BLG 506E – COMPUTER VISION

Project Progress Presentation

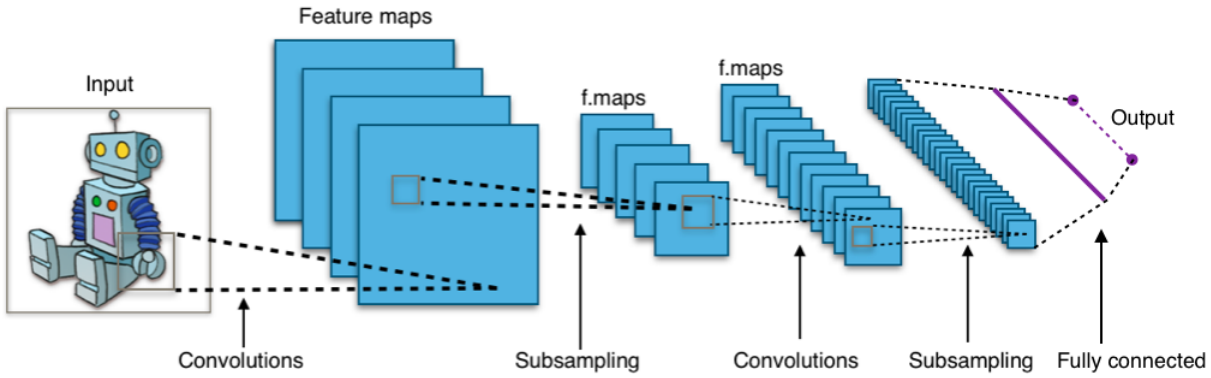
BURAK BOZDAĞ - 504211552

Recap: About the Project

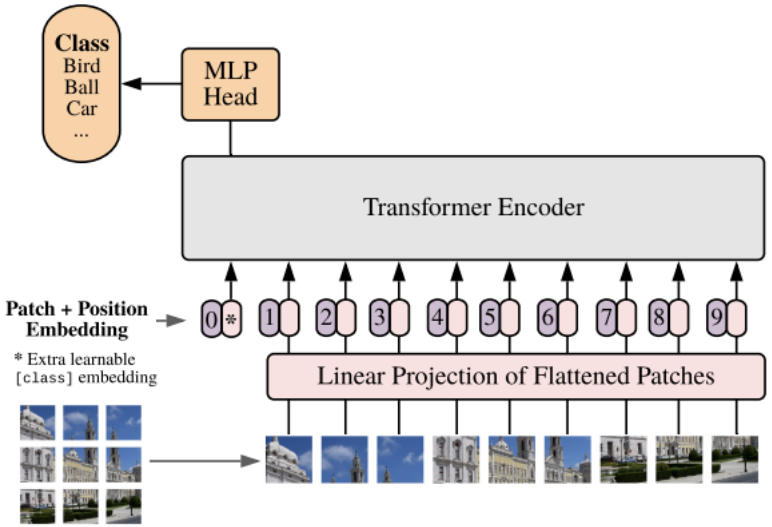
Classifying Chest X-Ray Images Using CNN and Transformer Based Architectures

Comparing CNN and transformer models for classifying patients as normal or infected

CNN



Vision Transformer (ViT)



Applied Processes and Methods

- TensorFlow, Keras
- Examining Dataset
- Data Augmentation
- ViT Evaluation

Test Bench

- AMD Radeon RX 6600 XT 8GB
- 16 GB RAM
- tensorflow-cpu==2.10
- tensorflow-directml-plugin
- Enable GPU Acceleration for TensorFlow 2 with tensorflow-directml-plugin [1]

