**Classification of Colorectal Cancer Tissue using Ensemble Stacking**

A thesis submitted in partial fulfillment of requirements for the award of the degree of

**Bachelor of Technology**

in

**Computer Science and Engineering**

by

Soumya Bera(18101104033)

Aryashree Mazumder (18101104050)

Arnab Mandal (18101104046)

Yugal Tamang (19101104088)

*Under the supervision*

of

Prof. Animesh Hazra

Assistant Professor, Department of Computer Science and Engineering



**Department of Computer Science and Engineering**

**JALPAIGURI GOVERNMENT ENGINEERING COLLEGE**

**June, 2022**

**BONAFIDE CERTIFICATE**

This is to certify that the project titled ***Classification of Colorectal Cancer Tissue using ensemble stacking*** is a bonafide record of the work done by

**Soumya Bera (18101104033)**

**Aryashree Mazumder (18101104050)**

**Arnab Mandal (18101104046)**

**Yugal Tamang (19101104088)**

in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering** of the **JALPAIGURI GOVERNMENT ENGINEERING COLLEGE** during the year 2021-2022.

—-------------------------------- —-------—---------------------------

Prof. Animesh Hazra Dr. Subhas Barman

Dept. of Computer Science and Engineering Dept. of Computer Science and Engineering

Project Supervisor Head of the Department

Project Viva-voce held on \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Acknowledgement**

The project work summarized in this report, explores the topic named: *Classification of Colorectal Cancer Tissue using Ensemble Stacking*  (A classification model for classifying 9 types of colorectal cancer from histology images).

We hereby thank our project supervisor Prof. Animesh Hazra. This documentation and paper writing would never have been more educative and efficient without the constant help and guidance of our project guide Prof. Animesh Hazra.

We would like to thank him for giving me the right guidance and encouraging me to complete the project within time. We also express my deepest and sincere gratitude to all our teachers for their kind comments and advice for my project. Last but not the least, we would also like to express my heartiest gratitude to our HoD, Prof. Subhas Barman and other faculties for their encouragement and kind suggestion.

Signature of students:

**Abstract**

Advancement in digital pathology has enabled deep learning based computer vision techniques for automated diagnosis and prognosis. The early diagnosis and prognosis of a cancer type have become a necessity in cancer research, as it can facilitate the subsequent clinical management of patients. Colorectal cancer (CRC) is the third most common form of cancer and is about 10% of all cases in the world. It is clinically important to classify and make an objective evaluation of colorectal cancer histological images. To test classification performance, current methodologies mainly rely on the use of various combinations of textual features and classifiers or transfer learning to classify different organizational kinds. However, classification remains difficult since histological images comprise a variety of tissue types and properties. In this study, I have proposed the best classification methodology based on the selected optimizer and modified the parameters of CNN method by adding dense layers. As a result, based on the use of artificial intelligence to classify colorectal cancer tissue, this method is believed to offer a lot of promise in terms of assisting clinicians in making clinical diagnoses and reducing the number of different evaluations.

*Keywords*: convolutional neural network; machine learning; deep learning; colorectal cancer.

**TABLE OF CONTENTS**

|  | **Title Page No.** |
| --- | --- |
|  | Abstract …………………………………………………………………… 4 |
|  | 1. Introduction ……………………………………………………………… 6-7 |
|  | 2. Related Work …………………………………………………………… 7-9 |
|  | 3. Proposed Methodology  3.1. Dataset Description ………………………………………………. 10  3.2. Classifications of Cancer Cells …………………………………… 10-11  3.3. Architecture Description …………………………………………… 11-15 |
|  | 4. Results and Discussion  4.1. Accuracy ………………………………………………………….. 15-18  4.2. Experimental Setup ……………………………………………… 19  4.3. Experimental Result ……………………………………………… 19 |
|  | Conclusion ………………………………………………………………… 19 |
|  | References ……………………………………………………………….. 20 |

