



1928

K. N. Toosi University  
of Technology

به نام خدا

پروژه پایانترم درس هوش مصنوعی

اعضا: دانیال خاکساری – سید طاها طیبی – امیرعلی یاوری  
موضوع پروژه: تشخیص بیماری ذات الریه با استفاده از عکس  
های X-Ray

تاریخ: زمستان 1404

لینک گیت هاب پروژه:

[YavariAmirali/AI-FinalProject-Tayebi: Medical Imaging AI Project - Pneumonia Detection using Chest X-Ray](https://github.com/YavariAmirali/AI-FinalProject-Tayebi: Medical Imaging AI Project - Pneumonia Detection using Chest X-Ray)

## لیست محتوا

1. تعریف دقیق مسئله (Problem Definition) ..... 3
2. مهم‌ترین اهداف و معیارهای پروژه ..... 3
3. توجیه انتخاب معماری و استراتژی مدل‌سازی (Model Justification) ..... 4
4. فاز اول: مدل پایه (Simple CNN Baseline) ..... 4
4. فاز دوم: یادگیری انتقالی با ResNet50 (Transfer Learning) ..... 4
5. دلایل انتخاب ResNet50: ..... 5
4. تحلیل اکتشافی داده‌ها (EDA): ..... 5
5. جدول پلن آزمایش‌ها ..... 9
6. خروجی ساختار مدل (Model Summary) ..... 10
7. معرفی دیتاست: ..... 11
8. جمع‌بندی: ..... 11

## 1. تعریف دقیق مسئله (Problem Definition)

هدف اصلی این پروژه، طراحی و پیاده‌سازی یک سامانه هوشمند مبتنی بر یادگیری عمیق برای تشخیص بیماری ذات‌الریه (Pneumonia) از روی تصاویر X-Ray قفسه سینه است. با توجه به ماهیت حساس پزشکی مسئله و نقش مستقیم آن در تصمیم‌گیری‌های درمانی، معیارهای ارزیابی مدل با دقت و اولویت‌بندی خاصی انتخاب شده‌اند.

## 2. مهم‌ترین اهداف و معیارهای پروژه

Recall (Sensitivity) بالاتر از ۹۵٪:

Recall به عنوان حیاتی‌ترین معیار ارزیابی در نظر گرفته شده است، زیرا در مسائل تشخیص پزشکی، هزینه‌ی False Negative (عدم تشخیص بیمار مبتلا) بسیار بالا بوده و می‌تواند منجر به تأخیر در درمان یا تهدید جان بیمار شود. هدف اصلی این پروژه کاهش حداکثری موارد False Negative و نزدیک شدن آن به صفر است.

Accuracy بالاتر از ۹۰٪:

این معیار به منظور سنجش عملکرد کلی مدل در تشخیص صحیح تصاویر سالم و بیمار مورد استفاده قرار می‌گیرد. اگرچه Accuracy به تنهایی معیار قابل اتکایی در داده‌های نامتوازن نیست، اما در کنار Recall و سایر معیارها می‌تواند دید مناسبی از عملکرد عمومی مدل ارائه دهد.

معیارهای تکمیلی پزشکی (Medical Metrics):

به منظور ارزیابی جامع‌تر، معیارهای Specificity و AUC (Area Under the ROC Curve) نیز محاسبه خواهند شد. Specificity بالا نشان‌دهنده توانایی مدل در تشخیص صحیح افراد سالم است و از تشخیص اشتباه فرد سالم به عنوان بیمار (False Positive) جلوگیری می‌کند؛ امری که می‌تواند از تحمیل فشار روانی، هزینه‌های غیرضروری و مداخلات درمانی اشتباه پیشگیری کند.

### 3. توجیه انتخاب معماری و استراتژی مدل سازی ( Model Justification)

برای دستیابی به اهداف فوق و کاهش ریسک خطا، فرآیند مدل سازی در این پروژه به صورت دو فاز مجزا طراحی شده است.

#### فاز اول: مدل پایه (Simple CNN Baseline)

در فاز نخست، از یک شبکه عصبی کانولوشنی ساده شامل سه لایه کانولوشن به عنوان مدل پایه استفاده می شود.

