Utillizing deep learning in medical imaging

Guided by: Roni Asulin

Colorectal Cancer Classification

Noam Koren & Shahar Siman Tov

מכון טכנולוגי חולון
Holon Institute of Technology

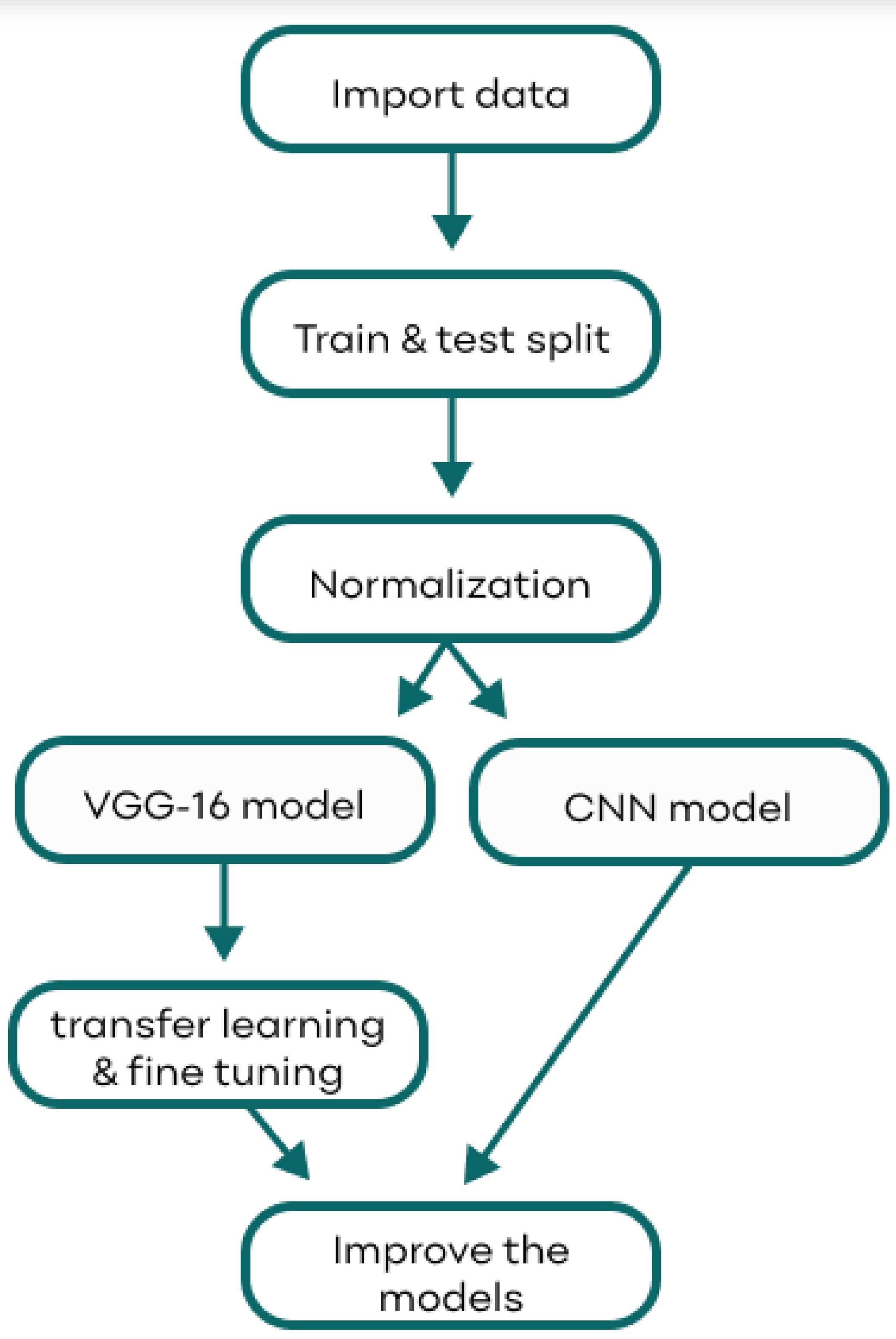
ABSTRACT

In this project, we investigated the histological dataset of the colon to develop and compare different deep learning models whose goal was to accurately classify the images into their respective labeled categories.

