STEVENS
INSTITUTE OF TECHNOLOGY
1870

Histopathological Cancer Detection

Varan Singh Rohila
Neeraj Lalwani
Lochan Basyal

Outline

- Introduction
- Description of Dataset
- Machine Learning Algorithms
- Multi-Layer Perceptron Network
- Convolution Model
- ResNet50 Model
- Inception Model
- Ensemble Model
- Model Comparison
- Implementation Details
- Future Direction
- Conclusion

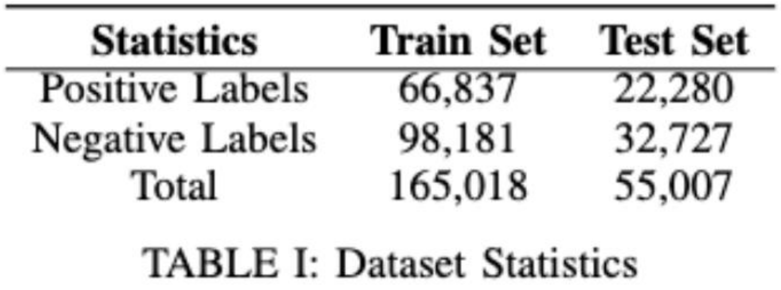
Introduction

Early diagnosis of the cancer cells is necessary for making an effective treatment plan and the health safety of a patient. Nowadays, doctors usually use a histological grade that pathologists determine by performing a semi-quantitative analysis of the histopathological and cytological features of hematoxylin-eosin (HE) stained histopathological images.

- This research contributes a potential classification model for cancer prognosis to efficiently utilize the valuable information underlying the HE-stained histopathological images. This work uses the PatchCamelyon benchmark datasets and trains them in a multi-layer perceptron and convolution model to observe the model's performance in terms of precision, Recall, F1 Score, Accuracy, and AUC Score. The evaluation result shows that the baseline convolution model outperforms the baseline MLP model. Also, this paper introduced ResNet50 and InceptionNet models with data augmentation where ResNet50 able to beat the state-of- the-art model. Furthermore, majority vote and concatenation ensemble were evaluated and provides the future direction of using transfer learning, and segmentation to understand the specific features.

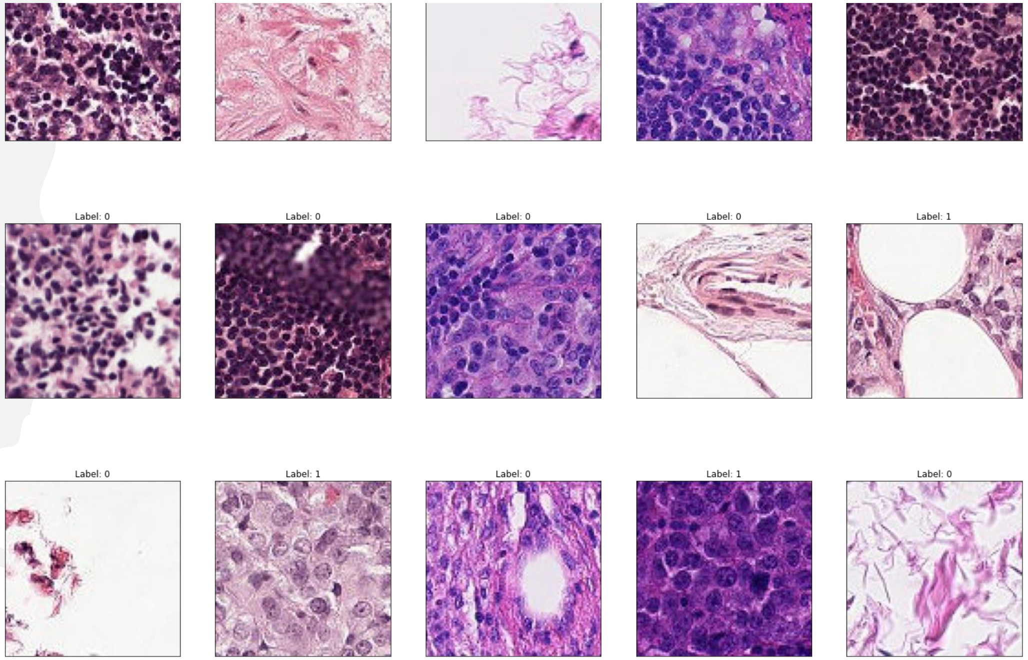
Description of Dataset

- The dataset we considered for to evaluate our approach is PatchCamelyon benchmark dataset [12]. It consists of 220,025 color images (96 x 96px) extracted from histopathologic scans of lymph node sections. Each image is annotated with a binary label indicating the presence of metastatic tissue. A positive label indicates that the center 32x32px region of a patch contains at least one pixel of tumor tissue. Tumor tissue in the outer region of the patch does not influence the label. The images from the data are shown in 1. The dataset is divided into the train, and test sets, with the test set being 25% of the dataset. The data statistics are tabulated in Table:



Statistics	Train Set	Test Set
Positive Labels	66,837	22,280
Negative Labels	98,181	32,727
Total	165,018	55,007