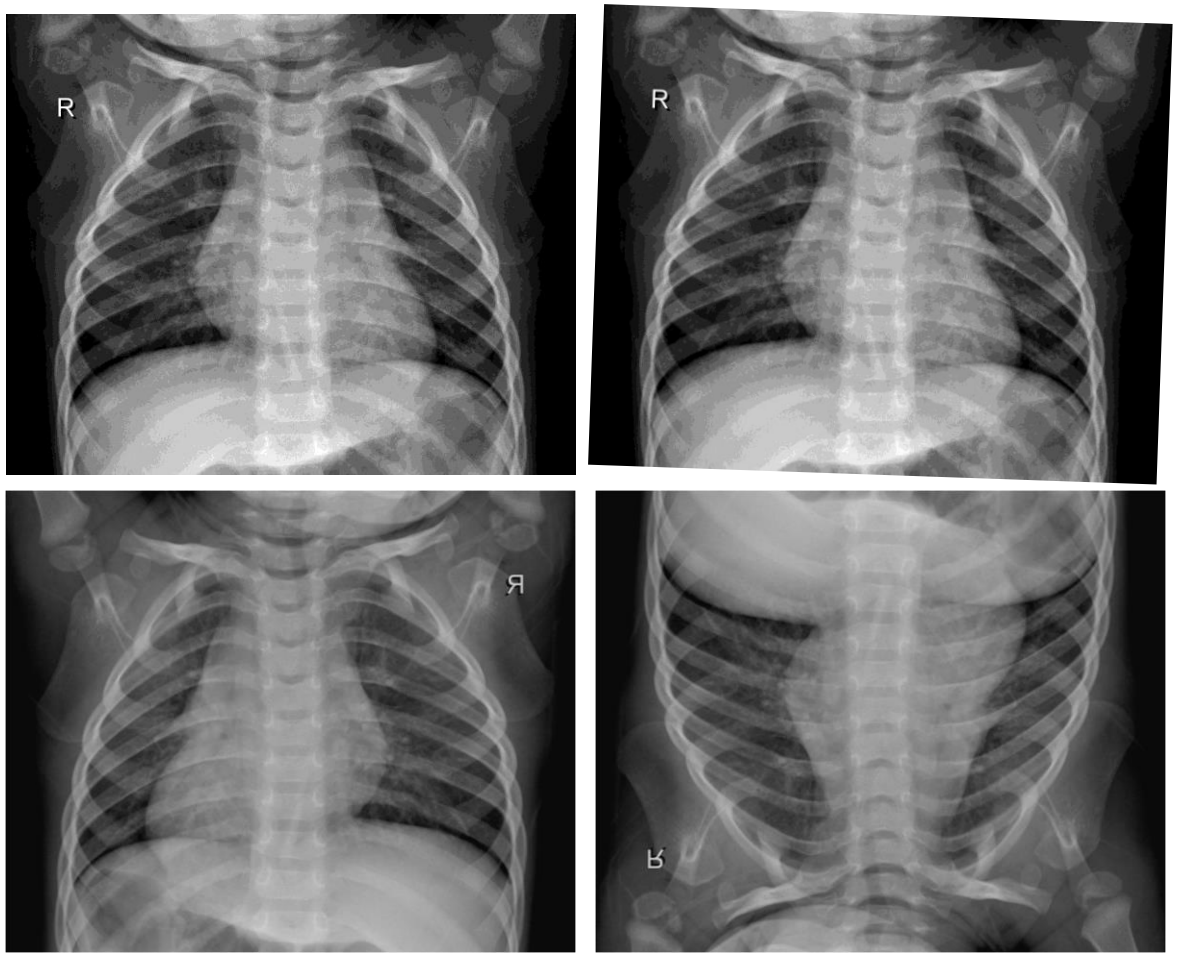
Dataset

- Chest X-Ray Images (Pneumonia) [2]
- 5856 JPEG images (1.15 GB)
- 5216 train
- 16 validation
- 624 test

```
(amd_gpu) PS C:\Users\Burak\Desktop\BLG506E-CV\Project> python .\chest_xray.py
2022-12-04 12:27:44.250691: I tensorflow/c/logging.cc:34] Successfully opened
2022-12-04 12:27:44.250771: I tensorflow/c/logging.cc:34] Successfully opened
2022-12-04 12:27:44.252886: I tensorflow/c/logging.cc:34] Successfully opened
2022-12-04 12:27:44.391692: I tensorflow/c/logging.cc:34] DirectML device enume
Found 5216 images belonging to 2 classes.
Found 16 images belonging to 2 classes.
Found 624 images belonging to 2 classes.
```

Data Augmentation

- Rescale = 1/255
- Zoom Range = 0.1
- Rotation Range = 0.2
- Horizontal-Vertical Flip
- 224x224 WxH



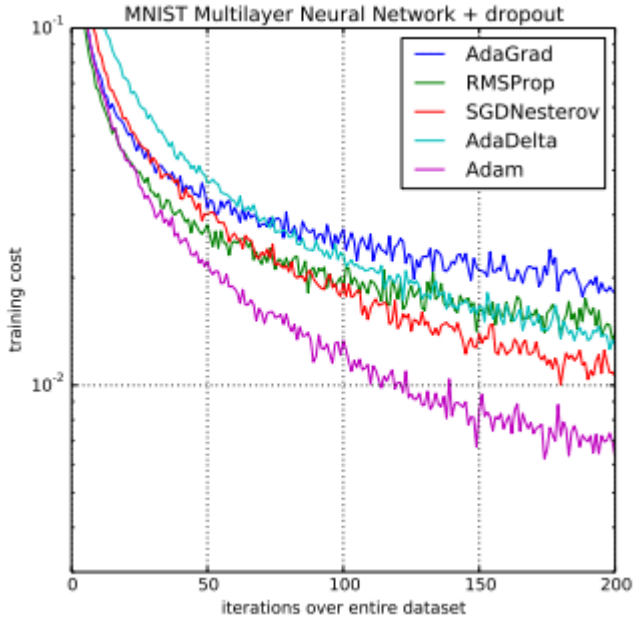
ViT Evaluation

- Callbacks:
  - Monitoring Validation Loss
  - Reduce LR
  - Early Stopping
  - Model Checkpoint
- ViT-B/16 architecture
- Layers: Input, Conv2D, Reshape, 12 x Transformer Encoders, Normalization, Lambda, Dense
- Pre-trained model with ImageNet 2012

