DESCRIPTION OF WORK

for

BLG 506E

COMPUTER VISION

COURSE PROJECT

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

Burak Bozdağ

504211552

17.11.2022

Table of Contents

1 EXECUTIVE SUMMARY

In the article titled “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”, it was emphasized that transformers applied directly to image patches and pre-trained on large datasets work really well on image classification [1]. In this project, it is aimed to examine that if classifying chest X-ray images as pneumonia, tuberculosis, COVID-19, etc. using transformer-based architectures gives better results than CNN architectures. This will be achieved by using a pre-trained model of popular CNN architectures such as AlexNet [2], GoogLeNet [3], etc. and ViT (Vision Transformer) stated in the related article [1]. Aside comparing models, fine-tuning models and hyperparameter optimization will also be performed.

2 INTRODUCTION

As the influence of the artificial intelligence (AI) and machine learning (ML) sector increased, many researches about this area have been showed up and also there are many ongoing projects and researches ahead. By realizing that the computer can learn from data, new algorithms of machine learning topic are found and developed. The data itself is not limited to just a bunch of numbers but also covers a whole area of computer vision (CV) world. In this world, the data is made of images that can be either classified (supervised) or unclassified (unsupervised).

Analysis of the medical data has been a huge topic itself. Examining and analyzing X-ray images has an importance in the medical domain. There are various work areas such as diagnosing radiology results with respect to the X-ray images of patients. There are so many projects out there for diagnosing and classifying different types of diseases in the literature.

Inspired by the popularity of the medical domain, I am proposing a new topic for comparing standard convolutional neural networks and vision transformers. It is also planned to apply fine-tuning, hyperparameter optimization, learning rate selection and regularization methods if needed.

In the rest of the document, there is given more detailed project explanations with the goals of project, impact of solution, state-of-the-art research articles and novel contributions. In the next section, scope of the project is included with the work breakdown structure and some information about the project. At next, assumptions for the entire project timeline are given. At the end of the document, a reference list is found after the deliverables and project schedule section.

3 PROJECT DESCRIPTION

In this project, chest X-ray images will be classified for different diseases such as pneumonia, tuberculosis, COVID-19, etc. Aside classifying data, a comparison between different CNN and transformer models will be made in order to find the best model for this scenario.

The dataset used in this project is planned to be prepared by collecting various public datasets from Kaggle website such as Chest X-Ray Images (Pneumonia) [4] from Paul Mooney, Tuberculosis (TB) Chest X-ray Database [5] and COVID-19 Radiography Database [6] from Tawsifur Rahman.

- Pneumonia dataset contains 5863 x-ray images in JPEG format which are categorized as pneumonia or normal.
- Tuberculosis dataset contains 700 tuberculosis images publicly accessible and 3500 normal images.
- COVID-19 dataset contains 3616 COVID-19 positive cases along with 10,192 Normal, 6012 Lung Opacity (Non-COVID lung infection), and 1345 Viral Pneumonia images in PNG format.

Merging these datasets, the final dataset size would be 31228 images.

3.1 Goals of Project

There are 2 main purposes of this project.

- Classifying X-ray images for diseases
- Comparing CNN and transformer models

3.2 Impact of Solution

This project will show difference between CNN and transformer models in the domain of X-ray images. At the end of the project, results will be given in a comparative manner in the final report.

3.3 SOTA

AlexNet [2] which published in 2012 is a state-of-the-art work that increased the popularity of convolutional neural network and dee learning models. This model's results are very good compared to existing machine learning and computer vision algorithms.

AlexNet includes 8 layers: 5 convolution and 3 fully connected layers. Comparing to the other algorithms, this model has more filter and convolution layers at each layer. There are various used functions such as max-pooling, dropout, augmentation, ReLU, SGD, etc. Rectified linear unit (ReLU) functions are placed after convolution layers.

