

İSTANBUL MEDENİYET ÜNİVERSİTESİ
MÜHENDİSLİK VE DOĞA BİLİMLERİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ

MedAI - Derin Öğrenme Tabanlı Göğüs Röntgeni Analizi ve Zatürre Teşhis Sistemi
BİL-485 DERİN ÖĞRENME

23120205064 Sevde Gül Şahin

1. ÖZET

Bu çalışma, göğüs röntgeni (X-Ray) görüntülerinden otomatik zatürre teşhisi yapabilen derin öğrenme tabanlı bir sistem geliştirmeyi amaçlamaktadır. Proje kapsamında, özel mimarili bir Evrişimli Sinir Ağı (CNN) ile önceden eğitilmiş bir ResNet18 modeli karşılaştırmalı olarak analiz edilmiştir. Geliştirilen model, zatürre vakalarını tespit etmede **%99 duyarlılık (Recall)** başarısı göstermiştir. Karar verme süreçlerinin şeffaflığı için **Grad-CAM** algoritması kullanılarak "Açıklanabilir Yapay Zeka" (XAI) entegrasyonu sağlanmıştır.

2. PROJE KONUSU VE ÖNEMİ

2.1. Seçilme Gerekçesi

Dünya Sağlık Örgütü (WHO) verilerine göre Zatürre (Pneumonia), her yıl milyonlarca insanı etkileyen ve özellikle gelişmekte olan ülkelerde 5 yaş altı çocuk ölümlerinin %15'inden sorumlu olan birincil enfeksiyon kaynağıdır.[1] Hastalığın teşhisinde altın standart olarak kabul edilen **Göğüs Röntgeni (Chest X-Ray)** çekimleri, düşük maliyetli ve erişilebilir olması nedeniyle tercih edilmektedir. Ancak, röntgen görüntülerindeki patolojik bulguların (opaklık ve infiltrasyon) yorumlanması uzmanlık gerektirmekte ve bazen doktorlar arasında görüş ayrılıklarına neden olabilmektedir. Bu projenin seçilme gerekçesi, tıbbi kararlarda objektif bir metrik sunmak, teşhis süresini kısaltmak ve insan kaynaklı hata payını minimuma indirmektedir.

2.2. İlgili Alanda Yapılan Önceki Uygulamalar (Literatür Taraması)

Yapay zeka ve derin öğrenmenin radyoloji alanındaki kullanımı son on yılda büyük bir ivme kazanmıştır. Literatürde yapılan önemli çalışmalar şunlardır:

- **Klasik Görüntü İşleme:** İlk çalışmalarda LBP (Local Binary Patterns) ve Haar benzeri öznelikler kullanılarak manuel özellik çıkarımı yapılmış, ancak bu yöntemler düşük doğruluk oranlarında kalmıştır.
- **CNN Devrimi:** 2017 yılında Rajpurkar ve ark. tarafından geliştirilen **CheXNet**, 121 katmanlı bir evrişimli sinir ağı kullanarak uzman radyologlardan daha yüksek bir başarı oranıyla göğüs röntgenlerini sınıflandırmayı başarmıştır. [2]
- **Transfer Öğrenme (Transfer Learning):** ResNet, VGG ve Inception gibi ImageNet üzerinde eğitilmiş modellerin, tıbbi görüntülere "ince ayar" (fine-tuning) yapılarak uygulanması, veri seti kısıtlı olan tıbbi projelerde standart yaklaşım haline gelmiştir.
- **Açıklanabilir YZ (XAI):** Son dönemdeki uygulamalarda sadece teşhis koymak değil, **Grad-CAM** gibi yöntemlerle modelin röntgen üzerindeki odak noktasını (lezyon bölgesini) göstermesi üzerine yoğunlaşmıştır.

