırmasında daha fazla yanlış pozitif üretildiğini göstermektedir.

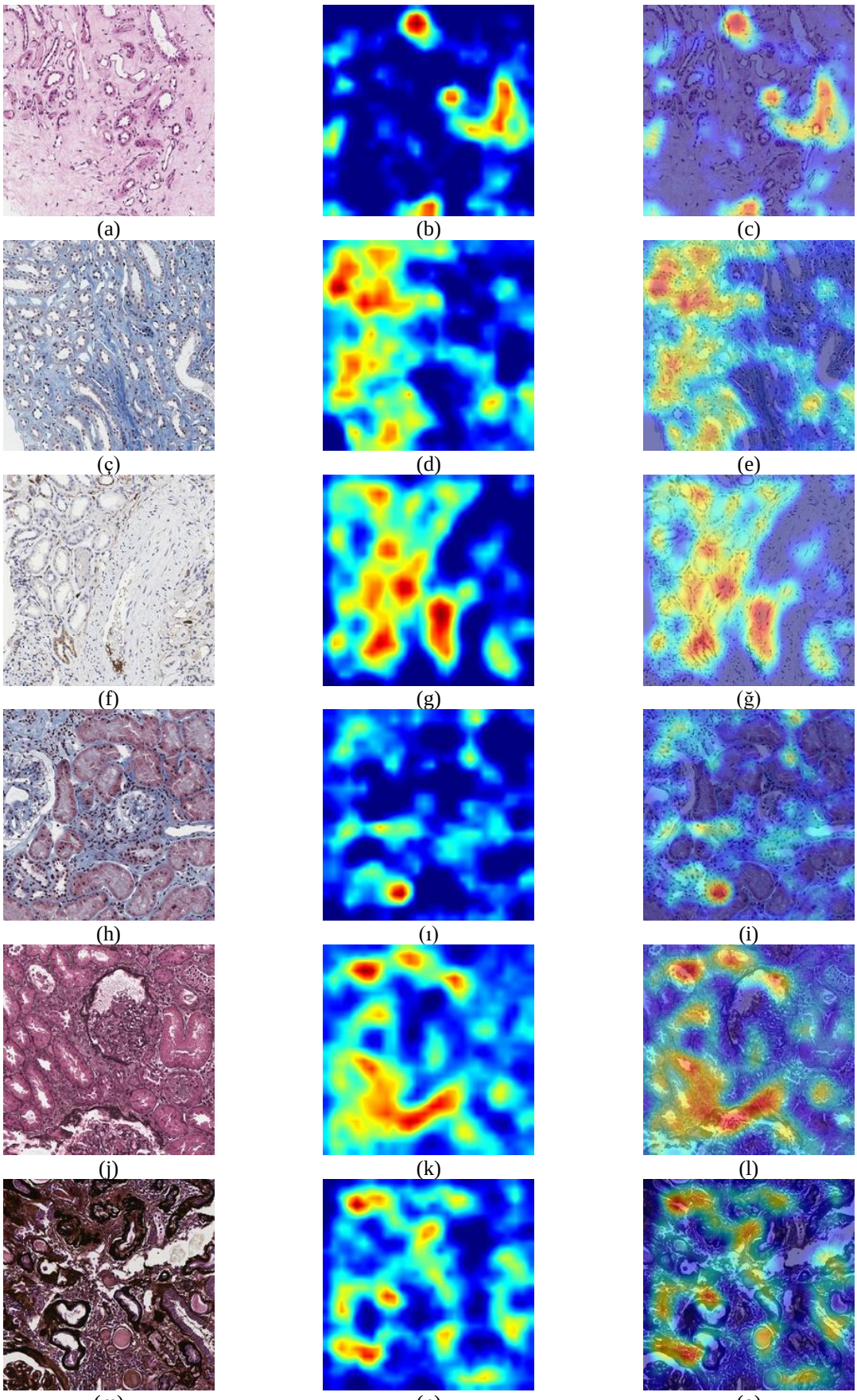
Şekil 5'te önerilen yaklaşıma ait (a) Evre 1 etiketli, (ç) gerçek etiketi Evre 1 olup yanlış tahmin edilmiş, (f) Evre 2 etiketli, (h) gerçek etiketi Evre 2 olup yanlış tahmin edilmiş, (j) Evre 3 etiketli, (m) gerçek etiketi Evre 3 olup yanlış tahmin edilmiş, (b, d, g, ı, k, n) ilgili görüntülerin ısı haritası, (c, e, ğ, i, l, o) örtüşme görüntüsü verilmektedir. Örtüşme görüntüleri incelendiğinde önerilen yaklaşımın evre etiketini belirlerken arka plan, boşluklar veya artefaktlar yerine böbrek yapısındaki çeşitli yapılardaki değişikliğe odaklandığı gözlemlenmektedir. Isı haritalarının her görüntüde farklı bölgelere yoğunlaşması, önerilen yaklaşımın her görüntü için etiket tahmininde farklı özellikler çıkardığını göstermektedir.



