Project Proposal

1. Problem Statement and Application

In Computer Vision (CV), Convolutional Neural Networks (CNNs) play a crucial role in processing image and video data, allowing machines to effectively understand and analyze visual information. Transfer learning improves performance in new tasks by using prior knowledge from a previous task to reduce the need for extensive data collection and optimize resource usage. Challenges include handling variability within classes, assessing transfer learning across domains, and tuning hyperparameters for CNN models. [4]. We expect to achieve better model performance through transfer learning in closely related domains compared to distantly related domains.

This project focuses on CNNs for real-world image classification tasks, with Task 1 involving CNN model selection, hyperparameter tuning, and evaluating the use of pretrained networks compared to training from scratch. Task 2 involves transfer learning, utilizing pre-trained CNN encoders to extract features from new datasets, subsequently integrating them into machine learning models for image classification.Feature extraction results are visualized with t-SNE, and model performance is assessed using confusion matrices and standard metrics.

2. Image Dataset Selection

Dataset 1 originated from the NCT Biobank (National Center for Tumor Diseases, Heidelberg, Germany) and the UMM pathology archive (University Medical Center Mannheim, Mannheim, Germany). Dataset 2 is related to a research paper [3]. Dataset 3 is from AFHQ (Animal Faces-HQ).

	D1	D2	D3
# Images	6000	6000	6000
Dim	224x224 pixels (RGB)	224x224 pixels (RGB)	224x224 pixels (RGB)
Format	TIFF	JPG	JPG
# classes	3	3	3
# images original D.	100000	6 dataset each containing 120000	16130
# classes original D.	9	3	3
Source	[2]	[5]	[1]

Table 1. Datasets description

3. Possible Methodology

As an initial preprocessing of the data we are going to bring all the images into the same dimensions (224x224x3) so the model can have a fair chance of learning all the information in all the images. Color-normalization is going to be used with the default mean and std of pyTorch for image normalization since these values come from ImageNet, and since we are going to compare our model with a pretrained model it would make the comparisons more fair.

In recent years, Residual Neural Networks (ResNet) have gained a lot of attention from researchers thanks to their ability to interpret medical pictures from various backgrounds (e.g., clinical diagnosis, metastatic evolution, and target selection for illnesses just to name a few) with high accuracy. Taking into consideration the above, all the documentation that is available about ResNet and the fact that it is one of the pretrained models available in PyTorch, we are going to use it in our project. The original architecture has 152 layers and since we don't have the computational resources to train given the time period of the project a simplified version of this architecture called ResNet 18 is going to be used.

For the first task, we are going to train ResNet 18 from scratch with the end goal of classifying 3 classes in the Colorectal cancer dataset, for this we need to tune some hyperparameters. We are going to consider the learning rate since it is crucial for the time of training and convergence time, batch size, kernel width and possibly weight decay. We will compare the results of our model with a ResNet 18 model that has already been trained on ImageNet.

For the second task, we are going to use the pretrained ResNet 18 model that is available in PyTorch and our trained model in two new datasets, and we are going to evaluate how good they are for transfer learning in two datasets. Also, KNN and Random Forest are going to be used to classify the extracted features by the CNN encoders.

For both tasks, we are going to compare the results for all the experiments using metrics such as: precision, recall, f1-score. Also, we are going to compare the experiments in time of convergence for both training and testing accuracy/loss. t-SNE is going to be applied to the output and extracted features for visualization and analysis.

We expect the performance of the models to be better when the transfer learning is in a "close-related" topic such as Colorectal and Prostate cancer classification, and worse when the transfer learning is to a more general topic such as Animal faces classification but still better than random feature selection. If we can confirm the above, even to a certain extent, then other researches may be able to use pretrained models and use transfer learning to transfer features information to another task to reduce greatly the training time and as a consequence the computational cost without having a severe compromise on the accuracy of their model.

