Deep Learning-Based Classification of Liver Cancer Histopathology Images Using Only Global Labels



**National Institute of Technology, Patna**

Computer Science and Engineering

CS5439 – Machine Learning

1. M Sampath Kumar (1806142)

2. M Vamsi (1806127)

3. Gaurav (1806146)

**Problem Statement:**

Liver cancer detection is time consuming process as it is mostly done manually. As of now it is the major causes of death all over the world. To reduce the time consumption for the tumor detection manually and accurate prediction of the tumor is to be made. The problem is to reduce the time consumption and accurate prediction of the tumor using the deep learning.

**Abstract:**

Histopathological image analysis (HIA) is a crucial step in cancer detection and done manually by the pathologists. As mentioned earlier we have to minimize time consumption and here we proposed an automatic HIA based on deep learning to improve the accuracy and efficiency of diagnosis. The model takes the “Whole Slide Image (WSI)” as an input and image preprocessing is done using OpenCV libraries. The patch level features of the WSI are extracted by “Transfer Learning” using ResNet50 and then combined with the “Multiple Instance Learning” to acquire the final image level features for classification. The proposed method can distinguish and classify liver histopathological images as abnormal or normal with high accuracy, thus providing support for the early diagnosis of liver cancer.

**Dataset:**

The dataset is extracted from the GDC portal which consists of slide images of liver tissues which are available in the SVS format. The slide images are large in size each image is of one billion pixels in size (30,000px \* 80,000px). The dataset consists of 170 slide images out of which 65 are normal and the rest are abnormal. The tissue slide images were obtained at a fixed zoom level of 0.5 µm/pixel, corresponding to slides scanned at 20X magnification. The histopathological images were randomly split into three datasets of 0.7 and 0.15, 0.15 corresponding to training, validation and testing dataset. The histopathological images cannot be directly fed to the CNN. Consequently, the WSIs were tiled into patches and reshaped to (224 × 224) pixels. All the patches from one histopathological image were included in the same dataset, ensuring that the network model used in this paper never encounters the same liver histopathological image across the training and testing processes. Most of the liver histopathological images contained 500 patches on an average; however, some histopathological images contained more than 2,000 patches also.

|  |  |  |
| --- | --- | --- |
| Slide Images | Normal | Abnormal |
| No. Of Images | 65 | 105 |

Data augmentation is done as we have limited image dataset which is imbalanced, so that we can prevent overfitting.

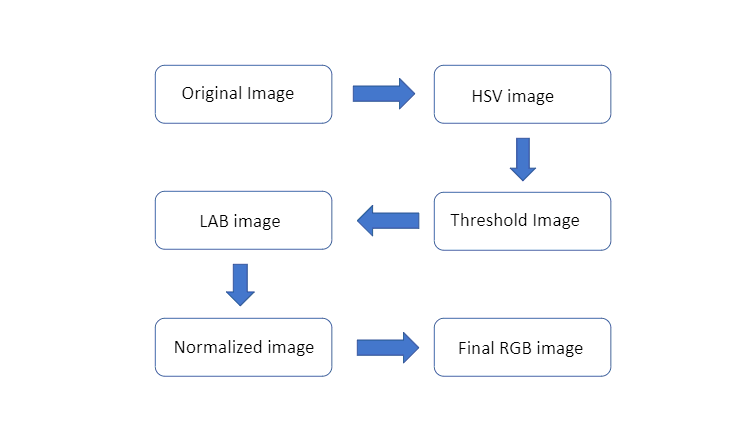
**WSI Preprocessing:**

As mentioned earlier the WSI’s are available in the SVS format initially they are converted to the PNG format and then preprocessed. First step includes the conversion of the BGR image, as the image is read by OpenCV library which reads the image in default as BGR, into HSV format. Then Otsu’s thresholding technique is applied for the hue and saturation layers of the HSV image. Thresholding technique is used for image segmentation where we change the pixels of an image which makes it easier to analyze the image. Then the masked layers H and S are merged to get the final threshold image. Then the image is converted to LAB (lightness, A channel, B channel) format and then the color normalization is applied. LAB format of an expresses colors as three values: L\* for the lightness from black (0) to white (100), a\* from green (−) to red (+), and b\* from blue (−) to yellow (+). CIELAB is designed to approximate human vision. Color normalization is done using the adaptive histogram equalization algorithm for the L layer of the LAB image. Color normalization is done to remove the stains of H&E on the tissue slide images. And then the resultant image is converted in to the RGB format.