DATASET

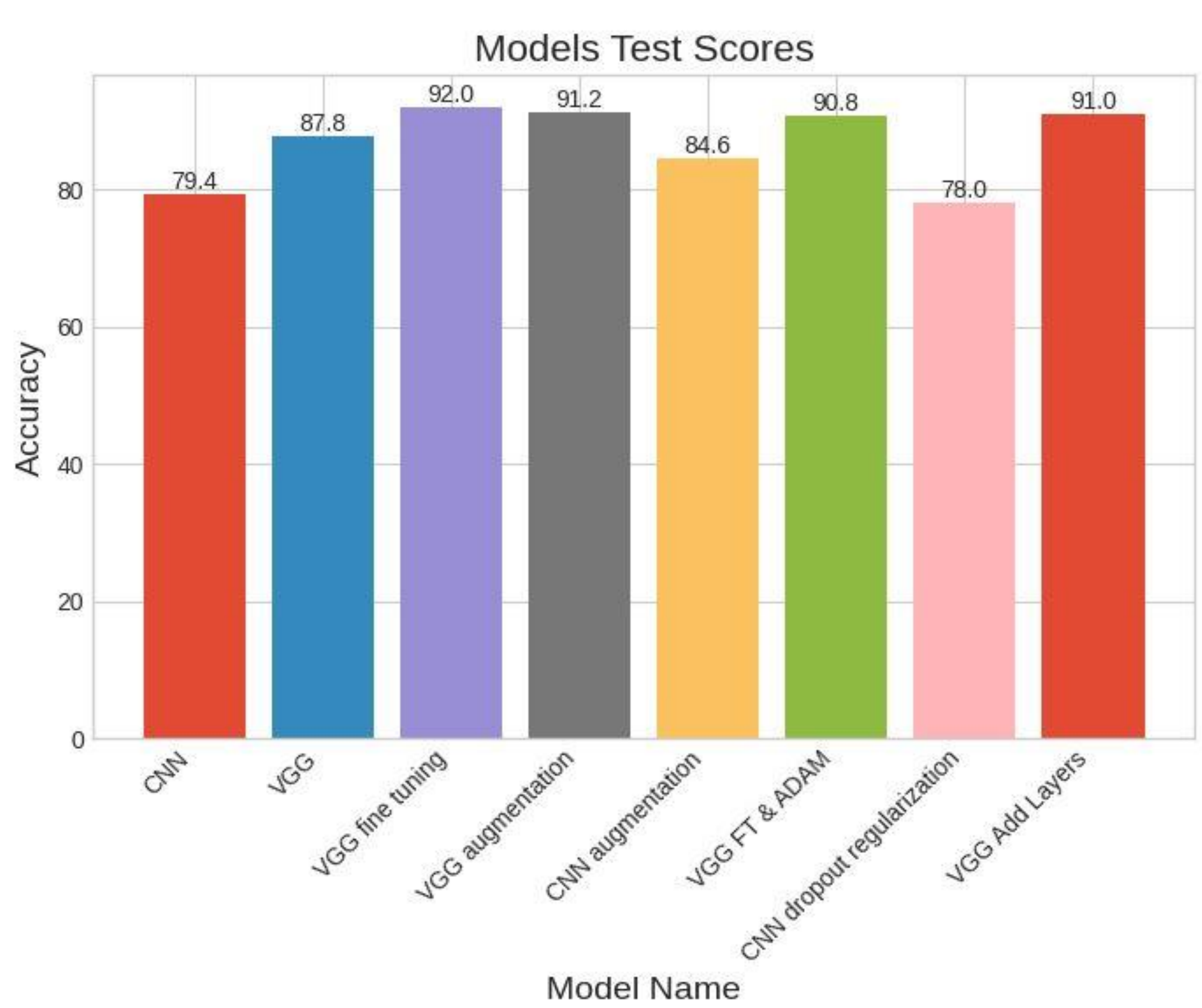
The “Colon Histology” dataset contains 5,000 images, each 150x150 pixels in size, depicting human colon tissues. These images are labeled using 8 different classes, representing different characteristics of the intestinal tissue.