TABLE I: Dataset Statistics



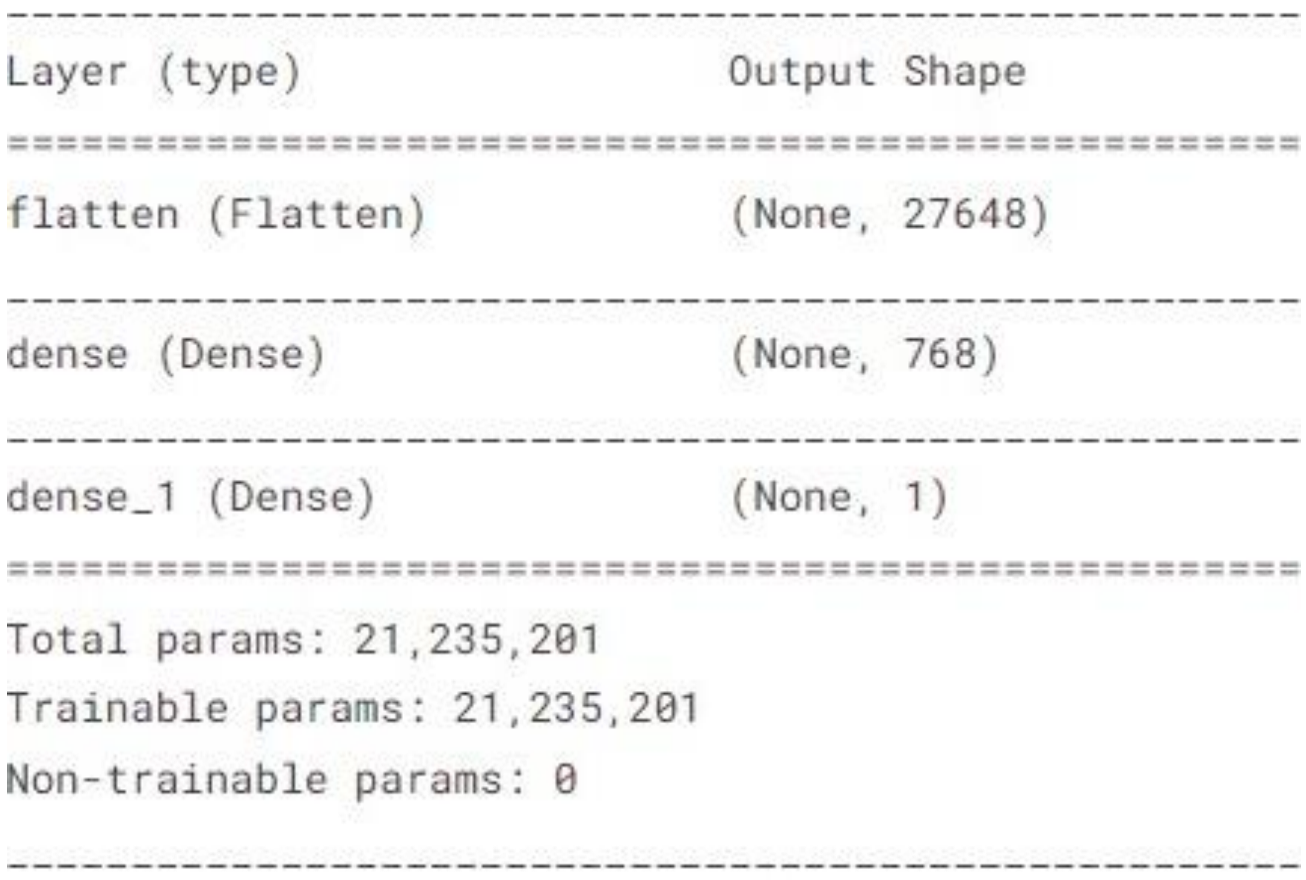
Examples from the dataset

Machine Learning Algorithms

- For our baseline models, we consider a multi-layer perceptron model and a simple convolutional model for image classification. The input of both these models is (96px, 96px, 3), where 3 corresponds to the RGB image channels. Briefly describing the baseline models below:
- Multi-Layer Perceptron Network
- Convolution Model
- ResNet50 Model
- Inception Model
- Majority Vote Ensemble Model
- Concatenation Ensemble Model