Şekil 3. Mevcut derin öğrenme yaklaşımları ve önerilen yaklaşımın karışıklık matrisleri ((a) InceptionV1, (b) ResNet18, (c) ResNeXt50, (ç) SqueezeNet v1.0, (d) ShuffleNetV2, (e) MobileNetV1, (f) EfficientNetB0, (g) Önerilen Yaklaşım).



Şekil 4. Önerilen yaklaşımın sınıf bazlı ROC eğrileri (a) ve PR eğrileri (b).



Şekil 5. Önerilen yaklaşımın evrelere ait orijinal (a, ç, f, h, j, m), Grad-CAM ısı haritası (b, d, g, ı, k, n) ve örtüşme (c, e, ğ, i, l, o) görüntüleri.

4. SONUÇ

Bu çalışmada patoloji görüntülerinde akut böbrek hasarı evresinin otomatik belirlenmesi için derin öğrenme tabanlı yeni bir yaklaşım önerilmektedir. Önerilen yaklaşım, KPMP veri tabanında mevcut olan etiketli verileri giriş olarak kullanmakta ve çıkış olarak akut böbrek hasarı evresini vermektedir. Çalışmanın literatüre sağladığı temel katkılar şu şekilde özetlenebilmektedir: (1) Patoloji görüntülerinden ABH evresinin otomatik belirlenmesine odaklanan ilk çalışmalardan biridir, (2) önerilen hiyerarşik ve çok ölçekli yaklaşım ile farklı boyama türlerinde yüksek performans gösteren bir akış sunulmaktadır, (3) klinik tanı süreçlerinde uzmanların iş yükünü azaltabilecek ve uzmana bağlı yorum değişikliklerini azaltarak nesnel bir değerlendirme sağlayacak bir yöntem sunulmaktadır. Önerilen yaklaşım ile elde edilen 0.7668 doğruluk ve 0.7252 makro F1-skoru performans sonuçları, literatürde yaygın kullanılan derin öğrenme modellerine göre üstünlüğünü kanıtlamaktadır. Ek olarak giriş görüntülerinin farklı boyama türlerini içermesine rağmen sonuçlarda elde edilen 5 katlı tabakalı çapraz doğrulamadaki metrik değerleri önerilen yaklaşımın veri dengesizliğine karşı dayanıklı ve kararlı bir yapıda olduğunu kanıtlamaktadır. Özellikle veri sayısı Evre 3'e göre az olan Evre 1 (REC: %73.3) ve Evre 2 (REC: %65.8) sınıflarında elde edilen sonuçlar, modelin karmaşık doku bozulmalarını ayırt etmedeki genelleme yeteneğini göstermektedir. Literatürdeki çalışmalar incelendiğinde akut böbrek hasarı evresinin çeşitli biyobelirteç sonuçları ile değerlendirildiği gözlemlenmektedir. Bu çalışma ise böbrek patoloji görüntüleri üzerinden evre belirlemesi yaparak patologların patolojik inceleme sürecini hızlandıracak ve patoloğa bağlı değişkenliği azaltarak objektif bir karar destek sistemi sunmaktadır. Sonuç olarak bu yaklaşım, dijital patoloji iş akışlarına entegre edilerek tanı süreçlerini standartlaştırma ve klinik değerlendirmelerde yüksek tutarlılık sağlama potansiyeline sahiptir.

Yazar Katkıları

Fatma Tuana DOĞU kavramsal tasarım, veri analizi ve makale taslağı yazımını üstlenmiştir; Mehmet ÖZTÜRK ve Hülya DOĞAN içerik eleştirisi ve danışmanlık sağlamıştır. Tüm yazarlar son onay ve tam sorumluluğu paylaşmıştır.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

Teşekkür

Bu çalışma Fatma Tuana DOĞU'nun Karadeniz Teknik Üniversitesi'nde yüksek lisans eğitimini tamamlamak üzere hazırladığı tez çalışmasının bir parçasıdır. Bu çalışma TÜBİTAK 2210-C Öncelikli Alanlara Yönelik Yurt İçi Yüksek Lisans Burs Programı (Proje No. 1649B022502850) ve Türkiye Sağlık Enstitüleri (TÜSEB) 2025 A Grubu Acil Ar-Ge Proje Destek Programı (Proje No. 46679) tarafından desteklenmiştir.