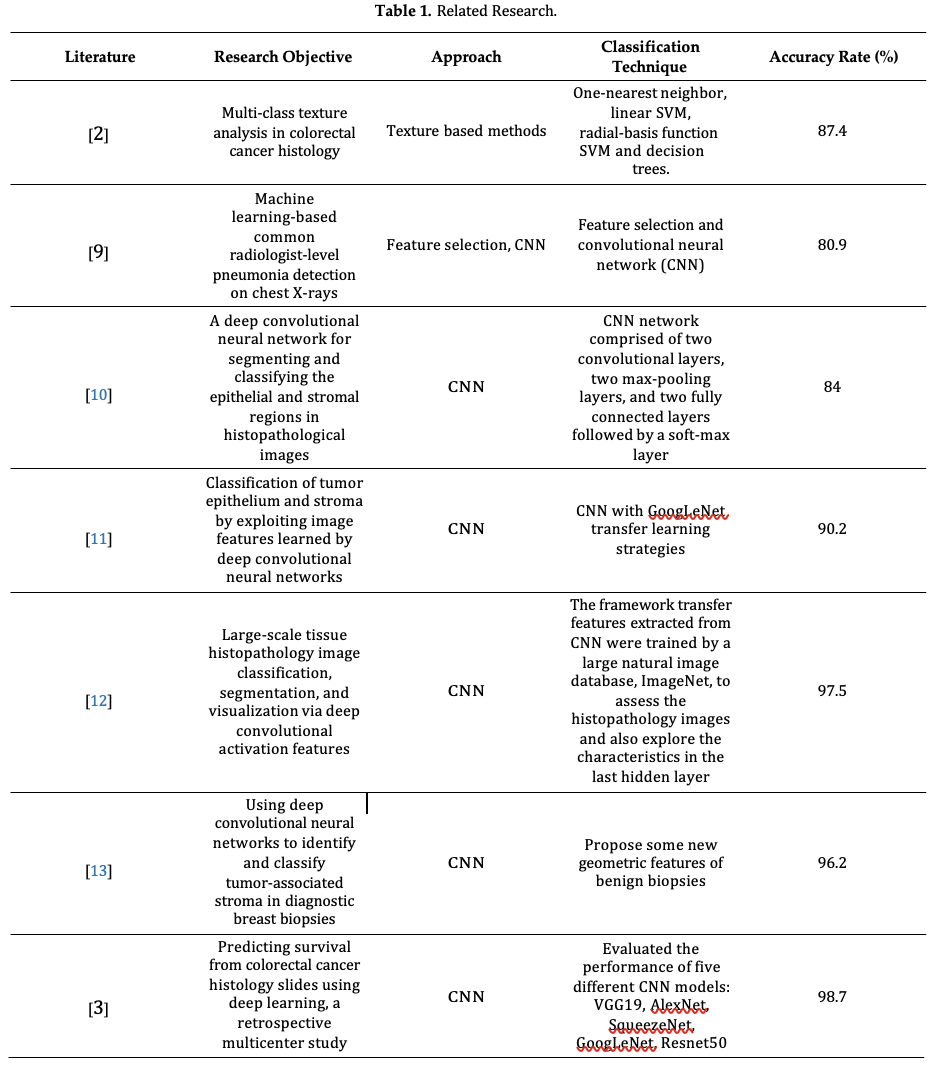
**1. Introduction**

Colorectal cancer is the third most prevalent malignancy diagnosed worldwide, and it is also the second largest cause of cancer death. Colorectal cancer affects about 145,600 people in the United States each year. These malignancies start on the colon's inner surface, or mucosal layer, and can spread to other organs by penetrating deeper layers of the colon. The sickness is lethal if left untreated. Flexible endoscopy, which involves visual inspection of the mucosal lining of the colon and rectum using an optical camera attached on the endoscope, is currently used for endoluminal screening or surveillance for colorectal malignancy. Biopsies of abnormal-looking regions are then taken for histologic examination. There are various flaws in the current standard of care for endoscopic screening. First, to guide biopsy site selection, this procedure depends on ocular detection of aberrant tissue. Early cancers are frequently missed because tiny or sessile lesions are difficult to identify with the naked eye. Second, visual endoscopy can only identify changes on the surface of the gut wall; while this is sufficient for screening, it severely limits the usefulness of endoscopic surveillance after some malignancies have been treated. Rectal tumours, in instance, can vanish completely from the mucosal surface while still containing nests of tumour cells beneath the mucosal barrier. Better imaging modalities and approaches are required to improve colorectal cancer screening and monitoring. As a result, the clinical importance of universal automatic classification of CRC pathological tissue slide images for fair appraisal is significant. Pathology slides include a tremendous quantity of data, which has been quantified over time using digital pathology and traditional machine learning techniques. Machine learning algorithms have been used in the past to judge cell categorization in histology slides of tumour tissue. Artificial intelligence-assisted classification of histopathology pictures not only enhances classification accuracy and efficiency, but also allows doctors to make faster clinical treatment decisions. The majority of the proposed experimental methods rely on manual feature labeling, which is a major flaw in traditional textual analysis methods. In recent years Deep learning has turned out to be the technique that is considered an advancement of machine learning because it employs numerous layers of neural networks to learn and extract higher-level information in order to eliminate the need for human interaction in image classification. In the field of deep learning, where a neural network may include dozens or hundreds of layers to learn including images with varied features, convolutional neural networks (CNN) recently showed effective results in image classification. To create advanced features, a convolutional layer consists of a small-sized kernel with applied weights to the inputs and guides them through an activation function as the output. When compared to a standard neural network, the key advantage of using CNN is that it reduces model parameters for more accurate results. With the features and advantages of CNN over other techniques this paper comprises the best performing CNN model with the highest classification accuracy rate.

**2. Related Work**

Some of the prior studies in relation to the automatic classification of histopathological images will be described and discussed in this section with a further explanation of how deep learning works. This will be followed by a presentation of the proposed method to conduct the current research.

Digital technology is currently used extensively to classify medical images, as evidenced by the results of several methods of histopathological image classification shown in Table 1. Kather [2] used a range of textual descriptors to analyze a multi-class problem of tumor epithelium and simple stroma in 5000 histological images. He proposed four classification methods: (1) the k-nearest neighbors algorithm (k-NN), (2) employ an SVM decision function in an attempt to classify all categories, (3) assemble decision tree models using the RUSBoost method, and (4) use a 10-fold cross validation to train the classifiers, without an explicit stratification approach. The results indicated that SVM was the best classification method, which achieved 87.4% accuracy over eight classes. Lately, the classification of tumor types has been found to be more accurate using the CNN classification method. Tsai [9] applied the CNN architecture of a deep learning technique to detect pneumonia from Chest X-rays and achieved an accuracy rate within 82.1% by using feature selection and the CNN.



Xu [10] used the CNN model and feature extraction approaches to compare two datasets of breast cancer and colorectal cancer. The two types of tissues in the histological images were epithelial (EP) and stromal (ST). He used automated segmentation or the classification of color features, which included intensive pixels in different color spaces, and analyzed the tumor microenvironment. In his study, Du [11] proposed that learning the basic features of CNN methods outperformed handcrafted features, and automatically distinguished the epithelial and stromal regions in the breast. In addition, he found that colorectal tumors could be distinguished from tumor tissue using a network architecture layer approach with results that were 84% accurate. Transfer learning is a methodology that consists of deep learning techniques to distinguish the features of leverage images. Du discussed the use of transfer learning methods to accurately distinguish breast or ovarian cancer from histological images and CNN for fine tuning the feature extractor of images. Additionally, he discussed how to distinguish high-level and low-level features inside the neural network. A deep neural network may have multiple layers, the first of which will learn the low-level features and then the more they progress toward the output layer, the more the layers will learn the high-level features. Du [11] also used a transfer learning approach with GoogLeNet and achieved 90.2% accuracy, suggesting the feasibility of using it to classify the tumor stroma ratio (TSR). Xu et al. [12] improved the activation features of the AlexNet model and proposed the characteristic of visualizing the neurons in the last hidden layer to classify and segment them. Trained by ImageNet, the framework successfully transferred the features extracted from the network into little histopathology images features for training and visualization and a test accuracy rate of 97.5% was reported. Bejnordi et al. [13] proposed deep convolutional neural networks with some new geometric features, and trained the algorithm networks to classify stroma images, including stroma, fat tissues, other situ lesions and to predict the stroma regions. Bejnordi analyzed the stroma between surrounding invasive cancer and situ lesions and achieved a 96.2% accuracy. Additionally, Kather [3] replaced the classification layer and the best accuracy rate was 98.7% with VGG19.