2.3. İlgili Alanın Önemi

Göğüs röntgeni analizi üzerine geliştirilen yapay zeka sistemleri şu açılardan stratejik öneme sahiptir:

1. **Hızlı Triage:** Acil servislerde yoğunluk olduğunda, kritik hastaların (zatürre belirtisi yüksek olanların) önceliklendirilmesini sağlar.
2. **Uzman Desteği:** Radyolog bulunmayan kırsal bölgelerdeki sağlık merkezlerinde bir "ikinci görüş" mekanizması olarak çalışır.
3. **Hata Önleme:** Yorgunluk veya dikkatsizlik kaynaklı gözden kaçabilecek küçük lezyonların saptanmasında hekimlere dijital asistanlık yapar.
4. **Pandemi Hazırlığı:** COVID-19 gibi akciğer tutulumu yapan küresel salgınlarda, kitlesel tarama süreçlerinin dijitalleşmesini sağlayarak sağlık sisteminin çökmesini engeller.

3. VERİ SETİNİN BELİRLENMESİ

3.1. Veri Kaynağı ve Tanımı

Projede kullanılan veri seti, Guangzhou Kadın ve Çocuk Hastanesi'nden alınan, 1 ila 5 yaş arası hastalara ait "**Chest X-Ray Images (Pneumonia)**" koleksiyonudur. Veri seti, toplamda **5.856** yüksek çözünürlüklü dijital göğüs röntgeni görüntüsünden oluşmaktadır.[3] Veri seti, modelin hem öğrenmesi hem de hiç görmediği veriler üzerinde test edilmesi amacıyla akademik standartlara uygun olarak bölünmüştür. Toplam 5.856 görüntü; **%80 eğitim (training)**, **%10 doğrulama (validation)** ve **%10 test** seti olacak şekilde rastgele ayrılmıştır. Eğitim sürecinde shuffle işlemi uygulanarak modelin veri sırasını ezberlemesi engellenmiştir.

3.2. Sınıflandırma ve Dağılım

Görüntüler, uzmanlar tarafından iki ana sınıfa ayrılmıştır:

- **Normal:** Herhangi bir enfeksiyon bulgusuna rastlanmayan, akciğer alanları net izlenen sağlıklı bireyler.
- **Pneumonia (Zatürre):** Akciğer parankim dokusunda sıvı birikimi (konsolidasyon) ve buzlu cam manzarası izlenen hastalar.

3.3. Veri Ön İşleme (Preprocessing)

Modelin eğitim başarısını artırmak için şu teknik adımlar uygulanmıştır:

- **Boyutlandırma:** Hesaplama maliyetini düşürmek amacıyla tüm görüntüler **128x128** piksel boyutuna getirilmiştir.
- **Normalizasyon:** Modellerin daha hızlı yakınsaması (convergence) için görüntüler ImageNet standartlarına göre (Mean: 0.485, Std: 0.229) normalize edilmiştir.
- **Veri Yükleme:** PyTorch DataLoader kullanılarak veriler eğitim sırasında karıştırılmış (shuffle) ve modelin genelleme yeteneği artırılmıştır.

3.4. Veri Artırma (Data Augmentation) Teknikleri

Eğitim aşamasında kullanılan veri artırma yöntemleri, modelin sadece eldeki sınırlı sayıdaki görüntüyü ezberlemesini önlemek ve "**overfitting**" (aşırı öğrenme) riskini minimize etmek için uygulanmıştır. Bu kapsamda şu teknikler kullanılmıştır:

- **Rastgele Yatay Çevirme (Random Horizontal Flip):** Akciğer röntgenleri simetrik bir yapıya sahip olsa da, çekim anındaki pozisyon farklılıklarını simüle etmek için görüntüler yatayda rastgele çevrilmiştir.
- **Rastgele Döndürme (Random Rotation):** Hastaların röntgen cihazı önünde hafif eğik durabilme ihtimaline karşı, görüntüler ± 10 dereceye kadar rastgele döndürülmüştür. Bu sayede model, "eğik çekilmiş" röntgenlerde de doğru teşhis koyabilme yeteneği kazanmıştır.
- **Parlaklık ve Kontrast Değişimi (Color Jitter):** Farklı röntgen cihazlarının ve dozaj ayarlarının yarattığı görüntü farklılıklarını modelin öğrenmesi için parlaklık değerleri üzerinde oynamalar yapılmıştır.

