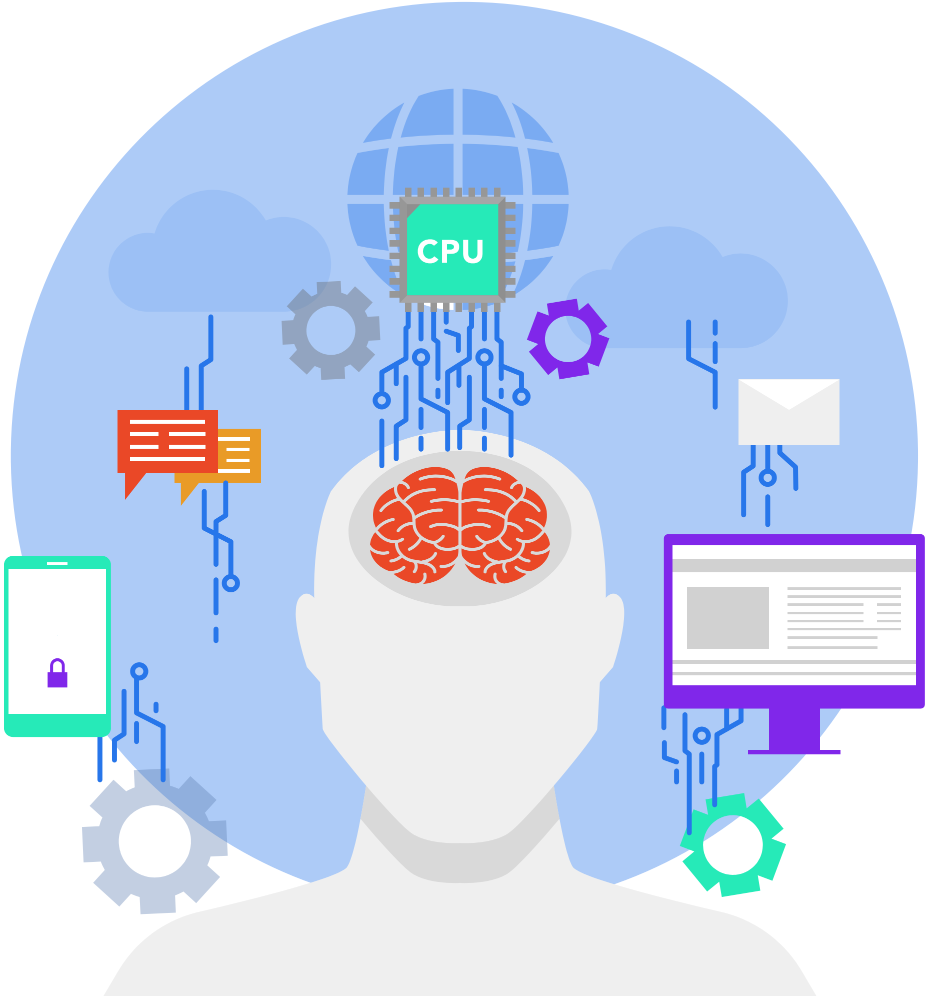


Lung Diseases Identification from X-Rays

Emmanouil Lykos
University of Piraeus-NCSR Demokritos

Introduction

- Deep Learning was firstly proposed at 60s however only recently was used, mostly because at that time computers were not that powerful to support them neither we had access to the amount of data that should be used to train them.
- Currently it has various applications on different fields. One of those fields is the medical one.
- Those applications vary from just take information from sensor data like heart rate to use specific models to assist the doctors to create new medicines.

- Also one application is to identify certain diseases from images, that can be identifying by watching an image, like melanoma or pneumonia.

- In this project we will present two Deep Learning approaches in order to identify various Lung Diseases, if any, from X-Ray Images.

Why Deep Learning and not Simple Machine Learning?

Deep Learning Advantages

- Minimal Preprocessing and Feature Extraction.
- Neural Networks can handle different structures of data without losing that information. For instance, we pass into them a whole image as it is without losing its topological information.
- Better Predictive performance.
- GPU optimization.

Deep Learning Disadvantages

- More Hyperparameter Tuning.
- More intensive computationally and memory-wise.

Preprocessing

- The basic preprocessing steps are the following:

    1. Normalization of image's pixels.
    2. Convert to RGB from Grayscale(if needed).
    3. Resize to 128x128 resolution.

- We also tried the following steps but they did not yielded better performance:

    1. Mask the images to contain only the object of interest.
    2. Try augmentation transforms.

Neural Network Architectures - CNNs

- Convolutional Neural Networks are the go-to architecture for models that need to handle images.
- The general architecture of those models for classification, is that they have some convolutional layers that work as feature extractors and then we apply some fully connected layers to get the prediction.

