

T.C.
ERZİNCAN BİNALİ YILDIRIM ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YAPAY ZEKA VE ROBOTİK ANABİLİM DALI

GÖĞÜS RÖNTGENİ GÖRÜNTÜLERİNDEN ZATÜRRE TANISINDA DERİN
ÖĞRENME TABANLI YAKLAŞIMLARIN PERFORMANS ANALİZİ

Çağlar ÇİLTAS

Danışman: Doç. Dr. Volkan KAYA

TEZ JÜRİ ÜYELERİ
Doç. Dr. Volkan KAYA
Dr. Öğr. Üyesi Canan TAŞTİMUR TEMİZ
Dr. Öğr. Üyesi İsmail ÇALIKUŞU

YÜKSEK LİSANS TEZİ
ERZİNCAN, 2026

© 2026 [Çağlar ÇİLTAS]. Tüm hakları saklıdır.

Kabul ve Onay Sayfası

Doçent Dr. Volkan KAYA danışmanlığında, Çağlar ÇİLTAŞ tarafından hazırlanan bu çalışma 16/07/2026 tarihinde aşağıdaki jüri tarafından Yapay Zeka ve Robotik Anabilim Dalı'nda Yüksek Lisans Tezi olarak kabul oybirliği ile kabul edilmiştir.

Başkan: Doç. Dr. Volkan KAYA

İmza:

Üye: Dr. Öğr. Üyesi Canan TAŞTİMUR TEMİZ

İmza:

Üye: Dr. Öğr. Üyesi İsmail ÇALIKUŞU

İmza:

Bu tez Enstitü Yönetim Kurulunun /07/2026 tarih ve/..... sayılı kararı ile onaylanmıştır.

Prof. Dr. Kemal Volkan ÖZDOKUR
Enstitü Müdür V.

Not: Bu tezde kullanılan özgün ve başka kaynaklardan yapılan bildirişlerin, şekil ve tabloların kaynak olarak kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

Bilimsel Etięe Uygunluk Sayfası

“GÖĞÜS RÖNTGENİ GÖRÜNTÜLERİNDEN ZATÜRRE TANISINDA DERİN ÖĞRENME TABANLI YAKLAŞIMLARIN PERFORMANS ANALİZİ” isimli “Yüksek Lisans” tezim tarafımca intihal tespit programı ile incelenmiştir. Buna göre tezimde bilimsel etik ihlali ve intihal olarak nitelendirilebilecek herhangi bir durum olmadığını taahhüt ederim.

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir biçimde elde edildiğini; aynı zamanda bu kural ve davranışların gerektirdiğı gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi beyan ederim. 16/07/2026.

(İmza)

Çağlar ÇİLTAŞ

ÖZET

GÖĞÜS RÖNTGENİ GÖRÜNTÜLERİNDEN ZATÜRRE TANISINDA DERİN ÖĞRENME TABANLI YAKLAŞIMLARIN PERFORMANS ANALİZİ

Çağlar ÇILTAŞ

Yüksek Lisans Tezi

Erzincan Binali Yıldırım Üniversitesi, Fen Bilimleri Enstitüsü,

Yapay Zeka ve Robotik Anabilim Dalı

Danışman: Doçent Dr. Volkan KAYA

2026, 45 sayfa

Göğüs röntgeni görüntülerinden zatürre tanısı, düşük maliyetli ve yaygın kullanılabilir olması nedeniyle klinik uygulamalarda önemli bir yere sahiptir. Bununla birlikte derin öğrenme tabanlı yöntemlerin yüksek performans elde edebilmesi genellikle büyük miktarda etiketli veriye ihtiyaç duymakta, bu durum ise tıbbi görüntü analizinde önemli bir kısıt oluşturmaktadır. Bu tezde, az etiketli veri koşullarında göğüs röntgeni görüntülerinden zatürre tanısına yönelik derin öğrenme ve geleneksel makine öğrenmesi yaklaşımlarının performansları kapsamlı olarak değerlendirilmiştir. Bu amaçla ConvNeXt-Tiny mimarisi temel alınarak geliştirilen Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) önerilmiş ve derin öznetelik temsili ile prototip tabanlı benzerlik hesaplaması tek bir karar mekanizmasında birleştirilmiştir. Ayrıca önerilen model; Xception, VGG16 ve ResNet50 mimarileri ile karşılaştırılmış, bunun yanında SIFT, KAZE ve HOG öznetelikleri kullanılarak Destek Vektör Makineleri, Lojistik Regresyon ve Aşırı Gradyan Artırma sınıflandırıcılarından oluşan geleneksel yaklaşımlar da aynı veri seti üzerinde analiz edilmiştir. Tüm deneyler Chest X-Ray (Pneumonia) veri seti kullanılarak gerçekleştirilmiş ve modeller doğruluk, kesinlik, duyarlılık ve F1-Skoru ölçütleri ile değerlendirilmiştir. Deneysel sonuçlar, önerilen HPFN modelinin %96,38 doğruluk, %96,23 kesinlik, %97,95 duyarlılık ve %97,08 F1-Skoru elde ederek hem karşılaştırılan derin öğrenme mimarilerinden hem de geleneksel öznetelik tabanlı yöntemlerden daha başarılı performans sergilediğini göstermiştir. Literatür karşılaştırmaları da önerilen yaklaşımın az etiketli veri koşullarında güvenilir, dengeli ve yüksek doğruluk sağlayan etkili bir bilgisayar destekli tanı yöntemi olduğunu ortaya koymuştur.

Anahtar Kelimeler: Zatürre tanısı, derin öğrenme, ConvNeXt-Tiny, hibrit prototip füzyon ağı.

ABSTRACT

PERFORMANCE ANALYSIS OF DEEP LEARNING APPROACHES FOR PNEUMONIA DIAGNOSIS USING CHEST X-RAY IMAGES

Çağlar ÇILTAŞ

Master's Thesis

Erzincan Binali Yıldırım University, Institute of Science and Technology,

Department of Artificial Intelligence and Robotics

Advisor: Assoc. Prof. Dr. Volkan KAYA

2026, 45 pages

Diagnosis of pneumonia from Chest X-Ray images holds a significant place in clinical practice due to its low cost and widespread availability. However, achieving high performance with deep learning-based methods generally requires a large amount of labeled data, which constitutes a significant limitation in medical image analysis. In this thesis, the performance of deep learning and traditional machine learning approaches for diagnosing pneumonia from Chest X-Ray images under conditions of low-label data is comprehensively evaluated. For this purpose, a Hybrid Prototype Fusion Network (HPFN) developed based on the ConvNeXt-Tiny architecture is proposed, and deep feature representation and prototype-based similarity calculation are combined in a single decision mechanism. Furthermore, the proposed model is compared with Xception, VGG16, and ResNet50 architectures, and traditional approaches consisting of Support Vector Machine, Logistic Regression, and Extreme Gradient Boosting classifiers using SIFT, KAZE, and HOG features are also analyzed on the same dataset. All experiments were performed using the Chest X-Ray (Pneumonia) dataset, and the models were evaluated using accuracy, precision, sensitivity, and F1-score measures. Experimental results showed that the proposed HPFN model achieved 96.38% accuracy, 96.23% precision, 97.95% sensitivity, and 97.08% F1-score, outperforming both the compared deep learning architectures and traditional feature-based methods. Literature comparisons also revealed that the proposed approach is an effective computer-aided diagnostic method providing reliable, stable, and high-accuracy data under low-label data conditions.

Keywords: Pneumonia diagnosis, deep learning, ConvNeXt-Tiny, hybrid prototype fusion network

TEŞEKKÜR

Yüksek lisans eğitimim ve tez çalışmam süresince bilgi ve deneyimlerinden yararlanma fırsatı bulduğum, akademik yaklaşımı ve yol gösterici desteğiyle çalışmamın gelişimine önemli katkılar sağlayan değerli danışmanım Doç. Dr. Volkan KAYA'ya en içten teşekkürlerimi sunarım.

Araştırma sürecinde karşılaştığım akademik ve teknik konularda görüş ve önerileriyle bana destek olan, bilgi birikimlerini paylaşarak çalışmama katkı sağlayan tüm hocalarıma da teşekkür ederim.

Bu uzun ve yoğun süreç boyunca her zaman yanımda olan, destekleriyle bana güç veren, anlayış ve sabırlarını eksik etmeyen başta eşim ve kız kardeşim olmak üzere kıymetli aileme sonsuz teşekkür ederim. Özellikle zorlayıcı çalışma dönemlerinde sağladıkları moral ve motivasyon, bu çalışmanın tamamlanmasında benim için en önemli destek olmuştur.

Çağlar ÇİLTAŞ

Temmuz, 2026

İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER.....	iv
TABLolar DİZİNİ.....	vi
ŞEKİLLER DİZİNİ	vii
SİMGELER VE KISALTMALAR DİZİNİ.....	viii
1. GİRİŞ	1
2. KAVRAMSAL ÇERÇEVE VE İLGİLİ ÇALIŞMALAR.....	4
2.1. Zatürre ve Göğüs Röntgeni Görüntülerinin Tanıdaki Yeri	4
2.2. Derin Öğrenme Tabanlı Tıbbi Görüntü Analizi	5
2.3. Az Etiketli Veri Problemi.....	7
2.4. Transfer Öğrenme Yaklaşımı	8
2.5. ConvNeXt Mimarisi	9
2.6. Az Örneklili Öğrenme ve Prototip Tabanlı Yaklaşımlar.....	10
2.7. Göğüs Röntgeni Görüntülerinde Zatürre Tespitine Yönelik Çalışmalar.....	11
2.8. Etik İlkeler ve Klinik Güvenilirlik	11
2.9. Çalışmanın Literatürdeki Yeri.....	12
3. YÖNTEM.....	14
3.1. Geleneksel Öznitelik Tabanlı Yaklaşımlar.....	16
3.1.1. Ölçekten bağımsız öznitelik dönüşümü.....	16
3.1.2. Kaze	17
3.1.3. Yönlendirilmiş gradyan histogramı	17
3.1.4. Destek vektör makineleri.....	18
3.1.5. Lojistik regresyon	18
3.1.6. Aşırı gradyan artırma.....	18
3.2 Derin Öğrenme Tabanlı Yaklaşımlar	19
3.2.1 Xception	19
3.2.2. Vgg16	20
3.2.3. ResNet50	20
3.2.4. Önerilen hibrit prototip füzyon ağı (HPFN)	21
4. BULGULAR.....	23

4.1. Veri Seti.....	23
4.2. Değerlendirme Metrikleri	25
4.3. Ön İşleme ve Önerilen Model Mimarisi.....	26
4.4. Deneysel Sonuçlar	27
5. TARTIŞMA VE SONUÇ	36
5.1. Öneriler.....	38
KAYNAKÇA.....	41

TABLÖLAR DİZİNİ

Tablo 1. Çalışmada kullanılan Chest X-Ray (Pneumonia) veri setinin sınıf dağılımı	24
Tablo 2. Derin öğrenme tabanlı modeller ile önerilen HPFN modelinin performans karşılaştırması.....	30
Tablo 3. Geleneksel öznitelik tabanlı yöntemler ile önerilen HPFN modelinin performans karşılaştırması.....	32
Tablo 4. Literatürde yer alan çalışmalar ile önerilen HPFN modelinin karşılaştırılması.....	34

ŞEKİLLER DİZİNİ

Şekil 1. Çalışmanın genel iş akış diyagramı.....	15
Şekil 2. Çalışmada kullanılan Chest X-Ray (Pneumonia) veri setinden örnek göğüs röntgeni görüntüleri.	24
Şekil 3. Önerilen HPFN modeline ait karışıklık matrisi.....	29
Şekil 4. Önerilen HPFN modeli ile derin öğrenme tabanlı modellerin performans karşılaştırması.....	31
Şekil 5. Geleneksel yöntemler ile önerilen HPFN modelinin doğruluk karşılaştırması.....	33
Şekil 6. Literatürde yer alan çalışmalar ile önerilen HPFN modelinin doğruluk değerlerinin karşılaştırılması.....	35

SİMGELER VE KISALTMALAR DİZİNİ

Kısaltmalar

Kısaltma	Açılımı
ABF	Alignment-Based Filtering
CNN	Convolutional Neural Network
ConvNeXt	Convolutional Network Next
CXR	Chest X-Ray
DoG	Difference of Gaussian
F1	F1-Score
FSL	Few-Shot Learning
GELU	Gaussian Error Linear Unit
HOG	Histogram of Oriented Gradients
HPFN	Hybrid Prototype Fusion Network
ImageNet	ImageNet Large Scale Visual Recognition Dataset
KAZE	KAZE Feature Detector
LR	Logistic Regression
PA	Posterior-Anterior
ReLU	Rectified Linear Unit
SIFT	Scale-Invariant Feature Transform
SVM	Support Vector Machine
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
ViT	Vision Transformer
XGBoost	Extreme Gradient Boosting

Simgeler

Simge	Açıklama
(x)	Giriş görüntüsü
(x_i)	i. giriş örneği
(y)	Gerçek sınıf etiketi
$(\hat{y})$	Tahmin edilen sınıf
(z)	Gömme (embedding) vektörü
(P_c)	c sınıfına ait prototip vektörü
(S_c)	c sınıfına ait eğitim örnekleri kümesi
$(f(x))$	Derin öznitelik çıkarım fonksiyonu
(W)	Ağırlık matrisi
(b)	Sapma (bias) terimi
(K)	Toplam karar ağacı sayısı
$(f_k(x))$	k. karar ağacının çıktısı
$(F(x))$	Artık (Residual) fonksiyonu
(σ)	Gauss filtresinin standart sapması
(∇)	Gradyan operatörü
$(\ \cdot\)$	Öklid normu
$(\cos(\cdot))$	Kosinüs benzerliği

1. GİRİŞ

Yapay zekâ teknolojilerindeki hızlı gelişmeler, görüntü işleme ve bilgisayarlı görü alanlarında önemli ilerlemelere olanak sağlamış ve bu gelişmeler sağlık alanındaki uygulamalara da doğrudan yansımıştır. Özellikle derin öğrenme tabanlı yöntemlerin görüntülerden ayırt edici özellikleri otomatik olarak öğrenebilmesi, tıbbi görüntü analizinde bilgisayar destekli tanı sistemlerinin geliştirilmesini mümkün hâle getirmiştir. Bu sistemler, hastalıkların erken teşhis edilmesine katkı sağlamalarının yanı sıra klinik karar verme süreçlerini destekleyerek tanı sürecinin daha güvenilir ve daha tutarlı şekilde yürütülmesine yardımcı olmaktadır (Rajpurkar ve ark., 2017; L. Wang ve ark., 2020).

Zatürre, akciğer dokusunun enfeksiyonu sonucu gelişen ve özellikle çocuklar, yaşlı bireyler ile bağışıklık sistemi zayıf olan hastalarda ciddi sağlık sorunlarına neden olabilen önemli solunum yolu hastalıklarından biridir. Dünya Sağlık Örgütü verileri, zatürrenin her yıl milyonlarca insanın yaşamını tehdit eden yaygın ölüm nedenlerinden biri olduğunu göstermektedir (World Health Organization, 2026). Hastalığın erken evrede doğru şekilde teşhis edilmesi, uygun tedavinin zamanında başlatılması açısından kritik öneme sahiptir. Bu nedenle zatürre tanısında hızlı ve güvenilir karar destek sistemlerine duyulan ihtiyaç her geçen gün artmaktadır.

Göğüs röntgeni (Chest X-Ray/CXR), düşük maliyetli olması, kısa sürede uygulanabilmesi ve akciğer dokusunu ayrıntılı olarak gösterebilmesi nedeniyle zatürre tanısında en yaygın kullanılan görüntüleme yöntemlerinden biridir. Bununla birlikte CXR görüntülerinin değerlendirilmesi uzman radyolog deneyimine bağlıdır. Görüntülerin manuel olarak incelenmesi zaman almakta, yoğun klinik iş yükü altında değerlendirme sürecini zorlaştırmakta ve farklı uzmanlar arasında yorum farklılıklarının ortaya çıkmasına neden olabilmektedir. Bu durum, bilgisayar destekli otomatik tanı sistemlerinin geliştirilmesini dikkat çekici bir araştırma konusu hâline getirmiştir.

Son yıllarda derin öğrenme tabanlı yöntemler, özellikle evrimsel sinir ağlarının (Convolutional Neural Networks-CNN) gelişmesiyle birlikte göğüs röntgeni görüntülerinden hastalık tespitinde dikkat çekici başarılar göstermiştir. Rajpurkar ve ark. tarafından geliştirilen CheXNet modeli, geniş ölçekli CXR 14 veri seti üzerinde eğitilmiş ve zatürre tespitinde ortalama bir radyolog seviyesinde performans elde etmiştir (Rajpurkar ve ark., 2017). Wang ve ark. (2020) tarafından önerilen COVID-Net modeli ise göğüs röntgeni görüntülerinden otomatik hastalık sınıflandırmasının yüksek doğrulukla gerçekleştirilebileceğini göstermiştir

(L. Wang ve ark., 2020). Bu çalışmalar, derin öğrenme tabanlı yaklaşımların tıbbi görüntü analizinde güvenilir karar destek sistemleri geliştirmek açısından önemli bir potansiyele sahip olduğunu ortaya koymaktadır.