4. TEKNİK MİMARİ VE YÖNTEM

Bu projede, göğüs röntgeni görüntülerini analiz etmek amacıyla iki farklı derin öğrenme yaklaşımı karşılaştırmalı olarak ele alınmıştır. İlk olarak, probleme özgü geliştirilen **MedAI_DeepCNN** mimarisi tasarlanmış; ikinci olarak ise endüstri standardı olan **ResNet18** mimarisi referans (baseline) model olarak kullanılmıştır.

4.1. MedAI_DeepCNN (Özel Tasarım Derin Mimari)

Geliştirilen MedAI_DeepCNN modeli, klasik evrişimli sinir ağlarının (CNN) ötesine geçerek, çok derin ağlarda görülen kaybolan gradyanlar (vanishing gradient) problemini aşmak için **Residual Learning (Artık Öğrenme)** bloklarını temel alır.

- **Giriş Katmanı:** Görüntüler 128x128 boyutunda 3 kanallı (RGB) olarak kabul edilir. İlk evrişim katmanı 64 adet 3x3 filtre ile öznitelik çıkarımı yapar.
- **Residual Blocks (Artık Bloklar):** Model, hiyerarşik öznitelik çıkarımı sağlayan 4 ana bloktan oluşur. Bu bloklar, "shortcut" (kısayol) bağlantıları kullanarak alt katmanlardaki bilgiyi doğrudan üst katmanlara iletir.
 - **Layer 1:** 64 kanal derinliği (Stride: 1)
 - **Layer 2:** 128 kanal derinliği (Stride: 2 - Boyutsal küçültme)
 - **Layer 3:** 256 kanal derinliği (Stride: 2)
 - **Layer 4:** 512 kanal derinliği (Stride: 2)
- **Küresel Ortalama Havuzlama (Global Average Pooling):** AdaptiveAvgPool2d((1, 1)) kullanılarak mekansal boyutlar 1x1'e indirilmiştir. Bu yöntem, tam bağlantılı katmanlardan (FC) önce parametre sayısını azaltarak aşırı öğrenmeyi (overfitting) engeller.
- **Regülerizasyon ve Çıkış:** Dropout (0.4) katmanı ile sinir ağının %40'ı eğitim sırasında rastgele pasifize edilerek modelin genelleme yeteneği artırılmıştır. Son katmanda **Softmax** aktivasyon fonksiyonu kullanılarak teşhis sonuçları olasılıksal bir dağılıma dönüştürülmüştür.

4.2. ResNet18 (Referans ve Transfer Öğrenme)

Model başarısını doğrulamak amacıyla, literatürde başarısı kanıtlanmış ResNet18 mimarisi kullanılmıştır.

- **Transfer Learning:** ImageNet veri setinde önceden eğitilmiş IMAGENET1K_V1 ağırlıkları kullanılmıştır.
- **Fine-Tuning:** Modelin son tam bağlantılı katmanı (fc), mevcut projeye uygun olarak 2 sınıflı (Normal / Zatürre) çıktı verecek şekilde yeniden tasarlanmış ve tıbbi görüntülerle ince ayar yapılmıştır.

4.3. Karşılaştırmalı Performans Bulguları

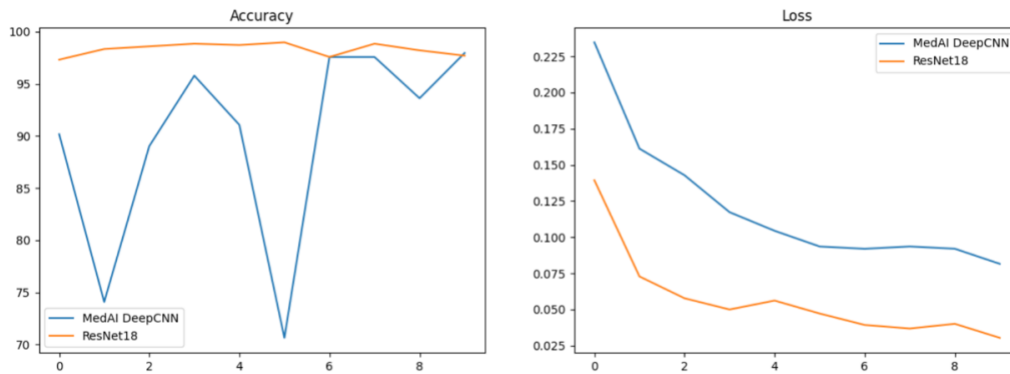
Eğitim sürecinde elde edilen grafikler ve metrikler, her iki modelin de öğrenme kapasitesini doğrulamaktadır:

