

# Kart Bilgisi ve Uygulama Bilgisi v1.0Kart Bilgisi ve Uygulama Bilgisi v1.0

Revizyon Kaydı

**Uygulama Adı:** LujainAI Akciğer Opasitesi Tespit Aracı

**Model Adı:** exp114\_model.pt

**Uygulama Dağıtım Durumu:** Prototip

**Versiyon:** 1.0

**Tarih:** 1-Ocak-2025

**Oluşturan:** Lujain

**Notlar:** Tanıtım amacıyla yayınlandı

## Bilinmeyen Sorunlar

**1-** Latans (tahmin süresi) her resim için yaklaşık 4 saniyedir. Bu, çıkarım işlemi için CPU kullanıldığı içindir. Çıkarım, bir GPU ile daha hızlı olacaktır.

**2-** Uygulama bazen birden fazla örtüşen sınırlama kutusu (bounding box) çıktısı veriyor. Bununla ilgili bir düzeltme yapmadım çünkü bu davranışın bir anlamı olabilir ki bunu tam olarak anlamıyorum, ancak bir radyolog için önemli olabilir. Sınırlama kutuları, sadece resme tıklanarak gizlenebilir.

**3-** Uygulama yanlış pozitifler (false positives) tahmin etmektedir. Uygulama, opasiteyi % 100 doğrulukla tespit etmemektedir.

-X-

## Özet

LujainAI Akciğer Opasitesi Tespit Aracı, göğüs röntgenlerinde opasiteleri tespit etmek için yapay zeka kullanan ücretsiz ve açık kaynaklı bir masaüstü uygulamasının kavramsal bir kanıtıdır.

Opasiteler, Tüberküloz (TB) ve Zatürre gibi akciğer hastalıklarının belirgin belirtileridir. Bu uygulama, göğüs röntgenlerini analiz eder ve opasitelerin etrafına sınırlama kutuları çizer. Radyologlar, bu ilgi alanlarını gözden geçirebilir ve nihai teşhisi koyarken klinik yargılarını kullanabilirler.

Yoksul ülkelerde radyolog eksikliği bulunmaktadır. 2015 yılında Ruanda, 12 milyonluk nüfusuna hizmet vermek için yalnızca 11 radyologa sahipti. Liberya ise, dört milyonluk nüfusu için iki pratik yapan radyologa sahipti. Bu uygulama, yüksek hacimli teşhis desteği sunmaktadır.

Radyologların röntgenleri hızla sınıflandırmalarına ve iş akışlarını hızlandırmalarına yardımcı olabilir.

Tahminler, Pytorch Faster R-CNN modeliyle yapılmaktadır. Model, dört göğüs röntgeni veri setinden alınan verilerle ince ayar yapılmıştır:

- TBX11K Tüberküloz veri seti
- Kaggle RSNA Zatürre Tespiti Yarışması
- Kaggle VinBigData Göğüs Röntgeni Anormallikleri Tespiti Yarışması
- Kaggle SIIM-FISABIO-RSNA COVID-19 Tespiti Yarışması

Uygulama, opasite sınırlama kutuları gösterse de, model aynı zamanda akciğerleri tespit etmek için de eğitilmiştir, yani model her iki akciğeri çevreleyen bir sınırlama kutusu tahmin etmektedir. Eğer model akciğerleri tespit edemezse, uygulama bir hata mesajı verir.

Model, %80/%20 eğitim-test bölünmesi ile doğrulandı. Ayrıca üç farklı örnek dışı veri seti üzerinde de test edilmiştir:

- Shenzhen ve Montgomery Tüberküloz veri setleri
- DA ve DB Tüberküloz veri setleri
- Çocuk Göğüs Röntgeni Zatürre veri seti

Bu örnek dışı veri setlerinde opasite sınırlama kutuları bulunmamaktadır. Bu nedenle, doğruluk kaba bir ölçüt olarak kullanılmıştır - eğer hedef pozitifse (örneğin TB pozitif) ve model bir sınırlama kutusu tahmin etmişse, model doğru tahmin yapmış kabul edilmiştir. Bu doğrulama yaklaşımı çok titiz değildir. Ancak, modelin yeteneklerini hızlıca değerlendirmek için basit ve hızlı bir yoldur.