ViT Evaluation

- Optimizer: Adam
- Loss: Binary Cross-Entropy
- Max. # of 50 Epochs

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$



```
652/652 [==============================] - 281s 410ms/step - loss: 0.1844 - accuracy: 0.9273 - val_loss: 0.4784 - val_accuracy: 0.9375 - lr: 1.0000e-04
Epoch 2/50
652/652 [==============================] - 264s 405ms/step - loss: 0.1078 - accuracy: 0.9615 - val_loss: 0.0347 - val_accuracy: 1.0000 - lr: 1.0000e-04
Epoch 3/50
652/652 [==============================] - 261s 400ms/step - loss: 0.0837 - accuracy: 0.9686 - val_loss: 0.5374 - val_accuracy: 0.8750 - lr: 1.0000e-04
Epoch 4/50
652/652 [==============================] - 261s 400ms/step - loss: 0.0868 - accuracy: 0.9663 - val_loss: 0.4417 - val_accuracy: 0.7500 - lr: 1.0000e-04
Epoch 5/50
652/652 [==============================] - 262s 401ms/step - loss: 0.0773 - accuracy: 0.9701 - val_loss: 0.4511 - val_accuracy: 0.8125 - lr: 1.0000e-04
Epoch 6/50
652/652 [==============================] - 265s 406ms/step - loss: 0.0648 - accuracy: 0.9766 - val_loss: 0.0164 - val_accuracy: 1.0000 - lr: 1.0000e-04
Epoch 7/50
652/652 [==============================] - 261s 399ms/step - loss: 0.0656 - accuracy: 0.9757 - val_loss: 0.5679 - val_accuracy: 0.8750 - lr: 1.0000e-04
Epoch 8/50
652/652 [==============================] - 262s 402ms/step - loss: 0.0659 - accuracy: 0.9762 - val_loss: 0.4500 - val_accuracy: 0.8125 - lr: 1.0000e-04
Epoch 9/50
652/652 [==============================] - 252s 386ms/step - loss: 0.0878 - accuracy: 0.9666 - val_loss: 0.1248 - val_accuracy: 1.0000 - lr: 1.0000e-04
Epoch 10/50
652/652 [==============================] - 252s 386ms/step - loss: 0.0601 - accuracy: 0.9781 - val_loss: 0.9072 - val_accuracy: 0.7500 - lr: 1.0000e-04
Epoch 11/50
652/652 [==============================] - ETA: 0s - loss: 0.0785 - accuracy: 0.9688
Epoch 11: ReduceLROnPlateau reducing learning rate to 2.4999999368446888e-05.
652/652 [==============================] - 252s 386ms/step - loss: 0.0785 - accuracy: 0.9688 - val_loss: 0.0600 - val_accuracy: 1.0000 - lr: 1.0000e-04
Epoch 12/50
652/652 [==============================] - 252s 386ms/step - loss: 0.0307 - accuracy: 0.9893 - val_loss: 0.2823 - val_accuracy: 0.8750 - lr: 2.5000e-05
Epoch 13/50
652/652 [==============================] - 252s 386ms/step - loss: 0.0310 - accuracy: 0.9904 - val_loss: 0.0616 - val_accuracy: 1.0000 - lr: 2.5000e-05
Epoch 14/50
652/652 [==============================] - 253s 387ms/step - loss: 0.0233 - accuracy: 0.9912 - val_loss: 0.1084 - val_accuracy: 0.9375 - lr: 2.5000e-05
Epoch 15/50
652/652 [==============================] - ETA: 0s - loss: 0.0234 - accuracy: 0.9906Restoring model weights from the end of the best epoch: 6.
652/652 [==============================] - 253s 387ms/step - loss: 0.0234 - accuracy: 0.9906 - val_loss: 0.1377 - val_accuracy: 0.9375 - lr: 2.5000e-05
Epoch 15: early stopping
2/2 [==============================] - 0s 139ms/step - loss: 0.0620 - accuracy: 1.0000
```

Results

- 0: Healthy
- 1: Pneumonia
- Accuracy: 0.93

```
              precision    recall  f1-score   support

           0       0.97      0.83      0.89       234
           1       0.91      0.98      0.94       390

    accuracy                           0.93       624
   macro avg       0.94      0.91      0.92       624
weighted avg       0.93      0.93      0.93       624

(amd_gpu) PS C:\Users\Burak\Desktop\BLG506E-CV\Project> python .\chest_xray.py
```

Remaining Work

- Setting up a traditional CNN model
- Comparing CNN and ViT
- Multi-class classification (viral-bacterial infection)
- Merging other datasets (Tuberculosis, COVID-19)

References

[1] “Enable GPU Acceleration for TensorFlow 2 with tensorflow-directml-plugin,” *DirectML Plugin for TensorFlow 2 | Microsoft Learn*, 2022. [Online]. Available: https://learn.microsoft.com/en-us/windows/ai/directml/gpu-tensorflow-plugin.

[2] D. S. Kermany, et al., *Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning*, 2018. [Online]. Available: https://www.cell.com/cell/fulltext/S0092-8674(18)30154-5.