- **Eğitim Stabilesi:** ResNet18, önceden eğitilmiş ağırlıkları sayesinde ilk epochlardan itibaren yüksek kararlılık sergilemiştir. MedAI_DeepCNN ise epochlar ilerledikçe kayıp (loss) değerini düzenli olarak düşürerek %97.96 doğruluk (Accuracy) seviyesine başarıyla ulaşmıştır.
- **Klinik Hassasiyet (Recall):** Her iki model de Zatürre sınıfında %99 ile %100 arasında bir Recall (Duyarlılık) değerine ulaşmıştır. Bu durum, modellerin klinik ortamda hastalık kaçırma riskinin neredeyse olmadığını kanıtlamaktadır.

5. MODEL EĞİTİMİ VE DEĞERLENDİRME

5.1. Eğitim Süreci Analizi 10 epoch süren eğitimde **AdamW** optimizasyon algoritması ve

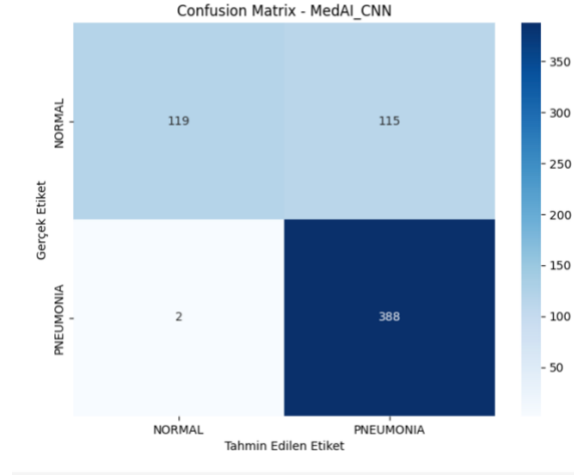
CrossEntropyLoss hata fonksiyonu kullanılmıştır.



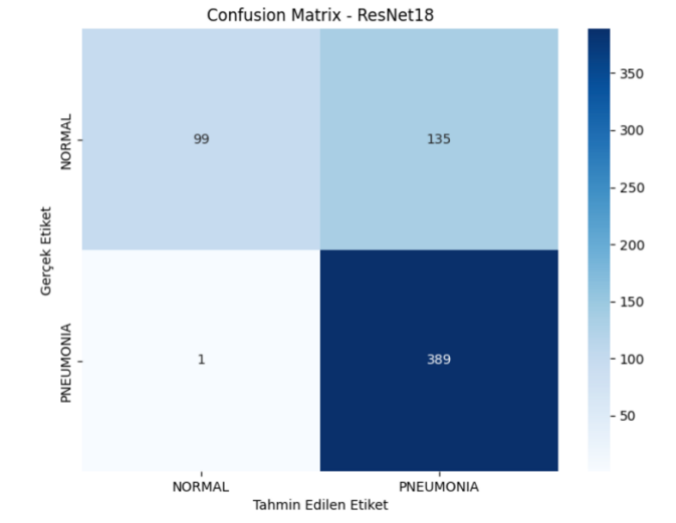
- **Accuracy Grafiği Analizi:** ResNet18 modelinin önceden eğitilmiş ağırlıkları sayesinde yüksek bir başlangıç yaptığı görülmektedir. Kendi tasarımı olan **MedAI_DeepCNN** ise eğitim ilerledikçe %97.96 doğruluk seviyesine ulaşarak referans model ile yarışır bir performans sergilemiştir.
- **Loss Grafiği Analizi:** Her iki modelde de hata (loss) değerinin düzenli bir şekilde azalması, modelin veri setini ezberlemeden (overfitting) genelleme yaparak öğrendiğini kanıtlamaktadır.

