Project Proposal

Group-I submission

1. Problem Statement and Application

This project focuses on a critical Computer Vision (CV) application that combines three distinct image classification challenges: Colorectal Cancer classification, Prostate Cancer classification, and Animal Faces classification. The primary objective is to develop and analyze Convolutional Neural Network (CNN) models tailored to address these unique image classification tasks. Accurate classification of colorectal and prostate cancers holds the potential to facilitate early detection and refine treatment strategies, potentially leading to life-saving outcomes. Animal face classification is crucial in wildlife conservation, ecological research, pet care, and other fields, aiding in understanding and protecting wild and domestic animals contributing to ecosystem well-being and human-animal interactions. The project also explores the knowledge transferability of pre-trained CNN models across medical and non-medical domains to optimize resource usage and reduce extensive data collection requirements. Key challenges include dealing with significant variability in visual characteristics within each class, assessing the effectiveness of transfer learning between medical and non-medical datasets, and identifying optimal hyperparameters for CNN models [3].

2. Image Dataset Selection

Dataset 1: Colorectal Cancer Classification
6,000 images in TIFF format, each with dimensions 224x224 pixels with 3 channels (RGB). The classes are Smooth Muscle (MUS), Normal Colon Mucosa (NORM), and Cancer-Associated Stroma (STR). Each class comprises 2,000 images. Dataset Origin: `https://zenodo.org/record/1214456`. Authors: Kather, Jakob Nikolas; Halama, Niels; Marx, Alexander (2018).
Dataset 2: Prostate Cancer Classification
6,000 images in JPG format, each with dimensions 300x300 pixels with 3 channels (RGB). The classes are gland, nongland, and tumor. Each class comprises 2,000 images. Dataset Origin: `https://zenodo.org/record/4789576`. Author: Tolkach, Yuri. (2021).
Dataset 3: Animal Faces Classification
6,000 images in JPG format, each with dimensions 512x512 pixels with 3 channels (RGB). The classes are Cats, Dogs, and Wild. Each class comprises 2,000 images. Dataset Origin: `https://www.kaggle.com/datasets/andrewmvd/animal-faces`. Paper: Choi et al. [2].

3. Possible Methodology

At the moment of this delivery, we are inclined to use the RestNet 18 architecture and train it with the Colorectal Cancer classification data. We plan to tune the learning rate and batch size (initially using the most common values or methods in the literature for this kind of task, e.g., the work of Hamida et al. [1]), and see how different loss functions behave with the data. The way we are going to do this is to train the model for two to four epochs with different loss functions and analyze the model performance in terms of the error, and pick the one that looks more promising (depending on the time it takes to train we may perform combinations of the hyperparameters and take the one that performance the best). For the second task, we plan on using the VGG16 architecture alongside KNN and Random Forest as ML models to classify the extracted features by the CNN encoders. We will compare the experiments' results using metrics such as precision, recall, and f1-score. Also, we will compare the experiments in time of convergence for both training and testing accuracy/loss. As Yosinski et al. explained in [4], 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 researchers may be able to use pre-trained models and use transfer learning to transfer feature information to another task to reduce the training time significantly and, as a consequence, the computational cost without having a severe compromise on the accuracy of their model.

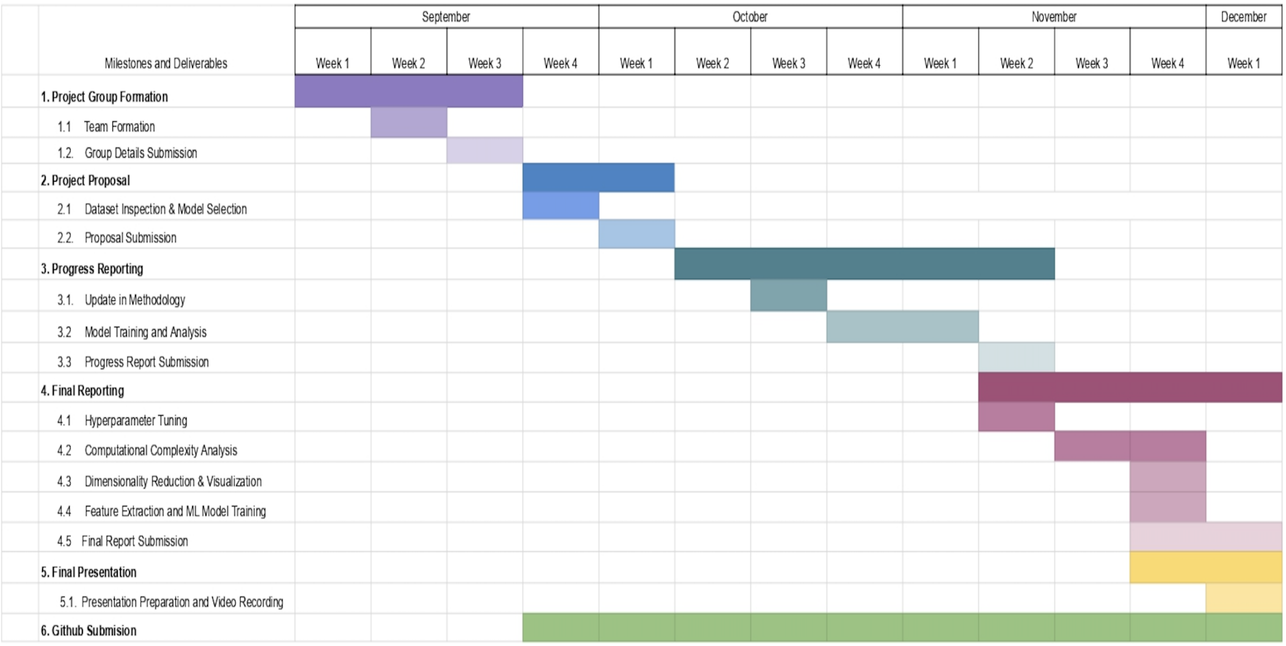

Figure 1. Gantt Chart

References

[1] A. Ben Hamida, M. Devanne, J. Weber, C. Truntzer, V. Derangère, F. Ghiringhelli, G. Forestier, and C. Wemmert. Deep learning for colon cancer histopathological images analysis. *Computers in Biology and Medicine*, 136:104730, 2021. 1

[2] 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

[3] 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

[4] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks?, 2014. 1