**3. Proposed Methodology**

After considering and testing over several pre-trained CNN architectures using pytorch we found out that the ResNet18 and EfficientNetB0 together when stacked using ensemble technique and applying Adam optimizer gives better accuracy to classify the colorectal cancer tissue.

**3.1. Dataset Description**

The open histological datasets of nine tissue classes from CRC-VAL-HE-7K were used to train, validate and test the models in this study. The datasets contain 7180 images, Kather et al. [3] created these images, which include 86 tissue slides stained with hematoxylin and eosin (H&E). The labels for the available data's histological images were taken from the NCT-UMM website. All of the images were 224 X 224 pixels (112 X 112 µm) in size, and they were fed into the model network in order for training, validation, and testing.

**3.2. Classifications of Cancer cells**

In the above mentioned dataset the nine classes are categorized as following (Figure 1)

(a) ADI: adipose tissue is mainly composed of adipocytes.

(b) BACK: histological image background.

(c) DEB: debris is widely used in histopathology and diagnoses.

(d) LYM: lymphocytes are the main type of cells found in the lymphatic system.

(e) MUC: mucus is produced by many tissues in the body, and acts as a protective force.

(f) MUS: smooth muscle.

(g) NORM TISSUE: tissues of colon mucosa.

(h) STR: stroma tissues associated with cancer.

(i) TUM: epithelium tissues of adenocarcinoma.

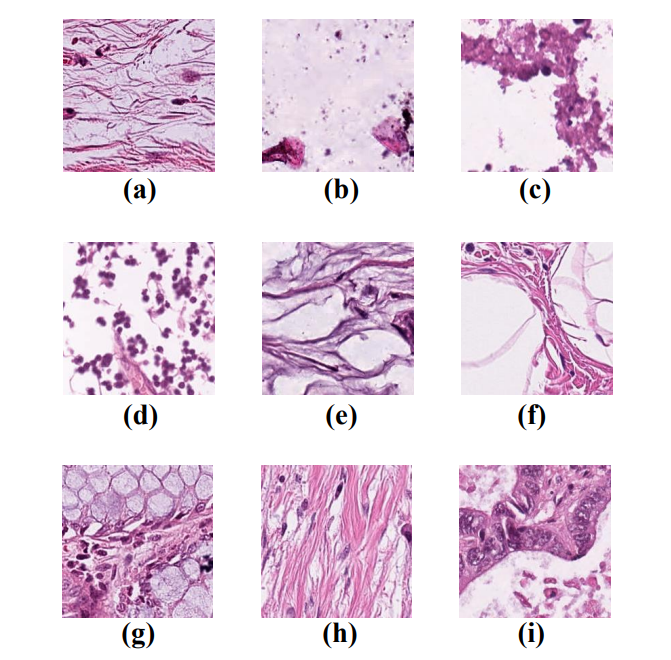


Figure 1. Example images of the nine tissue classes represented in the CRC-VAL-HE-7K dataset.

**3.3. Architecture Description**

The **ResNet18** architecture contains the following element (Figure 1):

* First there is a convolution layer with 7x7 kernel size and stride 2.
* After this there is the beginning of the skip connection. The input from here is added to the output that is achieved by 3x3 max pool layer and two convolution layers with kernel size 3x3, 64 kernels each. This was the **first residual block**.
* Then from here, the output of this residual block is added to the output of two convolution layers with kernel size 3x3 and 128 such filters. This constituted the **second residual block.**
* Then the **third residual block** involves the output of the second block through skip connection and the output of two convolution layers with filter size 3x3 and 256 such filters.
* The **fourth and final residual block** involves output of the third block through skip connections and output of two convolution layers with the same filter size of 3x3 and 512 such filters.
* Finally, average pooling is applied on the output of the final residual block and the received feature map is given to the fully connected layers followed by a softmax function to receive the final output.
* The output of each layer is shown in the diagram and input is changed in the skip connections according to that.

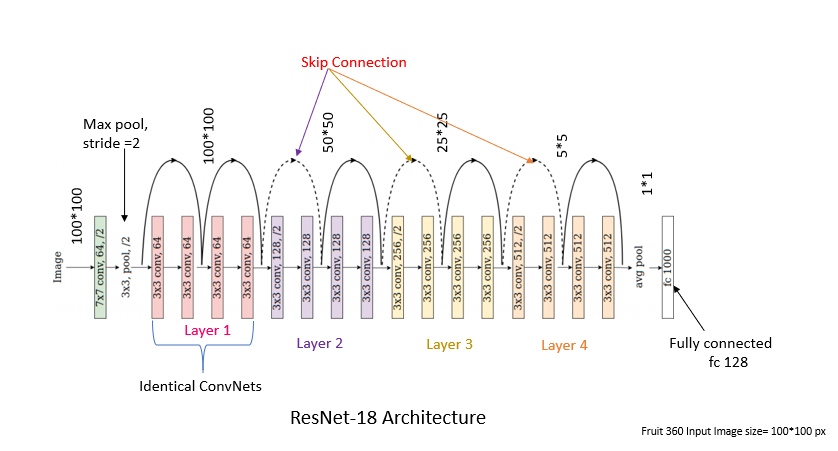


Figure 1: Architecture of ResNet50 model.