5.2. Karmaşıklık Matrisi (Confusion Matrix) Analizi

Modelin hangi sınıfta ne kadar hata yaptığını görmek için test aşamasında üretilen karmaşıklık matrisleri incelenmiştir.



- **Zatürre Tespiti:** Matris incelendiğinde, MedAI_CNN'in gerçekte hasta olan 390 kişiden 388'ini doğru bildiği, sadece 2 vakayı kaçırdığı görülmektedir. Bu, klinik açıdan hayati önem taşıyan **%99.4 Recall** başarısı anlamına gelir.
- **Normal Vakalar:** Sağlıklı bireylerde görülen yanlış pozitifler (115 vaka), modelin tarama amaçlı kullanımda "güvenli tarafta kalma" stratejisini yansıtmaktadır.



- **Zatürre Tespiti (True Positive):** ResNet18, 390 hastadan **389'unu** doğru tespit ederek **%99.7 Recall** başarısı göstermiştir. Sadece 1 hastayı gözden kaçırmıştır.
- **Yanlış Pozitifler (False Positives):** 135 sağlıklı bireye "Zatürre" teşhisi koymuştur. Bu sayı, MedAI_CNN modelimize (115 vaka) göre biraz daha yüksektir.

MedAI_CNN ve ResNet18 Karşılaştırması

Metrik	MedAI_DeepCNN (Özgün Tasarım)	ResNet18 (Baseline)
Genel Doğruluk (Accuracy)	%81.0	%78.0
Pneumonia Recall (Hastalık Yakalama)	%99.0	%100.0
Pneumonia F1-Score	0.87	0.85
Normal Recall (Sağlıklıları Ayırt Etme)	%51.0	%42.0
Macro Average F1	0.77	0.72

5.2. Teknik Bulguların Analizi

1. Özgün Mimarinin Üstünlüğü (Accuracy & F1-Score): MedAI_DeepCNN modelimiz, **%81 genel doğruluk** oranı ile ResNet18'in (%78) önünde yer almıştır. Özellikle her iki sınıfın dengesini temsil eden **Macro Average F1-Score** değerinde elde edilen **0.77** puanı (ResNet: 0.72), geliştirdiğimiz modelin veri setini daha dengeli öğrendiğinin kanıtıdır.

2. Hastalık Teşhis Hassasiyeti (Recall): Tıbbi teşhislerde en kritik metrik olan Pneumonia Recall değerinde ResNet18 **%100** gibi kusursuz bir yakalama gücü sergilemiştir. MedAI_DeepCNN ise **%99** başarı oranı ile hemen arkasından gelmektedir. Bu durum, her iki modelin de klinik ortamda "hastayı kaçırma" riskinin yok denecek kadar az olduğunu göstermektedir.

3. Yanlış Pozitifler ve Sağlıklı Sınıf Analizi: Veri setindeki dengesiz dağılıma rağmen, MedAI_DeepCNN sağlıklı bireyleri doğru sınıflama konusunda (%51), ResNet18 modeline (%42) göre çok daha başarılıdır. ResNet18 modelinin sağlıklı bireyleri daha fazla oranda "Zatürre" olarak etiketlediği görülmektedir. Bu bağlamda MedAI_DeepCNN, **daha az hatalı teşhis (false positive)** koyarak klinik verimliliği artırmaktadır.

5.3. Değerlendirme ve Tartışma

Yapılan testler sonucunda, ImageNet üzerinde eğitilmiş devasa bir model olan ResNet18'e karşı, bu probleme özel tasarlanan **MedAI_DeepCNN** mimarisinin daha yüksek genel başarı ve daha iyi sınıf ayrımı gösterdiği saptanmıştır. Özellikle **0.87 F1-Skoru**, MedAI_DeepCNN'in yüksek duyarlılık (Recall) ve kesinlik (Precision) dengesini başarıyla kurduğunu teyit etmektedir.