METHODOLOGY



RESULTS

Figure 1 shows the results of all the models developed in the project.



BEST MODEL RESULTS

The VGG-16 model with transfer learning and fine-tuning obtained the best results among all the models.

Figure 2 depicts the accuracy and loss of the model. The accuracy graph shows high performance on both the training and validation sets. The gap between the scores of the two sets is small, indicating that the model performs well in generalization and can classify unseen data effectively. The loss graph exhibits low and converging values, indicating the model's ability to provide reliable and consistent predictions.

This model utilized pre-training, enabling it to learn powerful and generalizable features useful for image-related tasks. Through transfer learning, we initialized the VGG-16 model with weights from ImageNet, providing a solid starting point to leverage its learned features for the colorectal histology dataset. Subsequently, fine-tuning further adapted the pre-trained features to enhance performance on our specific task.

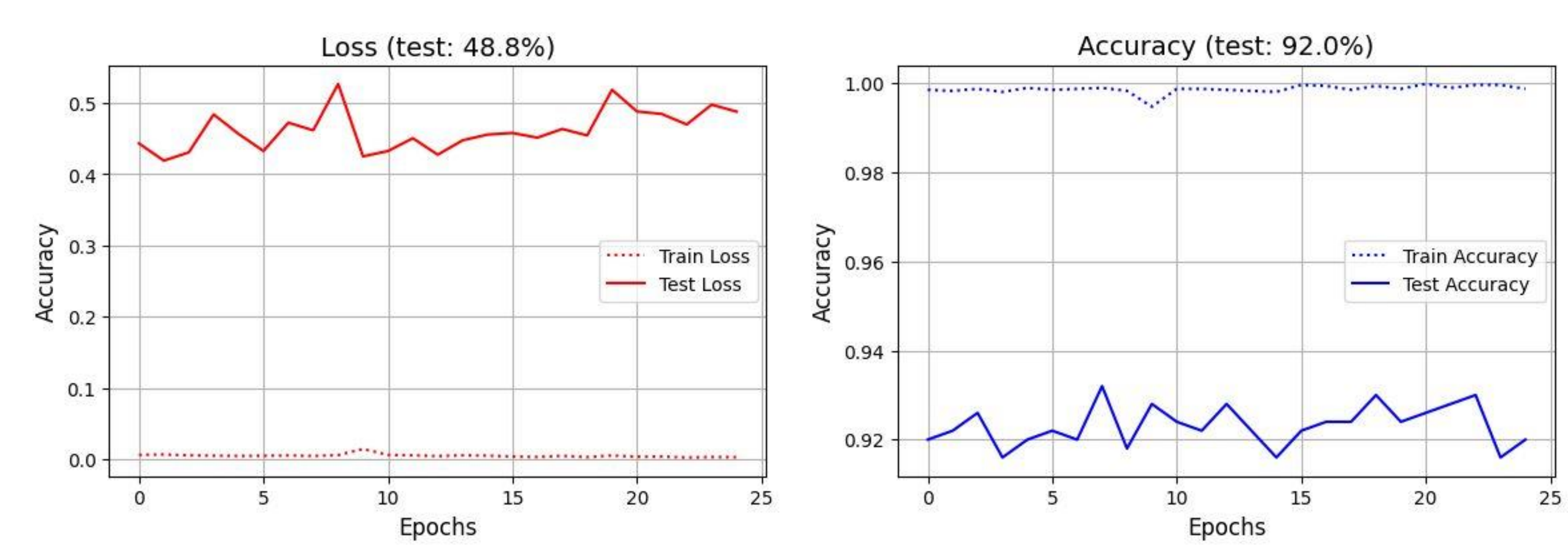


Figure 2: VGG-16 model results.

