# COMP 432 - Project Proposal

Anonymous submission

Paper ID

## 1. Introduction

This project sets out to define a machine learning model that will be able to perform classification tasks on three separate datasets: one for colorectal cancer diagnosis, another on prostate cancer, and a last one containing animal faces. This will be possible by using a technique called Learning Transferability: it is a model's ability to generalize knowledge from one dataset / task to another. It is imperative that this technique is studied seeing as models usually require large datasets in order to be trained, and training data is seldom easy to acquire. Hence, this technique allows researchers to bypass this requirement to some extent. However, models trained on one dataset will often face performance issues when applied to novel data (i.e. identifying cancer cells to animal recognition). Understanding transferability and proper application will make the model much more versatile and will reduce the need for extensive retraining. Among the several challenges of this technique, one major issue is the difference in data distributions: the characteristics of cancerous cells are wildly different from those in animal faces. These variations can lead to poor model performance if not properly relearned. The model may also face issues with overfitting to the original task, where it takes into account the noise from the training dataset, limiting its accuracy during the testing phase.

The goal of this project is to develop a machine learning model that can transfer knowledge between three distinct tasks: identifying prostate cancer, colorectal cancer, and animal face recognition. The use of these distinct datasets will help assess the model's ability to efficiently generalize. The project will involve testing the model on each dataset, and reporting on the challenges of transferability as the parameters are fine-tuned and the model's performance improves.

## 2. Methodology

This project will use ResNet as the underlying architecture for the CNN model. The PyTorch framework includes model builders for several versions of this architecture, but given the size of the datasets that will be used and the substantial computational resources that both RestNet-101 and 152 require, they will not considered when looking at possible options [3]. This leaves ResNet-18, 34, and 50 as possible choices. Experimenting with the pre-trained weights of each model [2] will help decide which out of the three to use.

To avoid overfitting, the use of regularization techniques such as Batch Augment, Cutout and RandomErasing will be crucial. This will artificially increase the size of the first dataset and introduce some random noise as well [8]. This is not required the second task of the project since one of the main advantages of transfer learning is that it needs less data. Finally, this project will utilize the Adam optimization algorithm, which has proven multiple times that it is capable of producing quality results when applied to Deep Learning problems [1].

For qualitative analysis, t-SNE can help visualize the extracted features of the data learned by the CNN models and see how each model clusters similar patterns [4]. When performing knowledge transfer from the CNN model trained on Colorectal Cancer to Prostate Cancer or Animal Faces datasets, t-SNE can depict how well the learned features generalize between the different datasets. Another tool is the Grad CAM method which produces a heat map that could illustrate what regions of colorectal cancer images (or prostate cancer/animal faces) are helping the most to the model's decision-making [6]. Also, it can help detect if the model is overfitting to useless details in the images.

For quantitative analysis, a confusion matrix will show the number of true positives, false positives, true negatives and false negatives for each class in the dataset [5]. This is particularly useful for knowledge transfer since it reveals if the model performs better on some classes over others. Other metrics can be calculated to evaluate the performance of the model such as precision, recall and F1 score [5]. Precision and recall are important metrics to consider because low scores on both will signal many false negatives and false positives respectively. Finally, the F1 score is the harmonic mean of precision and recall which will help balance both false positives and false negatives during the model's training and knowledge transfer phases [7].

# 3. Gantt Chart

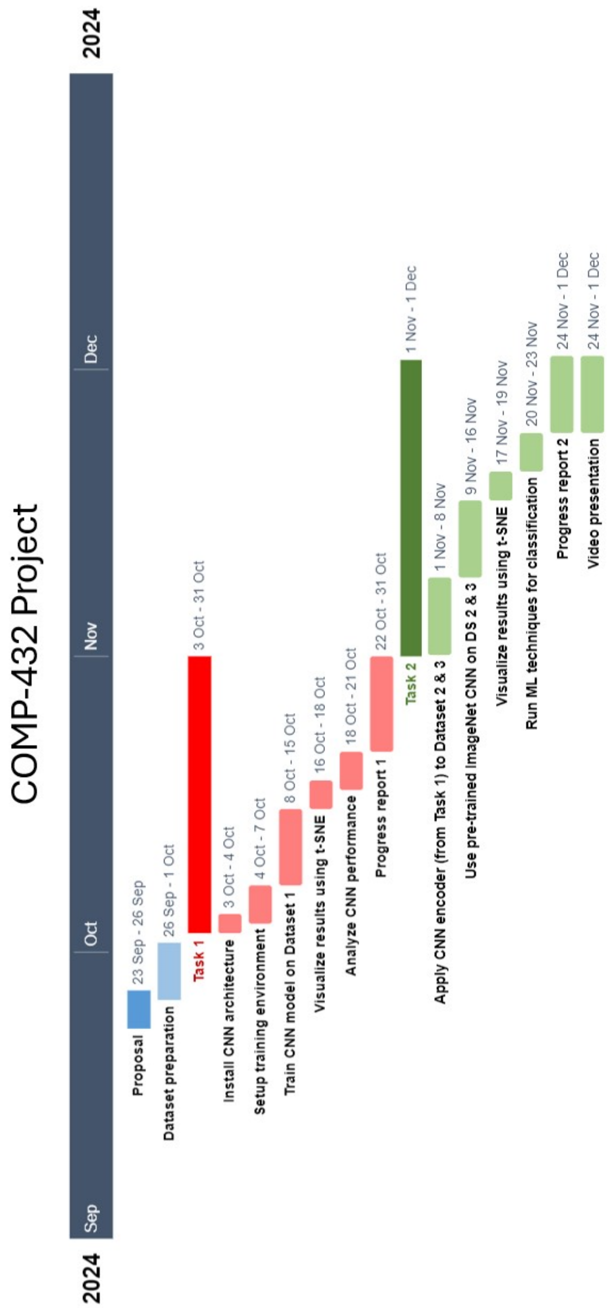


Figure 1. Estimated timeline for the project

## References

[1] Jason Brownlee. Gentle introduction to the adam optimization algorithm for deep learning, 2021. Accessed: 2024-09-26. 1

[2] Torch Contributors. Models and pre-trained weights, 2024. Accessed: 2024-09-26. 1

[3] Torch Contributors. Resnet, 2024. Accessed: 2024-09-26. 1

[4] Kemal Erdem. t-sne clearly explained, 2020. Accessed: 2024-09-26. 1

[5] Will Koehrsen. Precision and recall: How to evaluate your classification model, 2024. Accessed: 2024-09-25. 1

[6] Melanie. What is the grad cam method?, 2024. Accessed: 2024-09-25. 1

[7] Oona Rainio, Jarmo Teuho, and Riku Klén. Guide to transfer learning in deep learningevaluation metrics and statistical tests for machine learning, 2024. Accessed: 2024-09-26. 1

[8] Claudio Filipi Gonçalves Dos Santos and João Paulo Papa. Avoiding overfitting: A survey on regularization methods for convolutional neural networks. *ACM Computing Surveys*, 54 (10):1–25, 2022. Accessed: 2024-09-25. 1