The **EfficientNetB0** architecture contains the following element (Figure 2):

* First there is a convolution layer with 3x3 kernel size.
* After this there is one MBconv1 layer with kernel size 3X3.
* Then two layers of MBconv6 with kernel size 3X3.
* Then two layers of MBconv6 with kernel size 5X5.
* Then three layers of MBconv6 with kernel size 3X3.
* Then three layers of MBconv6 with kernel size 5X5.
* Then four layers of MBconv6 with kernel size 5X5.
* Then one MBconv6 layer with kernel size 3X3.
* Finally, average pooling is applied on the output of the final layer and the received feature map is given to the fully connected layers followed by a softmax function to receive the final output.
* The output of each layer is shown in the diagram.

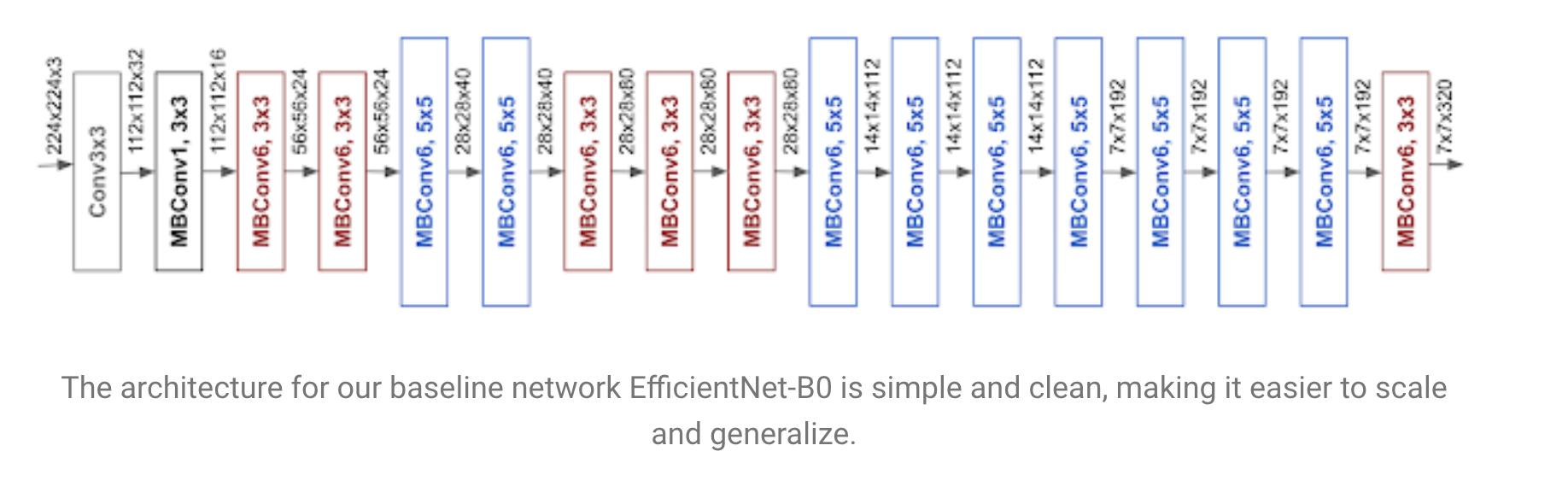
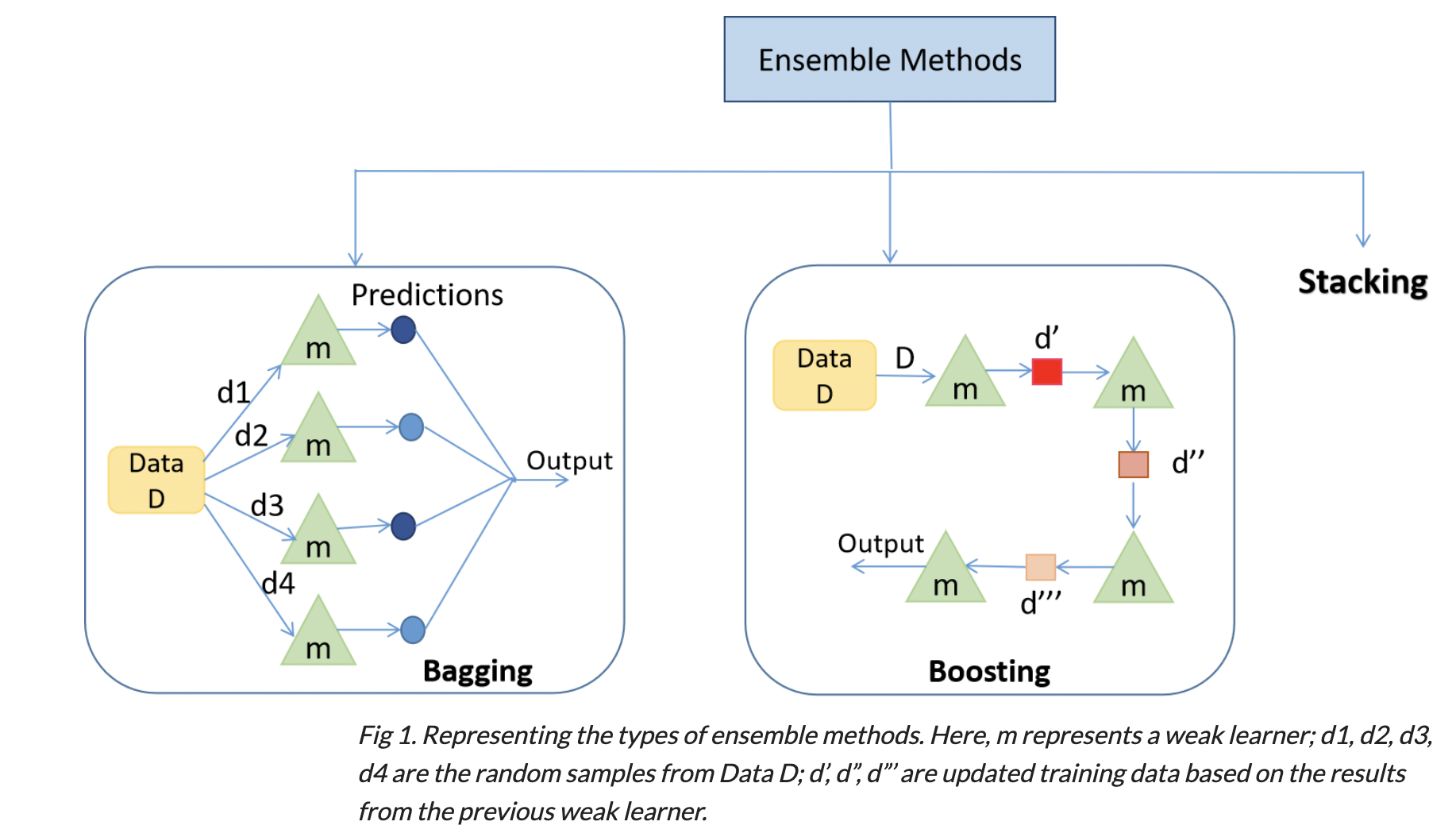