%20 doğrulama verisindeki sonuçlar: map@0.5: 0.776  
Doğruluk: 0.91

Örnek dışı veri setlerindeki doğruluk:

- Shenzhen ve Montgomery TB veri setleri: 0.84
- DA ve DB TB veri setleri: 0.85
- Çocuk Göğüs Röntgeni Zatürre veri seti: 0.83

Göğüs röntgenlerini okumak insanlar için zordur. Bir çalışmada (TBX11k makalesi), radyologların göğüs röntgenlerinde TB teşhisi koyarken %68.7 doğruluk sağladığı bulunmuştur. Bu sayıyı referans alarak, bu modelin test sonuçları oldukça iyi görünmektedir. Çocuk zatürre verilerindeki iyi performans, eğitim verilerinde büyük bir çocuk röntgeni sayısının bulunmaması nedeniyle şaşırtıcıdır.

Bu sonuçlar, bu opasite tespit uygulamasının TB ve Zatürre gibi akciğer hastalıklarının teşhisinde yardımcı olabileceğini göstermektedir. Bu uygulamayı denediğiniz için teşekkür ederiz. Lütfen geri bildiriminizi tartışma forumunda paylaşmaktan çekinmeyin:

<https://github.com/Luceyn-labed/LujainAI>

## **Hasta Veri Güvenliđi**

Bu bir masaüstü uygulamasıdır.

- Hasta verileri, kullanıcının bilgisayarından veya dizüstü bilgisayarından asla çıkmaz.
- Takip yapılmaz.
- Kod tamamen erişilebilir olduđu için kötü amaçlı yazılımlar için denetlenebilir.

## **Girdi**

Uygulama, göğüs röntgenlerini dicom, png ve jpg formatlarında kabul eder. Aynı anda birden fazla dosya gönderilebilir.

## **Çıktı**

Uygulama, tespit edilen opasitelerin etrafına sınırlama kutuları çizilmiş resimler üretir.

## **Veri Özeti**

Toplam resim sayısı: 17.283

Toplam opasite resmi sayısı: 10.983

Toplam opasite olmayan resim sayısı: 6.300

Bu veriler, dört göğüs röntgeni veri setinden örneklenmiştir:

- TBX11K Tüberküloz veri seti
  - Toplam resim sayısı: 2.299
  - Toplam opasite resmi sayısı: 799
  - Toplam opasite olmayan resim sayısı: 1.100
- Kaggle RSNA Zatürre Tespiti Yarışması
  - Toplam resim sayısı: 14.861
  - Toplam opasite resmi sayısı: 6.011
  - Toplam opasite olmayan resim sayısı: 2.700
- Kaggle VinBigData Göğüs Röntgeni Anormallikleri Tespiti Yarışması
  - Toplam resim sayısı: 2.752
  - Toplam opasite resmi sayısı: 1.322
  - Toplam opasite olmayan resim sayısı: 1.100
- Kaggle SIIM-FISABIO-RSNA COVID-19 Tespiti Yarışması
  - Toplam resim sayısı: 4.587
  - Toplam opasite resmi sayısı: 2.851
  - Toplam opasite olmayan resim sayısı: 1.400

## Veri Ön İşleme

Veri ön işleme, tüm resimlerin gri tonlamaya dönüştürülmesini, resimlerin kare olacak şekilde yastıklanmasını ve normalizasyonu içeriyordu. Faster R-CNN, tüm resimlerin aynı boyuta dönüştürülmesini gerektirmez.

## Eğitim ve doğrulama veri setlerinin oluşturulması

Yerel doğrulama verisi, %80/%20'lik bir bölünme ile oluşturuldu. Veri, her dört veri setinden de verilerin hem eğitim hem de doğrulama setlerine dahil edilmesini sağlamak için kaynak veriye göre katmanlı hale getirildi. Veri ayrıca hedefe göre katmanlı hale getirildi, yani opasite veya opasite olmayan olarak sınıflandırıldı.

## Doğrulama Sonuçları

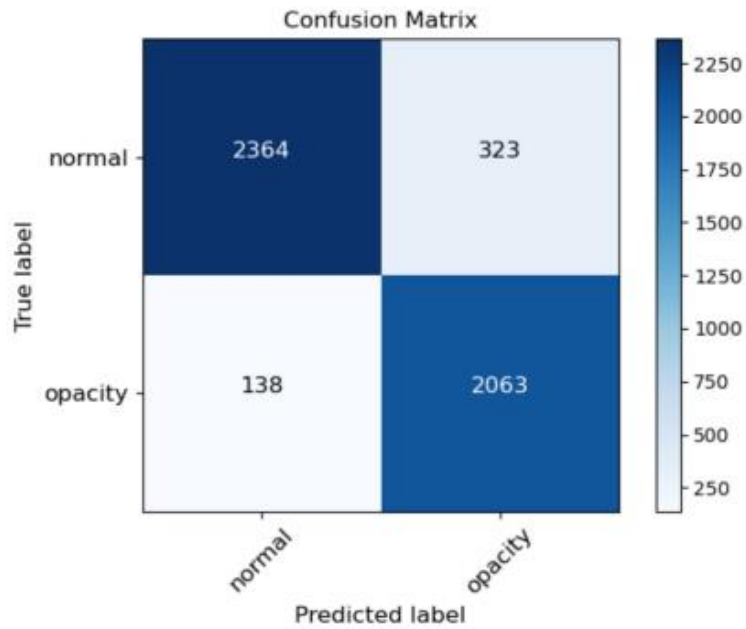
### 1- Yerel doğrulama (verinin %20'si)