دلایل انتخاب این رویکرد:

ایجاد یک Baseline برای مقایسه عملکرد مدل های پیشرفته تر

\* بررسی میزان قابل تفکیک بودن داده ها با یک مدل ساده

\* انتظار می رود این مدل به دلیل سادگی معماری و تعداد کم لایه ها، روی داده های پیچیده X-Ray دچار Underfitting شود (دقت پایین روی آموزش و تست). با این حال، این مدل به عنوان یک نقطه شروع (Baseline) عالی است تا در فاز دوم میزان بهبود مدل ResNet50 را نسبت به آن بسنجیم.

\* کمک به درک بهتر پیچیدگی ذاتی مسئله پیش از استفاده از معماری های عمیق تر

این مدل پایه نقش مرجع را ایفا می کند و عملکرد آن مبنایی برای سنجش میزان بهبود حاصل از به کارگیری یادگیری انتقالی خواهد بود

#### فاز دوم: یادگیری انتقالی با ResNet50 (Transfer Learning)

در فاز دوم، از معماری پیشرفته ResNet50 به همراه وزن های از پیش آموزش دیده بر روی دیتاست ImageNet استفاده می شود.

## دلایل انتخاب ResNet50:

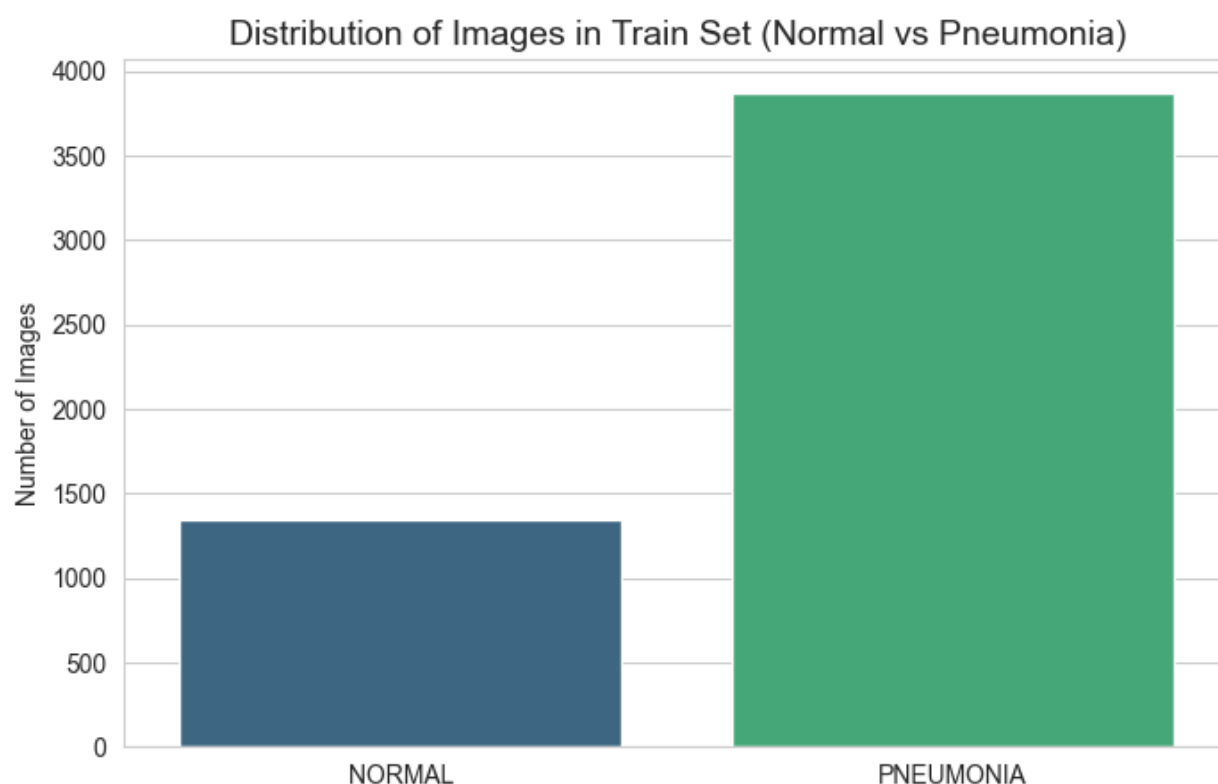
\* استفاده از Residual Connections که مشکل Vanishing Gradient را در شبکه‌های عمیق کاهش می‌دهد