As most of the WSI regions contains the background, the tissue regions are to be extracted. So, the image is converted into patches of size 224\*224, and the patches containing more than 50% of the tissue part are considered and given as an input to the ResNet50.

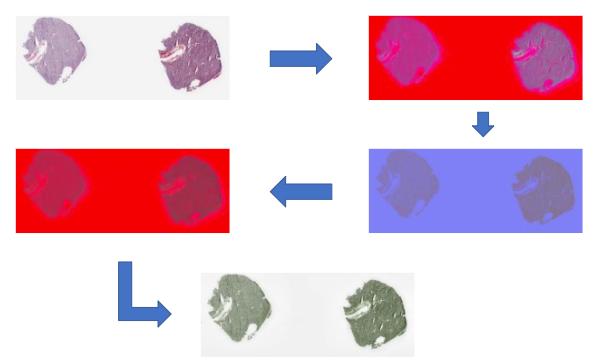
The whole image preprocessing is done using the OpenCV library which contains all the necessary functions that are required. There are also other libraries like pillow and openslide to perform preprocessing.

**Flow Chart:**



**Fig:** Flow chart of image processing of slide images

**Example:**

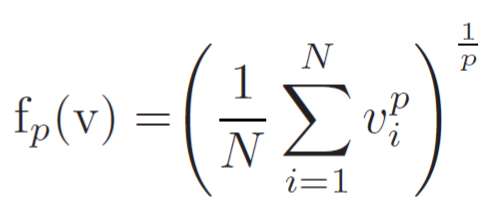


**Feature Extraction:**

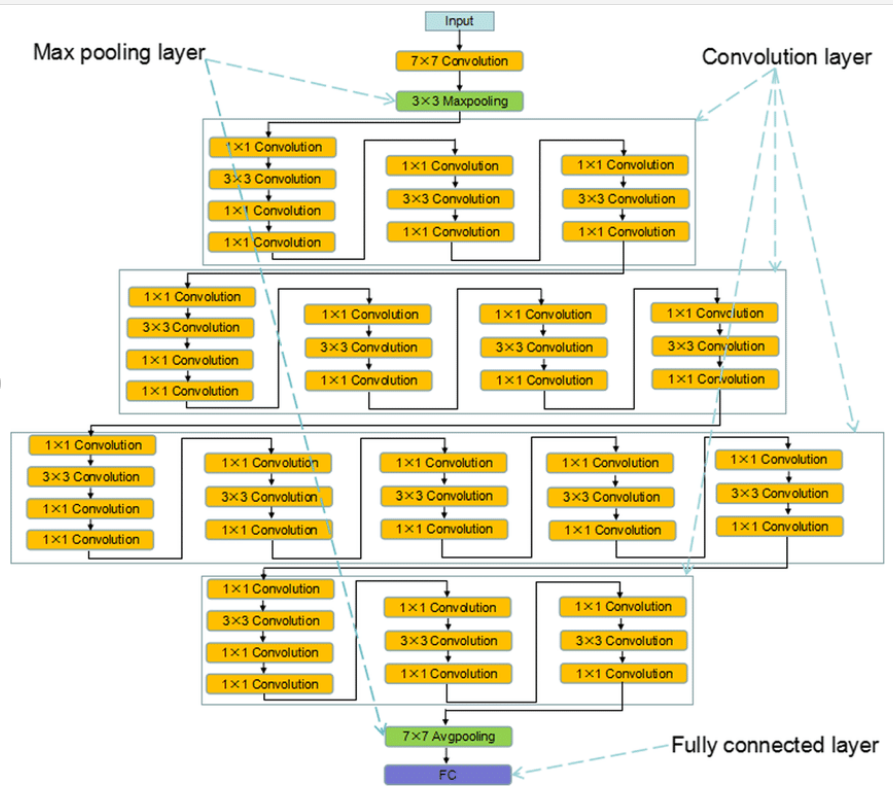
Feature Extraction aims to reduce the number of features in a dataset by creating new features from the existing ones (and then discarding the original features). These new reduced set of features should then be able to summarize most of the information contained in the original set of features.