Derin öğrenme alanındaki gelişmeler yalnızca klasik CNN mimarileri ile sınırlı kalmamış, Vision Transformer yaklaşımından esinlenerek geliştirilen ConvNeXt mimarisi de son yıllarda tıbbi görüntü analizinde başarılı sonuçlar vermeye başlamıştır. ConvNeXt mimarisi, modern tasarım ilkelerini evrimsel sinir ağları ile birleştirerek güçlü öznitelik temsili oluşturabilmektedir. Perera ve Thayasivam, Bhatti ve ark., Xia ve ark., ve Xiong ve ark. tarafından gerçekleştirilen çalışmalar, ConvNeXt tabanlı modellerin COVID-19 tanısı, glioma derecelendirmesi, Alzheimer hastalığının sınıflandırılması ve göğüs röntgeni görüntülerinin analizinde başarılı sonuçlar elde ettiğini göstermektedir (Bhatti ve ark., 2025; Perera ve Thayasivam, 2025; Xia ve ark., 2025; Xiong ve ark., 2025). Bu çalışmalar, ConvNeXt mimarisinin farklı tıbbi görüntüleme problemlerinde güçlü bir öznitelik çıkarıcı olarak kullanılabileceğini ortaya koymaktadır. Bununla birlikte derin öğrenme tabanlı modellerin kritik sınırlılıklarından biri, yüksek performans elde edebilmek için büyük miktarda ve dengeli etiketlenmiş veriye ihtiyaç duymalarıdır. Bu gereksinim, özellikle tıbbi görüntü analizinde veri toplama ve uzman etiketleme süreçlerinin maliyetli olması nedeniyle önemli bir problem oluşturmaktadır. Bu sorunun çözümüne yönelik olarak geliştirilen Az Örnekli Öğrenme (Few-Shot Learning-FSL) yaklaşımları, sınırlı sayıda eğitim örneği kullanılarak etkili sınıflandırma yapılmasını amaçlamaktadır. Bu alanda önerilen Prototip Ağları (Prototypical Networks), her sınıfı temsil eden prototip vektörleri oluşturarak yeni örnekleri bu prototiplere olan benzerlikleri üzerinden sınıflandırmaktadır (Snell ve ark., 2017). Daha sonraki bir çalışmada Riasat ve ark. (Riasat ve ark., 2025), prototip tabanlı yaklaşımların tıbbi görüntü analizinde başarılı sonuçlar sağlayabildiğini göstermiştir.

Ancak literatürde yer alan prototip tabanlı yöntemlerin büyük bölümü az örnekli öğrenme senaryoları için geliştirilmiştir. Tam veri kullanılarak gerçekleştirilen derin öğrenme modellerinde ise prototip tabanlı karar mekanizmalarının kullanımı oldukça sınırlıdır. Ayrıca klasik evrimsel sinir ağı modellerinde nihai sınıflandırma yalnızca softmax tabanlı parametrik sınıflandırıcı ile gerçekleştirilmekte olup, sınıflar arasındaki benzerlik ilişkilerinden doğrudan yararlanılmamaktadır. Bu durum, özellikle tıbbi görüntülerde birbirine görsel olarak benzeyen örneklerin sınıflandırılmasında karar mekanizmasının kararlılığını olumsuz etkileyebilmektedir.

Bu tezde, söz konusu eksikliğin giderilmesi amacıyla ConvNeXt-Tiny mimarisi temel alınarak Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) geliştirilmiştir. Önerilen modelde ConvNeXt-Tiny tarafından elde edilen derin öznitelikler kullanılarak sınıf prototipleri oluşturulmuş, test aşamasında prototip tabanlı benzerlik hesaplamaları ile softmax tabanlı sınıflandırma sonuçları hibrit karar mekanizması içerisinde birleştirilmiştir. Böylece derin öğrenmenin güçlü öznitelik temsil yeteneği ile prototip tabanlı benzerlik yaklaşımının yorumlanabilirliği aynı model içerisinde bir araya getirilmiştir.

Bu tez kapsamında ayrıca önerilen HPFN modeli; Xception, VGG16 ve ResNet50 derin öğrenme mimarileri ile karşılaştırılmış, bunun yanında Ölçekten Bağımsız Öznitelik Dönüşümü (Scale Invariant Feature Transform-SIFT), Doğrusal Olmayan Ölçek Uzayında Hızlandırılmış Öznitelikler (KAZE) ve Yönlendirilmiş Gradyan Histogramı (Histogram of Oriented Gradients-HOG) öznitelik çıkarım yöntemleri kullanılarak oluşturulan geleneksel makine öğrenmesi modelleri de değerlendirilmiştir. Tüm deneyler aynı Chest X-Ray (Pneumonia) veri seti üzerinde gerçekleştirilmiş ve elde edilen sonuçlar doğruluk, kesinlik, duyarlılık ve F1-Skoru ölçütleri kullanılarak analiz edilmiştir (Kermany ve ark., 2018). Böylece önerilen yaklaşımın hem geleneksel öznitelik tabanlı yöntemlere hem de yaygın kullanılan derin öğrenme mimarilerine göre performansı kapsamlı olarak değerlendirilmiştir.

2. KAVRAMSAL ÇERÇEVE VE İLGİLİ ÇALIŞMALAR

Bu bölümde, az etiketli veri ile zatürre tanısında derin öğrenme yaklaşımlarının performans analizine temel oluşturan kavramsal ve kuramsal çerçeve sunulmaktadır. Bu kapsamda öncelikle zatürrenin klinik önemi ve göğüs röntgeni görüntülerinin tanı sürecindeki yeri ele alınmaktadır. Daha sonra derin öğrenme tabanlı tıbbi görüntü analizi, az etiketli veri problemi, transfer öğrenme, ConvNeXt mimarisi, az örnekli öğrenme ve prototip tabanlı yaklaşımlar açıklanmaktadır. Son olarak, göğüs röntgeni görüntülerinden zatürre tespitine yönelik ilgili çalışmalar değerlendirilerek mevcut tez çalışmasının literatürdeki yeri ortaya konulmaktadır.

2.1. Zatürre ve Göğüs Röntgeni Görüntülerinin Tanıdaki Yeri

Zatürre, özellikle çocuklar, yaşlı bireyler ve bağışıklık sistemi zayıf olan hastalarda ciddi sağlık sorunlarına yol açabilen önemli solunum yolu hastalıklarından biridir (World Health Organization, 2026). Bununla birlikte zatürre; ateş, öksürük, takipne, solunum sıkıntısı ve göğüs ağrısı gibi klinik bulgularla ortaya çıkabilmektedir. Zatürrenin klinik belirtileri ve fizik muayene bulguları hastalar arasında değişkenlik gösterebilmekte ve klinik değerlendirme tek başına tanının güvenilir biçimde doğrulanması ya da dışlanması için yeterli olmayabilmektedir (Wootton ve Feldman, 2014). Erken tanı ve uygun tedavi, hastalığın neden olabileceği komplikasyonların azaltılmasında kritik öneme sahiptir. Günümüzde göğüs röntgeni, zatürrenin teşhisinde en yaygın kullanılan görüntüleme yöntemlerinden biridir. Ancak CXR görüntülerinin yorumlanması; anatomik yapıların üst üste binmesi, görüntü kalitesinde değişkenlik ve radyologların farklı görüşler bildirmesi nedeniyle zorluk oluşturmaktadır (Elemraid ve ark., 2014; Hopstaken ve ark., 2004). Bu nedenle son yıllarda bilgisayar destekli tanı sistemleri ile derin öğrenme tabanlı yaklaşımların kullanımı önemli ölçüde artmıştır. CXR görüntüleme yönteminin düşük maliyetli, hızlı ve yaygın erişilebilir olması, klinik uygulamada tercih edilmesini kolaylaştırmaktadır (X. Wang ve ark., 2019). Zatürreye bağlı konsolidasyon, infiltrasyon, interstisyel tutulum ve plevral sıvı gibi radyolojik bulgular CXR görüntüleri üzerinden değerlendirilebilmektedir. Ancak bu bulguların görünürlüğü hastalığın evresine, görüntü kalitesine, çekim pozisyonuna, yaş grubuna ve eşlik eden akciğer patolojilerine göre değişiklik gösterebilmektedir (Elemraid ve ark., 2014). Bu sistemler, göğüs röntgeni görüntülerinden otomatik öznitelik öğrenebilme yetenekleri sayesinde radyologlara karar verme sürecinde destek sağlayabilmektedir (Rajpurkar ve ark., 2017; L. Wang ve ark., 2020). CXR görüntülerinin yorumlanması uzmanlık gerektiren bir süreçtir. Zatürreye ilişkin radyolojik bulguların bazı olgularda silik veya diğer hastalıklarla benzer görünümde olması

tanısal değerlendirmeyi güçleştirebilmektedir (Hopstaken ve ark., 2004). Bu durum, CXR görüntülerinin yorumlanmasında standartlaştırılmış, yardımcı ve tekrarlanabilir karar destek sistemlerine duyulan gereksinimi artırmaktadır.

2.2. Derin Öğrenme Tabanlı Tıbbi Görüntü Analizi

Derin öğrenme, çok katmanlı yapay sinir ağları aracılığıyla veriden farklı soyutlama düzeylerinde temsil öğrenmeyi amaçlayan bir makine öğrenmesi yaklaşımıdır (LeCun ve ark., 2015). Görüntü sınıflandırma alanında derin öğrenmenin yaygınlaşmasında, büyük ölçekli görüntü veri kümeleri üzerinde eğitilen evrişimsel sinir ağlarının yüksek başarı göstermesi belirleyici olmuştur (Krizhevsky ve ark., 2012). Evrişimsel sinir ağları, görüntülerdeki yerel örüntüleri evrişim katmanları aracılığıyla işleyerek düşük düzeyli kenar ve doku özelliklerinden daha karmaşık görsel temsillere doğru hiyerarşik bir öğrenme süreci oluşturabilmektedir (LeCun ve ark., 2015). Bu özellik, tıbbi görüntü analizinde elle tasarlanmış özniteliklere duyulan gereksinimi azaltmakta ve modelin ayırt edici görüntü temsillerini doğrudan veriden öğrenmesine olanak sağlamaktadır (Shen ve ark., 2017).

Tıbbi görüntü analizi alanında derin öğrenme yöntemleri; görüntü sınıflandırma, lezyon tespiti, segmentasyon, kayıt işlemleri ve bilgisayar destekli tanı sistemleri gibi farklı görevlerde yaygın olarak kullanılmaktadır (Litjens ve ark., 2017). Göğüs röntgeni görüntüleri, akciğer alanı, kemik yapılar, yumuşak dokular ve patolojik opasiteler gibi çok sayıda anatomik ve radyolojik bileşeni aynı görüntü içinde barındırdığı için otomatik analiz açısından karmaşık bir veri türüdür. Bu nedenle CNN tabanlı modeller, göğüs röntgeni görüntülerinde normal akciğer görünümü ile hastalığa bağlı radyolojik örüntüleri ayırt etmek amacıyla sık kullanılan yöntemler arasında yer almaktadır (Rajpurkar ve ark., 2017).

Göğüs röntgenlerinden zatürre tespitine yönelik derin öğrenme modelleri, zatürre etiketiyle ilişkili radyografik örüntüleri görüntülerden otomatik olarak öğrenmek üzere geliştirilmektedir (Garg ve ark., 2019; Rajpurkar ve ark., 2017). Bununla birlikte infiltrasyon ve opasite gibi bulgular zatürreye özgü değildir; benzer radyolojik görünümüler atelektazi, pulmoner ödem, hemoraji veya diğer akciğer patolojilerinde de görülebilmektedir (Franquet ve ark., 2019). CXR veri kümelerinde radyoloji raporlarından türetilen etiketler belirsizlik ve etiket gürültüsü içerebilmektedir. CheXpert bu belirsizlikleri özel etiketleme stratejileriyle modellemekte, MIMIC-CXR ise görüntüleri serbest metin radyoloji raporlarıyla birlikte sunmaktadır. Model genellenebilirliğinin değerlendirilmesi için farklı kurum veya veri kümelerinde dış doğrulama

yapılması ayrıca gereklidir (Irvin ve ark., 2019; Johnson ve ark., 2019). Bu bağlamda ChestX-ray8/ChestX-ray14, CheXpert ve MIMIC-CXR gibi büyük ölçekli veri kümeleri, CXR tabanlı derin öğrenme modellerinin geliştirilmesi, karşılaştırılması ve genellenebilirliğinin incelenmesi için önemli bir araştırma zemini sağlamıştır (Irvin ve ark., 2019; Johnson ve ark., 2019; X. Wang ve ark., 2019).

ChestX-ray8/ChestX-ray14 veri seti, raporlardan doğal dil işleme yöntemleriyle elde edilen etiketlere sahip geniş ölçekli göğüs röntgeni görüntülerini içermesi nedeniyle CXR tabanlı derin öğrenme çalışmalarında önemli bir başlangıç noktası oluşturmıştır (X. Wang ve ark., 2019). CheXpert veri seti, radyoloji raporlarından çıkarılan etiketlerdeki belirsizlikleri ayrıca modellemeye çalışması nedeniyle CXR sınıflandırma araştırmalarında önemli bir referans olarak kabul edilmektedir (Irvin ve ark., 2019). MIMIC-CXR veri seti ise kimliksizleştirilmiş göğüs röntgeni görüntülerini serbest metin radyoloji raporlarıyla birlikte sunması bakımından tıbbi yapay zekâ araştırmaları için kapsamlı bir açık veri kaynağı sağlamıştır (Johnson ve ark., 2019). Bu veri kümeleri, derin öğrenme modellerinin yalnızca tekil hastalık sınıflandırması için değil, çok etiketli toraks patolojilerinin otomatik analizi için de kullanılmasına olanak tanımıştır.

Zatürre tespiti özelinde en çok bilinen çalışmalardan biri CheXNet modelidir. Rajpurkar ve arkadaşları, DenseNet-121 mimarisine dayalı CheXNet modelini ChestX-ray14 veri seti üzerinde eğiterek göğüs röntgeni görüntülerinden zatürre tespitini amaçlamıştır. Çalışmada model performansı radyolog değerlendirmeleriyle karşılaştırılmış ve modelin belirli test koşullarında F1-Skoru açısından radyolog ortalamasına yakın ya da onu aşan sonuçlar verdiği bildirilmiştir (Rajpurkar ve ark., 2017). Bu çalışma, CXR tabanlı zatürre tespitinde derin öğrenmenin potansiyelini göstermesi açısından önemli olmakla birlikte, sonuçların kullanılan veri seti, etiketleme yöntemi, test setinin yapısı ve dış doğrulama koşullarıyla birlikte değerlendirilmesi gerekmektedir.

Sonuç olarak, derin öğrenme tabanlı CXR analizi, zatürre gibi toraks hastalıklarının otomatik tespitinde güçlü bir yöntemsel temel sunmaktadır. Bununla birlikte tıbbi görüntülerde model başarısı yalnızca yüksek doğruluk veya F1-Skoru ile açıklanamamalıdır. Etiket kalitesi, hasta düzeyinde veri ayrımı, sınıf dengesizliği, dış doğrulama, veri setine özgü yanlılıklar ve klinik yorumlanabilirlik gibi faktörler modelin gerçek kullanım değerini doğrudan etkilemektedir (Irvin ve ark., 2019; Johnson ve ark., 2019). Bu nedenle az etiketli veri ile zatürre tanısına

yönelik çalışmalarda, derin öğrenme modelinin performansının yanı sıra veri verimliliği, genellenebilirlik ve klinik güvenilirlik boyutlarının birlikte değerlendirilmesi gerekmektedir.

2.3. Az Etiketli Veri Problemi

Derin öğrenme modelleri, yüksek performans elde edebilmek için çoğunlukla büyük miktarda etiketli veriye ihtiyaç duymaktadır (LeCun ve ark., 2015). Ancak tıbbi görüntüleme alanında yüksek kaliteli etiket üretimi, radyoloji uzmanlığı gerektirmesi nedeniyle zaman alıcı ve maliyetli bir süreçtir (Litjens ve ark., 2017). Göğüs röntgeni görüntülerinde doğru referans etiketlerin oluşturulması, yalnızca görüntünün teknik olarak değerlendirilmesini değil, klinik bağlamın ve radyolojik bulguların uzman yorumu ile ilişkilendirilmesini de gerektirmektedir (Irvin ve ark., 2019). Bu nedenle sınırlı uzman kaynağına sahip araştırma ortamlarında yeterli sayıda güvenilir etiketli CXR verisi oluşturmak güçleşmektedir.

Az etiketli veri problemi yalnızca görüntü sayısının sınırlı olmasıyla açıklanamaz. Tıbbi görüntü veri kümelerinde etiketlerin belirsiz, eksik veya zayıf biçimde üretilmiş olması da modelin öğrenme sürecini ve performansını etkileyebilmektedir (Cheplygina ve ark., 2019). ChestX-ray8/ChestX-ray14 veri setinde hastalık etiketlerinin radyoloji raporlarından doğal dil işleme yöntemleriyle çıkarılması, geniş ölçekli veri üretimini kolaylaştırmış; ancak bu yaklaşım zayıf etiketleme ve etiket gürültüsü sorunlarını da beraberinde getirmiştir (X. Wang ve ark., 2019). CheXpert veri setinde ise radyoloji raporlarından çıkarılan etiketlerdeki belirsizlikler ayrıca modellenmiş ve belirsiz etiketlerin CXR sınıflandırma performansı üzerindeki etkisi değerlendirilmiştir (Irvin ve ark., 2019). Bu durum, CXR tabanlı derin öğrenme çalışmalarında etiket kalitesinin model başarısı kadar önemli bir metodolojik unsur olduğunu göstermektedir.