\* توانایی استخراج ویژگی‌های سطح بالا و پیچیده که برای تشخیص الگوهای ظریف در تصاویر پزشکی ضروری هستند

\* بهره‌گیری از دانش پیش‌آموزته‌شده (Pre-trained Weights) که موجب افزایش سرعت همگرایی و بهبود عملکرد مدل، به‌ویژه در شرایط محدودیت داده، می‌شود

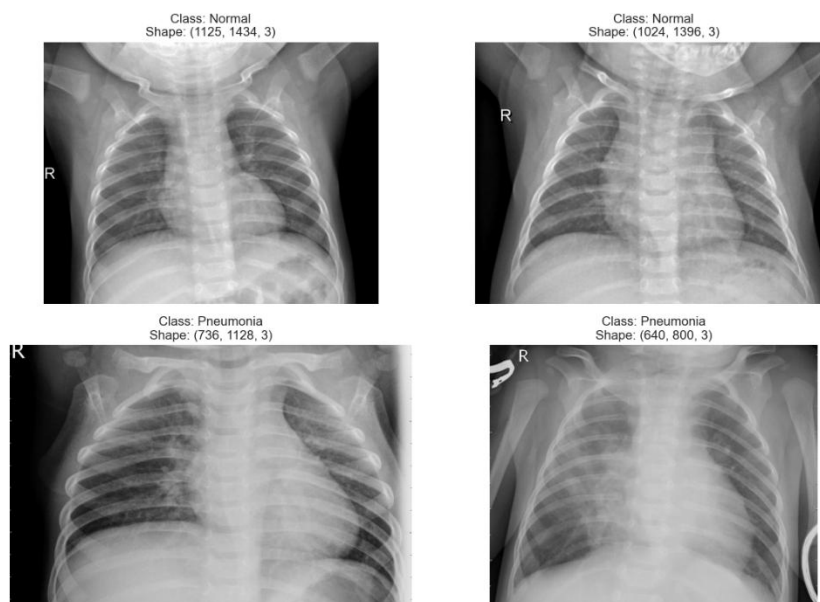
در این فاز، لایه‌های ابتدایی شبکه که ویژگی‌های عمومی بصری مانند لبه‌ها و بافت‌ها را استخراج می‌کنند، فریز (Freeze) شده و فرآیند Fine-tuning عمدتاً بر روی لایه‌های انتهایی انجام می‌گیرد تا مدل به‌طور مؤثر با داده‌های پزشکی X-Ray تطبیق یابد.

## 4. تحلیل اکتشافی داده‌ها (EDA):



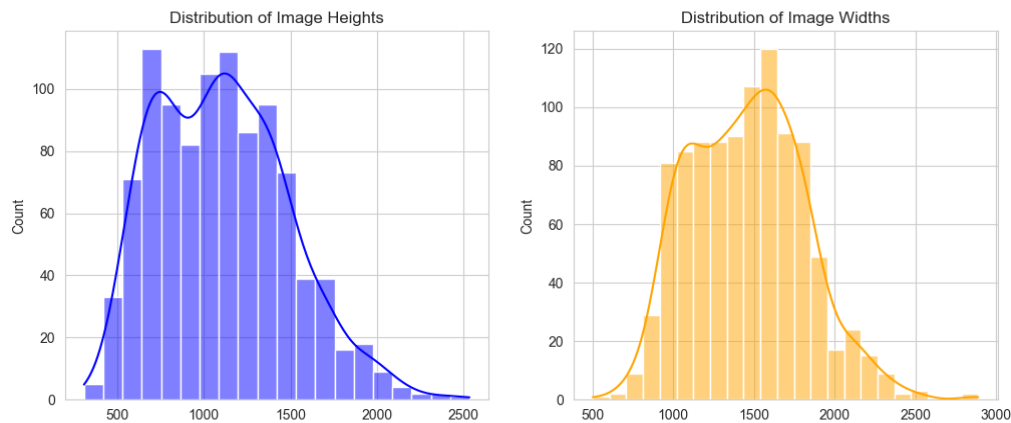
نمودار 1. عدم تعادل کلاس‌ها (Normal vs Pneumonia)

