Pneumonia Detection in Chest X-ray images

Wen Sun

Eastern Washington University

Introduction

- This project is to depict if a patient has pneumonia, one of the lung disease, by detecting chest X-rays.
- Three types of Pneumonia

Normal	No areas of abnormal opacification
Bacterial Pneumonia	Exhibit focal lobar consolidation
Viral Pneumonia	More diffuse “interstitial” pattern in both lungs

Dataset

- The dataset is organized into 2 folders (train, test), and contains the subfolders for each image category (pneumonia/normal)
- 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/normal)
- All chest x-ray images were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou

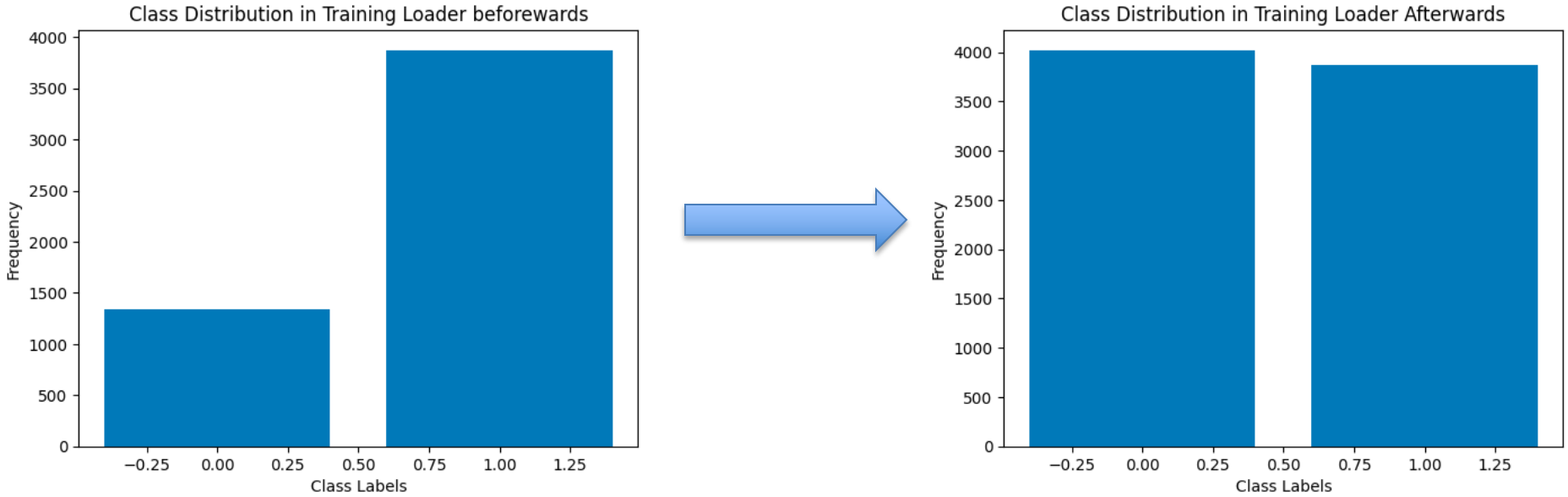
[1] Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X- Ray Images for Classification", Mendeley Data, v2

Data Preprocessing

Expand the training dataset and make sure it is balanced.

- Normal: 1341
- Pneumonia: 3875

Because we have enough images in the pneumonia case. Therefore, each image of only the normal (healthy) case was augmented twice.