Veri setine ilişkin bir diğer önemli sorun, sınıf dengesizliği ve alan kaymasıdır. Tıbbi görüntü veri kümelerinde bazı patolojiler diğerlerine göre daha az temsil edilebilmekte ve bu durum modelin az görülen sınıfları öğrenmesini güçleştirebilmektedir (Johnson ve ark., 2019). Ayrıca farklı cihazlar, görüntüleme protokolleri ve kurumlara ait veri dağılımları arasındaki farklılıklar, modelin yeni merkezlerdeki genelleme kapasitesini sınırlandırabilmektedir (Pooch ve ark., 2020). Bu nedenle az etiketli CXR çalışmalarında veri miktarının yanı sıra etiket güvenilirliği, sınıf dağılımı, hasta düzeyinde veri ayrımı ve dış doğrulama tasarımı birlikte değerlendirilmelidir (Roberts ve ark., 2021).

Az etiketli veri koşullarında geliştirilen modellerde aşırı öğrenme önemli bir risk olarak ortaya çıkmaktadır. Model, sınırlı sayıdaki eğitim örneğini öğrenmesine rağmen yeni görüntüler üzerinde düşük performans gösterebilir (W. Y. Chen ve ark., 2019). Bu nedenle az etiketli veriyle yapılan tıbbi görüntü sınıflandırma çalışmalarında veri artırma, transfer öğrenme, düzenlileştirme, çapraz doğrulama ve hasta düzeyinde veri ayrımı gibi yöntemler önem kazanmaktadır. Model değerlendirmesinde doğruluk oranının yanı sıra duyarlılık, özgüllük, kesinlik, F1-Skoru ve ROC-AUC gibi ölçütlerin birlikte raporlanması önerilmektedir (Irvin ve ark., 2019).

2.4. Transfer Öğrenme Yaklaşımı

Transfer öğrenme, bir kaynak görev veya veri seti üzerinde öğrenilen bilginin, ilişkili başka bir hedef görevde kullanılmasını amaçlayan bir makine öğrenmesi yaklaşımıdır (Pan ve Yang, 2010). Derin öğrenme bağlamında bu yaklaşım, çoğunlukla büyük ölçekli görüntü veri kümeleri üzerinde önceden eğitilmiş CNN modellerinin hedef görüntü sınıflandırma problemine uyarlanması şeklinde uygulanmaktadır (Kornblith ve ark., 2019). Bu yöntem, özellikle etiketli verinin sınırlı olduğu tıbbi görüntüleme problemlerinde modelin sıfırdan eğitilmesine kıyasla daha kararlı bir başlangıç noktası sağlayabilmektedir (Tajbakhsh ve ark., 2016).

Tıbbi görüntü analizinde transfer öğrenmenin yaygın kullanılmasının temel nedenlerinden biri, yüksek kaliteli etiketli tıbbi görüntü verisi oluşturma uzmanlık, zaman ve maliyet gerektirmesidir (Tajbakhsh ve ark., 2016). Ön-eğitilmiş CNN modelleri, erken katmanlarda kenar, doku ve basit şekil gibi genel görsel özellikleri öğrenebildiğinden, bu temsiller bazı tıbbi görüntü sınıflandırma görevlerine aktarılabilir (Shin ve ark., 2016). Bununla birlikte aktarım başarısı yalnızca kaynak veri setindeki performansa bağlı değildir; hedef veri setinin yapısı, görüntü modalitesi, sınıf ayrımı, veri miktarı ve ince ayar stratejisi de transfer öğrenme performansını belirlemektedir (Kornblith ve ark., 2019).

Transfer öğrenme yaklaşımı, VGG, ResNet, DenseNet, Xception ve EfficientNet gibi ön-eğitilmiş CNN mimarilerinin tıbbi görüntü sınıflandırma görevlerinde kullanılmasına olanak sağlamıştır. Shin ve arkadaşları, bilgisayarlı tomografi ve radyografi görüntülerinde ön-eğitilmiş CNN mimarilerinin bilgisayar destekli tanı problemlerine uyarlanabileceğini göstermiştir (Shin ve ark., 2016). Tajbakhsh ve arkadaşları, tıbbi görüntü analizinde yeterli ince ayar yapıldığında ön-eğitilmiş CNN modellerinin sıfırdan eğitilen modellere göre daha dayanıklı sonuçlar

verebildiğini bildirmiştir (Tajbakhsh ve ark., 2016). Kermany ve arkadaşları ise pediatrik göğüs röntgeni görüntüleri üzerinde derin öğrenme kullanarak normal ve zatürre sınıflarını ayırt etmeye yönelik bir uygulama sunmuştur (Kermany ve ark., 2018).

Bununla birlikte doğal görüntüler üzerinde ön-eğitilmiş modellerin tıbbi görüntülere aktarılması her zaman doğrudan klinik genellenebilirlik sağlamamaktadır. Doğal görüntüler ile radyolojik görüntüler arasında görüntü oluşum süreci, anatomik içerik, doku özellikleri ve patolojik örüntülerin görünümü bakımından belirgin farklılıklar bulunmaktadır (Tajbakhsh ve ark., 2016). ImageNet Large Scale Visual Recognition Dataset- Büyük Ölçekli Görsel Tanıma Veri Seti (ImageNet) üzerinde daha yüksek performans gösteren modellerin farklı hedef veri kümelerine aktarımında avantaj sağlayabildiği, ancak transfer başarısının hedef veri setinin niteliği ve görev özellikleriyle yakından ilişkili olduğu bildirilmiştir (Kornblith ve ark., 2019). Bu nedenle transfer öğrenme çalışmalarında model mimarisinin yanı sıra ön işleme yöntemi, dondurulan katman sayısı, ince ayar derinliği, öğrenme oranı, veri artırma stratejisi ve doğrulama tasarımı dikkatle belirlenmelidir (Shin ve ark., 2016).

2.5. ConvNeXt Mimarisi

ConvNeXt, evrimsel sinir ağlarının güçlü yönlerini korurken güncel derin öğrenme mimarilerindeki bazı tasarım ilkelerini CNN yapısına uyarlayan modern bir mimari ailesidir. Bu yaklaşımın geliştirilmesinde, derin CNN'lerin eğitimini kolaylaştıran artık bağlantılara dayalı ResNet mimarisi önemli bir başlangıç noktası oluşturmuştur (He ve ark., 2016). Görüntü sınıflandırma alanında Vision Transformer modellerinin güçlü performans göstermesi, evrimsel mimarilerin tasarım ilkelerinin yeniden değerlendirilmesine zemin hazırlamıştır (Dosovitskiy ve ark., 2021). ConvNeXt mimarisi ise standart ResNet yapısının aşamalı olarak modernleştirilmesiyle geliştirilmiş ve saf evrimsel bir ağın güncel görsel tanıma görevlerinde rekabetçi performans gösterebileceğini ortaya koymuştur. ConvNeXt mimarisinde büyük çekirdekli derinlik yönlü evrimseler, ters darboğaz yapısı, katman normalizasyonu ve GELU aktivasyonu gibi güncel tasarım bileşenleri kullanılmaktadır (Z. Liu ve ark., 2022).

Büyük çekirdekli evrimseler, modelin daha geniş uzamsal bağlamı değerlendirebilmesine katkı sağlarken, derinlik yönlü evrimseler hesaplama maliyetini sınırlayarak kanal bazlı özellik çıkarımını desteklemektedir (Chollet, 2017). Katman normalizasyonu ve GELU aktivasyonu gibi bileşenler ise modern derin öğrenme mimarilerinde eğitim kararlılığını ve temsil öğrenme kapasitesini artırmak amacıyla yaygın biçimde kullanılmaktadır (Ba ve ark., 2016; Hendrycks

ve Gimpel, 2023). Bu özellikleri nedeniyle ConvNeXt, yerel örüntü çıkarımı ile daha geniş bağlamsal temsil öğrenimini birlikte destekleyen güçlü bir CNN tabanlı omurga mimarisi olarak değerlendirilebilir (A. Liu ve Yue, 2023). ConvNeXt-Tiny, ConvNeXt ailesinin daha düşük parametrelili ve hesaplama maliyeti daha yönetilebilir üyelerinden biridir (A. Liu ve Yue, 2023). Bu özellik, sınırlı hesaplama kaynaklarının bulunduğu deneysel ortamlarda ve az etiketli veriyle yürütülen çalışmalarda ConvNeXt-Tiny'yi uygulanabilir bir omurga mimarisi hâline getirmektedir. Az etiketli veriyle tıbbi görüntü sınıflandırmada güçlü fakat aşırı karmaşık olmayan bir öznitelik çıkarıcı kullanılması, modelin öğrenme kapasitesi ile aşırı öğrenme riski arasında denge kurulması açısından önemlidir (W. Y. Chen ve ark., 2019; Tajbakhsh ve ark., 2016).

2.6. Az Örnekli Öğrenme ve Prototip Tabanlı Yaklaşımlar

Az örnekli öğrenme, modelin her sınıf için sınırlı sayıda etiketli örnekle sınıflandırma yapabilmesini amaçlayan bir öğrenme paradigmasıdır. Bu yaklaşım, etiketli veri toplamının zor ve maliyetli olduğu tıbbi görüntüleme alanı için önemlidir. Az örnekli öğrenme yöntemleri, sınırlı veri koşullarında modelin genelleme kapasitesini artırmayı hedeflemektedir (W. Y. Chen ve ark., 2019).

Prototypical Networks, az örnekli öğrenme alanında kullanılan temel yaklaşımlardan biridir. Bu yöntemde her sınıf, o sınıfa ait örneklerden elde edilen özniteliklerin ortalamasıyla oluşturulan bir prototip vektörü ile temsil edilir. Sorgu örneği, bu prototip vektörlerine olan uzaklığına göre sınıflandırılır (Snell ve ark., 2017). Bu yapı, sınıflandırma kararını doğrudan son katman ağırlıklarına değil, sınıfların temsil merkezlerine dayandırmaktadır.

Prototip tabanlı yaklaşımlar, az etiketli veri koşullarında sınıflar arası ayrımın daha açık biçimde modellenmesine yardımcı olabilir. Sınırlı sayıda görüntü bulunan durumlarda, klasik sınıflandırıcılar sınıf sınırlarını yeterince kararlı biçimde öğrenemeyebilir (W. Y. Chen ve ark., 2019). Prototip tabanlı yöntemler ise sınıf merkezlerine dayalı karar verdiği için temsil uzayındaki benzerlik ilişkilerini daha açık biçimde kullanabilmektedir (Snell ve ark., 2017). Bu özellik, tıbbi görüntü sınıflandırmada hem veri verimliliği hem de yorumlanabilirlik açısından avantaj sağlayabilir.

Göğüs röntgeni görüntülerinden zatürre tespitinde prototip tabanlı yaklaşımların kullanılması, normal ve zatürre sınıflarının derin öznitelik uzayında daha belirgin ayrıştırılmasına katkı

sunabilir. ConvNeXt-Tiny gibi güçlü bir öznitelik çıkarıcı mimari ile prototip tabanlı karar mekanizmasının birleştirilmesi, az etiketli veri koşullarında hem temsil öğrenme hem de benzerlik temelli sınıflandırma avantajlarını bir araya getirebilir. Bu nedenle bu tez kapsamında önerilen Hibrit Prototip Füzyon Ağı, ConvNeXt-Tiny tabanlı derin öznitelik çıkarımı ile prototip tabanlı karar verme sürecini bütünleştiren bir yapı olarak ele alınmaktadır.

2.7. Göğüs Röntgeni Görüntülerinde Zatürre Tespitine Yönelik Çalışmalar

Göğüs röntgeni görüntülerinden zatürre tespitine yönelik literatürde CNN tabanlı sınıflandırıcılar, transfer öğrenme yaklaşımları, topluluk öğrenme modelleri ve dikkat mekanizmalı yapılar kullanılmıştır. CheXNet çalışması, CXR görüntülerinden zatürre tespitinde derin öğrenme tabanlı yaklaşımların önemli örneklerinden biridir (Rajpurkar ve ark., 2017). Kermany ve arkadaşları, pediatrik CXR görüntülerinde derin öğrenme ile normal ve zatürre sınıflarının ayrıştırılabileceğini göstermiştir (Kermany ve ark., 2018). Mabrouk ve arkadaşları, farklı derin öğrenme mimarilerini bir araya getiren topluluk yaklaşımıyla zatürre sınıflandırmasına yönelik bir model önermiştir (Mabrouk ve ark., 2023). Bu çalışmalar, derin öğrenme tabanlı yaklaşımların göğüs röntgeni görüntülerinden zatürre tespitinde yüksek potansiyel taşıdığını göstermektedir. Ancak bildirilen performans değerleri, kullanılan veri setinin yapısına, etiket kalitesine, eğitim-test ayırımına ve dış doğrulama yapılıp yapılmadığına bağlı olarak değişebilmektedir (Irvin ve ark., 2019). Özellikle hasta düzeyinde ayırım yapılmayan veri bölme stratejileri, model performansının olduğundan yüksek görünmesine neden olabilir (Johnson ve ark., 2019). Bu nedenle literatürdeki sonuçlar değerlendirilirken model başarısının yanı sıra deneysel tasarımın güvenilirliği de dikkate alınmalıdır.

Zatürre tespiti çalışmalarında doğruluk oranı sık kullanılan bir performans ölçütüdür. Ancak sınıf dengesizliği bulunan veri kümelerinde doğruluk oranı tek başına yanıltıcı olabilir. Bu nedenle duyarlılık, özgüllük, kesinlik, F1-Skoru, ROC-AUC ve karışıklık matrisi gibi ölçütlerin birlikte raporlanması daha kapsamlı bir değerlendirme sağlar (Irvin ve ark., 2019). Klinik açıdan özellikle yanlış negatif sonuçlar önemli olduğundan, modelin zatürre olgularını kaçırma olasılığı dikkatle değerlendirilmelidir.

2.8. Etik İlkeler ve Klinik Güvenilirlik

Tıbbi görüntülerle yürütülen yapay zekâ çalışmalarında etik ilkeler araştırma tasarımının temel bileşenlerinden biri olmalıdır. Göğüs röntgeni görüntüleri doğrudan yüz görüntüsü içermese de

görüntü üzerinde yer alan hasta bilgileri, kurum işaretleri, tarih bilgileri veya meta veriler dolaylı kimliklendirme riski oluşturabilir (Johnson ve ark., 2019). Bu nedenle veri anonimleştirme veya kimliksizleştirme süreçlerinin dikkatle yürütülmesi gerekmektedir. Ayrıca etik kurul onayı, veri kullanım izni, güvenli veri saklama ve erişim kontrolü süreçleri açık biçimde tanımlanmalıdır.

Yapay zekâ tabanlı tanı sistemleri, klinik karar destek aracı olarak değerlendirilmelidir. Bu sistemler radyolog veya klinisyen kararının yerine geçmekten ziyade, karar verme sürecini destekleyen yardımcı araçlar olarak konumlandırılmalıdır. Az etiketli veriyle geliştirilen modellerde aşırı öğrenme, yanlılık, sınıf dengesizliği ve dış veri kümelerinde performans kaybı gibi riskler daha belirgin hâle gelebilir (Irvin ve ark., 2019). Bu nedenle model çıktılarının klinik bağlamdan bağımsız şekilde kesin tanı olarak yorumlanması uygun değildir.

Model performansının farklı hasta gruplarında, farklı cihazlarda ve farklı klinik merkezlerde benzer sonuçlar verip vermediği de etik ve klinik güvenilirlik açısından önemlidir. Bir modelin yalnızca belirli bir veri setinde yüksek başarı göstermesi, gerçek klinik ortamlarda aynı başarıyı göstereceği anlamına gelmez (Johnson ve ark., 2019). Bu nedenle geliştirilen modellerin dış doğrulama, hata analizi ve gerektiğinde açıklanabilir yapay zekâ yöntemleriyle desteklenmesi önerilmektedir.