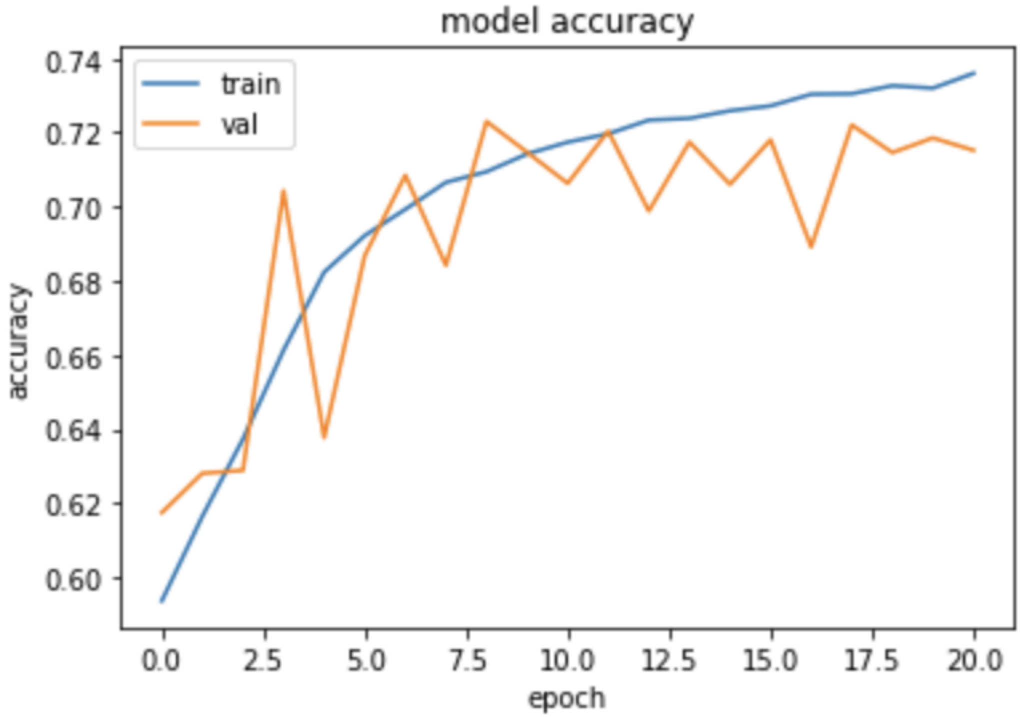
Multi-layer Perceptron Model

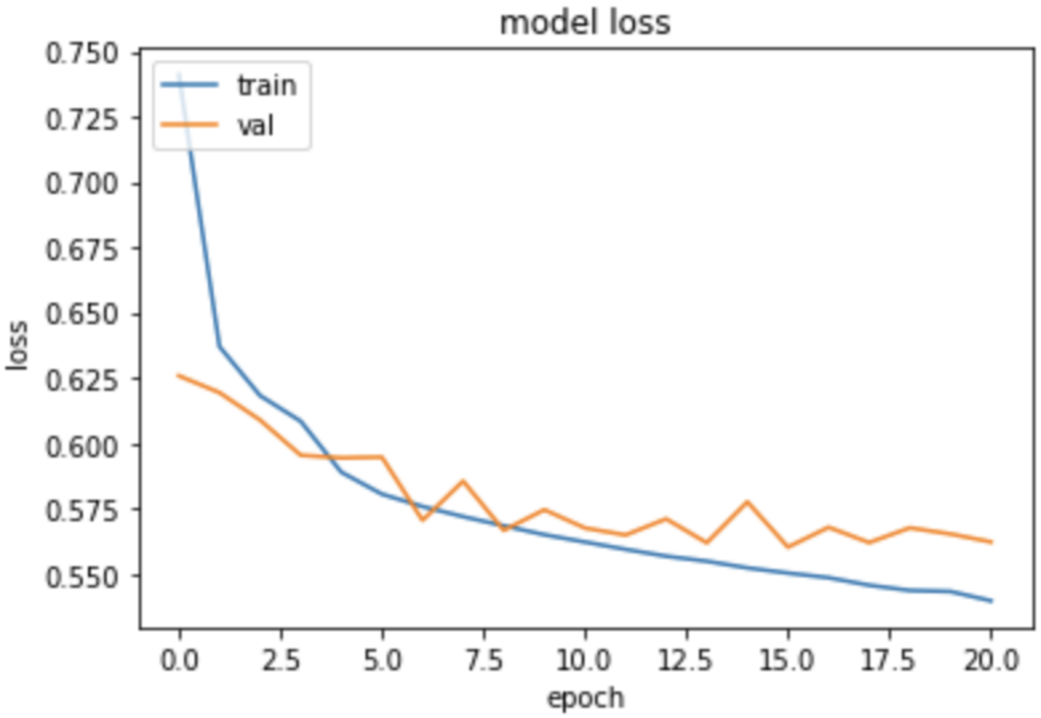
- Multi-layer Perceptron Model: Simple ANN models have been proven to perform well on image tasks such as MNIST Handwriting Benchmark Dataset. Hence, we consider a single-layer MLP model with 768 hidden nodes with relu activation. A single node with sigmoid activation outputs the predictions.



```
____________________________________________________
Layer (type)                 Output Shape
====================================================
flatten (Flatten)            (None, 27648)

____________________________________________________
dense (Dense)                (None, 768)

____________________________________________________
dense_1 (Dense)              (None, 1)
====================================================
Total params: 21,235,201
Trainable params: 21,235,201
Non-trainable params: 0
____________________________________________________
```