The tumor in the slide images is detected by some features, which cannot be given manually, we must initially extract the features of the images so that it can be useful in recognizing the tumor in slide images. The features are extracted by Transfer Learning, a deep learning method where a model developed for a task is reused as the starting point for a model on a second task. For this process ResNet50 is taken. ResNet50 is convolutional neural network and that is 50 layers. Here we load a pretrained version of the network trained on more than a million images from the imageNet database. The network can classify images into thousand categories. Here we don’t need the classification, so we remove the fully connected layer and make the trainable parameters to false, so that we can get a patch vectors that contains the necessary information to be classified. All the patches of a slide image are given as an input to the ResNet50 and patch-level feature vectors are obtained.

The image level feature vector is acquired by aggregating the patch level feature vectors of slide image obtained by transfer learning using the multiple instance learning which is used for the classification. The liver cancer classification accuracy can be increased by this feature aggregation model. The aggregation is done by p-norm pooling where we use p=3.



**Architecture of ResNet50:**

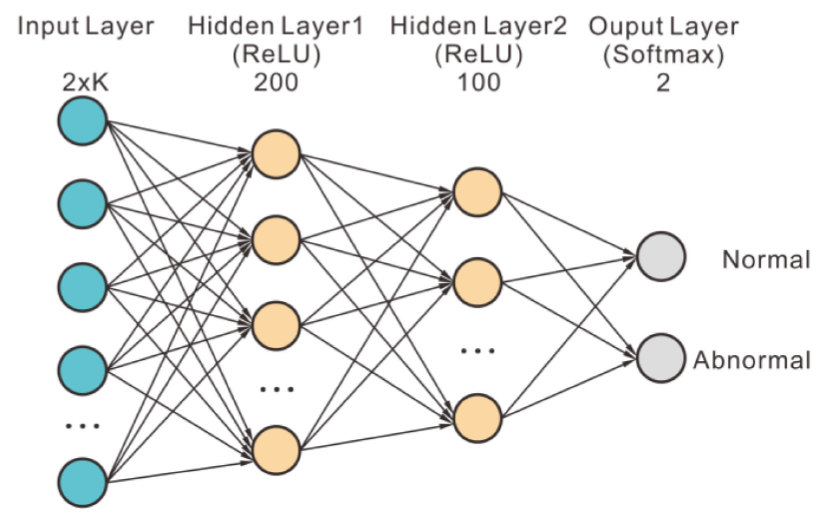


**Feature Selection:**

To obtain more discriminative features and eliminate redundant or unrelated features for classification, image-level feature descriptors are acquired via the MIL technique. However, rather than applying the image-level features obtained directly from the patch-level features, we used a small number of selected and representative features to classify liver cancer pathology images. It is easier and less costly to characterize histopathological images of liver cancer. For the label of histopathological image rather than the labels of the patches is given, the representative feature is selected from the aggregated features to distinguish the abnormal and normal regions of liver tissue. The final image level feature vector is of the size 2048\*1. Hence from the vector top k and the bottom k values are chosen. Here k=104.

**Neural Network Architecture:**

To better study the interactions between the top and bottom values, we adopted multilayer perceptron (shown in Fig) as the final classifier, although a CNN by itself is a good classifier. In our experiments, the multilayer perceptron consisted of two fully connected layers with 200 and 100 neurons, respectively and chose ReLU as the activation function. To perform the binary classification of liver cancer histopathological images, a layer with two neurons was set as the output layer with softmax as an activation function, to predict the input liver histopathological images as normal or abnormal.



**Training:**

L2 regularization and dropouts were used both at rates of 0.4 in the first fully connected layers of multi-layer perceptron. We adopted Adam optimizer with batch size of 24 and a learning rate of 0.001 to optimize binary cross entropy.

We trained an ensemble model of 10 models with different weights for achieving better classification performance. At last, the prediction of a histopathological image is the mean of predictions made by these 10 models.

**Performance measures:**

The validation set was used to calculate the classification performance of the model. By calculating the accuracy, we determined the true positives, false positives, false negatives and true negatives and evaluated the classification performance of our method using these performance measures. From these values precision and recall values are calculated. We used the Receiver Operating Characteristics (ROC curve) which is usually preferers to evaluate results obtained by medical image processing.