IoU metric: bbox

Average Precision (AP) @[ IoU=0.50:0.95 | area= all | maxDets=100 ] = 0.639

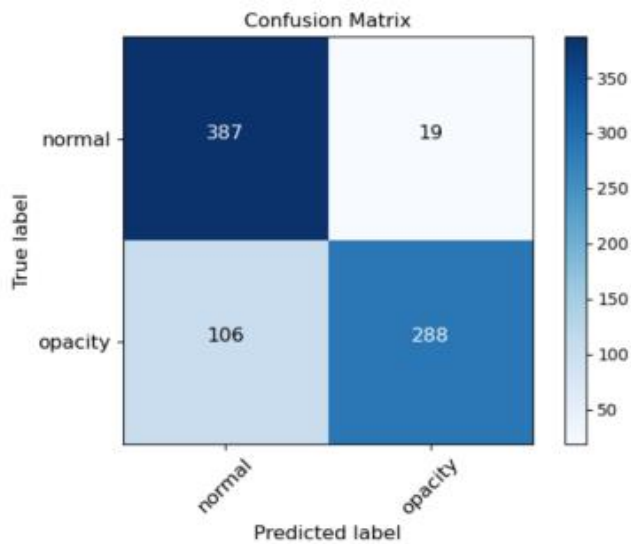
Average Precision (AP) @[ IoU=0.50 | area= all | maxDets=100 ] = 0.776

Average Precision (AP) @[ IoU=0.75 | area= all | maxDets=100 ] = 0.618



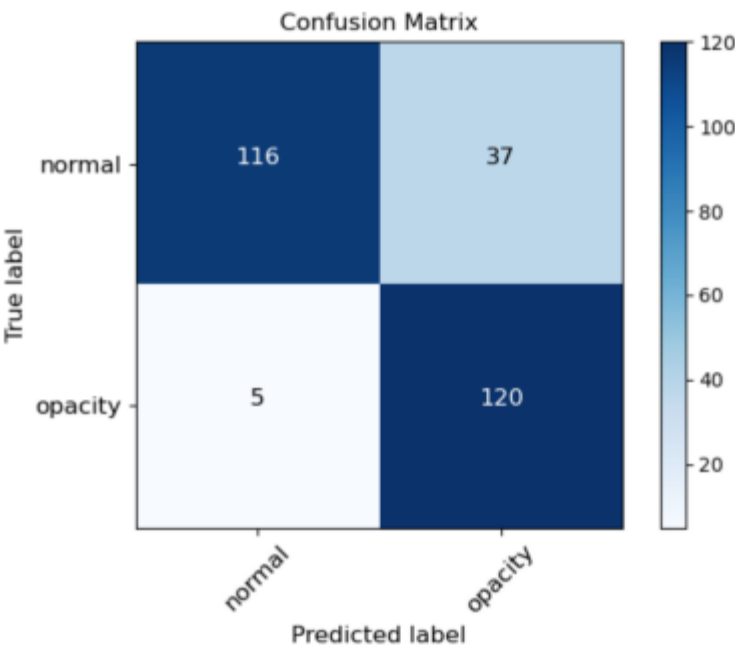
	precision	recall	f1-score	support
normal	0.94	0.88	0.91	2687
opacity	0.86	0.94	0.90	2201
accuracy			0.91	4888
macro avg	0.90	0.91	0.91	4888
weighted avg	0.91	0.91	0.91	4888