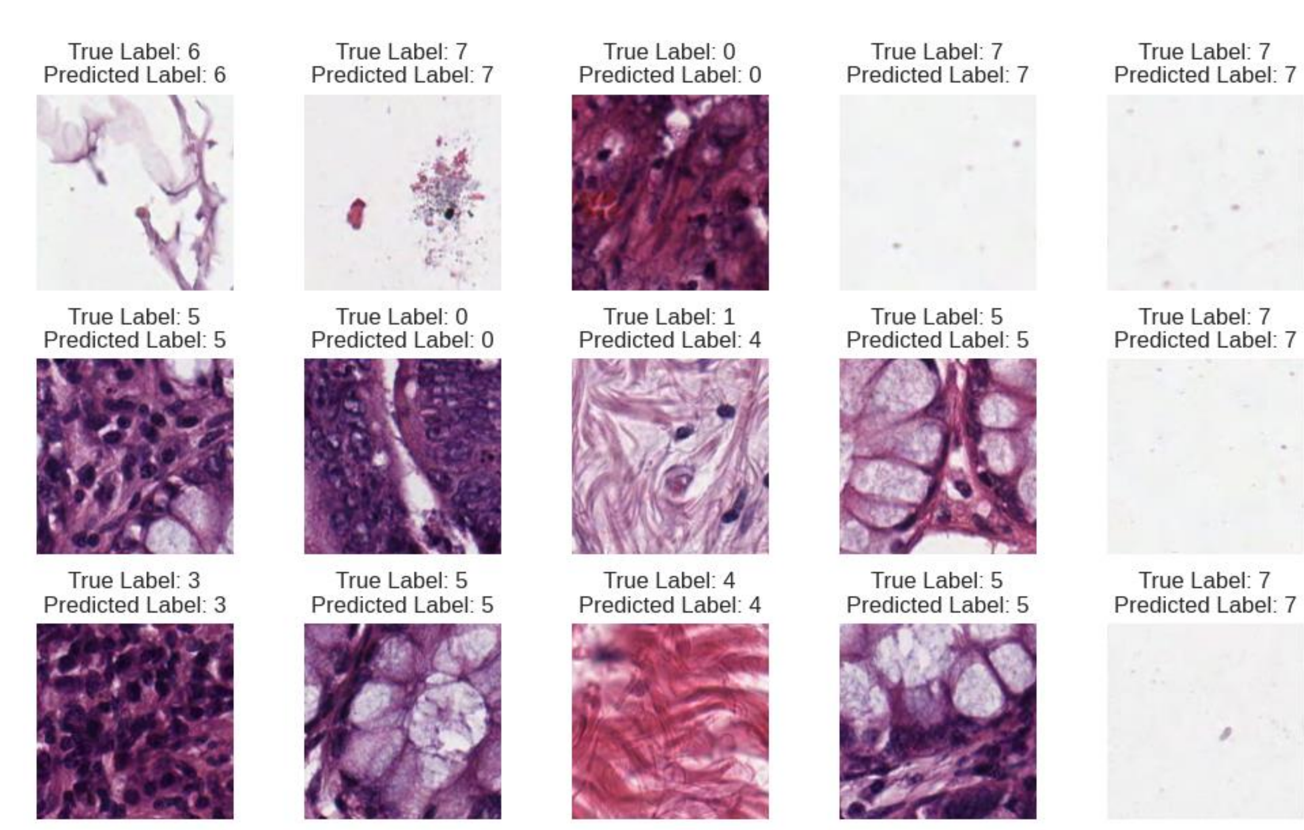


Figure 3: Classified test results.

CONCLUSIONS

Deep learning models that are pre-trained have significant potential in medical image classification tasks. There are numerous ways to enhance them, such as adding layers, exploring different architectures, optimizing, fine-tuning hyper-parameters, and using data augmentation. All these approaches can contribute to improving prediction and generalization results of the models. It can be inferred that the size of the dataset has a substantial impact on the model's success, emphasizing the importance of adapting the model architectures to the size and nature of the data.