نمودار میله‌ای که تعداد عکس‌های سالم و بیمار را نشان می‌دهد. همان‌طور که در نمودار مشخص است، تعداد عکس‌های بیمار (Pneumonia) بیشتر از عکس‌های سالم است. این Imbalanced بودن دیتاست می‌تواند باعث شود مدل به سمت کلاس اکثریت گرایش پیدا کند. به همین دلیل، در فاز دوم آموزش از تکنیک Weighting Class برای رفع این مشکل استفاده خواهیم کرد.



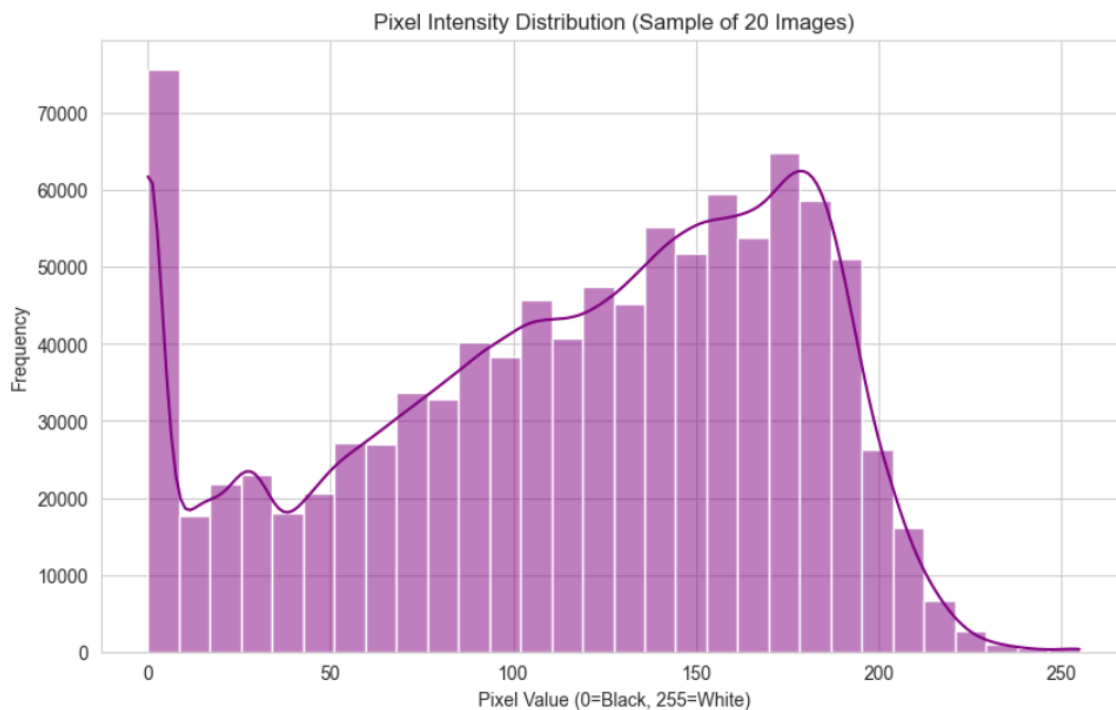
شکل ۱. نمونه تصاویر تصادفی از دو کلاس سالم و مبتلا به ذات‌الریه

بررسی چشمی (Visual Inspection) نمونه‌ها نشان می‌دهد که تصاویر از نوع Grayscale (تک‌کاناله) هستند و کیفیت آنها برای استخراج ویژگی توسط مدل‌های CNN مناسب است. ناحیه قفسه سینه در مرکز تصاویر قرار دارد، اما تفاوت‌هایی در کنتراست و روشنایی تصاویر دیده می‌شود که لزوم نرمال‌سازی را تایید می‌کند.



نمودار 2. پراکندگی سائز (ارتفاع و پهنا) تصاویر موجود در دیتاست

این نمودار نشان می‌دهد که تصاویر دیتاست دارای رزولوشن‌های متنوعی هستند و سائز یکسانی ندارند. از آنجا که معماری‌های استاندارد CNN مانند مدلی که طراحی خواهیم کرد و ResNet50 نیاز به ورودی با ابعاد ثابت دارند، تمامی تصاویر در مرحله پیش‌پردازش به سائز استاندارد  $224 \times 224$  پیکسل تغییر سائز (Resize) خواهند شد.