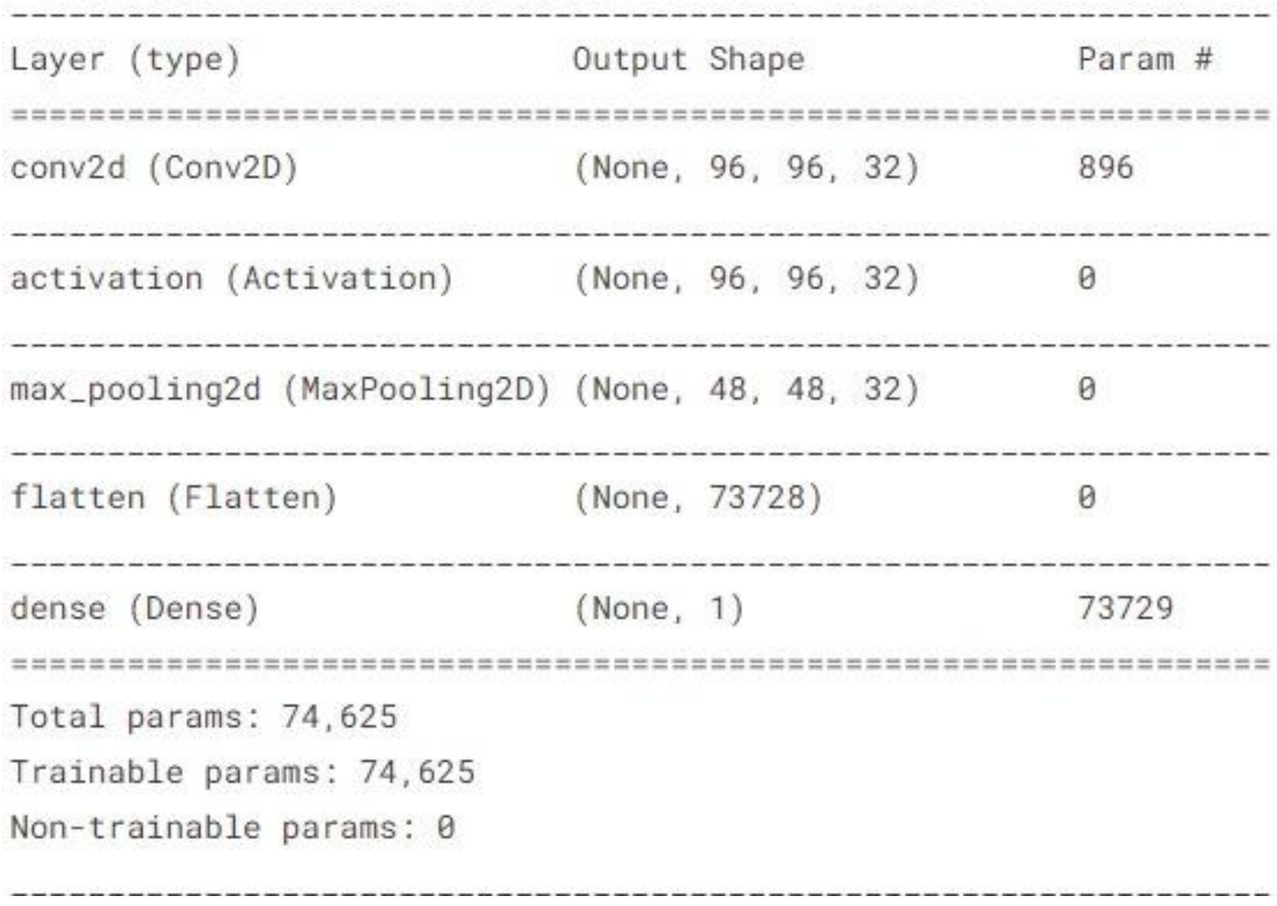
Multi-layer Perceptron Model





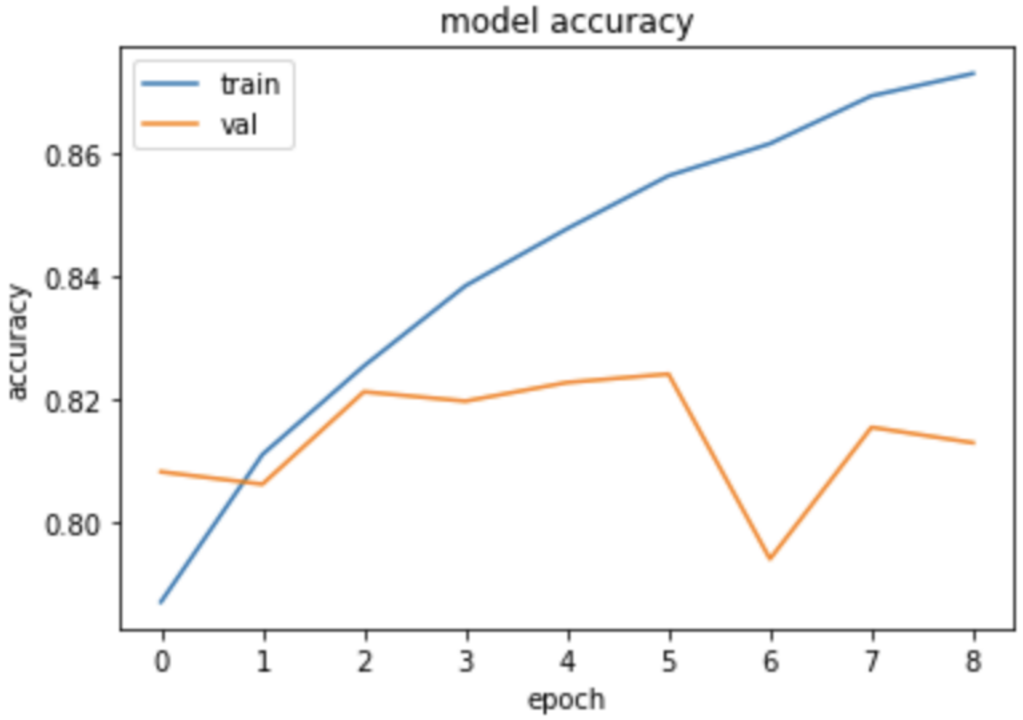
Convolution Model

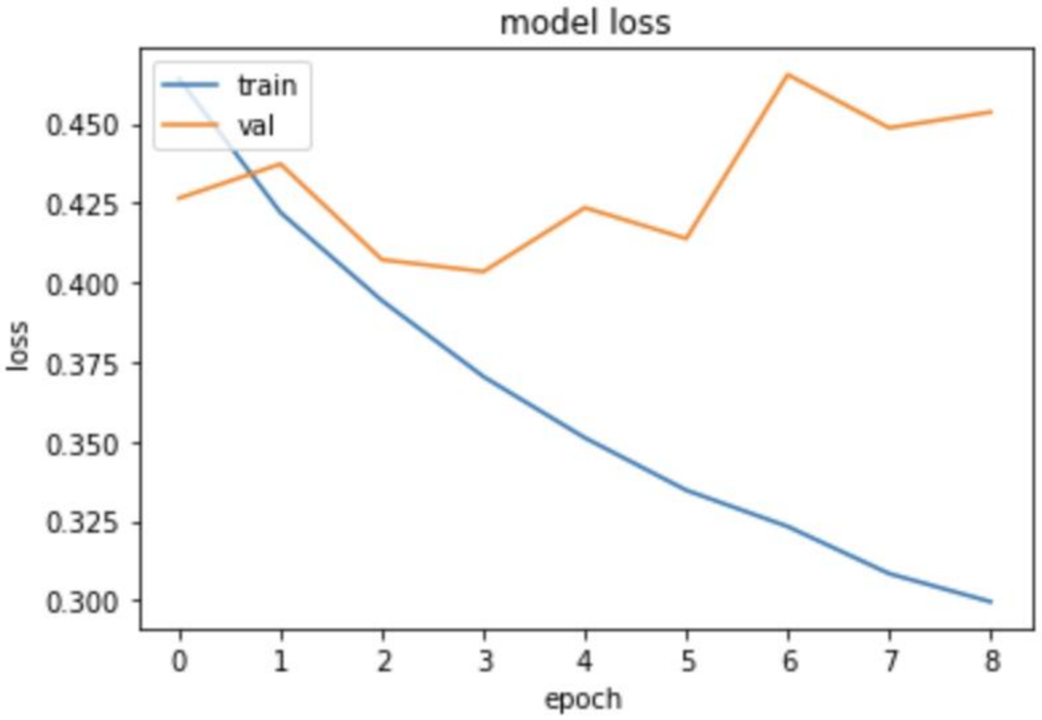
- Convolution Model: A convolution is the simple application of a filter to an input that results in activation. Repeated application of the same filter to an input results in a map of activations called a feature map, indicating the locations and strength of a detected feature in an input, such as an image. We consider a single-layer convolutional model with 32 filters, (3, 3) filter size with relu activation. After that, Max Pooling is applied with (2, 2) filter size. After flattening and adding a single output node with sigmoid activation, the output is calculated:



```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 96, 96, 32)        896

_________________________________________________________________
activation (Activation)      (None, 96, 96, 32)        0

_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 48, 48, 32)        0

_________________________________________________________________
flatten (Flatten)            (None, 73728)             0

_________________________________________________________________
dense (Dense)                (None, 1)                 73729
=================================================================
Total params: 74,625
Trainable params: 74,625
Non-trainable params: 0

_________________________________________________________________
```