2.9. Çalışmanın Literatürdeki Yeri

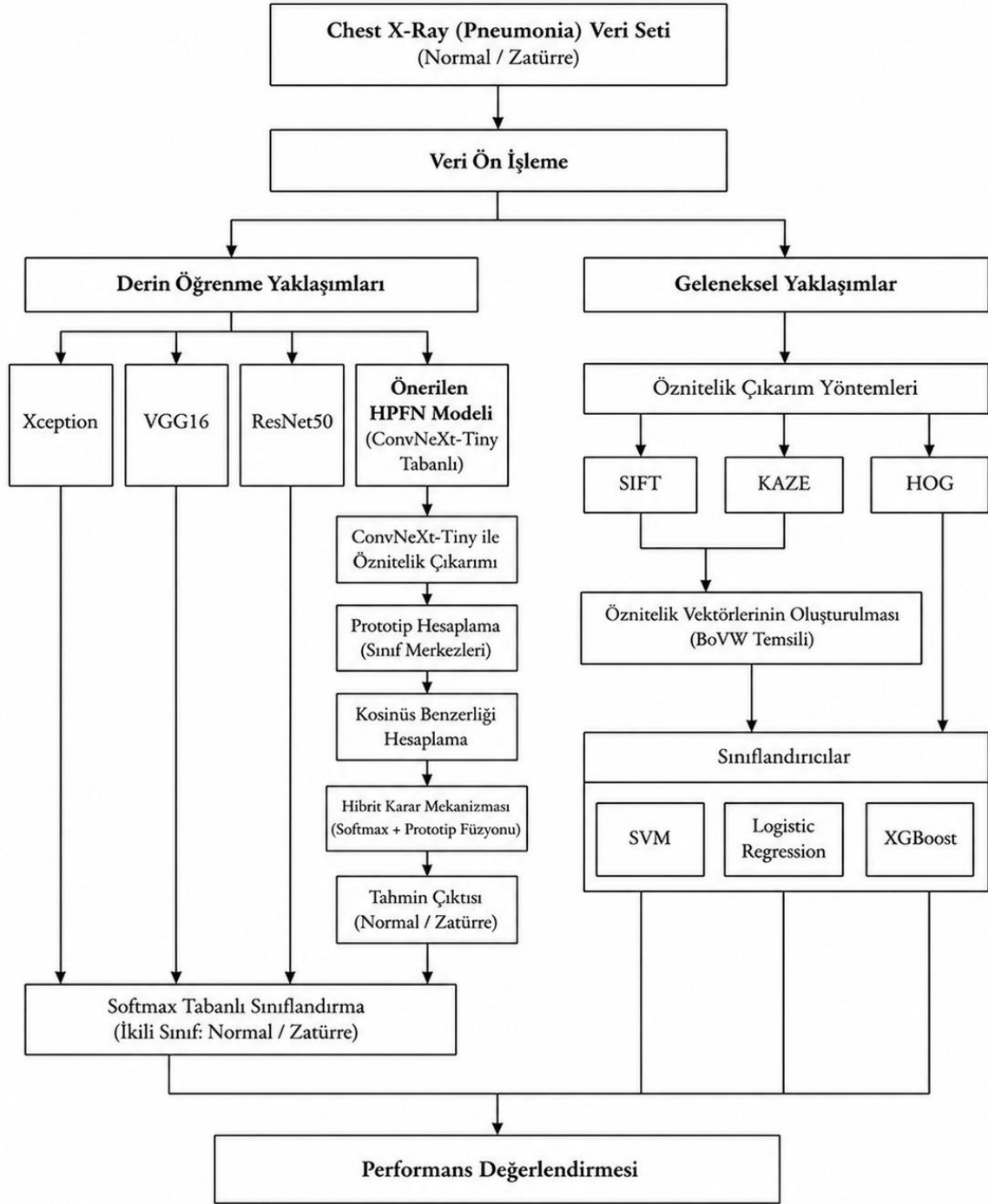
Mevcut literatür incelendiğinde, göğüs röntgeni görüntülerinden zatürre tespitinde derin öğrenme tabanlı yöntemlerin yaygın olarak kullanıldığı görülmektedir. Bununla birlikte çalışmaların önemli bir bölümü büyük ölçekli veri kümelerine, klasik transfer öğrenme yaklaşımlarına veya standart sınıflandırıcı katmanlarına dayanmaktadır (Kermany ve ark., 2018; Mabrouk ve ark., 2023; Rajpurkar ve ark., 2017). Az etiketli tıbbi görüntü sınıflandırma çalışmalarında modelin sınırlı sayıdaki örnekten genellenebilir temsiller öğrenebilmesi temel bir yöntemsel sorun olarak öne çıkmaktadır (Cheplygina ve ark., 2019). Bu soruna yönelik yaklaşımlardan biri, güçlü bir CNN tabanlı öznetelik çıkarıcı ile benzerlik temelli karar mekanizmalarını birlikte kullanmaktır. Modern CNN mimarileri, görüntülerden ayırt edici derin temsiller elde etmek için güçlü bir omurga yapı sunarken, prototip tabanlı yöntemler sınıf merkezlerine dayalı daha veri-verimli bir sınıflandırma yaklaşımı sağlamaktadır (Z. Liu ve ark., 2022; Snell ve ark., 2017). Bununla birlikte literatürde, göğüs röntgenlerinden zatürre tespitinde CNN ve transfer öğrenme tabanlı yaklaşımların yaygın olduğu; buna karşılık ConvNeXt-Tiny

tabanlı öznitelik çıkarımı ile prototip temelli karar mekanizmasını doğrudan bütünleştiren çalışmaların daha sınırlı biçimde ele alındığı görülmektedir.

Bu tez çalışmasında, az etiketli veri koşullarında göğüs röntgeni görüntülerinden zatürre tanısına yönelik derin öğrenme yaklaşımlarının performansının analiz edilmesi amaçlanmaktadır. Zatürre tespitinde derin öğrenme modellerinin CXR görüntülerinden patolojik örüntüleri ayırt edebilme potansiyeli daha önce farklı model mimarileriyle gösterilmiştir (Kermany ve ark., 2018; Rajpurkar ve ark., 2017). Bu çalışmada ise ConvNeXt-Tiny tabanlı derin öznitelik çıkarımı ile prototip tabanlı karar mekanizmasını birleştiren Hibrit Prototip Füzyon Ağı yaklaşımı önerilmektedir. Önerilen yaklaşım, modern CNN temelli temsil öğrenme kapasitesi ile az örnekli öğrenmede kullanılan sınıf prototipi mantığını bir araya getirerek az etiketli CXR verilerinde sınıflandırma performansını ve prototip temelli karar mekanizmasının katkısını değerlendirmeyi hedeflemektedir.

3. YÖNTEM

Bu tez kapsamında, az etiketli veri kullanılarak göğüs röntgeni görüntülerinden zatürre tanısına yönelik farklı derin öğrenme ve geleneksel makine öğrenmesi yaklaşımlarının performansları karşılaştırmalı olarak incelenmiştir. Çalışmada ilk olarak Chest X-Ray (Pneumonia) veri seti kullanılarak veri ön işleme işlemleri gerçekleştirilmiş ve görüntüler derin öğrenme modellerinin eğitimine uygun hale getirilmiştir. Daha sonra ConvNeXt-Tiny tabanlı önerilen Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) geliştirilmiş ve aynı veri seti üzerinde Xception, VGG16 ve ResNet50 derin öğrenme mimarileri ile karşılaştırılmıştır. Ayrıca SIFT, KAZE ve HOG öznitelik çıkarım yöntemleri kullanılarak elde edilen öznitelikler Destek Vektör Makineleri (Support Vector Machine-SVM), Lojistik Regresyon (Logistic Regression-LR) ve Aşırı Gradyan Artırma (Extreme Gradient Boosting-XGBoost) sınıflandırıcıları ile değerlendirilmiş ve geleneksel yaklaşımların performansları analiz edilmiştir. Son aşamada ise tüm modeller doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1-Skoru performans ölçütleri kullanılarak değerlendirilmiş, elde edilen sonuçlar aynı veri seti üzerinde gerçekleştirilen literatür çalışmaları ile karşılaştırılmıştır.



Şekil 1. Çalışmanın genel iş akış diyagramı

Çalışmada izlenen metodolojiye ilişkin genel iş akış diyagramı Şekil 1’de sunulmuştur. Uygulanan yöntemler; veri ön işleme, derin öğrenme tabanlı yaklaşımlar, geleneksel öznitelik tabanlı yaklaşımlar ve performans değerlendirme aşamalarından oluşmaktadır. İş akış diyagramında özetlenen yöntemler aşağıdaki alt başlıklarda ayrıntılı olarak açıklanmıştır.

3.1. Geleneksel Öznitelik Tabanlı Yaklaşımlar

Bu çalışmada, derin öğrenme tabanlı yaklaşımların performansını karşılaştırmalı olarak değerlendirebilmek amacıyla geleneksel öznitelik tabanlı yöntemlerden de yararlanılmıştır. Bu kapsamda göğüs röntgeni görüntülerinden SIFT, KAZE ve HOG yöntemleri kullanılarak öznitelikler çıkarılmıştır. Elde edilen öznitelik vektörleri ise Destek Vektör Makineleri (SVM), Lojistik Regresyon (LR) ve Aşırı Gradyan Artırma (XGBoost) sınıflandırıcıları kullanılarak sınıflandırılmıştır. Böylece el yapımı özniteliklere dayalı geleneksel yaklaşımların sınıflandırma performansları, derin öğrenme tabanlı yöntemler ile aynı veri seti üzerinde karşılaştırmalı olarak değerlendirilmiştir.

3.1.1. Ölçekten bağımsız öznitelik dönüşümü

Ölçekten Bağımsız Öznitelik Dönüşümü (SIFT), görüntülerde ölçek ve dönme değişimlerinden büyük ölçüde etkilenmeden ayırt edici yerel anahtar noktaların belirlenmesini sağlayan öznitelik çıkarım yöntemlerinden biridir (Lowe, 1999). Yöntem; ölçek uzayının oluşturulması, aday anahtar noktaların belirlenmesi, anahtar noktaların hassas konumlarının hesaplanması, yön tayini ve öznitelik tanımlayıcılarının oluşturulması olmak üzere temel aşamalardan oluşmaktadır.

Ölçek uzayı, görüntünün farklı ölçeklerde Gauss filtresi ile bulanıklaştırılması sonucu elde edilir.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

Denklem 1’de $I(x, y)$ giriş görüntüsünü, $G(x, y, \sigma)$ iki boyutlu Gauss fonksiyonunu ve $L(x, y, \sigma)$ ise ölçek uzayındaki görüntüyü göstermektedir.

Anahtar noktaların belirlenmesinde Gauss Farkı (Difference of Gaussian, DoG) yaklaşımı kullanılmaktadır (Denklem 2).

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2)$$

Elde edilen yerel anahtar noktalar için yön bilgisi hesaplanmakta ve her anahtar nokta çevresindeki gradyan dağılımı kullanılarak 128 boyutlu SIFT tanımlayıcısı oluşturulmaktadır. Bu çalışmada elde edilen SIFT öznitelikleri daha sonra makine öğrenmesi tabanlı sınıflandırıcılara giriş olarak kullanılmıştır.

3.1.2. Kaze

KAZE, doğrusal olmayan ölçek uzayında anahtar nokta çıkarımı gerçekleştiren bir öznitelik çıkarım yöntemidir (Fernández Alcantarilla ve ark., 2012). SIFT ve SURF yöntemlerinden farklı olarak doğrusal olmayan difüzyon filtreleri kullanması sayesinde görüntü kenarlarını koruyarak daha ayırt edici yerel öznitelikler elde edebilmektedir.

Doğrusal olmayan ölçek uzayı aşağıdaki difüzyon denklemi ile tanımlanmaktadır (Denklem 3).

$$\frac{\partial L}{\partial t} = \text{div}(c(x, y, t)\nabla L) \quad (3)$$

Burada L ölçek uzayındaki görüntüyü, $c(x, y, t)$ difüzyon katsayısını, ∇L gradyan operatörünü ve $\text{div}(\cdot)$ diverjans işlemini ifade etmektedir.

Anahtar noktalar doğrusal olmayan ölçek uzayında belirlenmekte ve her anahtar nokta için yerel tanımlayıcı oluşturulmaktadır. Bu çalışmada elde edilen KAZE öznitelikleri sınıflandırma aşamasında kullanılmıştır.

3.1.3. Yönlendirilmiş gradyan histogramı

Yönlendirilmiş Gradyan Histogramı (HOG), görüntülerdeki kenar ve gradyan yönelimlerinin dağılımını kullanarak öznitelik çıkaran bir yöntemdir (Dalal ve Triggs, 2005). HOG yaklaşımında görüntü küçük hücrelere bölünmekte, her hücre için gradyan büyüklüğü ve yönü hesaplanmakta ve yönelim histogramları oluşturulmaktadır. Daha sonra komşu hücreler bloklar halinde normalize edilerek aydınlatma değişimlerinin etkisi azaltılmaktadır.

Gradyan büyüklüğü Denklem 4' te verilen ifade ile hesaplanmaktadır.

$$M(x, y) = \sqrt{G_x^2 + G_y^2} \quad (4)$$

Gradyan yönü ise Denklem 5' te verilen ifade ile hesaplanmaktadır.

$$\theta(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (5)$$

Hesaplanan histogramlar birleştirilerek görüntüyü temsil eden HOG öznitelik vektörü oluşturulmaktadır. Bu öznitelikler çalışmada makine öğrenmesi tabanlı sınıflandırıcıların giriş verisi olarak kullanılmıştır.

3.1.4. Destek vektör makineleri

Destek vektör makineleri, farklı sınıflara ait örnekleri maksimum marjin ile ayıran optimum hiper düzlemi belirlemeye dayanan denetimli bir sınıflandırma yöntemidir (Cortes ve Vapnik, 1995). İkili sınıflandırma problemi aşağıdaki karar fonksiyonu ile ifade edilmektedir (Denklem 6).

$$f(x) = w^T x + b \quad (6)$$

Burada w hiper düzlemin ağırlık vektörünü, b ise sabit terimi göstermektedir.

3.1.5. Lojistik regresyon

Lojistik regresyon, giriş özniteliklerini kullanarak bir örneğin belirli bir sınıfa ait olma olasılığını hesaplayan doğrusal tabanlı bir sınıflandırma yöntemidir (Hosmer ve ark., 2013). Modelde doğrusal birleşim sonucu elde edilen değer sigmoid fonksiyonu ile 0 ile 1 arasında bir olasılık değerine dönüştürülmektedir. Bu olasılık değeri Denklem (7)' de verilen sigmoid fonksiyonu kullanılarak hesaplanmaktadır.

$$P(y = 1 | x) = \frac{1}{1+e^{-z}} \quad (7)$$

Burada $P(y = 1 | x)$ giriş örneğinin pozitif sınıfa ait olma olasılığını, z ise özniteliklerin ağırlıklı toplamını ifade etmektedir. Doğrusal birleşim Denklem (8)' de verilmiştir.

$$z = w^T x + b \quad (8)$$

Burada w ağırlık vektörünü, x giriş öznitelik vektörünü ve b ise sabit terimi göstermektedir.

3.1.6. Aşırı gradyan artırma

Aşırı Gradyan Artırma (Extreme Gradient Boosting, XGBoost), karar ağaçlarının ardışık olarak oluşturulduğu gradyan artırma tabanlı bir topluluk öğrenme yöntemidir (T. Chen ve Guestrin, 2016). Modelin nihai tahmini, oluşturulan tüm karar ağaçlarının çıktılarının birleştirilmesiyle elde edilmektedir. Bu işlem Denklem (9)'da verilen ifade ile gösterilmektedir.

$$\hat{y} = \sum_{k=1}^K f_k(x) \quad (9)$$

Burada $\hat{y}$ model tarafından üretilen nihai tahmini, $f_k(x)$ k . karar ağacının giriş örneği için ürettiği çıktıyı ve K ise modelde yer alan toplam karar ağacı sayısını ifade etmektedir. Böylece

her yeni karar ağacı, önceki ağaçların hatalarını azaltacak şekilde modele katkı sağlamakta ve nihai sınıflandırma sonucu tüm ağaçların ortak çıktısı kullanılarak elde edilmektedir.

3.2 Derin Öğrenme Tabanlı Yaklaşımlar

Bu çalışmada, göğüs röntgeni görüntülerinden zatürre tespiti için derin öğrenme tabanlı yaklaşımlar kullanılmıştır. Karşılaştırmalı değerlendirme amacıyla Xception, VGG16 ve ResNet50 mimarileri kullanılmış, ayrıca ConvNeXt-Tiny mimarisi temel alınarak Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) geliştirilmiştir. Tüm modeller aynı veri seti üzerinde eğitilmiş ve aynı değerlendirme ölçütleri kullanılarak analiz edilmiştir. Önerilen HPFN modeli, ConvNeXt-Tiny tabanlı derin öznitelik çıkarımı ile prototip tabanlı benzerlik hesaplamasını bir araya getiren hibrit bir sınıflandırma yaklaşımı sunmaktadır. Bu bölümde öncelikle karşılaştırmada kullanılan derin öğrenme mimarileri özetlenmiş, daha sonra önerilen HPFN modelinin yapısı ayrıntılı olarak açıklanmıştır.

3.2.1 Xception

Xception, derin evrişimsel sinir ağlarında hesaplama maliyetini azaltırken öznitelik çıkarım başarısını artırmak amacıyla geliştirilen bir mimaridir. Bu mimarinin temel özelliği, standart evrişim işlemleri yerine ayrılabilir evrişim (Depthwise Separable Convolution- Derinlemesine Ayrılabilir Konvolüsyon) yapısını kullanmasıdır. Bu yaklaşımda önce her giriş kanalı üzerinde bağımsız uzamsal evrişim işlemi gerçekleştirilmekte, ardından kanal bilgileri 1×1 noktasal evrişim ile birleştirilmektedir.

Derinlik yönündeki evrişim işlemi Denklem (10)' da verilmiştir.

$$Y = X * K_d \quad (10)$$

Burada X giriş öznitelik haritasını, K_d derinlik yönündeki evrişim çekirdeğini ve Y ise elde edilen ara öznitelik haritasını ifade etmektedir.

Derinlik yönündeki evrişim sonrasında kanal bilgilerinin birleştirilmesi amacıyla noktasal evrişim uygulanmaktadır. Bu işlem Denklem (11)' de gösterilmiştir.

$$Z = Y * K_p \quad (11)$$

Burada K_p noktasal evrişim çekirdeğini, Z ise elde edilen çıktı öznitelik haritasını göstermektedir.

Bu çalışmada Xception modeli, önerilen HPFN modelinin performansını değerlendirmek amacıyla karşılaştırma modeli olarak kullanılmıştır.