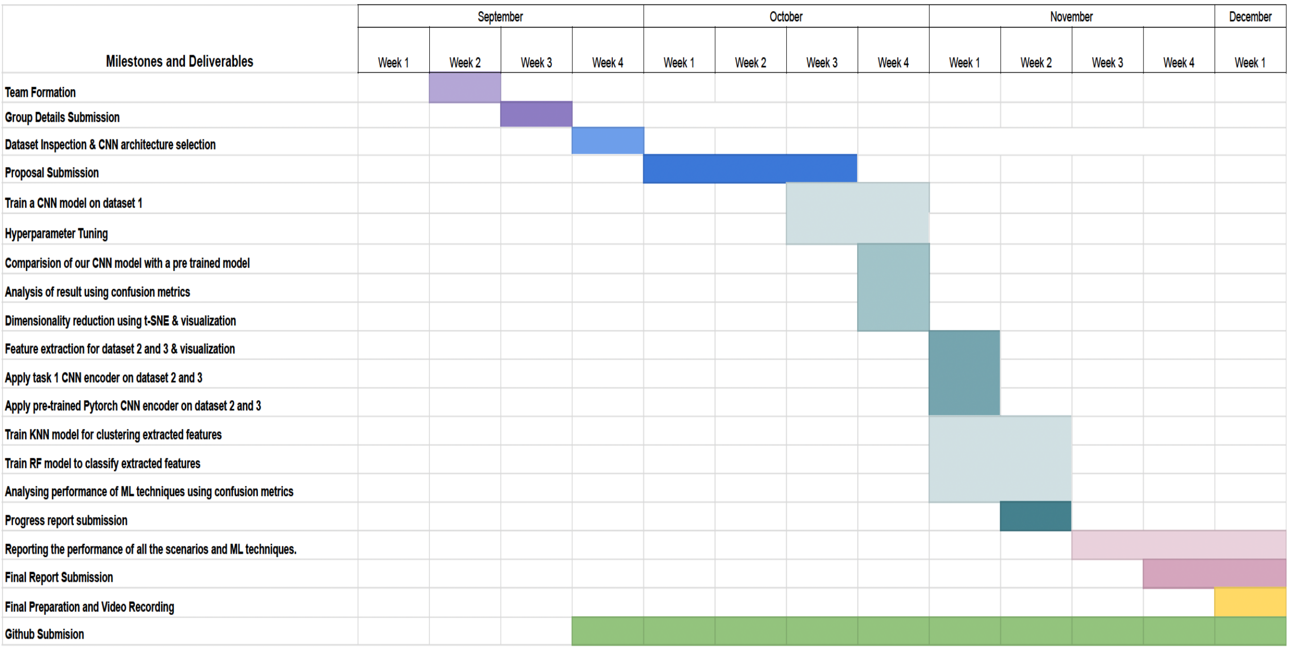

Figure 1. Gantt Chart

References

[1] Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. Stargan v2: Diverse image synthesis for multiple domains. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 1

[2] Jakob Nikolas Kather, Niels Halama, and Alexander Marx. 100,000 histological images of human colorectal cancer and healthy tissue, May 2018. 1

[3] Birgid Markiefka, Alexey Pryalukhin, Wolfgang Hulla, Andrey Bychkov, Junya Fukuoka, Anant Madabhushi, Viktor Achter, Lech Nieroda, Reinhard Büttner, Alexander Quaas, and Yuri Tolkach. Quality control stress test for deep learning-based diagnostic model in digital pathology. *Modern Pathology*, 34, 06 2021. 1

[4] Abdulrahman Abbas Mukhlif, Belal Al-Khateeb, and Mazin Abed Mohammed. An extensive review of state-of-the-art transfer learning techniques used in medical imaging: Open issues and challenges. *Journal of Intelligent Systems*, 31(1):1085–1111, 2022. 1

[5] Yuri Tolkach. Datasets digital pathology and artifacts, part 1, June 2021. 1