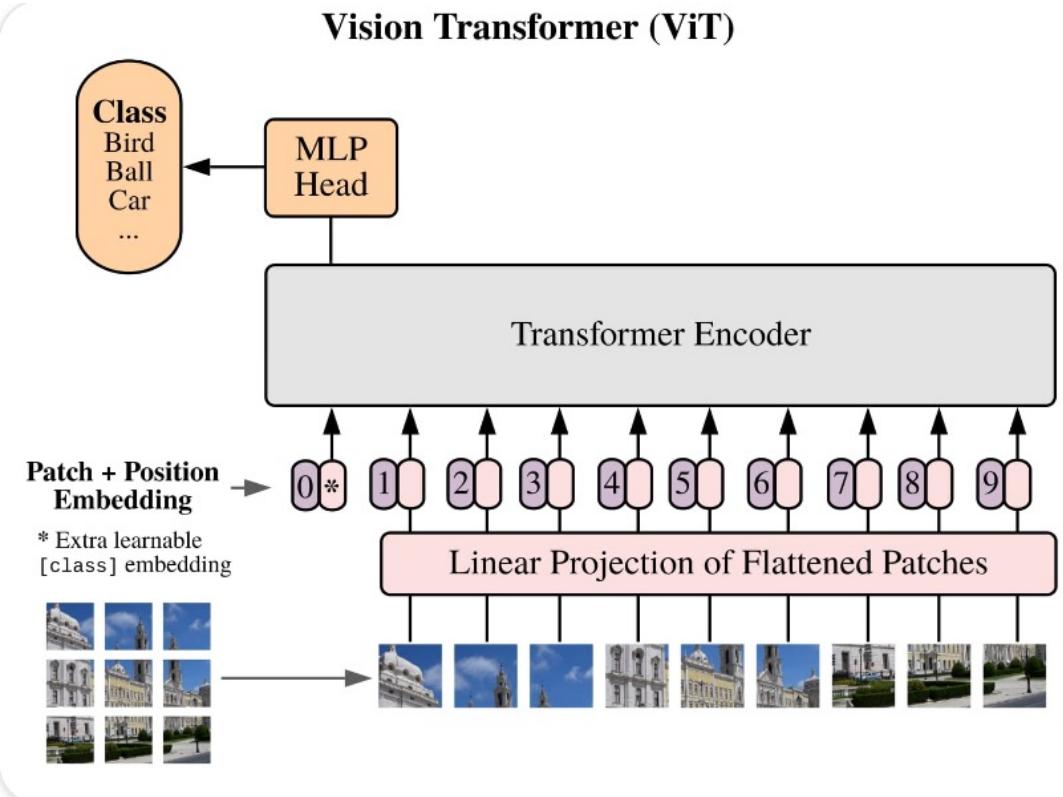


Data Augmentation

- To avoid overfitting, we want to increase the diversity of the training dataset. This is particularly useful in image classification tasks to improve the robustness and generalization of the model. Common augmentation techniques include transformations like rotations, translations, scaling, flipping, and color adjustments.

```
train_transform = transforms.Compose([
    transforms.Resize((224, 224)),  # Resize images to a fixed size
    transforms.ToTensor(),  # Convert images to PyTorch tensors
    #Data augumentation to vary images, to avoid overfitting problems
    transforms.RandomRotation(20),
    transforms.Normalize(mean=image_mean, std=image_std)  # Normalize images
])
```