3.2.2. Vgg16

VGG16, ardışık küçük boyutlu (3×3) evrişim katmanlarından oluşan derin evrişimsel sinir ağı mimarisidir (Simonyan ve Zisserman, 2015). Mimaride elde edilen öznetelikler tam bağlantılı katmanlar aracılığıyla sınıflandırılmaktadır. Basit ve düzenli yapısı sayesinde görüntü sınıflandırma problemlerinde yaygın olarak kullanılan temel derin öğrenme modellerinden biridir.

Bir evrişim katmanındaki çıktı Denklem (12)' de verilen eşitlik ile ifade edilmektedir.

$$x_l = f(W_l * x_{l-1} + b_l) \quad (12)$$

Burada x_{l-1} önceki katmanın çıktısını, W_l evrişim filtrelerini, b_l sapma terimini, $f(\cdot)$ aktivasyon fonksiyonunu ve x_l ise mevcut katmanın çıktısını göstermektedir.

Bu çalışmada VGG16 modeli, önerilen HPFN modeli ile karşılaştırmalı performans değerlendirmesi amacıyla kullanılmıştır.

3.2.3. ResNet50

ResNet50, derin ağlarda eğitim sırasında ortaya çıkan gradyan kaybolması problemini azaltmak amacıyla artık (Residual) öğrenme yaklaşımını kullanan bir evrişimsel sinir ağı mimarisidir (He ve ark., 2016). Ağ içerisinde bulunan artık bağlantılar sayesinde giriş bilgisi sonraki katmanlara doğrudan aktarılabilmekte ve daha derin ağların etkin şekilde eğitilebilmesi sağlanmaktadır.

Artık öğrenme yapısı Denklem (13)'de verilen ifade ile gösterilmektedir.

$$y = F(x) + x \quad (13)$$

Burada x giriş öznetelik haritasını, $F(x)$ artık fonksiyonunu ve y ise blok çıkışı ifade etmektedir.

Bu çalışmada ResNet50 mimarisi, önerilen HPFN modelinin performansını karşılaştırmak amacıyla kullanılan temel derin öğrenme modellerinden biri olarak değerlendirilmiştir.

3.2.4. Önerilen hibrit prototip füzyon ağı (HPFN)

Bu çalışmada, göğüs röntgeni görüntülerinin Normal ve Zatürre sınıflarına ayrılması amacıyla ConvNeXt-Tiny mimarisi (Z. Liu ve ark., 2022) temel alınarak Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) geliştirilmiştir. Önerilen model, derin öznitelik çıkarımı ile prototip tabanlı benzerlik yaklaşımını aynı yapı içerisinde birleştiren hibrit bir sınıflandırma mimarisidir. Model; ConvNeXt-Tiny tabanlı öznitelik çıkarımı, sınıf prototiplerinin oluşturulması, prototip benzerliklerinin hesaplanması ve hibrit karar mekanizması olmak üzere dört temel aşamadan oluşmaktadır.

Modelde öznitelik çıkarıcı olarak ConvNeXt-Tiny mimarisi kullanılmıştır. Başlangıç ağırlıkları için ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklardan yararlanılmıştır. ConvNeXt-Tiny mimarisinin özgün sınıflandırıcı katmanı kaldırılarak yerine 768 boyutlu öznitelikleri 512 boyutlu gömme uzayına dönüştüren doğrusal katman, GELU aktivasyon fonksiyonu ve bırakma olasılığı 0,3 olan Dropout katmanından oluşan yeni bir yapı eklenmiştir. Son aşamada ise 512 boyutlu gömme vektörü tam bağlantılı çıkış katmanına aktarılmakta ve sınıflandırma çıktısı üretilmektedir. Böylece model hem sınıflandırma çıktısını hem de prototip tabanlı karar mekanizmasında kullanılacak gömme vektörlerini oluşturmaktadır.

Klasik evrişimsel sinir ağı modellerinde nihai sınıflandırma yalnızca softmax tabanlı parametrik sınıflandırıcı kullanılarak gerçekleştirilmektedir. Önerilen HPFN modelinde ise buna ek olarak öğrenilen gömme uzayında oluşturulan sınıf prototiplerine dayalı parametrik olmayan ikinci bir karar mekanizması kullanılmaktadır. Model eğitimi tamamlandıktan sonra aynı sınıfa ait tüm gömme vektörlerinin ortalaması alınarak her sınıf için bir prototip vektörü oluşturulmaktadır. Böylece her sınıf, gömme uzayında tek bir temsil vektörü ile ifade edilmektedir.

Her sınıfa ait prototip vektörü Denklem (14)'te verilen ifade kullanılarak hesaplanmaktadır.

$$P_c = \frac{1}{|S_c|} \sum_{x_i \in S_c} f(x_i) \quad (14)$$

Burada P_c , c sınıfına ait prototip vektörünü, S_c ilgili sınıfa ait eğitim örnekleri kümesini, $f(x_i)$ ise i . görüntüden elde edilen 512 boyutlu gömme vektörünü ifade etmektedir.

Prototiplerin oluşturulmasının ardından her test görüntüsüne ait gömme vektörü ile tüm sınıf prototipleri arasındaki benzerlik kosinüs uzaklığı kullanılarak hesaplanmaktadır. En küçük kosinüs uzaklığına sahip sınıf, prototip tabanlı tahmin sonucu olarak belirlenmektedir.

İki vektör arasındaki kosinüs benzerliği Denklem (15)'te verilmiştir.

$$\text{Cos}(z, P_c) = \frac{z^T P_c}{\|z\| \|P_c\|} \quad (15)$$

Burada z , test görüntüsüne ait gömme vektörünü, P_c ilgili sınıfın prototip vektörünü, $\|\cdot\|$ ise Öklid normunu göstermektedir. Kosinüs benzerliği değeri arttıkça test örneğinin ilgili sınıf prototipine olan benzerliği de artmaktadır.

Son aşamada ağın softmax tabanlı tahmini ile prototip tabanlı tahmin hibrit karar mekanizması kullanılarak birleştirilmektedir. Her iki karar mekanizmasının aynı sınıfı tahmin etmesi durumunda ortak sınıf etiketi doğrudan kabul edilmektedir. Tahminlerin farklı olması durumunda ise nihai sınıf etiketi olarak softmax tabanlı tahmin kullanılmaktadır. Böylece derin öznitelik temsili ile prototip tabanlı benzerlik yaklaşımı birlikte değerlendirilerek daha tutarlı bir sınıflandırma mekanizması oluşturulmaktadır.

4. BULGULAR

Bu bölümde, az etiketli veri kullanılarak zatürre tanısına yönelik gerçekleştirilen deneysel çalışmalardan elde edilen bulgular sunulmaktadır. İlk olarak çalışmada kullanılan veri seti tanıtılarak veri dağılımına ilişkin bilgiler verilmiştir. Ardından, model performanslarının değerlendirilmesinde kullanılan performans ölçütleri açıklanmış ve görüntülerin eğitim sürecine hazırlanması amacıyla gerçekleştirilen veri ön işleme işlemleri ile önerilen model mimarisi sunulmuştur. Son bölümde ise gerçekleştirilen deneyler sonucunda elde edilen sınıflandırma performansları doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1-Skoru gibi performans ölçütleri kullanılarak sunulmuş, ayrıca elde edilen sonuçlar farklı derin öğrenme yaklaşımları ve literatürde yer alan çalışmalarla karşılaştırılarak değerlendirilmiştir. Böylece, az etiketli veri koşullarında derin öğrenme tabanlı yöntemlerin zatürre tanısındaki performansları kapsamlı bir şekilde analiz edilmiştir.

4.1. Veri Seti

Bu çalışmada, zatürre tanısına yönelik derin öğrenme modellerinin performansının değerlendirilmesinde, Kaggle platformunda yayımlanan ve Kermany ve ark. (2018) tarafından geliştirilen Chest X-Ray (Pneumonia) veri seti kullanılmıştır. Veri seti, Normal ve Pneumonia olmak üzere iki sınıfa ait posterior-anterior (PA) göğüs röntgeni görüntülerinden oluşmaktadır. Görüntüler, eğitim (train) ve test (test) olmak üzere iki ayrı alt küme halinde sunulmuş olup, her alt küme kendi içerisinde sınıflara göre ayrı klasörlerde düzenlenmiştir.

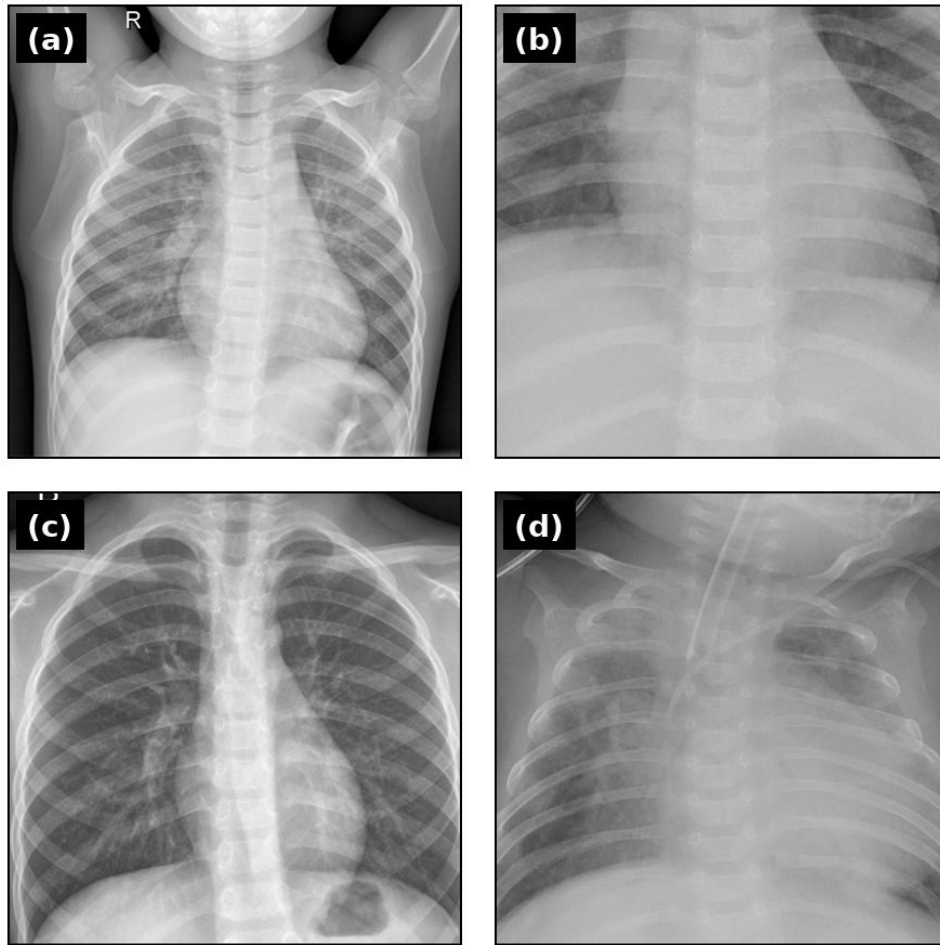
Çalışmada veri setinin orijinal eğitim ve test ayrımı korunmuş, herhangi bir ek rastgele bölme veya çapraz doğrulama işlemi uygulanmamıştır. Eğitim kümesindeki görüntüler modelin eğitilmesi amacıyla kullanılırken, test kümesindeki görüntüler modelin performansının değerlendirilmesinde kullanılmıştır.

Çalışmada kullanılan veri setinde eğitim kümesi 1.349 normal ve 3.883 zatürre görüntüsü olmak üzere toplam 5.232 görüntüden oluşmaktadır. Test kümesinde ise 234 normal ve 390 zatürre görüntüsü bulunmakta olup toplam 624 görüntü yer almaktadır. Böylece çalışmada kullanılan veri seti toplam 5.856 göğüs röntgeni görüntüsünü içermektedir. Veri setine ait sınıf dağılımı Tablo 1'de verilmiştir.

Tablo 1. Çalışmada kullanılan Chest X-Ray (Pneumonia) veri setinin sınıf dağılımı

Veri seti	Normal	Zatürre	Toplam
Eğitim (Train)	1349	3883	5232
Test	234	390	624
Toplam	1583	4273	5856

Çalışmada kullanılan veri setine ait eğitim ve test kümelerinden seçilen normal ve zatürre sınıflarına ilişkin örnek göğüs röntgeni görüntüleri Şekil 2'de sunulmuştur.



Şekil 2. Çalışmada kullanılan Chest X-Ray (Pneumonia) veri setinden örnek göğüs röntgeni görüntüleri. (a) Eğitim kümesi–Normal, (b) Eğitim kümesi–Zatürre, (c) Test kümesi–Normal ve (d) Test kümesi–Zatürre.

Şekil 2'de çalışmada kullanılan Chest X-Ray (Pneumonia) veri setinin eğitim ve test kümelerinden seçilen örnek görüntüler sunulmuştur. Görüntüler incelendiğinde, normal sınıfa ait örneklerde akciğer alanlarının homojen bir görünüm sergilediği, zatürre sınıfına ait örneklerde ise akciğer bölgelerinde farklı yoğunluklarda opasitelerin bulunduğu görülmektedir.

4.2. Değerlendirme Metrikleri

Bu çalışmada önerilen modelin sınıflandırma performansının değerlendirilmesinde sınıflandırma problemlerinde yaygın olarak kullanılan doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1-Skoru (F1-Score) performans ölçütleri kullanılmıştır. Bu ölçütlerin hesaplanmasında doğru pozitif (True Positive, TP), yanlış pozitif (False Positive, FP), yanlış negatif (False Negative, FN) ve doğru negatif (True Negative, TN) değerlerinden yararlanılmıştır. Çalışmada Pneumonia sınıfı pozitif sınıf, Normal sınıfı ise negatif sınıf olarak kabul edilmiştir. Buna göre TP, zatürre görüntülerinin doğru şekilde zatürre olarak sınıflandırılmasını; FP, normal görüntülerin yanlışlıkla zatürre olarak sınıflandırılmasını; FN, zatürre görüntülerinin normal olarak sınıflandırılmasını; TN ise normal görüntülerin doğru şekilde normal olarak sınıflandırılmasını ifade etmektedir.

Model performansının genel doğruluğunu belirlemek amacıyla doğruluk, pozitif olarak tahmin edilen örneklerin ne kadarının gerçekten pozitif olduğunu belirlemek amacıyla kesinlik, gerçek pozitif örneklerin ne kadarının doğru olarak tespit edildiğini belirlemek amacıyla duyarlılık ve kesinlik ile duyarlılık değerlerinin dengeli bir değerlendirmesini sağlayan F1-Skoru kullanılmıştır. Bu performans ölçütleri birlikte değerlendirilerek modelin sınıflandırma başarısı kapsamlı bir şekilde analiz edilmiştir. Özellikle F1-Skoru, tıbbi görüntü sınıflandırma çalışmalarında yanlış sınıflandırmaların önemli olduğu durumlarda dengeli bir performans değerlendirmesi sağlamaktadır.

Çalışmada kullanılan performans ölçütleri aşağıda verilen eşitlikler kullanılarak hesaplanmıştır.

$$Doğruluk = \frac{TP+TN}{TP+FP+TN+FN} \quad (16)$$

$$Kesinlik = \frac{TP}{TP+FP} \quad (17)$$

$$Duyarlılık = \frac{TP}{TP+FN} \quad (18)$$

$$F1 - score = 2 * \frac{(Kesinlik \times Duyarlılık)}{(Kesinlik + Duyarlılık)} \quad (19)$$

Bu performans ölçütleri model performansının kapsamlı bir şekilde değerlendirilmesini sağlamaktadır. Özellikle F1-Skoru, duyarlılık ve kesinlik değerlerini birlikte dikkate alması nedeniyle yanlış sınıflandırma maliyetinin yüksek olduğu tıbbi görüntü analizlerinde önemli bir performans ölçütü olarak değerlendirilmektedir.

4.3. Ön İşleme ve Önerilen Model Mimarisi

Bu çalışmada, zatürre tespitine yönelik olarak ConvNeXt-Tiny tabanlı Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) geliştirilmiştir. Önerilen yöntem; veri ön işleme ve veri artırma, ConvNeXt-Tiny tabanlı öznitelik çıkarımı, prototip hesaplama ve kosinüs uzaklığına dayalı benzerlik eşleştirmesi ile prototip ve softmax tahminlerinin birleştirilmesi olmak üzere dört temel aşamadan oluşmaktadır. Bu hibrit yaklaşım, derin özniteliklerin ayırt edici gücü ile prototip tabanlı karar mekanizmasını bir araya getirerek zatürre sınıflandırmasının doğruluğunu ve karar tutarlılığını artırmayı amaçlamaktadır.

Çalışmada kullanılan Chest X-Ray (Pneumonia) veri seti, Kaggle platformunda sunulan orijinal eğitim ve test ayrımı korunarak kullanılmıştır. Tüm giriş göğüs röntgeni görüntüleri 224×224 piksel boyutuna yeniden ölçeklendirilmiş ve ImageNet veri setine ait ortalama ([0.485, 0.456, 0.406]) ve standart sapma ([0.229, 0.224, 0.225]) değerleri kullanılarak normalize edilmiştir. Modelin genelleme yeteneğini artırmak amacıyla eğitim aşamasında rastgele yatay çevirme, parlaklık/kontrast/doygunluk değişimi (0.2), $\pm 10^\circ$ rastgele döndürme ve 0.8–1.0 ölçek aralığında rastgele yeniden boyutlandırılmış kırpma işlemleri uygulanmıştır. Eğitim kümesi yalnızca model parametrelerinin öğrenilmesi ve prototiplerin oluşturulması amacıyla kullanılmış, herhangi bir ek rastgele bölme veya çapraz doğrulama işlemi uygulanmamıştır. Model eğitimi tamamlandıktan sonra test kümesi, performans değerlendirmesi, öznitelik çıkarımı ve prototip tabanlı çıkarım işlemlerinde kullanılmıştır.