Figure 2: Architecture of EfficientNetB0 model.

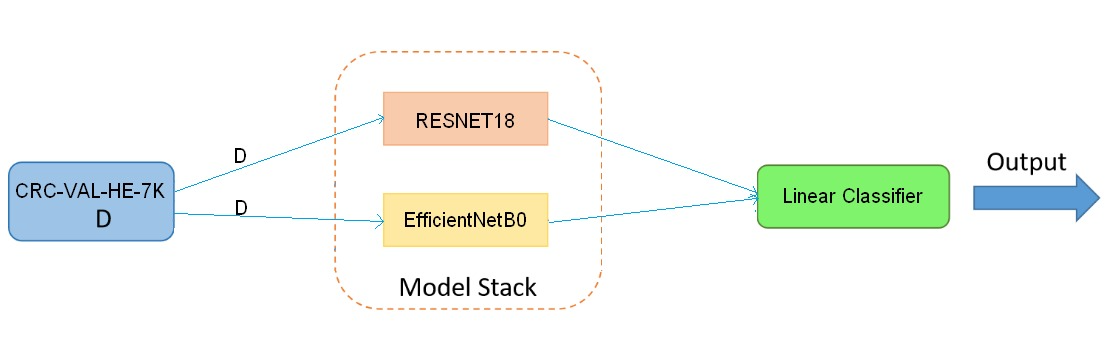
**Ensembled Architecture:**

Ensemble techniques are the methods that use multiple learning algorithms or models to produce one optimal predictive model. The model produced has better performance than the base learners taken alone. Other applications of ensemble learning also include selecting the important features, data fusion, etc. Ensemble techniques can be primarily classified into Bagging, Boosting, and Stacking.



1. **Bagging**: Bagging is mainly applied in supervised learning problems. It involves two steps, i.e., bootstrapping and aggregation. Bootstrapping is a random sampling method in which samples are derived from the data using the replacement procedure. The first step in bagging is bootstrapping, where random data samples are fed to each base learner. The base learning algorithm is run on the samples to complete the procedure. In Aggregation, the outputs from the base learners are combined. The goal is to increase the accuracy while reducing variance to a large extent.
2. **Boosting:** It is an ensemble method in which each predictor learns from preceding predictor mistakes to make better predictions in the future. The technique combines several weak base learners that are arranged in a sequential manner such that weak learners learn from the previous weak learner’s errors to create a better predictive model. Hence one strong learner is formed through significantly improving the predictability of models.
3. **Stacking**: While bagging and boosting used homogenous weak learners for ensemble, Stacking often considers heterogeneous weak learners, learns them in parallel, and combines them by training a meta-learner to output a prediction based on the different weak learner’s predictions. A meta learner inputs the predictions as the features and the target being the ground truth values in data D, it attempts to learn how to best combine the input predictions to make a better output prediction.

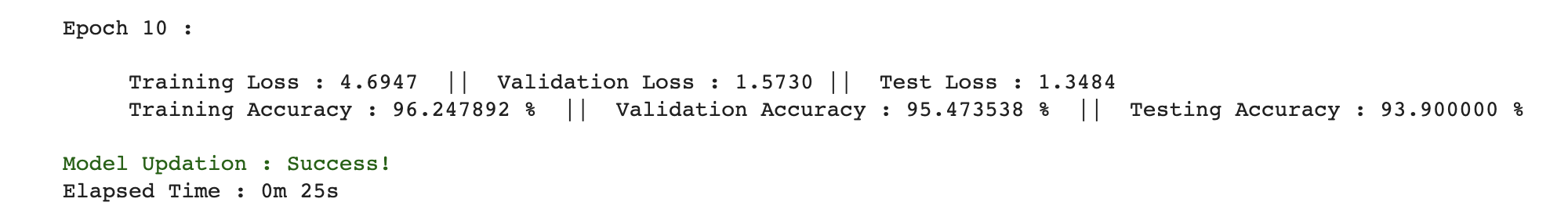
In stacking, an algorithm takes the outputs of sub-models as input and attempts to learn how to best combine the input predictions to make a better output prediction.