- The convolutional layers apply a number of filters to the given image and in that way we manage to exploit the topology of an image in order to extract new features.

- Also, after the convolution it is a common approach to apply max pooling, which means that we get the essential information from each part of the given image.

Handcrafted CNN

- Our first approach to our problem was to create a Handcrafted Convolutional Neural Network with the following architecture:
  - Has 5 convolution layers with output channels (32, 64, 128, 256, 512) and kernel sizes (5, 5, 3, 3, 3) followed by a Batch Normalization layer that centers and scales the outputs.
  - Each layer has a Max Pooling 2D layer of kernel size equal to 2.
  - Those layers result to 512 features thus we add a fully connected one with output neurons equal to the number of classes.

- This network uses the Adam optimizer and all layers are using the ReLU activation function , except the last one which uses Softmax Activation.

Transfer Learning

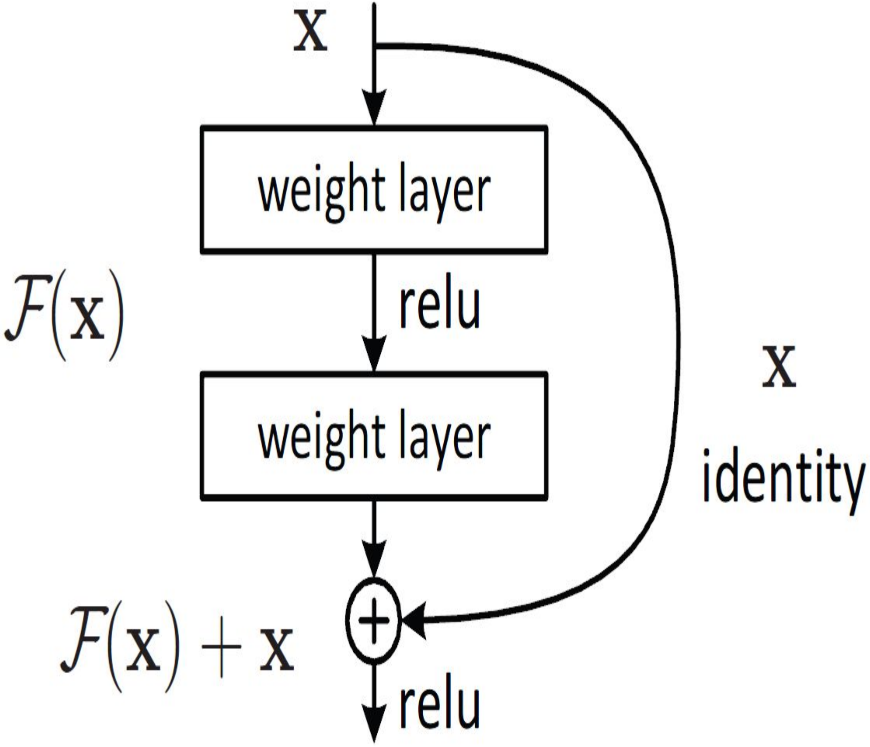
- Transfer Learning is the methodology where we want to take an approach to solve some problem to incorporate them into another one.

- In Deep Learning we want to take some layers of a network that excelled on some task along with the pretrained weights and incorporate them to our architecture.

- We can handle those layers with the following two ways:
    - Freeze those layers, thus, they will not get updated during training and they will keep the initial pretrained weights.
    - Just initialize the weights of those layers with the pretrained ones and start training from there.
    - Combine both of the above approaches.

ResNet model

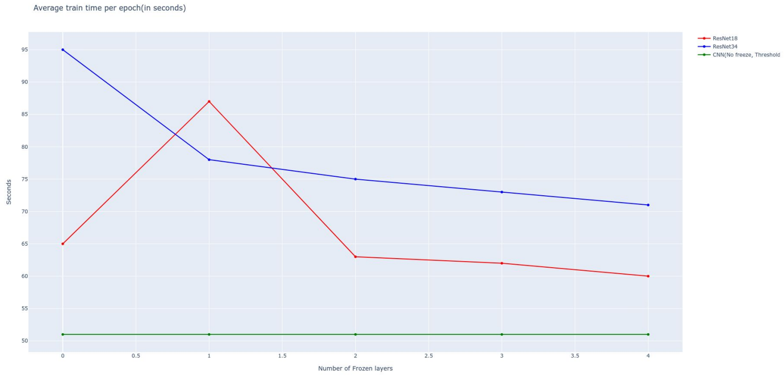
- ResNet model uses residual connections in order to address the degradation problem.
- In terms of general architecture it contains some pooling and convolution layers, grouped in 4 blocks. Then, they have a fully connected layer used for classification.
- This architecture was the winner of ImageNet Challenge.
- In our approach we use the all the layers of different ResNet versions along with their winner pretrained weights but we replace the last layer with a fully connected layer that outputs the number of classes that we want to identify.