ViT

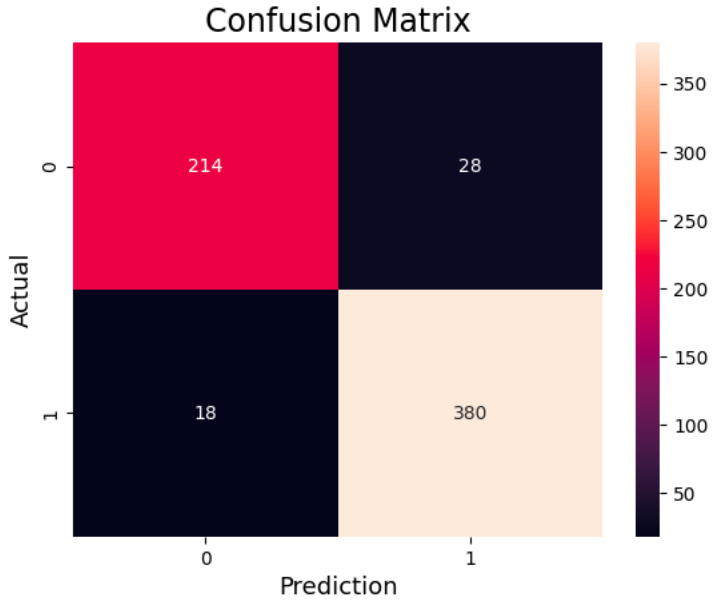


- ViT (Vision Transformer) has recently emerged as a competitive alternative to CNN by almost x4 in terms of computational efficiency and accuracy in image recognition computer vision task.
- ViT model architecture was introduced in a research paper published as a conference paper at ICLR 2021 titled “An image is Words: Transformers for Image Recognition at Scale”
- ViT models were pre-trained on the ImageNet and ImageNet-21k datasets.
- Used ViTForImageClassification

Result

[1482/1482 16:56, Epoch 3/3]

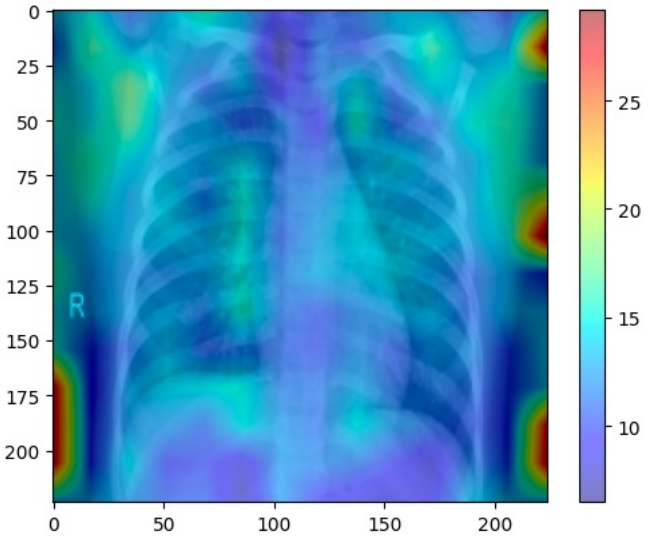
Epoch	Training Loss	Validation Loss	Accuracy
1	No log	0.209665	0.928125
2	0.130500	0.463210	0.887500
3	0.047500	0.309084	0.923438



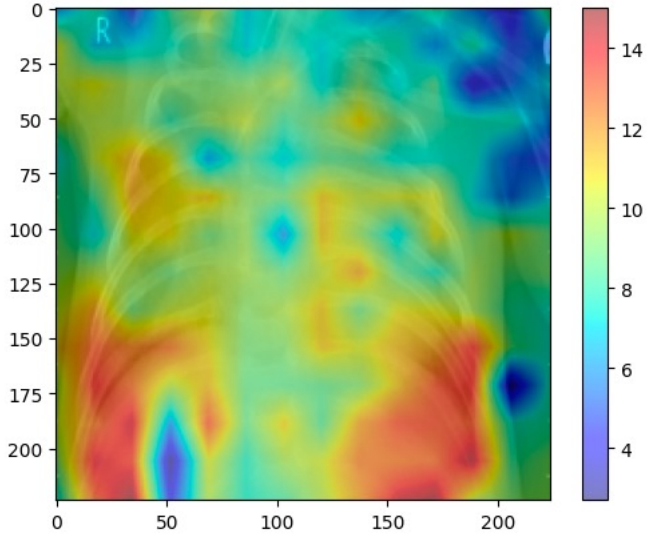
	precision	recall	f1-score	support
0	0.92	0.88	0.90	242
1	0.93	0.95	0.94	398
accuracy			0.93	640
macro avg	0.93	0.92	0.92	640
weighted avg	0.93	0.93	0.93	640

Visualization

Normal



pneumonia



Demo