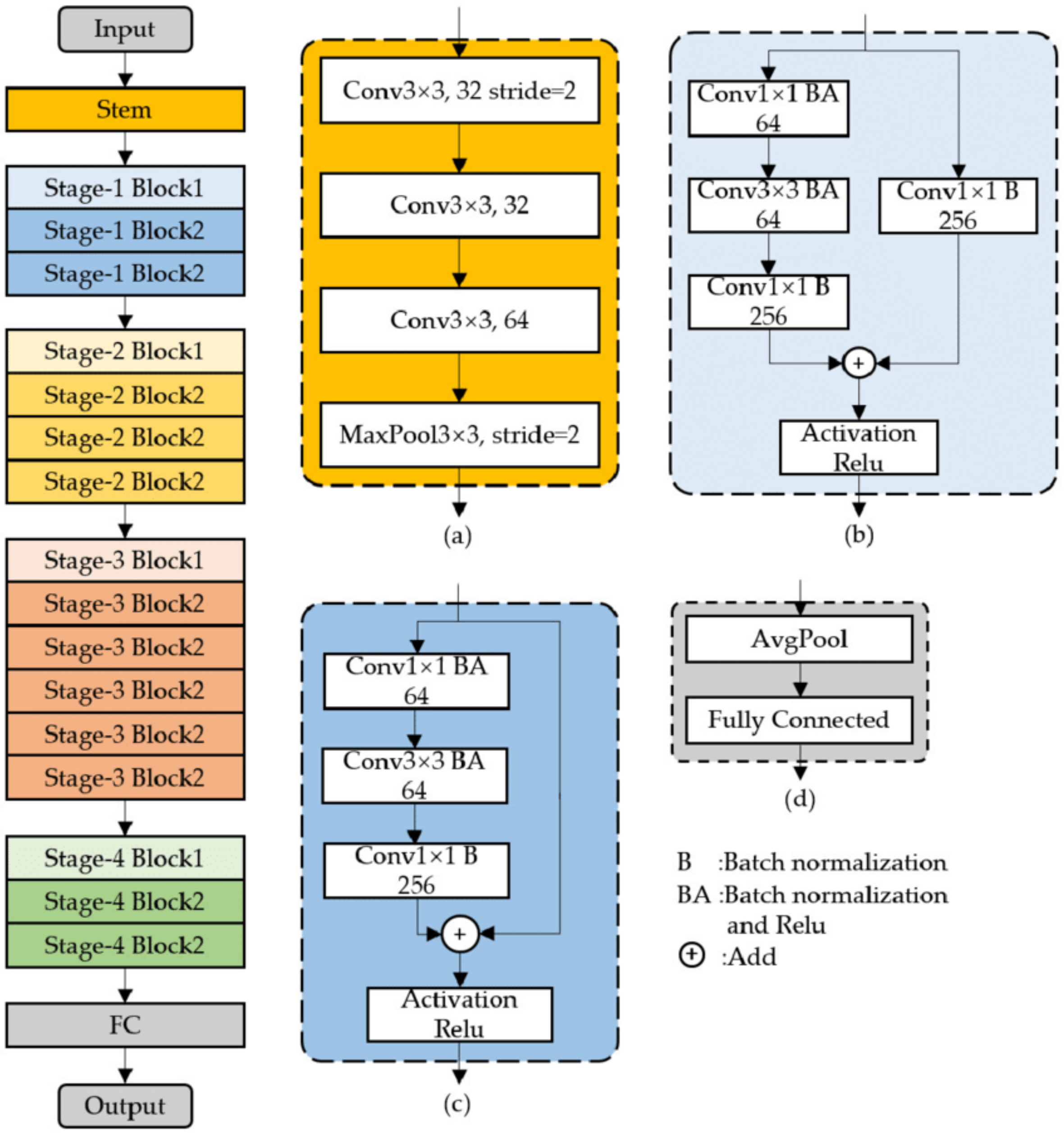
Convolution Model



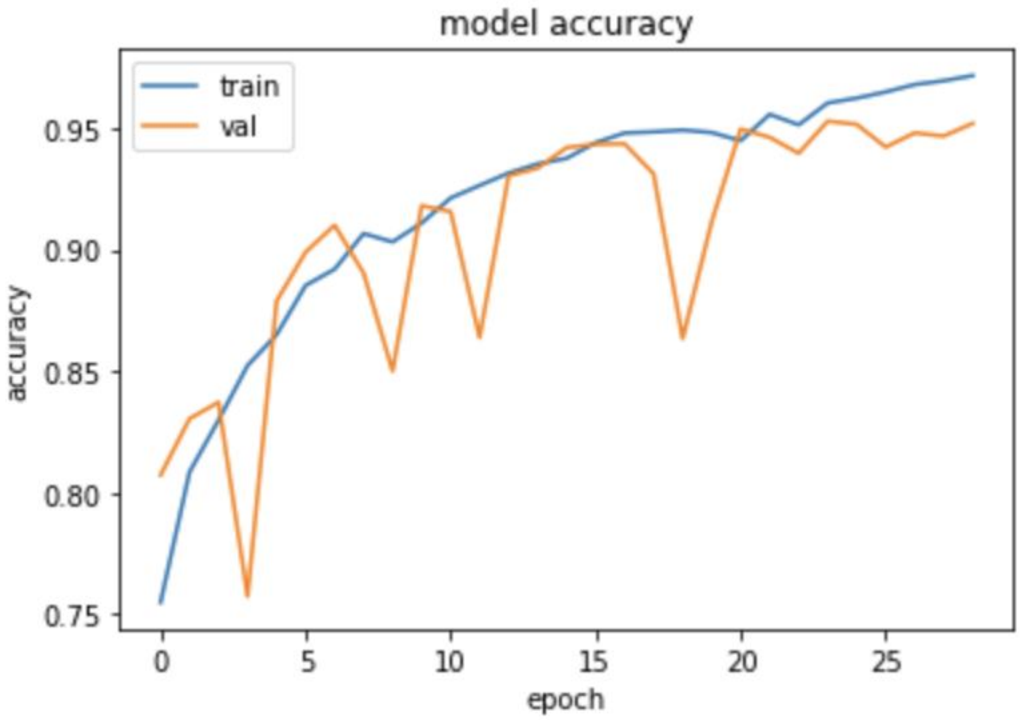


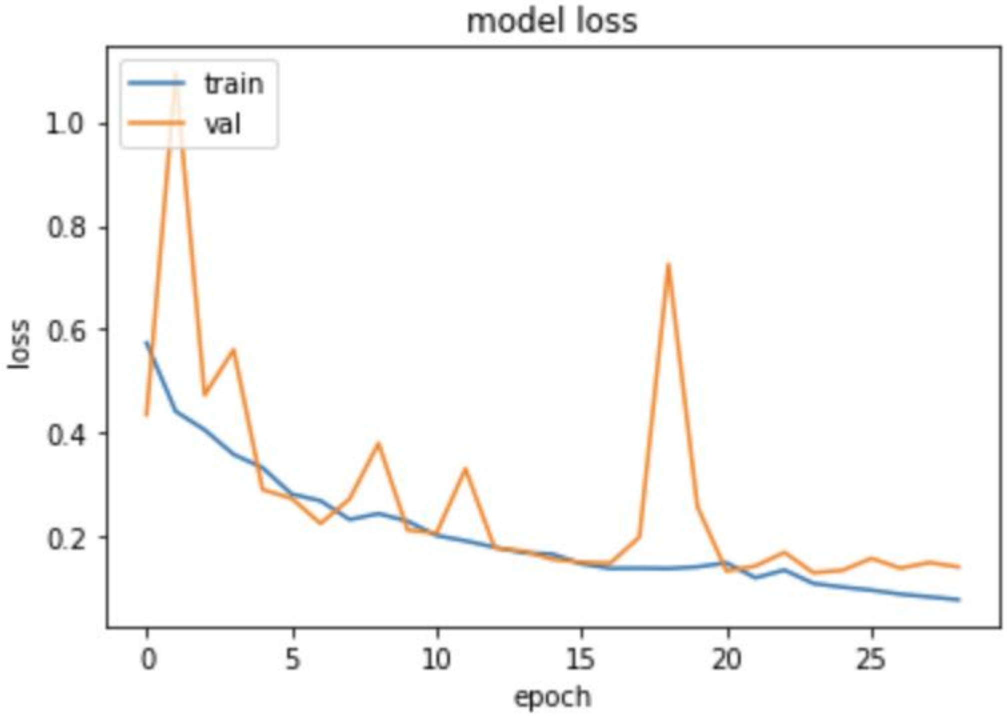
ResNet50 Model

- We utilize data augmentation with the ResNet50 model. To the output of the ResNet50 model, we add a Global Max Pooling layer and Global Average Pooling Layer. Finally, these three outputs are concatenated and attached to a single output node. A dropout layer of 0.2 is also added to the model.