KAYNAKÇA

[1] World Health Organization. (2024, Aug.). The top 10 causes of death. WHO [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] S. Hacking and V. Bijol, "Deep learning for the classification of medical kidney disease: a pilot study for electron microscopy," Ultrastruct. Pathol., vol. 45, no. 2, pp. 118–127, Mar. 2021.

[3] P. Chagas, L. Souza, I. Araújo, N. Aldeman, A. Duarte, M. Angelo, and L. Oliveira, "Classification of glomerular hypercellularity using convolutional features and support vector machine," Artif. Intell. Med., vol. 103, pp. 101808, Mar. 2020.

[4] Q. Zheng, R. Yang, H. Xu, J. Fan, P. Jiao, X. Ni, and X. Liu, "A weakly supervised deep learning model and human–machine fusion for accurate grading of renal cell carcinoma from histopathology slides," Cancers, vol. 15, no. 12, pp. 3198, Jun. 2023.

[5] S. Shubham, N. Jain, V. Gupta, S. Mohan, M. M. Ariffin, and A. Ahmadian, "Identify glomeruli in human kidney tissue images using a deep learning approach," Soft Comput., vol. 27, no. 5, pp. 2705–2716, Mar. 2023.

[6] N. Bouteldja, B. M. Klinkhammer, R. D. Bülow, P. Droste, S. W. Otten, S. F. von Stillfried, and D. Merhof, "Deep learning–based segmentation and quantification in experimental kidney histopathology," J. Am. Soc. Nephrol., vol. 32, no. 1, pp. 52–68, Jan. 2021.

[7] T. T. U. Nguyen, A. T. Nguyen, H. Kim, Y. J. Jung, W. Park, K. M. Kim, and W. Kim, "Deep-learning model for evaluating histopathology of acute renal tubular injury," Sci. Rep., vol. 14, no. 1, pp. 9010, Apr. 2024.

[8] Y. Zheng, C. A. Cassol, S. Jung, D. Veerapaneni, V. C. Chitalia, K. Y. Ren, and V. B. Kolachalama, "Deep-learning–driven quantification of interstitial fibrosis in digitized kidney biopsies," Am. J. Pathol., vol. 191, no. 8, pp. 1442–1453, Aug. 2021.

[9] J. N. Marsh, T. C. Liu, P. C. Wilson, S. J. Swamidass, and J. P. Gaut, "Development and validation of a deep learning model to quantify glomerulosclerosis in kidney biopsy specimens," JAMA Netw. Open, vol. 4, no. 1, pp. e2030939, Jan. 2021.

[10] C. F. Juang, Y. W. Chuang, G. W. Lin, I. F. Chung, and Y. C. Lo, "Deep learning-based glomerulus detection and classification with generative

morphology augmentation in renal pathology images," Comput. Med. Imaging Graph., vol. 115, pp. 102375, Jul. 2024.

[11] N. Tomašev et al., "A clinically applicable approach to continuous prediction of future acute kidney injury," Nature, vol. 572, no. 7767, pp. 116–119, Aug. 2019.

[12] S. Yue, S. Li, X. Huang, J. Liu, X. Hou, Y. Zhao, D. Niu, Y. Wang, W. Tan, and J. Wu, "Machine learning for the prediction of acute kidney injury in patients with sepsis," J. Transl. Med., vol. 20, no. 1, pp. 215, May. 2022.

[13] J. He, J. Lin, and M. Duan, "Application of machine learning to predict acute kidney disease in patients with sepsis associated acute kidney injury," Front. Med., vol. 8, pp. 792974, Dec. 2021.

[14] Kidney Precision Medicine Project. (2026, Feb.). KPMP Atlas Repository [Online]. Available: https://atlas.kpmp.org/repository/

[15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. Miami, USA, 2009.

[16] C. Szegedy et al., "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9. Boston, USA, 2015.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. Las Vegas, USA, 2016.

[18] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, (2017, Nov.). "Aggregated residual transformations for deep neural networks." arXiv [Online]. Available: https://arxiv.org/abs/1611.05431.

[19] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. (2016, Feb.). "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size." arXiv [Online]. Available: https://arxiv.org/abs/1602.07360.

[20] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "ShuffleNet V2: Practical guidelines for efficient CNN architecture design," in Proceedings of the European Conference on Computer Vision (ECCV), pp. 116–131. Munich, Germany, 2018.

[21] A. G. Howard et al. (2017, Apr.). "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv [Online]. Available: https://arxiv.org/abs/1704.04861

[22] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proceedings of the International Conference on Machine Learning (ICML), pp. 6105–6114. Long Beach, USA, 2019.