## 2- Shenzhen ve Montgomery veri setleri



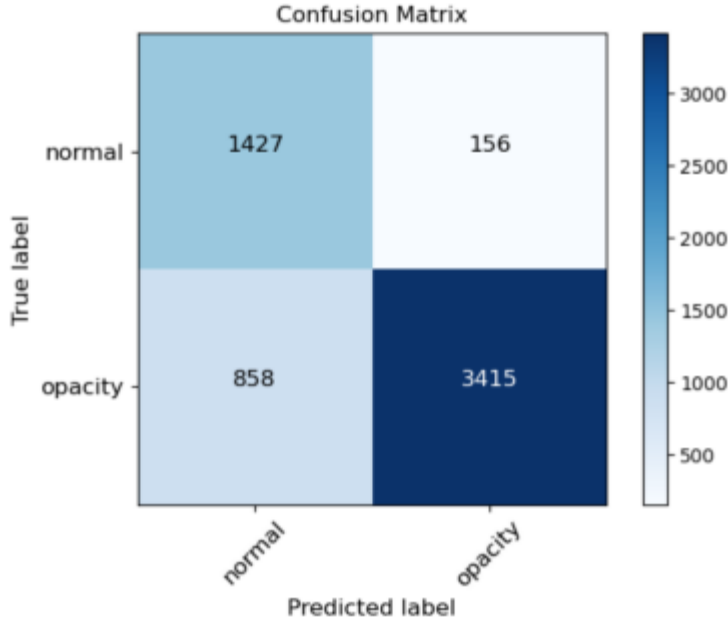
	precision	recall	f1-score	support
normal	0.78	0.95	0.86	406
opacity	0.94	0.73	0.82	394
accuracy			0.84	800
macro avg	0.86	0.84	0.84	800
weighted avg	0.86	0.84	0.84	800

3- DA ve DB veri setleri



	precision	recall	f1-score	support
normal	0.96	0.76	0.85	153
opacity	0.76	0.96	0.85	125
accuracy			0.85	278
macro avg	0.86	0.86	0.85	278
weighted avg	0.87	0.85	0.85	278

#### 4- Çocuk Göğüs Röntgeni Zatürre veri seti



	precision	recall	f1-score	support
normal	0.62	0.90	0.74	1583
opacity	0.96	0.80	0.87	4273
accuracy			0.83	5856
macro avg	0.79	0.85	0.80	5856
weighted avg	0.87	0.83	0.83	5856

#### Donanım

- 1 x P100 GPU
- Kaggle üzerinde eğitildi
- Eğitim süresi yaklaşık 7 saat (8 dönem)

#### Diğer Bilgiler

1- Pratikte, kullanıcılar farklı kalitelardaki resimleri gönderebilir. Bu resimler farklı türdeki röntgen makineleriyle çekilmiş olabilir ya da cep telefonu kameralarıyla çekilmiş röntgen filmlerinin fotoğrafları olabilir. Bu ve diğer gerçek dünya faktörleri modelin doğruluğunu azaltabilir.

- 2- Uygulama, pediatrik görüntüler üzerinde iyi test sonuçları verse de, eğitim verisinde çok fazla pediatrik görüntü bulunmamaktadır. Gerçek dünyada, bu uygulama pediatrik veriler üzerinde güçlü sonuçlar vermeyebilir.
- 3- Uygulama, bir GPU kullanacak şekilde kolayca yapılandırılabilir.
- 4- Doğruluğu artırmak ve aşırı uyumu (overfitting) azaltmak için veri artırma (data augmentation) kullandım.

### **Dokümantasyon**

Tüm kod GitHub üzerinde mevcuttur:

<https://github.com/Luceyn-Iabed/LujainAI>

### **Lisans**

Uygulama kodu MIT Lisansı altında mevcuttur. Ancak, eğitim verilerinin bazıları ticari kullanım için lisanslı olmadığı için, eğitilmiş model ticari amaçlarla kullanılamaz.

### **İletişim**

E-posta: [luceyniabed2002@gmail.com](mailto:luceyniabed2002@gmail.com)

### **İşte veritabanları ve makalelere tam bağlantılar:**

- **Shenzhen ve Montgomery Veri Setleri**  
İndirme (GitHub): <https://github.com/Luceyn-Iabed/LujainAI>  
Makale: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4256233/>
- **VinDr-PCXR Veri Seti**  
İndirme: <https://physionet.org/content/vindr-pcxr/1.0.0/>  
Makale: <https://www.medrxiv.org/content/10.1101/2022.03.04.22271937v1.full-text>
- **Tuberculosis (TB) Chest X-ray Veri Tabanı**  
İndirme (Kaggle): <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>  
Makale: <https://ieeexplore.ieee.org/document/9224622>
- **SIIM-FISABIO-RSNA COVID-19 Tespiti**  
İndirme (Kaggle): <https://www.kaggle.com/competitions/siim-covid19-detection/data>  
Makale: <https://osf.io/532ek/>