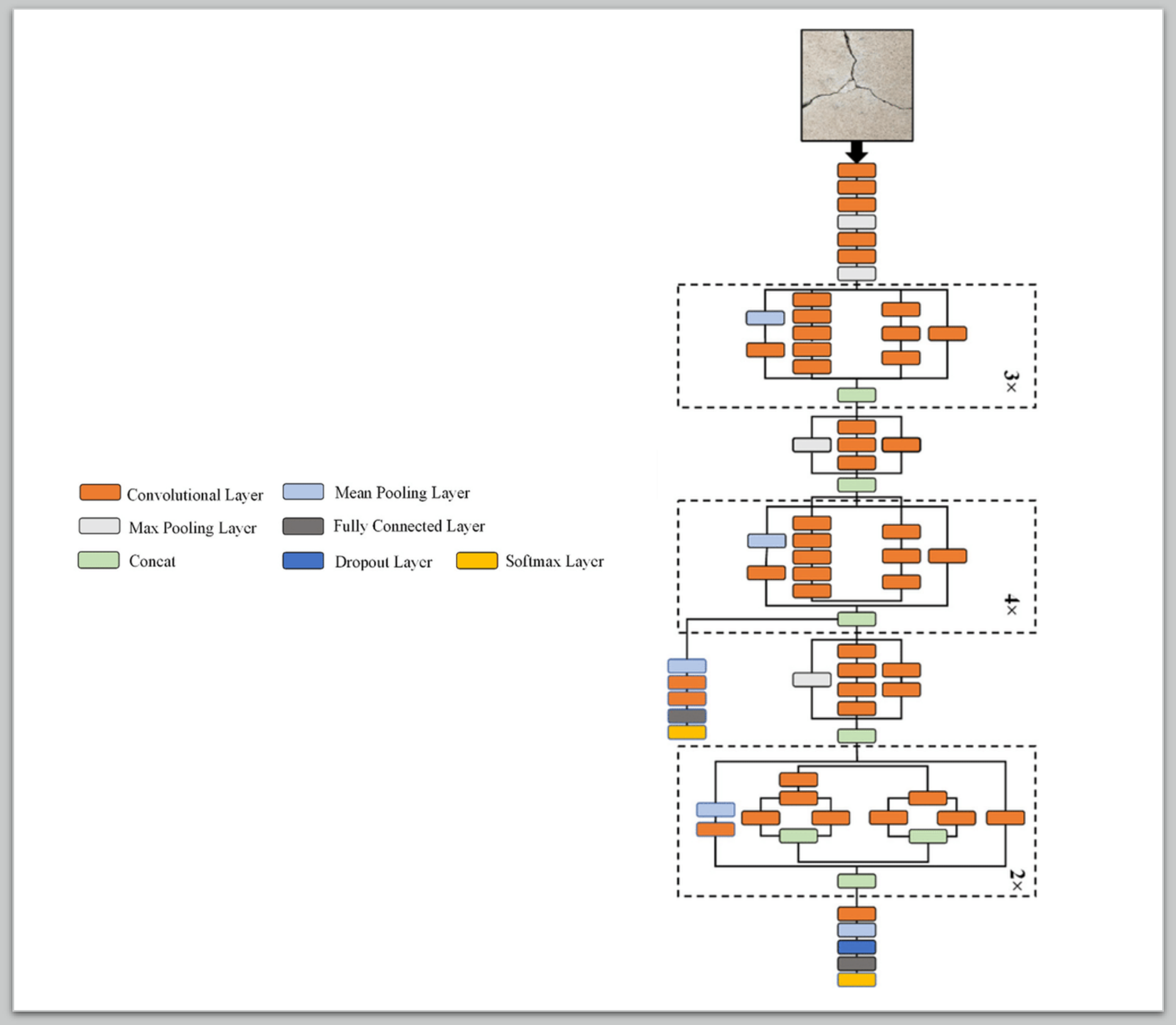


ResNet50 Model

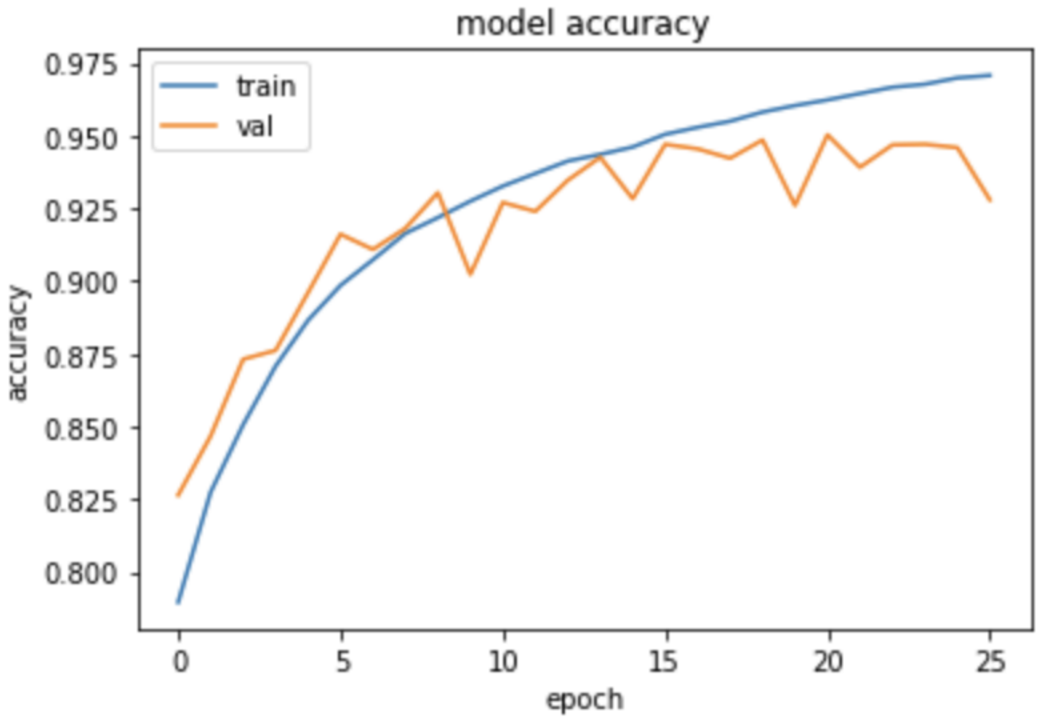


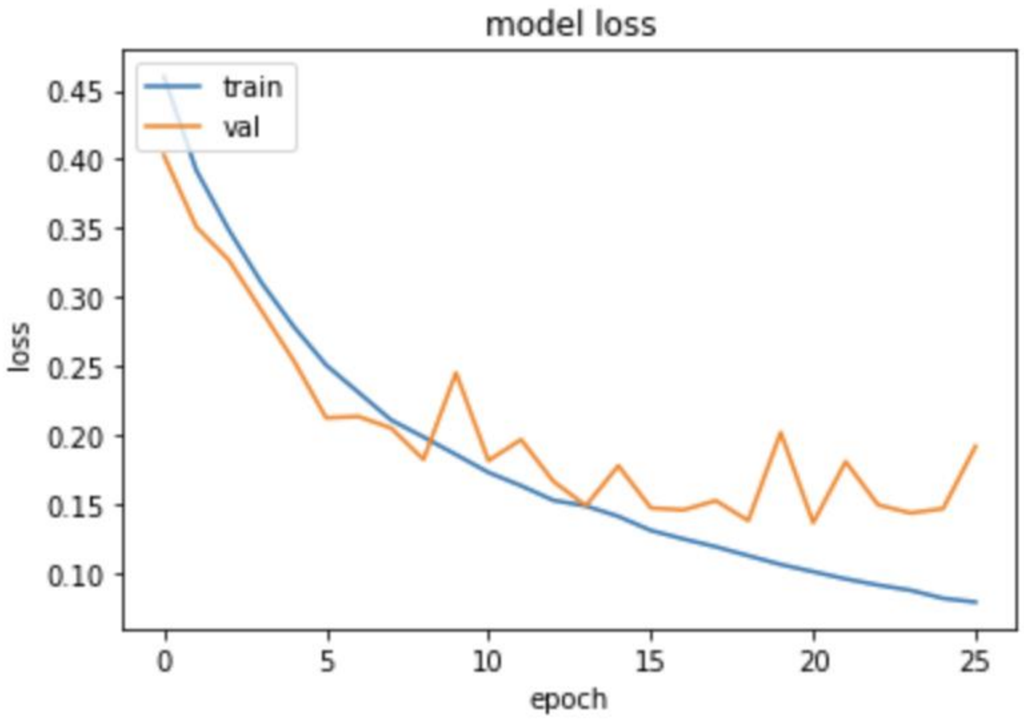


InceptionV3 Model



InceptionV3 Model





Ensemble Model

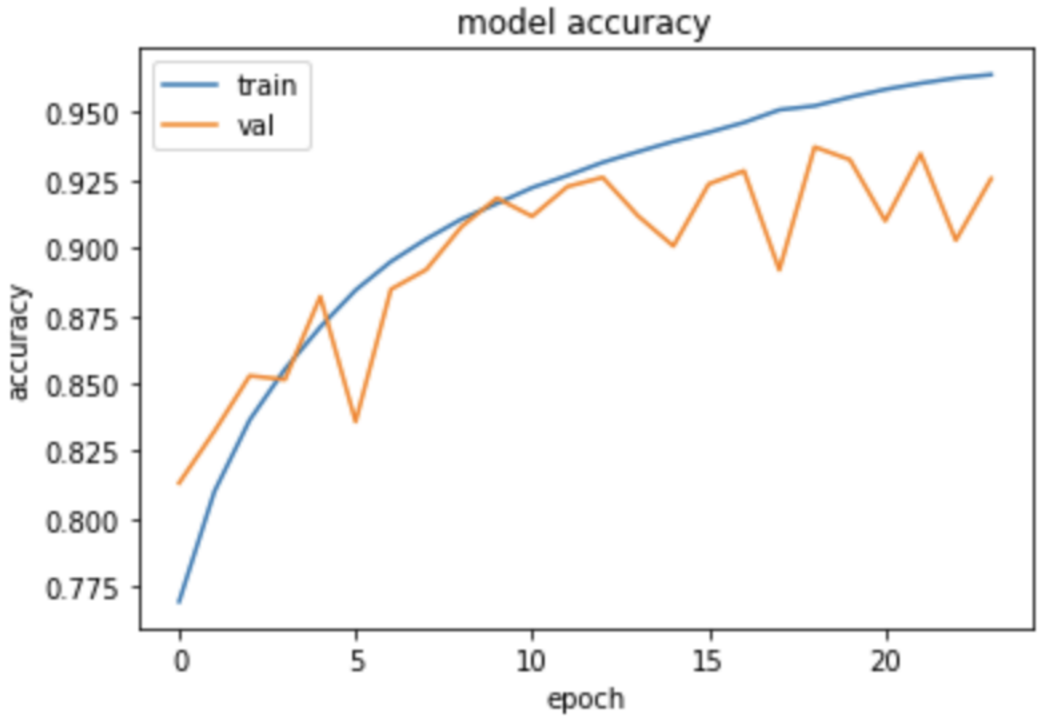
- **Majority Vote Ensemble Model**

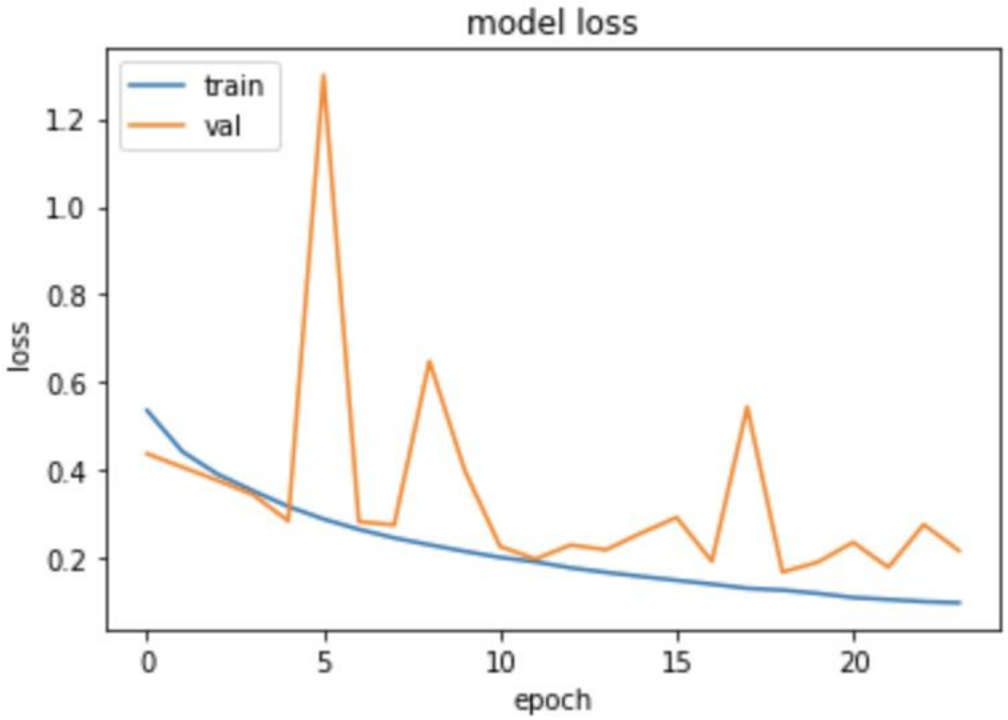
We combine the ResNet50 and Inception model with majority voting. This is a hard ensemble.

- **Concatenation Ensemble Model**

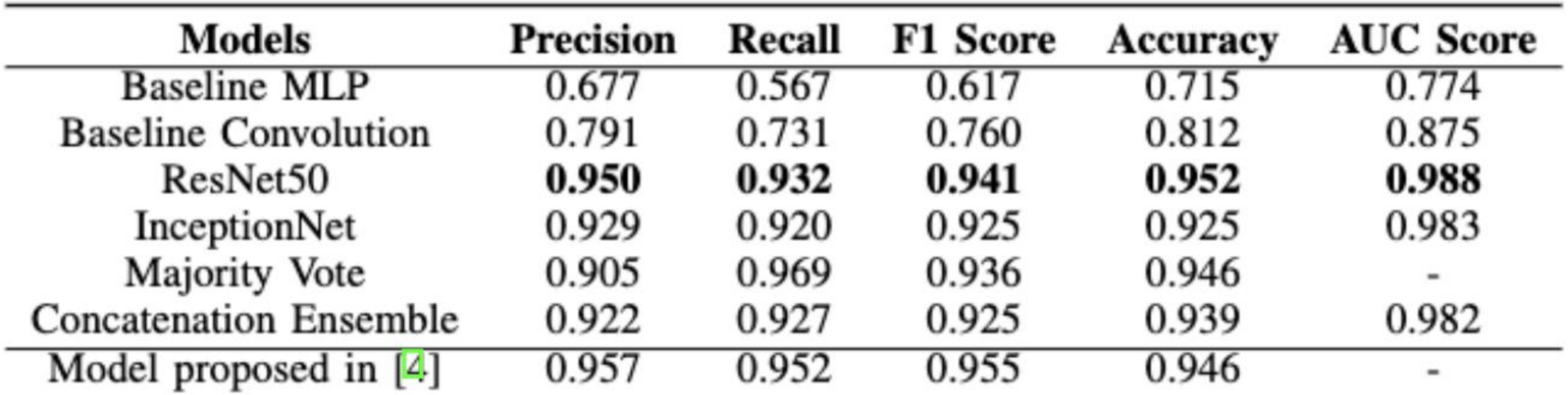
We combine the ResNet50 and Inception models using concatenation and train them together as a joint Neural Network. This is a soft ensemble.

Concatenation Ensemble Model





Model Comparison



Models	Precision	Recall	F1 Score	Accuracy	AUC Score
Baseline MLP	0.677	0.567	0.617	0.715	0.774
Baseline Convolution	0.791	0.731	0.760	0.812	0.875
ResNet50	**0.950**	**0.932**	**0.941**	**0.952**	**0.988**
InceptionNet	0.929	0.920	0.925	0.925	0.983
Majority Vote	0.905	0.969	0.936	0.946	-
Concatenation Ensemble	0.922	0.927	0.925	0.939	0.982
Model proposed in [4]	0.957	0.952	0.955	0.946	-

TABLE II: Baseline Results

Implementation Details