Önerilen HPFN modeli, yüksek temsil gücü ve verimli yapısıyla bilinen ConvNeXt-Tiny mimarisi temel alınarak geliştirilmiştir. Modelin başlangıç ağırlıkları için önceden eğitilmiş (pretrained) DEFAULT ağırlıkları kullanılmıştır. ConvNeXt-Tiny mimarisinin özgün sınıflandırıcı katmanı kaldırılarak yerine özelleştirilmiş bir gömme (embedding) ve sınıflandırma yapısı eklenmiştir. Bu yapı; 768 boyutlu öznitelikleri 512 boyutlu gömme uzayına dönüştüren doğrusal bir katman, GELU aktivasyon fonksiyonu ve 0.3 bırakma olasılığına sahip Dropout katmanından oluşmaktadır. Son aşamada ise 512 boyutlu gömme vektörünü hedef sınıf sayısına dönüştüren tam bağlantılı bir çıkış katmanı kullanılmıştır. Böylece model, hem sınıflandırma çıktıları hem de değerlendirme aşamasında kullanılacak ara gömme vektörlerini üretmektedir.

Klasik evrimsel sinir ağlarında sınıflandırma işlemi yalnızca softmax tabanlı parametrik sınıflandırıcı ile gerçekleştirilirken, önerilen HPFN modelinde öğrenilen gömme uzayındaki sınıf prototiplerini kullanan parametrik olmayan ikinci bir karar mekanizması da

kullanılmaktadır. Modelin eğitim sürecinde ConvNeXt-Tiny mimarisi uçtan uca eğitilerek ayırt edici öznitelikler öğrenilmekte, eğitim tamamlandıktan sonra ise her sınıfa ait eğitim örneklerinin gömme vektörlerinin ortalaması alınarak sınıf prototipleri oluşturulmaktadır. Böylece her sınıf için gömme uzayında bir prototip vektörü elde edilmekte ve bu prototipler kosinüs uzaklığına dayalı benzerlik hesaplamalarında sınıf temsilcisi olarak kullanılmaktadır. Model eğitimi tamamlandıktan sonra eğitim ve test kümelerindeki tüm görüntüler için 512 boyutlu derin öznitelik vektörleri elde edilmiştir. Ardından eğitim kümesindeki aynı sınıfa ait tüm gömme vektörlerinin ortalaması alınarak sınıf prototipleri hesaplanmıştır. Test aşamasında ise her test görüntüsüne ait öznitelik vektörü ile tüm sınıf prototipleri arasındaki kosinüs uzaklıkları SciPy kütüphanesinin `cdist()` fonksiyonu kullanılarak hesaplanmıştır. En küçük kosinüs uzaklığına sahip sınıf, prototip tabanlı tahmin olarak belirlenmiştir. Son aşamada softmax tabanlı tahmin ile prototip tabanlı tahmin birleştirilmiştir. Her iki tahminin aynı olması durumunda ortak sınıf etiketi doğrudan kabul edilmiş, farklı olması durumunda ise softmax tabanlı tahmin korunmuştur. Böylece önerilen hibrit karar mekanizması ile sınıflandırma kararlarının tutarlılığının artırılması amaçlanmıştır.

Bu bölümde önerilen Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) modelinin deneysel performansı sunulmakta ve elde edilen sonuçlar farklı derin öğrenme modelleri, geleneksel öznitelik tabanlı yaklaşımlar ve literatürde aynı veri seti üzerinde gerçekleştirilen çalışmalar ile karşılaştırılmaktadır. Tüm değerlendirmeler, önceki bölümde açıklanan veri seti, veri ön işleme adımları ve performans ölçütleri kullanılarak gerçekleştirilmiştir. Böylece önerilen yöntemin hem derin öğrenme tabanlı hem de geleneksel sınıflandırma yaklaşımları karşısındaki başarısı kapsamlı olarak incelenmiştir.

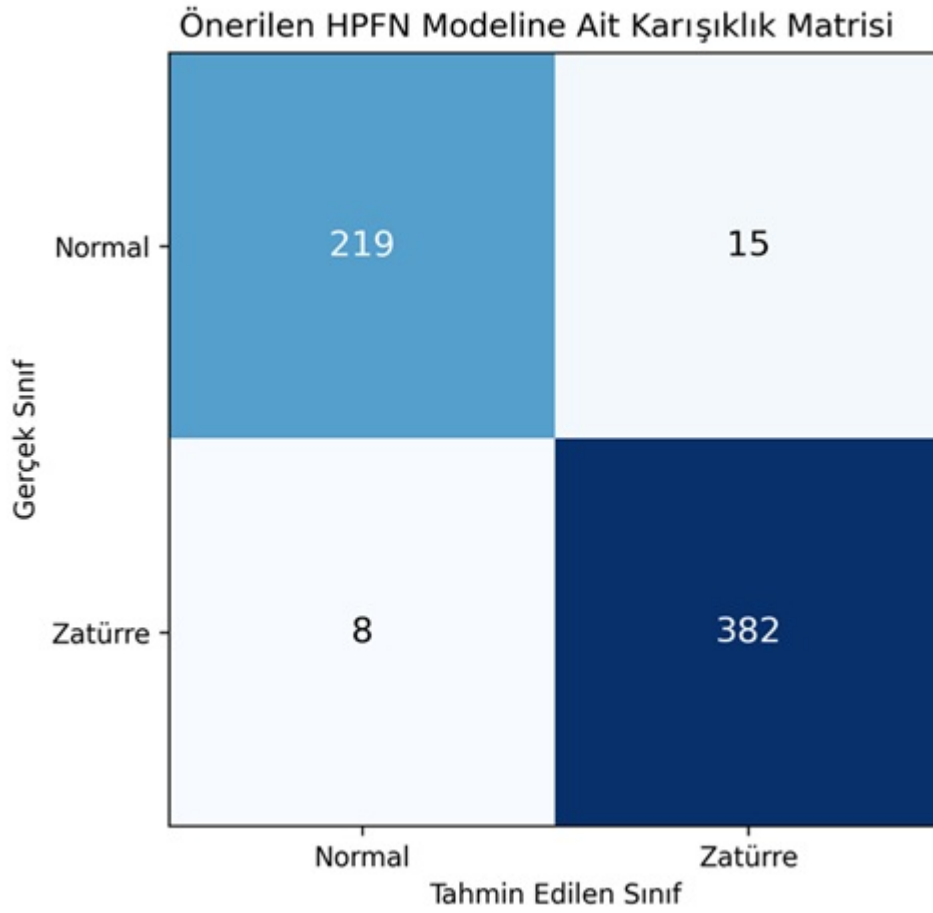
4.4. Deneysel Sonuçlar

Bu bölümde, az etiketli veri kullanılarak zatürre tanısına yönelik bu tez kapsamında gerçekleştirilen deneysel çalışmalardan elde edilen bulgular sunulmaktadır. Çalışma kapsamında önerilen Hibrit Prototip Füzyon Ağı (Hybrid Prototype Fusion Network-HPFN) modelinin performansı değerlendirilmiş, ayrıca farklı derin öğrenme mimarileri ile geleneksel öznitelik çıkarım yöntemlerinin sınıflandırma başarımları karşılaştırmalı olarak incelenmiştir. Bu doğrultuda Xception, VGG16 ve ResNet50 derin öğrenme modelleri ile SIFT, KAZE ve HOG öznitelik çıkarım yöntemleri kullanılarak Destek Vektör Makineleri, Lojistik Regresyon ve Aşırı Gradyan Artırma sınıflandırıcıları üzerinde deneysel değerlendirmeler gerçekleştirilmiştir. Bunun yanı sıra önerilen yöntemin performansı, aynı veri seti üzerinde

literatürde raporlanan çalışmalar ile karşılaştırılarak kapsamlı bir analiz yapılmıştır. Böylece tez kapsamında gerçekleştirilen tüm deneysel çalışmalar birlikte değerlendirilmiş ve farklı yaklaşımların az etiketli veri koşullarındaki zati sınıflandırma performansları ayrıntılı olarak incelenmiştir.

Deneysel değerlendirmeler kapsamında ilk olarak önerilen HPFN modelinin sınıflandırma performansı incelenmiştir. Modelde, ImageNet üzerinde önceden eğitilmiş ConvNeXt-Tiny mimarisi kullanılarak giriş göğüs röntgeni görüntülerinden 512 boyutlu derin öznitelik vektörleri elde edilmiştir. Sınıflandırıcı katmanı yerine doğrusal gömme katmanı, GELU aktivasyon fonksiyonu ve Dropout katmanından oluşan yeni bir yapı kullanılmış, ardından tam bağlantılı çıkış katmanı ile sınıflandırma gerçekleştirilmiştir. Eğitim tamamlandıktan sonra en yüksek doğrulama başarımını sağlayan model ağırlıkları kullanılarak eğitim ve test kümelerine ait öznitelikler çıkarılmış, eğitim kümesindeki aynı sınıfa ait öznitelik vektörlerinin ortalaması alınarak sınıf prototipleri oluşturulmuştur. Test aşamasında her görüntü için elde edilen öznitelikler ile sınıf prototipleri arasındaki kosinüs uzaklığı hesaplanmış ve prototip tabanlı tahminler elde edilmiştir. Son aşamada ise prototip tabanlı karar mekanizması ile ağırlıklı softmax tabanlı çıktıları hibrit karar mekanizması kullanılarak birleştirilmiş ve nihai sınıflandırma sonuçları oluşturulmuştur.

Gerçekleştirilen deneyler sonucunda önerilen HPFN modeli test kümesi üzerinde %96,38 doğruluk, 0,9623 Kesinlik, 0,9795 Recall ve 0,9708 F1-Skoru elde etmiştir. Elde edilen sonuçlar, ConvNeXt-Tiny tabanlı derin özniteliklerin prototip tabanlı benzerlik yaklaşımı ile birlikte kullanılmasının sınıflandırma performansını artırdığını göstermektedir. Modelin sınıflandırma başarısının ayrıntılı olarak değerlendirilebilmesi amacıyla oluşturulan karışıklık matrisi Şekil 3'de verilmiştir.



Şekil 3. Önerilen HPFN modeline ait karışıklık matrisi.

Şekil 3 incelendiğinde modelin Normal ve Zatürre sınıflarını yüksek doğrulukla ayırt edebildiği görülmektedir. Tahminlerin büyük çoğunluğu doğru sınıflandırılmış olup yanlış sınıflandırmaların önemli bir kısmı zatürre sınıfına ait görüntülerin normal sınıfı olarak tahmin edilmesinden oluşmaktadır. Karışıklık matrisi genel olarak değerlendirildiğinde önerilen sınıflandırma yaklaşımının kararlı ve güvenilir bir performans sergilediği görülmektedir.

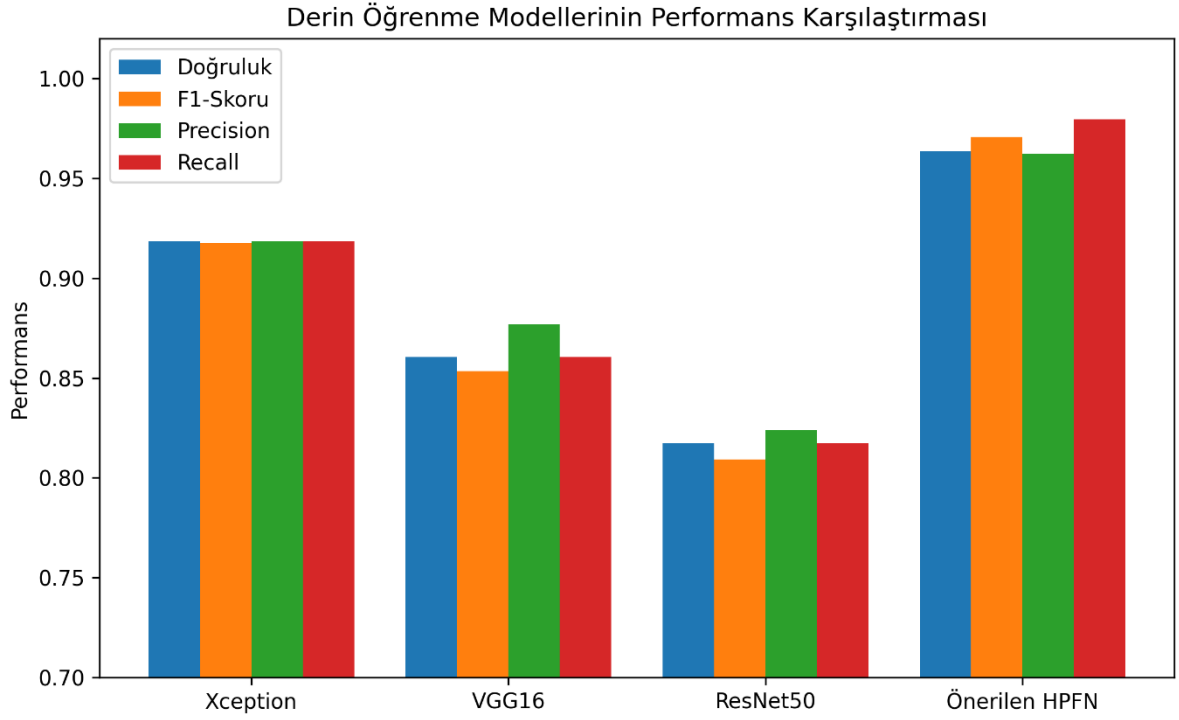
Önerilen yöntemin performansının farklı derin öğrenme mimarileri karşısındaki başarısını değerlendirmek amacıyla HPFN modeli, Xception, VGG16 ve ResNet50 modelleri ile karşılaştırılmıştır. Elde edilen sonuçlar doğruluk, kesinlik, duyarlılık ve F1-Skoru ölçütleri kullanılarak değerlendirilmiş olup karşılaştırmalı performans değerleri Tablo 2'de sunulmaktadır.

Tablo 2. Derin öğrenme tabanlı modeller ile önerilen HPFN modelinin performans karşılaştırması

Model	Doğruluk	F1-Skoru	Kesinlik	Duyarlılık
Xception	0,9183	0,9176	0,9184	0,9183
VGG16	0,8606	0,8534	0,8770	0,8606
ResNet50	0,8173	0,8091	0,8239	0,8173
Önerilen HPFN modeli	0,9638	0,9708	0,9623	0,9795

Tablo 2 incelendiğinde, önerilen HPFN modelinin karşılaştırmaya dâhil edilen tüm derin öğrenme tabanlı modellerden daha yüksek performans sergilediği görülmektedir. Xception modeli %91,83 doğruluk oranı ile temel modeller arasında en başarılı sonucu elde ederken, VGG16 ve ResNet50 modelleri sırasıyla %86,06 ve %81,73 doğruluk değerlerine ulaşmıştır. Buna karşılık önerilen HPFN modeli %96,38 doğruluk, 0,9623 kesinlik, 0,9795 duyarlılık ve 0,9708 F1-Skoru ile tüm performans ölçütlerinde en yüksek başarıyı göstermiştir. Özellikle yüksek Duyarlılık ve F1-Skoru değerleri, önerilen yöntemin zatürre vakalarını doğru tespit etme başarısının ve dengeli sınıflandırma performansının arttığını göstermektedir. Bu sonuçlar, ConvNeXt-Tiny mimarisinden elde edilen derin özneteliklerin prototip tabanlı karar mekanizması ile birleştirilmesinin sınıflandırma performansına önemli katkı sağladığını ortaya koymaktadır.

Derin öğrenme tabanlı modellerden elde edilen performans sonuçlarının daha açık bir şekilde değerlendirilebilmesi amacıyla doğruluk, Kesinlik, Duyarlılık ve F1-Skoru değerleri grafiksel olarak da sunulmuştur. Önerilen HPFN modeli ile karşılaştırmada kullanılan temel derin öğrenme mimarilerine ait performans karşılaştırması Şekil 3'te verilmiştir.



Şekil 4. Önerilen HPFN modeli ile derin öğrenme tabanlı modellerin performans karşılaştırması.