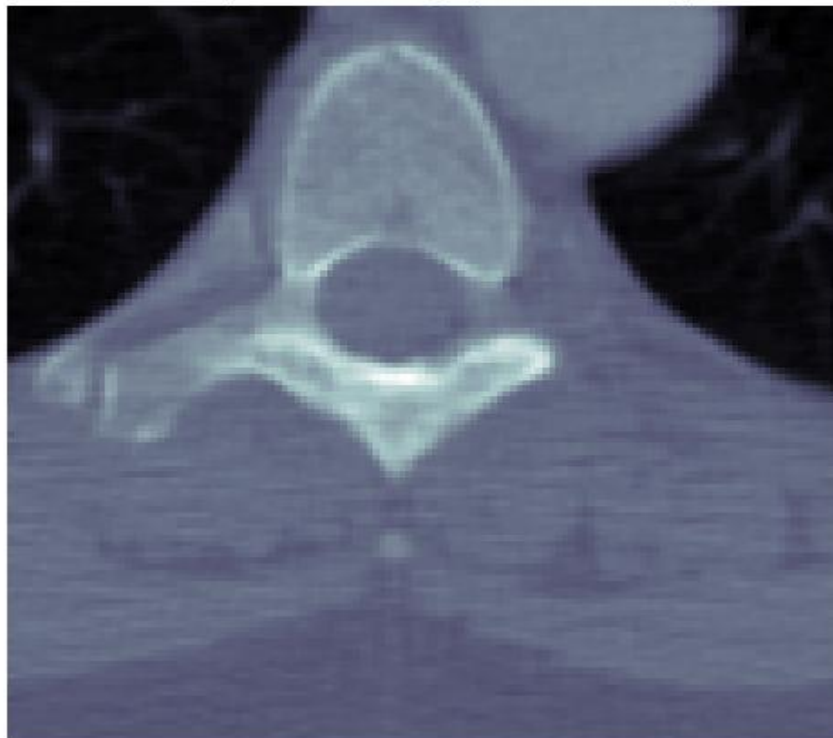


نمودار 3. هیستوگرام توزیع شدت روشنایی پیکسل‌ها

توزیع شدت روشنایی نشان می‌دهد که مقادیر پیکسل‌ها در بازه  $[0, 255]$  قرار دارند. برای کمک به همگرایی سریع‌تر مدل و جلوگیری از مشکلات گرادیان، در مرحله پیش‌پردازش از

تکنیک Normalization استفاده کرده و تمامی مقادیر پیکسل‌ها را به بازه  $[0, 1]$  مقیاس‌دهی (Scale) خواهیم کرد.

Sample DICOM Image (Loaded Locally)



شکل 2. نمونه ای از عکس های DICOM

عکس های پزشکی معمولا با پسوند DICOM ذخیره میشوند . عکس بالا نمونه ای از همین تصاویر هست که ضرورت تشخیص عکس ها با پسوند DICOM را علاوه بر باقی پسوند ها ضروری میسازد.

Original X-Ray (Low Contrast)



After Histogram Equalization (High Contrast)



نمودار 9 : Histogram equalization



در تصاویر پزشکی X-Ray گاهی اوقات کنتراست تصاویر پایین است و تشخیص دقیق عفونت دشوار میشود به همین دلیل ما از تکنیک histogram equalization استفاده میکنیم تا کنتراست تصاویر متعادل شوند.