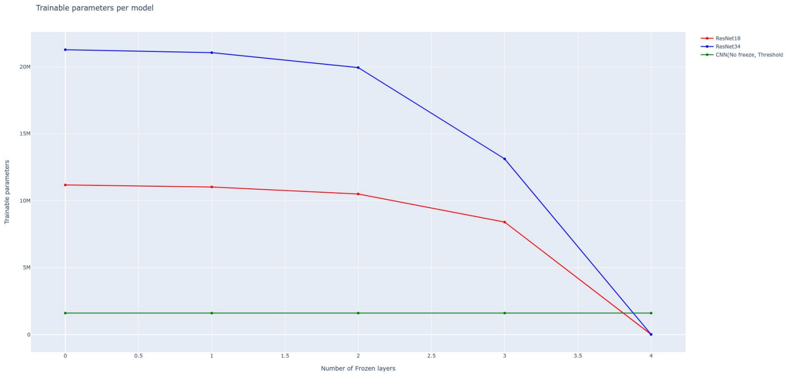


Experimental Evaluation

- We used the COVID-19 Radiography Dataset that contains 20000 images from Lung X-Rays where each one falls in these categories:
  - Healthy Lungs
  - COVID-19
  - Lung Opacity
  - Viral Pneumonia
- We split in a stratified way the dataset into train-test set and we cure the imbalance by using weighted random sampler.
- All the networks were trained for 100 epochs with patience equal to 20 and batch size equal to 128. We trained the handcrafted CNN, ResNet18 and ResNet34.
- We will evaluate the approaches with respect to how many ResNet blocks we will freeze using the following metrics:
  - Micro-Recall
  - Training epochs needed to converge.
  - Training time per epoch.
  - Trainable parameters.

Experimental Results - Time and Memory Efficiency





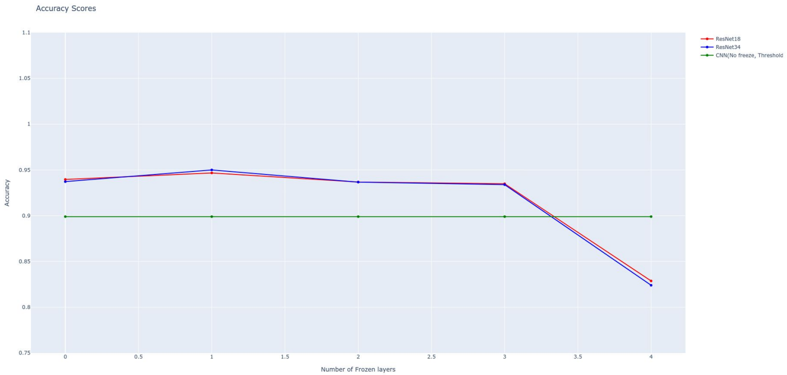
- The above shows the training time per epoch and the other the number of trainable parameters with respect to how many layers we freeze.
- We can see that ResNet34 uses double the parameters than the ResNet18 and the handcrafted CNN uses 10-times less the parameters than ResNet18.
- The above is reflected to the training time per epoch, where even the trainable parameters are less than handcrafted CNN they need more time to train because the calculations for inference are done during training.
- At inference time we still get through the froze layers.

Experimental Results - Inference Efficiency

- We can see that the best performance, in terms of Micro-Recall, is yielded when we freeze only the first block of ResNet.

- Between ResNet18 and ResNet34 we do not see much improvement thus, considering the previous slide , we can see that the extra computational cost does not yield better performance.

- Finally, we can see that ResNet models are only 5% better than our handcrafted CNN.



Conclusion & Further Work

- We presented two approaches in order to address the Lung Disease Detection from X-Rays task.
- We saw that if we want a more lightweight version for our model we can use our proposed handcrafted architecture because it is 10-times less complex than ResNet and has 90% recall which is only 5% worse performance than ResNet.
- But, if we care only about inference performance we should use ResNet18 and freeze the first block which seems to have 95% recall.

- In terms of further work we can use the pretrained models and use the feature vectors to train a Machine Learning algorithm.
- Afterwards, we can use autoencoders to work as feature extractors in order to have variable sized feature vectors.

Thank you for your time!!!

Any Questions ???