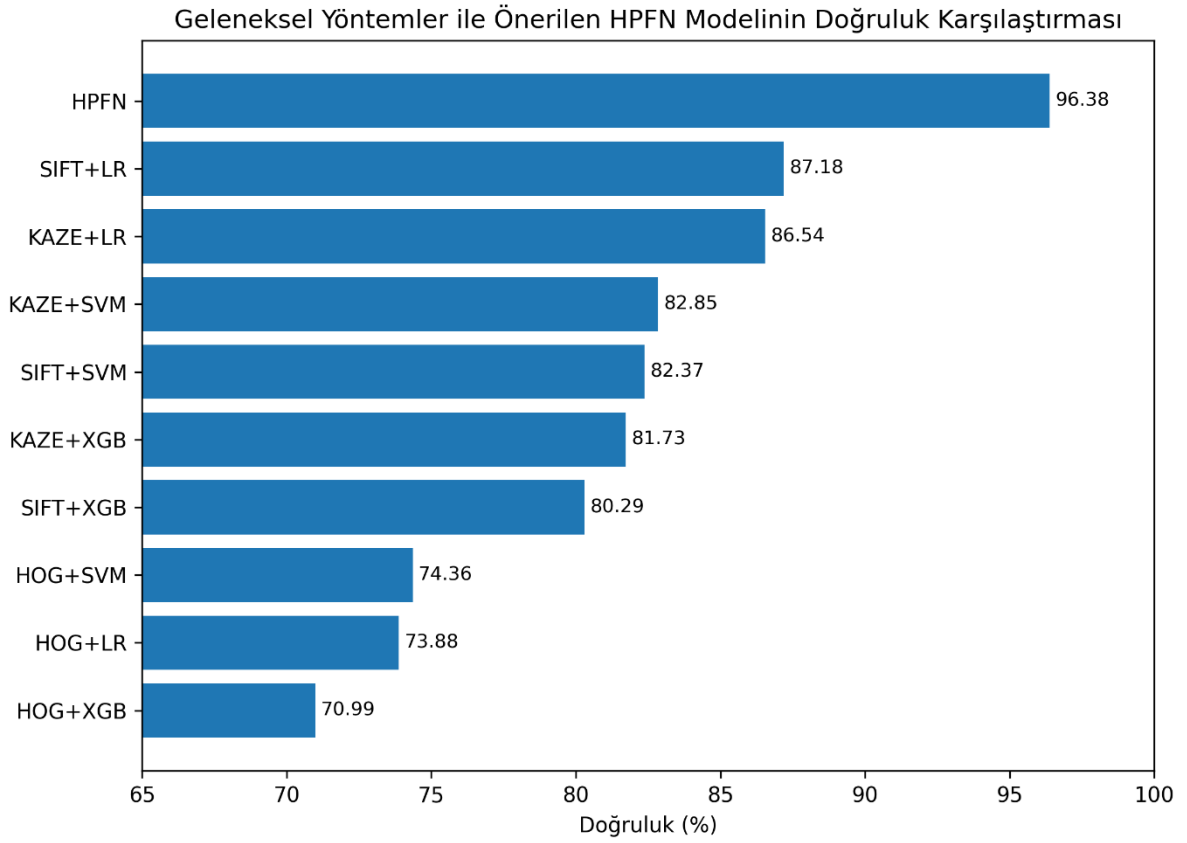
Şekil 4 incelendiğinde, önerilen HPFN modelinin tüm performans ölçütlerinde diğer derin öğrenme modellerinden daha yüksek değerlere ulaştığı görülmektedir. Özellikle doğruluk ve F1-Skoru bakımından elde edilen üstün performans, derin öznitelik temsili ile prototip tabanlı karar mekanizmasının birlikte kullanılmasının sınıflandırma başarısını artırdığını göstermektedir. Elde edilen bulgular, önerilen hibrit yaklaşımın yalnızca yüksek doğruluk sağlamakla kalmayıp aynı zamanda farklı performans ölçütleri açısından da dengeli ve kararlı bir sınıflandırma gerçekleştirdiğini ortaya koymaktadır.

Önerilen yöntemin etkinliğinin daha kapsamlı olarak değerlendirilebilmesi amacıyla, derin öğrenme tabanlı modellerin yanı sıra geleneksel öznitelik çıkarım yöntemleri ile makine öğrenmesi sınıflandırıcılarının birlikte kullanıldığı yaklaşımlar da incelenmiştir. Bu kapsamda aynı göğüs röntgeni veri setinden SIFT, KAZE ve HOG öznitelikleri çıkarılmış; elde edilen öznitelikler Destek Vektör Makineleri, Lojistik Regresyon ve Aşırı Gradyan Artırma sınıflandırıcıları kullanılarak değerlendirilmiştir. Elde edilen sonuçlar Tablo 3'te sunulmaktadır.

Tablo 3. Geleneksel öznitelik tabanlı yöntemler ile önerilen HPFN modelinin performans karşılaştırması.

Öznitelik	Sınıflandırıcı	Doğruluk	F1-Skoru	Kesinlik	Duyarlılık
SIFT	SVM	0,8237	0,8080	0,8602	0,8237
	Logistic Regression	0,8718	0,8658	0,8867	0,8718
	XGBoost	0,8029	0,7834	0,8401	0,8029
KAZE	SVM	0,8285	0,8152	0,8572	0,8285
	Logistic Regression	0,8654	0,8588	0,8806	0,8654
	XGBoost	0,8173	0,8010	0,8518	0,8173
HOG	SVM	0,7436	0,7009	0,8095	0,7436
	Logistic Regression	0,7388	0,6937	0,8067	0,7388
	XGBoost	0,7099	0,6538	0,7697	0,7099
Önerilen HPFN modeli (ConvNeXt-Tiny Fusion)		0,9638	0,9708	0,9623	0,9795

Tablo 3'te verilen sonuçlar incelendiğinde, geleneksel öznitelik çıkarım yöntemleri kullanılarak elde edilen doğruluk değerlerinin %70,99 ile %87,18 arasında değiştiği görülmektedir. En yüksek performans SIFT öznitelikleri ile Logistic Regression sınıflandırıcısının birlikte kullanıldığı modelde elde edilirken (%87,18), bunu KAZE + Logistic Regression (%86,54) yaklaşımı takip etmiştir. HOG tabanlı yöntemler ise diğer geleneksel yaklaşımlara kıyasla daha düşük performans göstermiştir. Buna karşılık önerilen HPFN modeli %96,38 doğruluk ve 0,9708 F1-Skoru ile tüm geleneksel yöntemlerden belirgin şekilde daha yüksek performans elde etmiştir. Bu bulgular, ConvNeXt-Tiny tabanlı derin özniteliklerin, SIFT, KAZE ve HOG gibi el yapımı özniteliklere kıyasla göğüs röntgeni görüntülerindeki ayırt edici bilgileri daha etkin temsil edebildiğini göstermektedir. Ayrıca prototip tabanlı karar mekanizmasının softmax tabanlı sınıflandırma ile birlikte kullanılması, modelin özellikle duyarlılık ve F1-Skoru değerlerinde daha başarılı sonuçlar elde etmesine katkı sağlamıştır. Şekil 4'te geleneksel öznitelik tabanlı yöntemler ile önerilen HPFN modelinin doğruluk değerleri karşılaştırmalı olarak sunulmaktadır.



Şekil 5. Geleneksel yöntemler ile önerilen HPFN modelinin doğruluk karşılaştırması

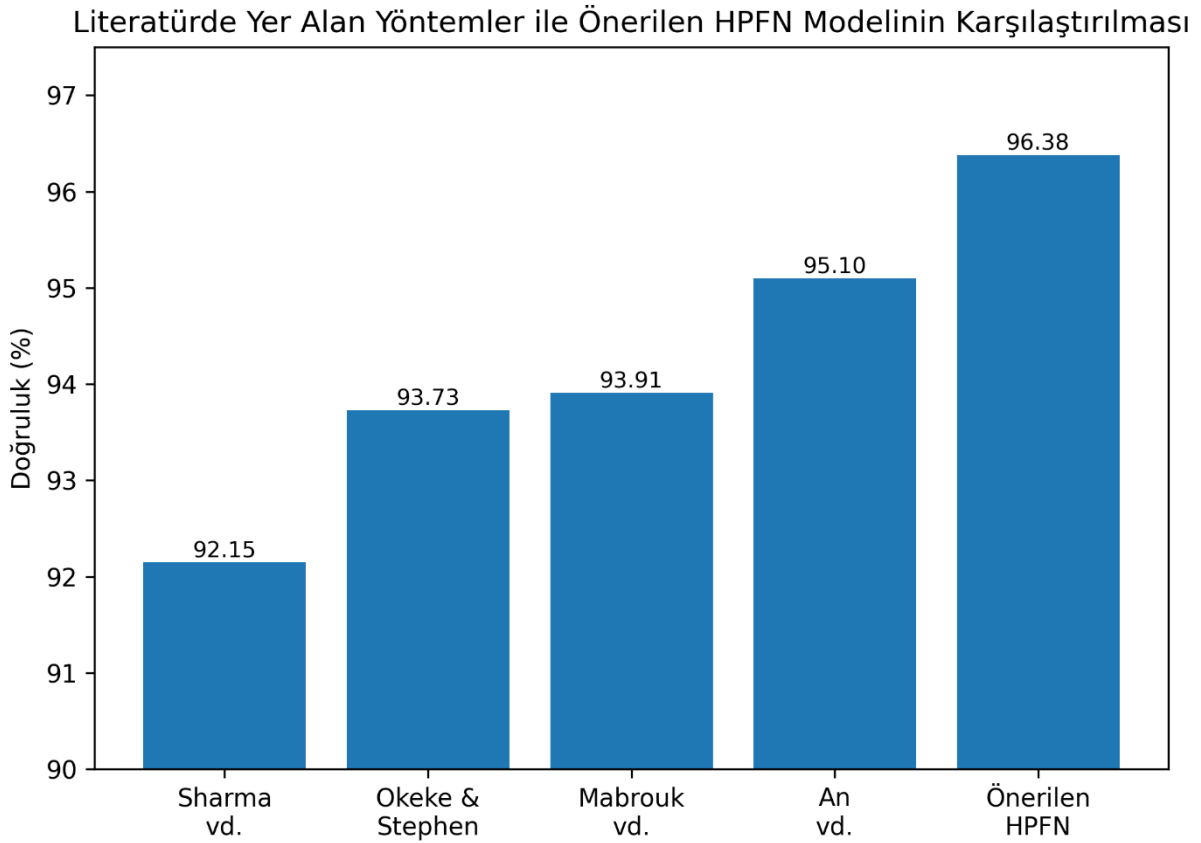
Şekil 5 incelendiğinde, önerilen HPFN modelinin tüm geleneksel yaklaşımlardan daha yüksek doğruluk değeri elde ettiği açıkça görülmektedir. Geleneksel yöntemler arasında en başarılı sonuç SIFT öz nitelikleri ile Logistic Regression sınıflandırıcısının birlikte kullanıldığı yaklaşımda elde edilirken, HOG tabanlı yöntemlerin daha düşük doğruluk değerlerine sahip olduğu görülmektedir. Bu sonuçlar, ConvNeXt-Tiny tabanlı derin öz niteliklerin göğüs röntgeni görüntülerindeki ayırt edici bilgileri el yapımı öz niteliklere kıyasla daha etkin temsil edebildiğini ve önerilen HPFN modelinin geleneksel yaklaşımlara göre daha başarılı bir sınıflandırma performansı sunduğunu göstermektedir.

Son olarak önerilen HPFN modelinin başarısı, aynı Chest X-Ray (Pneumonia) veri seti kullanılarak literatürde gerçekleştirilen çalışmalar ile karşılaştırılmıştır. Karşılaştırma sonuçları Tablo 4'te sunulmaktadır.

Tablo 4. Literatürde yer alan çalışmalar ile önerilen HPFN modelinin karşılaştırılması.

Metot	Doğruluk (%)
Mabrouk ve ark. (2019) (Mabrouk ve ark., 2023)	93,91
Stephen ve ark. (2019) (Stephen ve ark., 2019)	93,73
An ve ark. (2024) (An ve ark., 2024)	95,19
Sharma ve Guleria (2023) (Sharma ve Guleria, 2023)	92,15
Önerilen HPFN Yöntemi	96,38

Tablo 4 incelendiğinde, önerilen HPFN modelinin %96,38 doğruluk değeri ile literatürde aynı veri seti üzerinde gerçekleştirilen birçok çalışmadan daha yüksek performans sergilediği görülmektedir. Mabrouk ve ark. ile Stephen ve ark. tarafından önerilen yöntemler yaklaşık %94 doğruluk seviyesinde kalırken, önerilen yöntem bu çalışmalardan daha yüksek sınıflandırma başarısı elde etmiştir. Ayrıca An ve ark. tarafından önerilen dikkat mekanizması tabanlı derin öğrenme modeli ile benzer düzeyde performans göstermesi, önerilen yaklaşımın yüksek doğruluk sağlayabildiğini ortaya koymaktadır.



Şekil 6. Literatürde yer alan çalışmalar ile önerilen HPFN modelinin doğruluk değerlerinin karşılaştırılması

Şekil 6 incelendiğinde, önerilen HPFN modelinin literatürde karşılaştırmaya dâhil edilen yöntemler arasında en yüksek doğruluk değerine ulaştığı görülmektedir. Grafik, önerilen yöntemin diğer çalışmalar ile arasındaki performans farkını görsel olarak ortaya koymaktadır. Özellikle Mabrouk ve ark., Stephen ve ark. ile Sharma ve ark. tarafından önerilen yöntemlere göre daha yüksek doğruluk elde edildiği, An ve ark. tarafından önerilen yöntem ile ise yakın performans sergilendiği görülmektedir. Elde edilen bulgular, önerilen HPFN modelinin aynı veri seti üzerinde literatürde raporlanan güncel yöntemlerle rekabet edebilecek düzeyde başarılı bir sınıflandırma performansı sunduğunu göstermektedir.

5. TARTIŞMA VE SONUÇ

Bu tez kapsamında, az etiketli veri koşullarında göğüs röntgeni görüntülerinden zatürre tanısına yönelik farklı derin öğrenme yaklaşımlarının ve geleneksel öznitelik tabanlı yöntemlerin sınıflandırma performansları kapsamlı olarak değerlendirilmiştir. Çalışmada, ConvNeXt-Tiny tabanlı HPFN önerilmiş; bunun yanında Xception, VGG16 ve ResNet50 derin öğrenme modelleri ile SIFT, KAZE ve HOG öznitelik çıkarım yöntemlerinin Destek Vektör Makineleri, Lojistik Regresyon ve Aşırı Gradyan Artırma sınıflandırıcıları ile oluşturduğu modeller aynı veri seti üzerinde karşılaştırılmıştır. Ayrıca elde edilen sonuçlar, aynı Chest X-Ray (Pneumonia) veri setini kullanan literatürdeki çalışmalar ile birlikte değerlendirilmiştir. Böylece tez kapsamında gerçekleştirilen tüm deneysel çalışmalar ortak bir çerçevede analiz edilmiştir.

Karşılaştırılan temel derin öğrenme mimarileri arasında Xception modeli %91,83 doğruluk değeri ile en başarılı temel model olurken, VGG16 ve ResNet50 modelleri sırasıyla %86,06 ve %81,73 doğruluk değerleri elde etmiştir. Buna karşılık ConvNeXt-Tiny tabanlı HPFN modeli %96,38 doğruluk, 0,9623 Kesinlik, 0,9795 Duyarlılık ve 0,9708 F1-Skoru ile tüm derin öğrenme modellerinden daha yüksek performans göstermiştir. Bu sonuçlar, ConvNeXt-Tiny mimarisinden elde edilen derin özniteliklerin prototip tabanlı karar mekanizması ile birlikte kullanılmasının sınıflandırma başarısını artırdığını ortaya koymaktadır.

Geleneksel öznitelik tabanlı yöntemlerin performansları incelendiğinde, SIFT, KAZE ve HOG öznitelikleri kullanılarak oluşturulan modellerin doğruluk değerlerinin %70,99 ile %87,18 arasında değiştiği belirlenmiştir. Geleneksel yöntemler arasında en yüksek doğruluk değeri %87,18 ile SIFT ve Logistic Regression yaklaşımında elde edilirken, bunu %86,54 doğruluk değeri ile KAZE ve Logistic Regression modeli izlemiştir. HOG tabanlı yöntemler ise tüm değerlendirmelerde daha düşük performans göstermiştir. Elde edilen bulgular, derin öğrenme tabanlı özniteliklerin göğüs röntgeni görüntülerindeki ayırt edici bilgileri el yapımı özniteliklere göre daha etkili temsil edebildiğini göstermektedir.

Çalışmada kullanılan HPFN'nin sonuçları, aynı Chest X-Ray (Pneumonia) veri setini kullanan önceki çalışmalarda raporlanan sonuçlarla karşılaştırıldığında yüksek ve rekabetçi bir performans sunduğunu göstermektedir. Bununla birlikte, farklı araştırmalardaki doğruluk değerlerinin doğrudan karşılaştırılması dikkatle yapılmalıdır. Eğitim-test bölme oranları, veri artırma işlemleri, hasta düzeyinde ayırım, görüntü ön işleme adımları, transfer öğrenme

stratejileri, sınıf dağılımları ve hiperparametre optimizasyonu çalışmalara göre değişebilmektedir. Bu nedenle HPFN'nin literatürdeki tüm yöntemlerden kesin olarak üstün olduğu sonucuna varmak yerine, aynı veri seti üzerinde bildirilen sonuçlar arasında yüksek performanslı bir yaklaşım olduğu sonucuna varılabilir.

Önerilen HPFN modeli yalnızca doğruluk bakımından değil, kesinlik, duyarlılık ve F1-Skoru bakımından da en yüksek performansı elde etmiştir. Özellikle 0,9795 duyarlılık değeri, zatürre vakalarının doğru tespit edilmesinde modelin yüksek duyarlılığa sahip olduğunu göstermektedir. Aynı şekilde 0,9708 F1-Skoru değeri, kesinlik ve duyarlılık arasında dengeli bir performans sağlandığını ortaya koymaktadır. Oluşturulan karışıklık matrisi de modelin normal ve zatürre sınıflarını yüksek doğrulukla ayırt edebildiğini göstermektedir.