- The models are trained on the train set for 50 epochs and then evaluated on the test set. Adam is the optimizer for both models, with a 0.0005 learning rate for the baseline MLP and 0.001 for the baseline convolution model. For the ResNet50, the Inception model, and the ensemble model, are trained with a 0.00003 learning rate. Early stopping with patience 5 is used to avoid overfitting. For data augmentation, random vertical and horizontal flips are applied to the training data with a re-scale of 1/255.

Future Direction

In this work, we evaluate the performances of baseline models on the Patch-Cameleyon dataset. The classification performances can be further increased by using the following:

Transfer Learning: Transfer learning refers to the technique of applying already learned knowledge from one application area to another with shared weights and retraining some layers of the model.

Segmentation: In medical image applications, segmentation extracts useful information about the subject, i.e., the tissue, which can help the model to specific features and ensures the model's focus.

Conclusion

- Automatic detection of cancer tissues can help doctors identify tumor tissues and act accordingly. This paper proposes and evaluates baseline models for detecting cancer tissue in lymph nodes. The baseline convolution model achieves an accuracy of 81.2%, which is quite good for a single-layer model indicating bigger, better models equipped with other techniques such as data augmentation, transfer learning, and segmentation can achieve state-of-the-art performance on this benchmark dataset. As we advance with that idea, ResNet50 and InceptionNet models were trained with data augmentation. Finally, their majority and Concatenation ensemble were also evaluated on the dataset. Extensive experimentation showcases that the ResNet50 model trained with proper data augmentation is able to beat the state-of-the-art model. In the future, transfer learning and segmentation techniques can be utilized to create a more general and efficient model.

- ZOOM LINK: https://stevens.zoom.us/rec/share/xw_cbtr6sGqg6a6Jot9h28oz77Yudlbv84FM9wTFauNIX8JKqT3Z_raoCY2DtfXM.6jftc5tbtnLRh6WU?startTime=1671657836000

THANK **YOU**

Stevens Institute of Technology
1 Castle Point Terrace, Hoboken, NJ 07030