The AlexNet structure is given below:

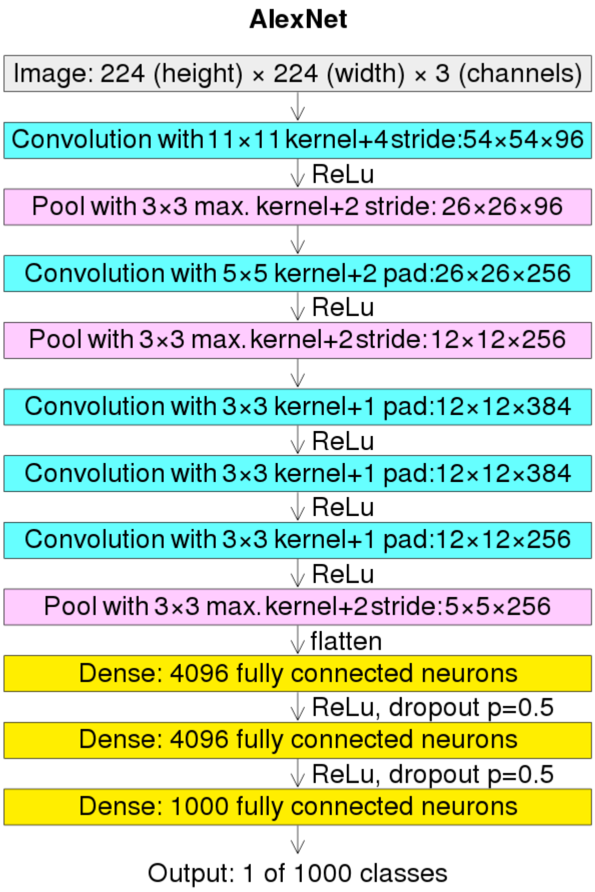

Figure 1. AlexNet Structure

Even though transformer architectures are a standard for natural language processing (NLP) projects, its applicability to computer vision projects were limited. In computer vision, convolutional neural networks seem like the best way to solve classification problems.

In the related study [1], it is stated that the dependence to the CNN models is not necessary. Transformer structures can also be adapted to classification tasks and will give very good results. For example, some popular public datasets such as ImageNet, CIFAR-100, VTAB, etc. trained with vision transformers have given better results compared to the state-of-the-art CNN models.

The vision transformer structure is given below:

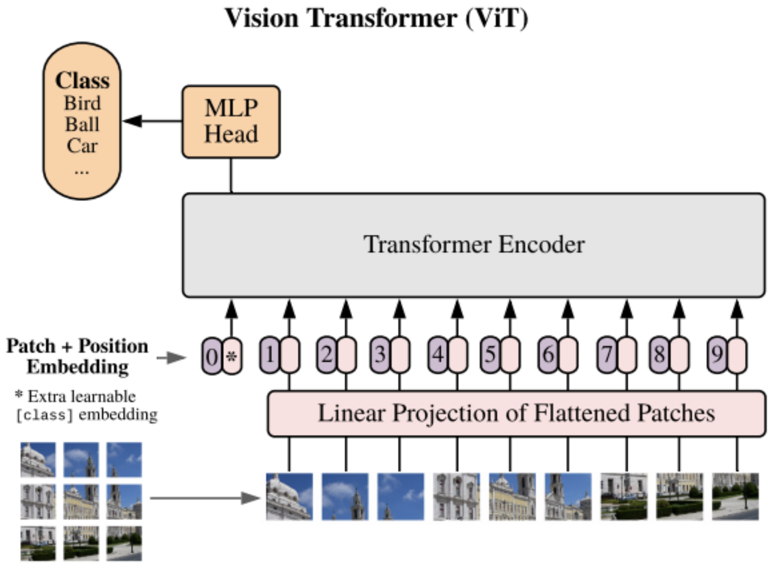

Figure 2. Vision Transformer (ViT)

Similar works about my study area include but not limited to:

- Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning: Demonstrating the general applicability of their AI system for diagnosis of pneumonia using chest X-ray images [7].
- Reliable Tuberculosis Detection using Chest X-ray with Deep Learning, Segmentation and Visualization: They have detected tuberculosis reliably from the chest X-ray images using image pre-processing, data augmentation, image segmentation, and deep-learning classification techniques [8].
- Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images: A novel U-Net model was proposed and compared with the standard U-Net model for lung segmentation [9].