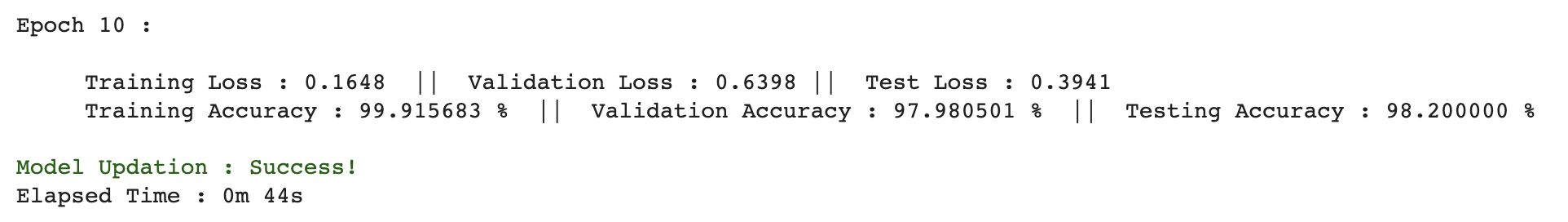
**4. Results and Discussion**

In this section, we have discussed in detail about the experimental setup for the model along with best results and comparison study to support the proposed method.

**4.1. Accuracy**  
Accuracy result of ResNet18 :-



Accuracy result of Stacked Ensemble ResNet18 and EfficientNetB0 model :-



**Loss Graph :**

The graph is plotted with the y axis as the loss and x axis as the steps.

Steps here are calculated with the size of the dataset, epoch and batch size.

* Train Loss



Total steps = (Training dataset size/ batch size) x epoch

Training dataset is 66% of the total data which is equal to 4738.

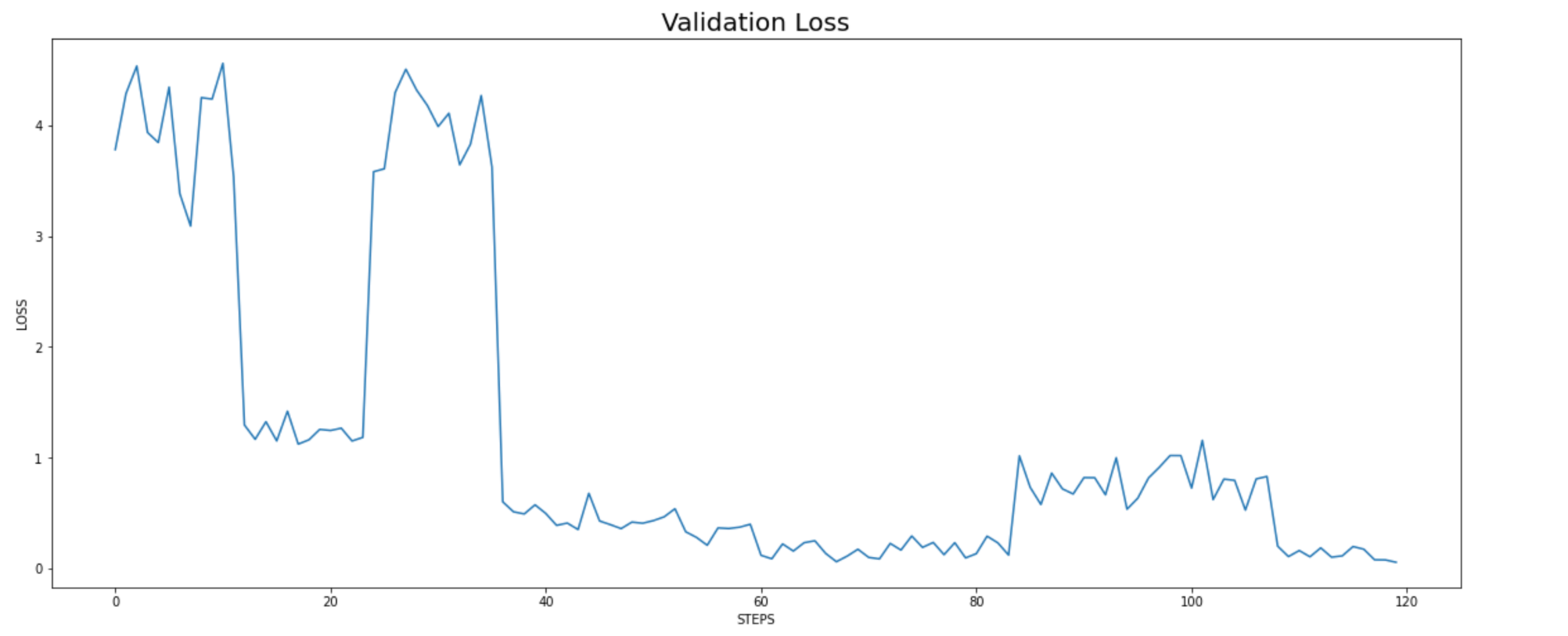
Batch size = 128

Epoch = 10

Therefore, total steps = (4308/128) x 10 = 370

For the Train loss graph, the loss is lowest in the 10th epoch.

* Validation Loss



Total steps = (Validation dataset size/ batch size) x epoch

Validation dataset is 20% of the total data which is equal to 1436.