## 5. جدول پلن آزمایش ها

ویژگی (Feature)	مدل فاز اول (Baseline)	مدل فاز دوم (Final Phase)
نام مدل	Simple CNN (دست ساز)	ResNet50 (Transfer Learning)
معماری	۳ لایه کانولوشن ساده + MaxPooling	۵۰ لایه عمیق با وزن های ImageNet
ورودی داده	تغییر سایز (Resize) و نرمال سازی ساده	پیش پردازش Data + ImageNet Augmentation
مدیریت عدم تعادل	ندارد (None)	Class Weighting (وزن دهی کلاس ها)
هدف آزمایش	بررسی امکان پذیری (Feasibility) و ایجاد خط مبنا	دستیابی به Recall > 95% و قابلیت اطمینان بالا
معیارهای ارزیابی	Accuracy, Recall	Specificity, AUC, Accuracy, Recall

**تحلیل معیارها:** در فاز دوم پروژه، علاوه بر معیارهای Accuracy و Recall، دو معیار Specificity و AUC نیز به دقت اندازه گیری خواهند شد

دلیل پزشکی: در تشخیص بیماری ها، فقط پیدا کردن بیماران (Recall) کافی نیست؛ بلکه باید مطمئن شویم افراد سالم به اشتباه بیمار تشخیص داده نمی شوند. (False Positive) تشخیص

اشتباه یک فرد سالم به عنوان بیمار، می‌تواند هزینه‌های روانی سنگین و درمان‌های غیرضروری را به بیمار تحمیل کند. بنابراین معیار Specificity در کنار Recall برای ما حیاتی است

## 6. خروجی ساختار مدل (Model Summary)

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 224, 224, 32)	896
max_pooling2d (MaxPooling2D)	(None, 112, 112, 32)	0
conv2d_1 (Conv2D)	(None, 112, 112, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 56, 56, 64)	0
conv2d_2 (Conv2D)	(None, 56, 56, 128)	73856
max_pooling2d_2 (MaxPooling2D)	(None, 28, 28, 128)	0
flatten (Flatten)	(None, 100352)	0
dense (Dense)	(None, 64)	6422592
dense_1 (Dense)	(None, 1)	65

Total params: 6,515,905

Trainable params: 6,515,905

Non-trainable params: 0

## 7. معرفی دیتاست:

برای آموزش و ارزیابی مدل هوش مصنوعی در این پروژه، از مجموعه داده معتبر Chest X-Ray Images (Pneumonia) موجود در پلتفرم Kaggle استفاده شده است. این مجموعه داده شامل 5863 تصویر رادیوگرافی قفسه سینه (X-Ray) است که در دو دسته اصلی Normal (سالم) و Pneumonia (مبتلا به ذات‌الریه) طبقه‌بندی شده‌اند. تصاویر متعلق به بیماران اطفال 1 تا 5 سال بوده و در مرکز پزشکی زنان و کودکان گوانگژو جمع‌آوری شده‌اند.

ساختار داده‌ها در سه پوشه مجزا به شرح زیر تقسیم‌بندی شده است:  
داده‌های آموزش (Train): 5,216 تصویر (برای یادگیری وزن‌های مدل).

داده‌های آزمون (Test): 624 تصویر (برای ارزیابی نهایی).

داده‌های اعتبارسنجی: 16 تصویر که به دلیل تعداد کم، در فاز آموزش با بخشی از داده‌های Train ترکیب و مجدداً Split خواهند شد.

با وجود پیاده‌سازی ماژول خواندن استاندارد DICOM، به دلیل اینکه دیتاست مرجع (Kaggle) فرمت‌ها را به JPEG فشرده کرده است، آموزش مدل روی فرمت JPEG انجام می‌شود؛ اما سیستم قابلیت استقرار در بیمارستان با DICOM را دارد.

لینک دیتاست: [Chest X-Ray Images \(Pneumonia\)](#)

## 8. جمع‌بندی:

استراتژی دوبرحله‌ای اتخاذ شده در این پروژه، با ترکیب یک مدل پایه ساده و یک مدل عمیق مبتنی بر یادگیری انتقالی، امکان ارزیابی دقیق عملکرد و دستیابی به استانداردهای مورد انتظار در کاربردهای پزشکی را فراهم می‌کند. تمرکز اصلی بر Recall بالا در کنار کنترل False Positive، این سامانه را به گزینه‌ای قابل اتکا برای کاربردهای تشخیصی در حوزه سلامت تبدیل می‌کند.

همچنین معیارهای Specificity و AUC نیز اندازه‌گیری می‌شوند، زیرا تشخیص اشتباه فرد سالم به عنوان بیمار (False Positive) هزینه‌های روانی و درمانی غیرضروری تحمیل می‌کند.