The F1-score is calculated from the recall and precision values obtained above, which ranges from 0 to 1 and rewards algorithms that maximize both the precision and recall simultaneously rather than favoring one over the other was also computed for the binary classification of liver cancer histopathological images using the Python library sklearn library.

**Effect of K value selection:**

The value of K ranges from 10 to 300. As we increase the value of K from 10 to 104 the accuracy of the model is increased gradually and from later on as the value of K increases the accuracy of the model got decreases eventually.

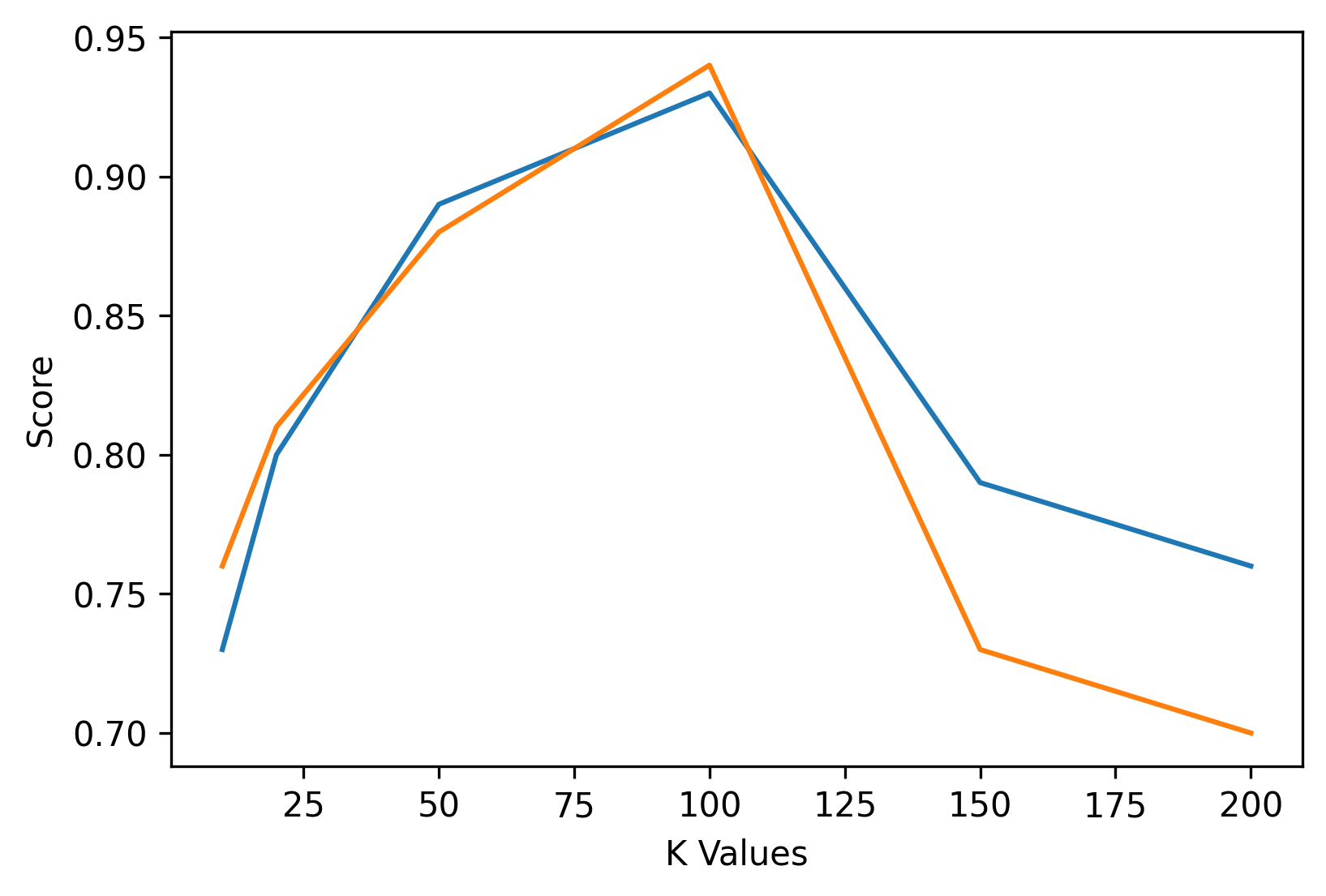


Fig: Blue line indicates the accuracy and orange line indicates the F1 score.