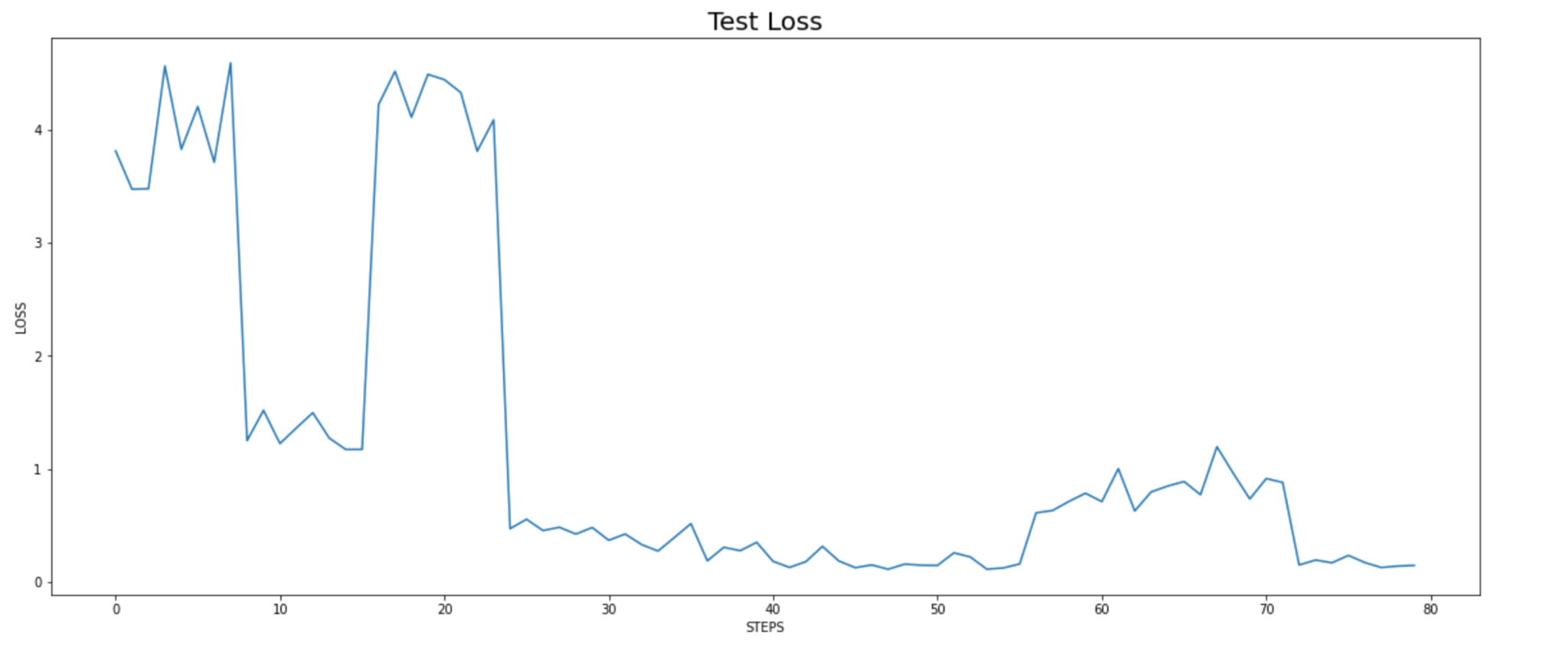
Batch size = 128

Epoch = 10

Therefore, total steps = (1436/128) x 10 = 113

For the validation loss graph, the loss is lowest in the 10th epoch.

* Test Loss



Total steps = (Testing dataset size/ batch size) x epoch

Testing dataset is 14% of the total data which is equal to 1005.

Batch size = 128

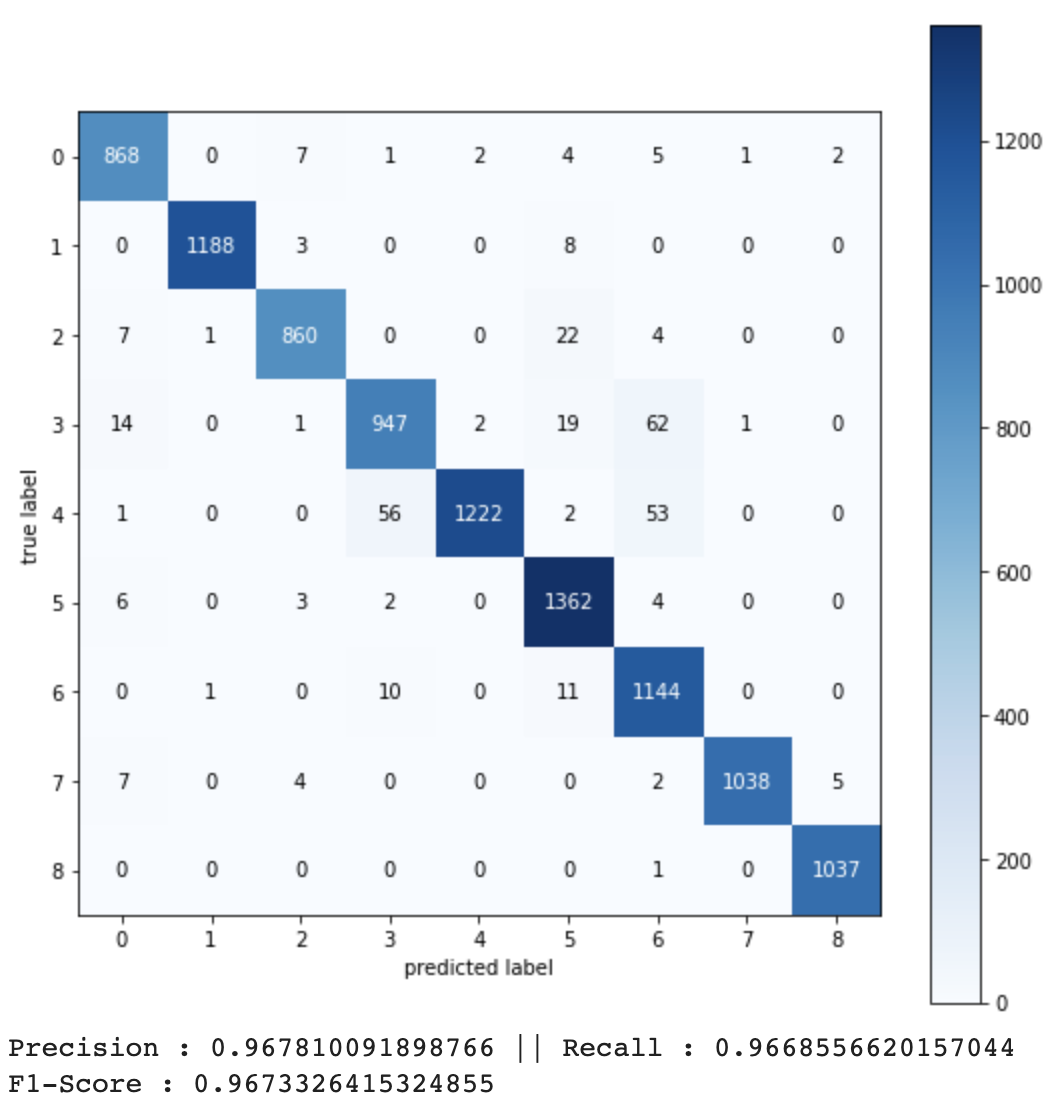
Epoch = 10

Therefore, total steps = (1005/128) x 10 = 79

For the Test loss graph, the loss is lowest in the 10th epoch.