3.3.1 Novel contributions

This project will lead us that if CNN models are enough and sufficient for detecting diseases from X-ray images or if transformer models can be also used for this purpose. Also, trials for multiple disease diagnosis will be made which means that there will be not only one disease to classify but other diseases (different public datasets are planned to be merged).

3.4 Risk Assessment

Possible Risk	Risk Reason	Contingency Plans
Different resolutions on dataset	Dataset	Adjust the brightness/contrast of images
Limited data	Dataset	Apply different fine-tuning approach
Low model accuracy	Training strategy or model building errors	Check the model, apply hyper parameter optimization, use different learning rate selection
Some samples may have poor quality	Dataset	Reduce resolution for each image
Training takes very long time	Model or technical	Train on GPU, change GPU model, train parallel on GPU, freeze more layers, change hyper parameters
Overfitting	Model	Use regularization, dropout, separate train, validation and test datasets

Table 1 Risk Assessment Table

4 PROJECT SCOPE

This SOW shall apply to the tasks, services and terms detailed below:

4.1 Work Breakdown Structure (WBS)

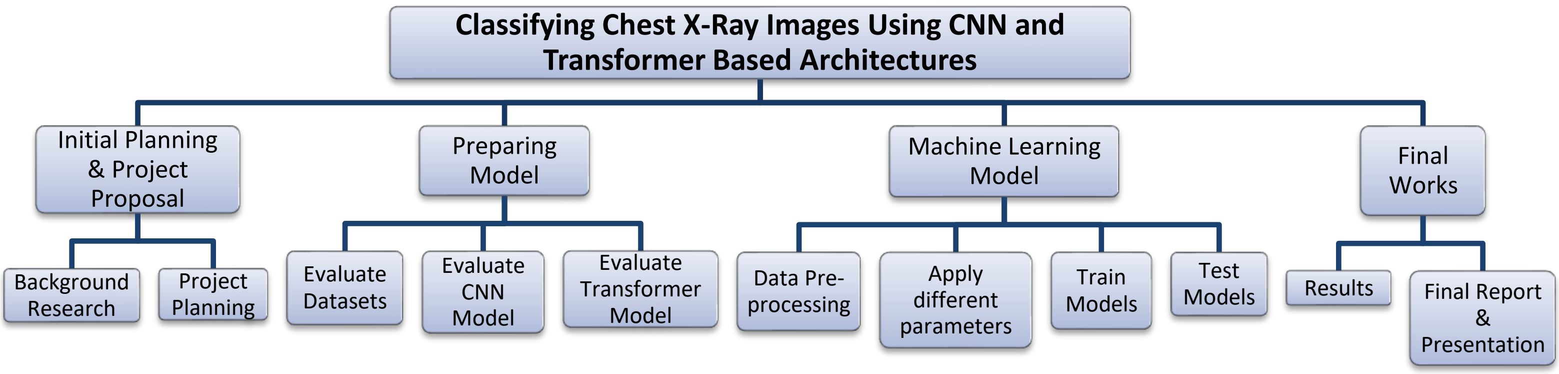

Figure 1 Example Work Breakdown Structure (WBS)

4.2 Work Packages

WP 1	**Initial Planning & Project Proposal**		
Start Date		End Date	14.11.2022
Objectives: This work package will cover determining the problem and the project subject, doing research on the subject, determining the project schedule and the project management.			
Tasks ☐ *Background Research* ☐ *Project Planning*			
Deliverables and Milestones: D1: Project proposal and presentation			

WP 2	**Preparing Model**		
Start Date	14.11.2022	End Date	28.11.2022
Objectives: This work package will cover preliminary preparations to create models for the problem. In this section, all technical features of the models to be used will be decided.			
Tasks ☐ *Evaluate Datasets* ☐ *Evaluate CNN Model* ☐ *Evaluate Transformer Model*			
Deliverables and Milestones: M2: All technical features of the models decision D2: Decision of the project process			

WP 3	**Machine Learning Model**		
Start Date	28.11.2022	End Date	19.12.2022
Objectives: This work package will cover creating models, pre-processing dataset, training models and testing models.			