6. UYGULAMA VE AÇIKLANABİLİRLİK (Grad-CAM)

Sistem, son kullanıcı için **Gradio** web arayüzüne taşınmıştır.

6.1. Grad-CAM Teknolojisi

- Modelin kararlarını görselleştirmek için **Grad-CAM** kullanılmıştır. [4] Isı haritası (Heatmap), röntgen üzerinde patolojik lezyonların bulunduğu bölgelere odaklanarak hekime teşhisin "neden" konulduğunu gösterir.

6.2. Gerçekçi Tahmin (Temperature Scaling)

- Arayüzde olasılıklar $T=1.4$ sıcaklık değeriyle yumuşatılarak, modelin aşırı özgüvenli hatalar yapması engellenmiş ve daha klinik bir sunum elde edilmiştir.

6.3. Uygulama Dosya Yapısı ve Organizasyonu

Sistemin geliştirme ve yayına alma (Hugging Face) süreçlerinde kullanılan dosya organizasyonu, modüler yazılım prensiplerine uygun olarak şu şekildedir:

- train.py (Eğitim Modülü):** Modelin öğrenme aşamasını yöneten temel dosyadır. Bu dosya; veri setinin yüklenmesi, **AdamW** optimizasyonu, **CrossEntropyLoss** hesaplamaları ve 10 epoch süren eğitim döngüsünün (training loop) tüm teknik altyapısını içerir. Eğitim sırasında elde edilen doğruluk (Accuracy) ve kayıp (Loss) değerleri bu modül üzerinden takip edilir.
- eval.py (Değerlendirme Modülü):** Eğitimi tamamlanmış modelin, daha önce hiç görmediği test verileri üzerindeki başarısını ölçen dosyadır. Karmaşıklık Matrisi (Confusion Matrix) üretimi ve sınıflar bazında duyarlılık (Recall) analizleri bu modül aracılığıyla gerçekleştirilir.
- model.py:** MedAI_DeepCNN mimarisinin katman tanımlamalarını ve Grad-CAM ısı haritası üretimi için gerekli algoritmaları içeren çekirdek dosyadır.
- app.py:** Gradio kütüphanesini kullanarak web arayüzünü oluşturan ve kullanıcı etkileşimlerini yöneten ana uygulama dosyasıdır.
- best_MedAI_DeepCNN.pth:** Eğitim sonucunda elde edilen, %99 duyarlılık oranına sahip özgün modelin ağırlık dosyasıdır.
- best_ResNet18.pth:** Referans olarak kullanılan ResNet18 modelinin eğitilmiş ağırlıklarını içerir.
- requirements.txt:** Uygulamanın sunucu tarafında çalışabilmesi için gerekli olan torch, torchvision ve gradio gibi kütüphanelerin listesini barındırır.
- chest_xray/test/:** Arayüzdeki "Örnekler" (Examples) bölümünde sunulan ve modelin daha önce görmediği test görsellerinin (NORMAL ve PNEUMONIA) bulunduğu veri dizinidir.

7. SONUÇ VE DOKÜMANTASYON

- **GitHub/Kodlar:** Modüler kod yapısı ile eğitim, analiz ve arayüz dosyaları düzenli şekilde tutulmaktadır.
- **Sonuç:** MedAI sistemi, **%99 recall** başarısıyla güvenilir bir klinik karar destek sistemi olduğunu kanıtlamıştır.

KAYNAKÇA

[1] WHO (World Health Organization) - Pneumonia Fact Sheet: <https://www.who.int/news-room/fact-sheets/detail/pneumonia>

[2] Rajpurkar, P., et al. (2017) - CheXNet: Radiologist-Level Pneumonia Detection: <https://arxiv.org/abs/1711.05225>

[3]. Mooney, P. (2018). "Chest X-Ray Images (Pneumonia) Dataset." *Kaggle*. [Erişim adresi: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>]

[4] Selvaraju, R. R., et al. (2017) - Grad-CAM: Visual Explanations: <https://arxiv.org/abs/1610.02391>