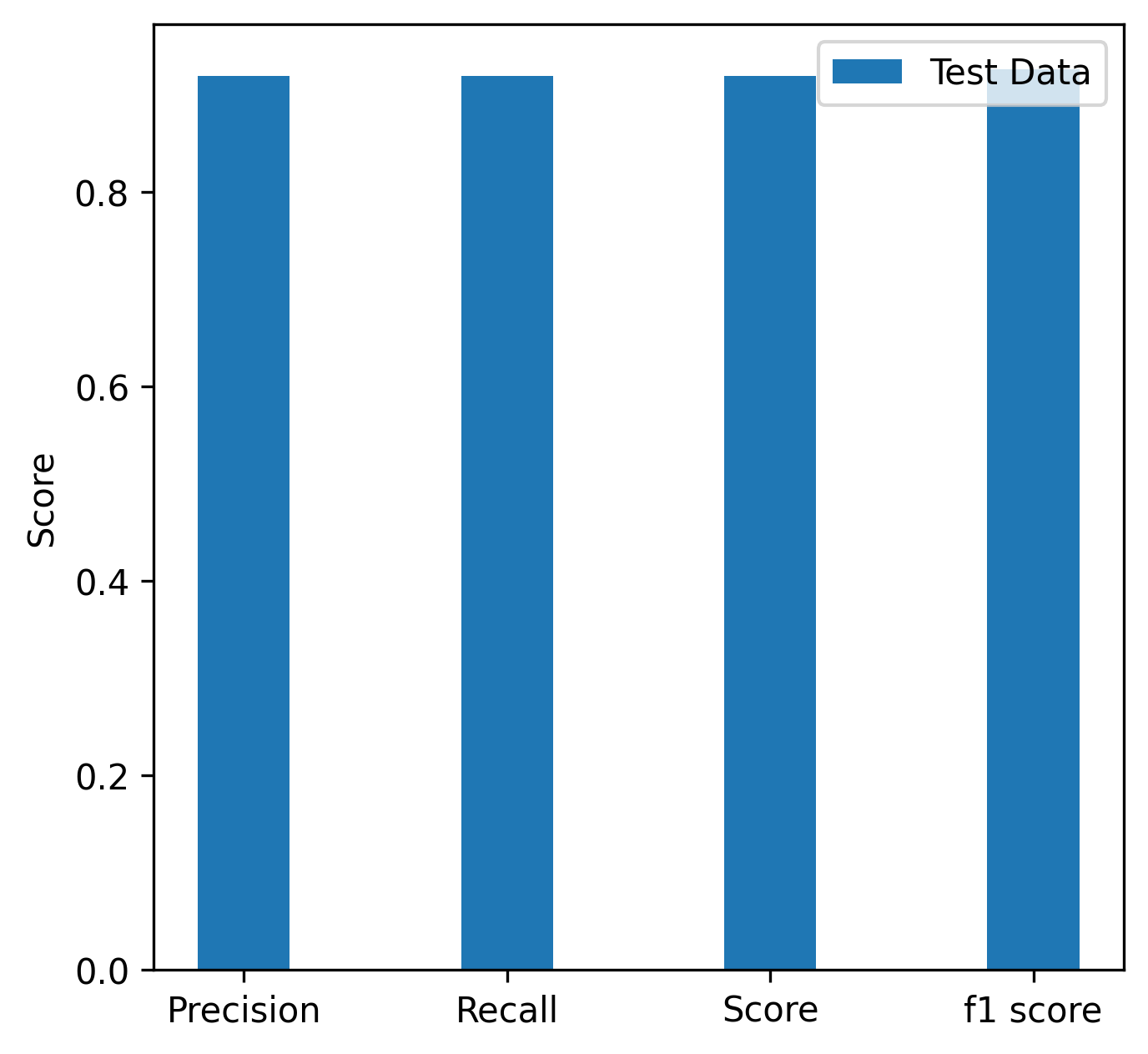
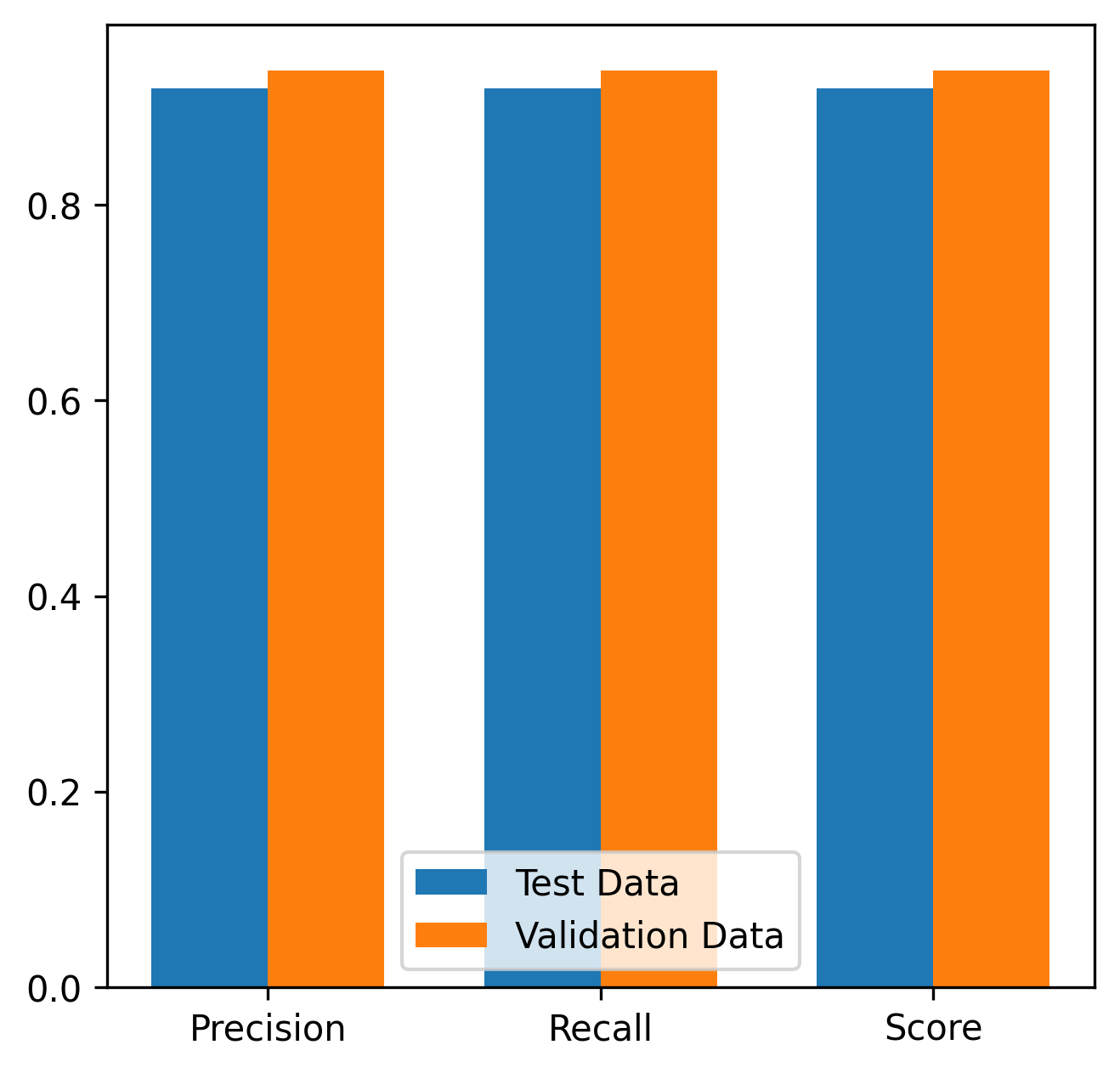


Fig 1. Blue bar indicates test data scores and orange line indicates

Scores for the validation data.

**Confusion Matrices:**

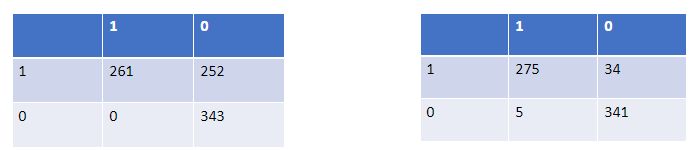


Table 1: Confusion matrix for test data.

Table 2: Confusion matrix for validation data.

**Results:**

Final accuracy of the model is 94.8%.

Precision is 0.889

Recall is 0.982

F1 score is 0.933

--THE END--