Tasks
☐ *Data Pre-processing* ☐ *Apply different parameters* ☐ *Train Models* ☐ *Test Models*
Deliverables and Milestones:
M3.1: Creation of the models M3.2: Model training M3.3: Model testing D3: Progress presentation

WP 4	**Final Works**		
Start Date	19.12.2022	End Date	26.12.2022
Objectives: This work package will cover collecting, evaluating, comparing the results, preparing a final report with all the information obtained.			
Tasks ☐ *Results* ☐ *Final Report & Presentation*			
Deliverables and Milestones: M4.1. Collecting and comparing results M4.2: Comment and conclude about the work D4.1: Final Report D4.2: Final Presentation			

4.3 Out of Scope

The following are considered OUT OF SCOPE for this contract:

- Detection of anomaly regions on X-ray images
- Detecting viral-bacterial infection areas

5 ASSUMPTIONS

Project assumptions typically revolve around constraints such as time, hardware, and scope. The assumptions of this project can be listed as follows:

1. All the resources needed to complete the project, both information and material, will be accessible.
2. While working on the project, I will have the resources needed to complete tasks on time, from the necessary equipment, software, electricity throughout the project life.
3. All equipment and hardware will remain operational throughout the project cycle.
4. The overall scope of the project will not change throughout project life cycle. However, tasks may undergo minor changes during detailing.
5. Training of the models will not exceed the duration of the project.
6. The dataset will be in the quality, amount and format suitable for the project. The images used will be labeled in the appropriate format.

6 MILESTONES and DELIVERABLES

6.1 Deliverables and Milestone Tables

Deliverable (D)	Description	Date	Milestone (M)
D1	Project proposal and presentation	14.11.2022	-
D2	Decision of the project process	05.12.2022	M2
D3	Progress presentation	19.12.2022	M3.2
D4.1	Final report	26.12.2022	M4.1
D4.2	Final presentation	26.12.2022	M4.2

Table 2 Deliverable Table

6.2 Project Schedule (Gantt Chart)

Weeks	W1	W2	W3	W4	W5	W6	W7
WP 1	D1						
WP 2		M2	D2				
WP 3			M3.1	M3.2	M3.3	D3	
WP 4						M4.1 M4.2	D4.1 D4.2

Figure 2 Gantt chart

7 References

[1] A. Dosovitsky et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale", 2021. Available: https://openreview.net/forum?id=YicbFdNTTy.

[2] Krizhevsky, Alex & Sutskever, Ilya & Hinton, Geoffrey. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems. 25.10.1145/3065386.

[3] C. Szegedy et al., "Going Deeper with Convolutions", arXiv.org, 2014. [Online]. Available: https://arxiv.org/abs/1409.4842.

[4] "Chest X-Ray Images (Pneumonia)", Kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia. [Accessed: 17- Nov- 2022].

[5] "Tuberculosis (TB) Chest X-ray Database", Kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset. [Accessed: 17- Nov- 2022].

[6] "COVID-19 Radiography Database", Kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database. [Accessed: 17- Nov- 2022].

[7] Kermany, Daniel S et al. "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning." Cell vol. 172,5 (2018): 1122-1131.e9. doi:10.1016/j.cell.2018.02.010

[8] T. Rahman et al., "Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization," in IEEE Access, vol. 8, pp. 191586-191601, 2020, doi: 10.1109/ACCESS.2020.3031384.

[9] Rahman, Tawsifur et al. "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images." Computers in biology and medicine vol. 132 (2021): 104319. doi:10.1016/j.compbiomed.2021.104319