Literatür karşılaştırmaları da önerilen yöntemin başarısını desteklemektedir. HPFN modeli %96,38 doğruluk değeri ile literatürde önerilen yöntemlerden daha yüksek performans göstermiştir. Bu bulgular, önerilen HPFN modelinin aynı veri seti üzerinde değerlendirilen yöntemler arasında en yüksek doğruluk değerini elde ettiğini ve literatürde raporlanan güncel yaklaşımlardan daha başarılı bir sınıflandırma performansı sunduğunu göstermektedir.

Bu çalışmanın önemli katkılarından biri, derin öğrenme modelleri ile geleneksel öznitelik tabanlı yöntemlerin aynı veri seti ve ortak değerlendirme ölçütleri altında karşılaştırılmasıdır. Bu tasarım, farklı model ailelerinin göreceli performanslarının daha tutarlı biçimde incelenmesine olanak sağlamıştır. İkinci katkı, ConvNeXt-Tiny mimarisinin yalnızca standart bir softmax sınıflandırıcıyla kullanılmaması, sınıf prototiplerinin de karar sürecine katılmasıdır. Üçüncü katkı ise prototip tabanlı sınıflandırma mantığının derin öznitelikler ile standart sınıflandırma çıktılarının birleştirildiği hibrit bir yapı içerisinde uygulanmasıdır. Bu yönleriyle çalışma, sınırlı etiketli tıbbi görüntü verilerinde temsil öğrenme ile benzerlik temelli karar vermeyi bütünleştiren bir yaklaşım sunmaktadır.

Sonuç olarak bu tez, ConvNeXt-Tiny tabanlı derin özniteliklerin prototip benzerlikleri ve softmax çıktılarıyla hibrit biçimde birleştirilmesinin göğüs röntgeni görüntülerinden zatürre sınıflandırmasında etkili bir yaklaşım olduğunu göstermiştir. Önerilen HPFN modeli, çalışmada karşılaştırılan temel derin öğrenme mimarileri ve geleneksel öznitelik tabanlı yöntemler arasında en yüksek doğruluk, kesinlik, duyarlılık ve F1-Skorunu elde etmiştir. Bulgular, parametrik sınıflandırma ile sınıf merkezlerine dayalı benzerlik bilgisinin birlikte kullanılmasının, özellikle sınırlı etiketli veri koşullarında sınıflandırma performansını

geliştirebileceğine işaret etmektedir. Bununla birlikte modelin klinik karar destek süreçlerinde kullanılabilirliğine ilişkin güçlü bir sonuca ulaşılabilmesi için bağımsız ve çok merkezli veri setlerinde dış doğrulama, ayrıntılı hata analizi, kalibrasyon değerlendirmesi ve açıklanabilirlik analizleri yapılmalıdır.

5.1. Öneriler

Bu tez kapsamında gerçekleştirilen çalışmalar yalnızca Chest X-Ray (Pneumonia) veri seti üzerinde değerlendirilmiştir. Gelecek çalışmalarda önerilen yaklaşımın farklı göğüs röntgeni veri setleri üzerinde test edilmesi, yöntemin farklı veri kaynaklarındaki performansının değerlendirilmesine katkı sağlayabilir. Gelecek çalışmalarda HPFN modelinin CheXpert, MIMIC-CXR, ChestX-ray14 ve farklı kurumlara ait bağımsız göğüs röntgeni veri setlerinde doğrulanması, özellikle farklı cihazlar, hasta grupları ve görüntüleme protokollerinden elde edilen verilerle yapılacak dış doğrulamalar, modelin genellenebilirliği hakkında daha güvenilir kanıt sağlayacaktır.

Bu çalışmada sınıflandırma performansı doğruluk, Precision, Recall ve F1-Skoru ölçütleri kullanılarak değerlendirilmiştir. Model performansının yalnızca doğruluk, kesinlik, duyarlılık ve F1-Skoruyla sınırlandırılmaması; özgüllük, ROC eğrisi, AUC, pozitif ve negatif öngörü değerleri, Matthews korelasyon katsayısı, dengeli doğruluk ve kalibrasyon ölçütleriyle de değerlendirilmesi gerekmektedir. Performans değerleri için güven aralıklarının raporlanması ve modeller arasındaki farkların uygun istatistiksel testlerle incelenmesi, daha kapsamlı bir performans analizi sağlayabilir.

HPFN mimarisinin katkısını daha açık biçimde ortaya koymak amacıyla çalışmalar yapılmalı, yalnızca ConvNeXt-Tiny ve softmax sınıflandırıcı, yalnızca prototip tabanlı sınıflandırıcı, farklı füzyon katsayıları, farklı prototip oluşturma stratejileri ve alternatif benzerlik ölçütleri karşılaştırılmalıdır. Böylece performans artışının ConvNeXt omurgasından, prototip mekanizmasından veya hibrit füzyon yapısından ne ölçüde kaynaklandığı belirlenebilir.

Gelecek araştırmalarda ConvNeXt ailesinin farklı boyutları ile EfficientNet, DenseNet, Vision Transformer ve Swin Transformer gibi güncel mimariler HPFN yapısına entegre edilebilir. Bunun yanında Mahalanobis uzaklığı ve öğrenilebilir metrik fonksiyonları gibi farklı benzerlik ölçütlerinin prototip tabanlı karar sürecine etkisi incelenebilir.

Önerilen HPFN modeli ConvNeXt-Tiny mimarisi temel alınarak geliştirilmiştir. Gelecekte ConvNeXt ailesinin farklı sürümleri veya diğer güncel derin öğrenme mimarileri kullanılarak benzer prototip tabanlı karar mekanizmasının performansı incelenebilir. Bu çalışmada prototip tabanlı karar mekanizması ile softmax tabanlı tahminler hibrit olarak birleştirilmiştir. Farklı karar füzyonu stratejilerinin veya farklı benzerlik ölçütlerinin kullanılması yöntemin sınıflandırma performansına etkisinin araştırılmasına katkı sağlayabilir.

Gelecek çalışmalarda SimCLR, MoCo, BYOL, DINO ve masked image modeling gibi self-supervised ön eğitim yaklaşımlarının HPFN yapısıyla bütünleştirilmesi değerlendirilebilir. Model öncelikle geniş miktarda etiketlenmemiş göğüs röntgeni görüntüsü üzerinde ön eğitime tabi tutulabilir, ardından daha küçük etiketli bir veri seti kullanılarak zatürre sınıflandırması için ince ayarlanabilir. Self-supervised ön eğitimden elde edilen özniteliklerin sınıf prototipleri oluşturmak için kullanılması, prototiplerin sınıf içi değişkenliği daha iyi temsil etmesini ve sınıflar arası ayrımın güçlendirilmesini sağlayabilir. Bu yaklaşımın ImageNet tabanlı transfer öğrenmeyle karşılaştırılması, tıbbi görüntüye özgü ön eğitimin katkısını ortaya koyabilir. Self-supervised learning çalışmalarında yalnızca nihai sınıflandırma performansı değil, kullanılan etiket oranına göre performans değişimi de değerlendirilmelidir. Örneğin modelin veri setinin %1, %5, %10, %25 ve %50'si etiketli olacak şekilde ayrı senaryolarda test edilmesi, yöntemin etiket verimliliğini daha açık biçimde gösterebilir. Böylece HPFN'nin gerçekten az etiketli koşullarda avantaj sağlayıp sağlamadığı, tam veri kullanılarak elde edilen tek bir performans değerinden daha güçlü biçimde ortaya konulabilir.

Tıbbi görüntüleme alanında geliştirilen foundation models da gelecek çalışmalar açısından önemli bir araştırma yönü sunmaktadır. Geniş ve heterojen veri kaynakları üzerinde önceden eğitilen foundation models, farklı görev ve veri setlerine sınırlı ek eğitimle uyarlanabilen genel amaçlı temsiller sağlayabilmektedir. Göğüs röntgenlerine özgü foundation modellerden elde edilen özniteliklerin HPFN mimarisine aktarılması, küçük veri kümelerinde sıfırdan eğitim veya yalnızca ImageNet tabanlı transfer öğrenmeye kıyasla daha güçlü bir başlangıç temsili sağlayabilir.

Son olarak, önerilen yaklaşımın yalnızca zatürre tespiti ile sınırlı kalmayıp farklı akciğer hastalıklarının veya çok sınıflı göğüs röntgeni sınıflandırma problemlerinin çözümünde de değerlendirilmesi yöntemin farklı tıbbi görüntüleme uygulamalarındaki kullanılabilirliğinin incelenmesine olanak sağlayabilir. Bununla birlikte, klinik kullanıma geçmeden önce model

ıktılarının radyolog deęerlendirmeleriyle karşılařtırıldıęı ileriye dnk, ok merkezli ve gerek yařam kořullarını yansıtan alıřmaların yrtlmesi gerekmektedir.

KAYNAKÇA

- An, Q., Chen, W., & Shao, W. (2024). A Deep Convolutional Neural Network for Pneumonia Detection in X-ray Images with Attention Ensemble. *Diagnostics*, 14(4), 1–23.
<https://doi.org/10.3390/diagnostics14040390>
- Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016, July 21). Layer Normalization. *ArXiv Preprint*.
<http://arxiv.org/abs/1607.06450>
- Bhatti, Y. I., Aslam, N., Naeem, A., & Mazhar, F. (2025). Enhanced Brain Tumor Grade Classification with ConvNext: Performance Analysis. *Journal of Computing & Biomedical ...*, 09(01).
<https://www.jcbi.org/index.php/Main/article/view/1027%0Ahttps://www.jcbi.org/index.php/Main/article/download/1027/715>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen, W. Y., Liu, Y. C., Kira, Z., Wang, Y. C. F., & Huang, J. B. (2019). A closer look at few-shot classification. *International Conference on Learning Representations*.
- Cheplygina, V., de Bruijne, M., & Pluim, J. P. W. (2019). Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. *Medical Image Analysis*, 54, 280–296. <https://doi.org/10.1016/J.MEDIA.2019.03.009>
- Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1251–1258. <http://arxiv.org/abs/1610.02357>
- Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. In *Machine Learning* (Vol. 20). Kluwer Academic Publishers. <https://doi.org/10.1007/BF00994018>
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, I*, 886–893. <https://doi.org/10.1109/CVPR.2005.177>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). an Image Is Worth 16X16 Words: Transformers for Image Recognition At Scale. *ICLR 2021 - 9th International Conference on Learning Representations*.
- Elemraid, M. A., Muller, M., Spencer, D. A., Rushton, S. P., Gorton, R., Thomas, M. F.,

- Eastham, K. M., Hampton, F., Gennery, A. R., Clark, J. E., & Group, on behalf of the N. E. of E. P. R. I. S. (2014). Accuracy of the Interpretation of Chest Radiographs for the Diagnosis of Paediatric Pneumonia. *PLoS ONE*, 9(8), e106051. <https://doi.org/10.1371/journal.pone.0106051>
- Fernández Alcantarilla, P., Bartoli, A., & Davison, A. J. (2012). *LNCS 7577 - KAZE Features*.
- Franquet, T., Chung, J. H., Hodler, J., Kubik-Huch, R. A., & Schulthess, G. K. von. (2019). *Imaging of Pulmonary Infection*. (J. Hodler, R. A. Kubik-Huch, & G. K. von Schulthess (eds.); pp. 65–77). https://doi.org/10.1007/978-3-030-11149-6_7
- Garg, M., Prabhakar, N., Gulati, A., Agarwal, R., & Dhooria, S. (2019). Spectrum of imaging findings in pulmonary infections. Part 1: Bacterial and viral. *Polish Journal of Radiology*, 84, 205–213. <https://doi.org/10.5114/pjr.2019.85812>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Hendrycks, D., & Gimpel, K. (2023, June 6). Gaussian Error Linear Units (GELUs). *ArXiv Preprint*. <http://arxiv.org/abs/1606.08415>
- Hopstaken, R. M., Witbraad, T., van Engelshoven, J. M. A., & Dinant, G. J. (2004). Inter-observer variation in the interpretation of chest radiographs for pneumonia in community-acquired lower respiratory tract infections. *Clinical Radiology*, 59(8), 743–752. <https://doi.org/10.1016/J.CRAD.2004.01.011>
- Hosmer, D. W., Lemeshow, J. S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. Wiley Series in Probability and Statistics.
- Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghighi, B., Ball, R., Shpanskaya, K., Seekins, J., Mong, D. A., Halabi, S. S., Sandberg, J. K., Jones, R., Larson, D. B., Langlotz, C. P., Patel, B. N., Lungren, M. P., & Ng, A. Y. (2019). CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*. <http://arxiv.org/abs/1901.07031>
- Johnson, A. E. W., Pollard, T. J., Berkowitz, S. J., Greenbaum, N. R., Lungren, M. P., Deng, C. ying, Mark, R. G., & Horng, S. (2019). MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Scientific Data*, 6(317). <https://doi.org/10.1038/s41597-019-0322-0>
- Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L.,

- McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., ... Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, 172(5), 1122–1131.e9. <https://doi.org/https://doi.org/10.1016/j.cell.2018.02.010>
- Kornblith, S., Shlens, J., & Le, Q. V. (2019). Do Better ImageNet Models Transfer Better? *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2661–2671. <http://arxiv.org/abs/1805.08974>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25. <https://doi.org/10.1016/B978-0-12-374105-9.00493-7>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. <https://doi.org/10.1038/nature14539>
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42(1), 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- Liu, A., & Yue, H. (2023). Facial Expression Recognition Based on CNN-LSTM. *Proceedings of the 2023 7th International Conference on Electronic Information Technology and Computer Engineering*. <https://doi.org/10.1145/3650400.3650480>
- Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A ConvNet for the 2020s. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2022-June*, 11966–11976. <https://doi.org/10.1109/CVPR52688.2022.01167>
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of the IEEE International Conference on Computer Vision*, 2, 1150–1157. <https://doi.org/10.1109/iccv.1999.790410>
- Mabrouk, A., Redondo, R. P. D., Dahou, A., Elaziz, M. A., & Kayed, M. (2023). *Pneumonia Detection on chest X-ray images Using Ensemble of Deep Convolutional Neural Networks*. <https://doi.org/10.3390/app12136448>
- Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359. https://doi.org/10.1007/978-981-15-5971-6_83
- Perera, I., & Thayasivam, U. (2025). *A Novel Multi-branch ConvNeXt Architecture for Identifying Subtle Pathological Features in CT Scans*. <http://arxiv.org/abs/2510.09107>

- Pooch, E. H. P., Ballester, P. L., & Barros, R. C. (2020). *Can we trust deep learning models diagnosis? The impact of domain shift in chest radiograph classification*.
https://doi.org/10.1007/978-3-030-62469-9_7
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). *CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning*.
<http://arxiv.org/abs/1711.05225>
- Riasat, K., Jamil, A., Al-Otaibi, S., Zeb, S., Riasat, S., & Kanwal, S. (2025). Tuberculosis detection using few shot learning. *Scientific Reports*, 15(1).
<https://doi.org/10.1038/s41598-025-97803-9>
- Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A. I., Etmann, C., McCague, C., Beer, L., Weir-McCall, J. R., Teng, Z., Gkrania-Klotsas, E., Ruggiero, A., Korhonen, A., Jefferson, E., Ako, E., Langs, G., Gozaliasl, G., ... Schönlieb, C.-B. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence*, 3(3), 199–217. <https://doi.org/10.1038/s42256-021-00307-0>
- Sharma, S., & Guleria, K. (2023). A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks. *Procedia Computer Science*, 218, 357–366. <https://doi.org/10.1016/j.procs.2023.01.018>
- Shen, D., Wu, G., & Suk, H.-I. (2017). Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering*, 19(1), 221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298.
<https://doi.org/10.1109/TMI.2016.2528162>
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.
- Snell, J., Swersky, K., & Zemel, R. S. (2017). *Prototypical Networks for Few-shot Learning*.
<http://arxiv.org/abs/1703.05175>
- Stephen, O., Sain, M., Maduh, U. J., & Jeong, D.-U. (2019). An Efficient Deep Learning

- Approach to Pneumonia Classification in Healthcare. *Journal of Healthcare Engineering*, 1–7. <https://doi.org/10.1155/2019/4180949>
- Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging*, 35(5), 1299–1312. <https://doi.org/10.1109/TMI.2016.2535302>
- Wang, L., Lin, Z. Q., & Wong, A. (2020). COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-76550-z>
- Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2019). ChestX-ray: Hospital-Scale Chest X-ray Database and Benchmarks on Weakly Supervised Classification and Localization of Common Thorax Diseases. In *Advances in Computer Vision and Pattern Recognition* (pp. 369–392). https://doi.org/10.1007/978-3-030-13969-8_18
- Wootton, D., & Feldman, C. (2014). The diagnosis of pneumonia requires a chest radiograph (x-ray)—yes, no or sometimes? *Pneumonia*, 5(S1), 1–7. <https://doi.org/10.15172/pneu.2014.5/464>
- World Health Organization, W. (2026, August 1). *Pneumonia*. https://www.who.int/health-topics/pneumonia#tab=tab_1.
- Xia, J., Yin, Y., & Li, X. (2025). *An Efficient Medical Image Classification Method Based on a Lightweight Improved ConvNeXt-Tiny Architecture*. 1–12. <http://arxiv.org/abs/2508.11532>
- Xiong, K., Tu, Y., Rao, X., Zou, X., & Du, Y. (2025). Multi-Label Disease Detection in Chest X-Ray Imaging Using a Fine-Tuned ConvNeXtV2 with a Customized Classifier. *Informatics*, 12(3). <https://doi.org/10.3390/informatics12030080>