**Confusion Matrix:**

* Testing



In the confusion Matrix above, there have been some misclassifications in some of the classes like **LYM**, **DEB**, **MUC** and **NORM**. The data images of the classes look similar at times and for that reason there can be misclassifications. Sometimes the quality of images are low which might lead to misclassification as well. Noise classification is another reason that might lead to misclassification.

**4.2. Experimental Setup**

In this study, we used the Colab to train and test. In Experiment we, compared the accuracy rate of three training network optimizer methods: the stochastic gradient descent with momentum (SGDM), the root mean square propagation (RMSProp), which utilizes the magnitude of recent gradients to normalize the gradients, and the adaptive moment estimation (Adam), which is an optimization algorithm that can be used for a classical stochastic gradient descent, over the two individual network architectures : ResNet18 and EfficientNetB0 and the stacked ensemble model. In addition, parts of the parameters in the network layers were modified. The approach is used to identify the accuracy rate of the colorectal cancer tissue types from the histological images in open dataset CRC-VAL-HE-7K.

The image dataset was split into three data stores: 66% into training data and 20% into validation and 14% into testing.

**4.3. Experimental Result**

After training on CRC-VAL-HE-7K, the stacked ensemble model performed the best with accuracy rate. The Adam network optimizer had the highest accuracy rate with batch size of 128, epoch of 10, learning rate of 0.0001.

**5. Conclusion**

In this study, different deep learning models and texture based methods for recognising colorectal cancer tissue using CNN were compared. The stacking ensemble model of ResNet18 network model and EfficientNetB0 network model with improved parameters and modified layers showed a better result with accuracy of 98.2%. Colorectal Cancer histological images were used as experimental datasets. Modification of deep learning networks can increase the accuracy of detection of cancer tissues which in turn can enhance doctor’s critical thinking and patient’s proper diagnosis.

**6. References**

1. Egeblad, M.; Nakasone, E.S.; Werb, Z. Tumors as Organs: Complex Tissues that Interface with the Entire Organism. Dev. Cell

2010, 18, 884–901.

2. Kather, J.N.; Weis, C.-A.; Bianconi, F.; Melchers, S.M.; Schad, L.R.; Gaiser, T.; Marx, A.; Zöllner, F. Multi-class texture analysis in

colorectal cancer histology. Sci. Rep. 2016, 6, 27988.

3. Kather, J.N.; Krisam, J.; Charoentong, P.; Luedde, T.; Herpel, E.;Weis, C.-A.; Gaiser, T.; Marx, A.; Valous, N.A.; Ferber, D.; et al.

Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. PLoS Med.

2019, 16, e1002730.

4. Gurcan, M.N.; Boucheron, L.E.; Can, A.; Madabhushi, A.; Rajpoot, N.M.; Yener, B. Histopathological Image Analysis: A Review.

IEEE Rev. Biomed. Eng. 2009, 2, 147–171.

5. Zhang, S.; Metaxas, D. Large-scale medical image analytics: Recent methodologies, applications and future directions. Med. Image

Anal. 2016, 33, 98–101.

6. Zhang, X.; Su, H.; Yang, L.; Zhang, S. Fine-grained histopathological image analysis via robust segmentation and large-scale

retrieval. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12

June 2015; pp. 5361–5368.

7. Janowczyk, A.; Madabhushi, A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use

cases. J. Pathol. Inform. 2016, 7, 29.

8. Korbar, B.; Olofson, A.M.; Miraflor, A.P.; Nicka, C.M.; Suriawinata, M.A.; Torresani, L.; Suriawinata, A.A.; Hassanpour, S. Deep

learning for classification of colorectal polyps on whole-slide images. J. Pathol. Inform. 2017, 8.

9. Tsai, M.J.; Tao, Y.H. Machine Learning Based Common Radiologist-Level Pneumonia Detection on Chest X-rays. In Proceedings

of the 2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS), Gold Coast, Australia,

16–18 December 2019.

10. Xu, J.; Luo, X.;Wang, G.; Gilmore, H.; Madabhushi, A. A Deep Convolutional Neural Network for segmenting and classifying

epithelial and stromal regions in histopathological images. Neurocomputing 2016, 191, 214–223.

11. Du, Y.; Zhang, R.; Zargari, A.; Thai, T.C.; Gunderson, C.C.; Moxley, K.M.; Liu, H.; Zheng, B.; Qiu, Y. Classification of tumor

epithelium and stroma by exploiting image features learned by deep convolutional neural networks. Ann. Biomed. Eng. 2018, 46,

1988–1999.

12. Xu, Y.; Jia, Z.; Wang, L.-B.; Ai, Y.; Zhang, F.; Lai, M.; Chang, E.I.-C. Large scale tissue histopathology image classification,

segmentation, and visualization via deep convolutional activation features. BMC Bioinform. 2017, 18, 281.

13. Bejnordi, B.E.; Mullooly, M.; Pfeiffer, R.M.; Fan, S.; Vacek, P.M.; Weaver, D.L.; Herschorn, S.; Brinton, L.A.; Van Ginneken, B.;

Karssemeijer, N.; et al. Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic

breast biopsies. Mod. Pathol. 2018, 31, 1502–1512.

14. Bowles, M. Machine Learning in Python: Essential Techniques for Predictive Analysis; John Wiley & Sons: Hoboken, NJ, USA, 2015.

15. Hubel, D.H.;Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. J. Physiol.

1962, 160, 106–154.