Image Classification of Ischemic Stroke Blood Clot Origin using Stacked EfficientNet-B0, VGG19 and ResNet-152

by

SURYA PULIGUNDLA	***411958***
MASABATHULA V S RAGHAVENDRA RAO	***411950***
NIKHIL SAI EKKALA	***411918***

Under the guidance of

Dr. SRILATHA CHEBROLU

DEPARTMENT OF

COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY ANDHRA PRADESH

TADEPALLIGUDEM-534102, INDIA

APRIL-2023

Image Classification of Ischemic Stroke Blood Clot Origin using Stacked EfficientNet-B0, VGG19 and ResNet-152

Submitted in partial fulfillment of the requirements

of the degree

of

Bachelor of Technology

by

SURYA PULIGUNDLA	***411958***
MASABATHULA V S RAGHAVENDRA RAO	***411950***
NIKHIL SAI EKKALA	***411918***

Supervisor:

Dr. SRILATHA CHEBROLU

DEPARTMENT OF

COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY ANDHRA PRADESH

TADEPALLIGUDEM-534102, INDIA

APRIL-2023

APPROVAL SHEET

This project work entitled **"Image Classification of Ischemic Stroke Blood Clot Origin using Stacked EfficientNet-B0, VGG19 and ResNet-152"** by "P. Surya, bearing Roll No:411958, M. Raghavendra Rao, bearing Roll No:411950, E. Nikhil Sai, bearing Roll No:411918" is approved for the degree of Bachelor of Technology in the department of Computer Science and Engineering.

Examiners

_____

_____

_____

Supervisor(s)

Dr. Srilatha Chebrolu

_____

Chairman

_____

Date:

Place: Tadepalligudem

DECLARATION

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

P.Surya	M.Raghavendra Rao	E.Nikhil Sai
411958	411950	411918
Date:	Date:	Date:

CERTIFICATE

It is certified that the work contained in the thesis titled **"Image Classification of Ischemic Stroke Blood Clot Origin using Stacked EfficientNet-B0, VGG19 and ResNet-152"** by "P. Surya, bearing Roll No:411958, M. Raghavendra Rao, bearing Roll No:411950, E. Nikhil Sai, bearing Roll No:411918" has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

Dr.Srilatha Chebrolu
Computer Science and Engineering
N.I.T. Andhra Pradesh
April, 2023

ACKNOWLEDGMENTS

We would like to express our utmost gratitude to our supervisor and mentor, Dr. Srilatha Chebrolu, who was a continual source of inspiration. She pushed us to think imaginatively and urged us to do this homework without hesitation. Her vast knowledge, extensive experience, and professional competence in Deep Learning enabled us to successfully accomplish this project. This endeavor would not have been possible without her help and supervision. We could not have asked for a finer supervisor in our studies. This initiative would not have been a success without the contributions of each and every individual. We were always there to cheer each other on, and that is what kept us together until the end. I am deeply grateful to everyone who has contributed to the successful completion of this project.We would also like to express our gratitude to the faculty of the Department of Computer Science and Engineering, for providing such an opportunity for us and for their support.

P.Surya	M.Raghavendra Rao	E.Nikhil Sai
411958	411950	411918
Date:	Date:	Date:

ABSTRACT

Stroke continues to be the second-leading cause of mortality globally. Over 700,000 Americans suffer from an ischemic stroke every year as a result of a blood clot clogging a brain artery. The chances of the patient surviving a second stroke where recurrent strokes account for 23% of all incidents, are decreased. However, if pathologists can identify the cause of a stroke, which affects the therapeutic management after a stroke event, they may be able to reduce the likelihood of repeat strokes. Whole slide digital pathology images capture the blood clot of the patient who has suffered from ischemic stroke. The objective of this work is to classify the blood clot origin in the case of ischemic stroke. It's a binary classification problem where the classes are Cardioembolism and Large artery atherosclerosis. Deep learning models can be applied to classify the stroke origin efficiently. This work proposes a stacked deep learning model with VGG19, ResNet-152, EfficientNet-B0 to classify the stroke origin. Experiments have been conducted on Mayo Clinic - Strip AI dataset from Kaggle. Comparisons have been made with these individual models and also combinations of these models and found that the proposed model is achieving lower loss value.

In chapter 1, consists of the introduction for our project. Chapter 2 consists of the different papers that were primarily used for this project. Chapter 3 consists of proposed methodology**.** In chapter 4, Experimental setup and implementation details were mentioned. Chapter 5 discusses the ‘Results and Discussion’ of our implementation. Chapter 6 provides the conclusion.

TABLE OF CONTENTS

Contents

List of Figures

List of Tables

List of Abbreviations

WSI	Whole Slide Image
AIS	Acute Ischemic Stroke
CNN	Convolutional Neural Network
CE	Cardioembolism
LAA	Large Artery Atherosclerosis
VGG	Visual Geometry Group
RGBA	Red Green Blue Alpha
FC	Fully Connected
SE	Squeeze-Excitation
RELU	Rectified Linear Unit
AI	Artificial Intelligence
ADAM	Adaptive Moment Estimation

Chapter 1

Introduction

Artificial Intelligence (AI) algorithms can be trained to distinguish tissue patterns and certain cell types and to complement the analysis using entire slide images in order to decrease their hands-on time. AI can also correctly measure the abundance of specific cell types in tissue sections, aiding pathologists in diagnosis and image analysis in order to discover the best course of action. With advancements in deep learning for computer vision, there is a surge in the cognification of devices to impart intelligence while tackling tasks which reduces the need for human interventions in such tasks and promotes a better living. The medical images modalities that are widely used for training deep neural networks include digital radiography, ultrasound, magnetic resonance imaging, computed tomography and Whole slide images (WSI) [5].

Whole slide imaging allows you to digitally scan and archive entire slides in high detail and the images of every field of view on the whole microscope slide are captured which are simultaneously stitched together to produce a single digital image file. These complete WSI can be archived and documented, shared for consultation, and used in instruction. Changjiang Zhou, et al., [31] performed histopathology identification and colorectal cancer localisation by weakly supervised deep learning using global labels. The authors have combined a cell-level framework and image-level framework and observed better classification results. The work by Wei Ba, et al., [4] demonstrates a diagnosis classification system for gastritis using deep learning algorithms using WSIs which achieved high accuracy by pre-highlighting the different gastritis regions. The study by Ching-Wei Wang, et al., [28] highlights the automatic bone marrow WSI analysis without manual intervention by developing a powerful and totally autonomous framework for deep learning that is hierarchical for Bone marrow nucleated differential count WSI analysis in seconds. Wei Ba, et al., [3] has developed a deep learning algorithm and analyzed its performance in discriminating melanoma from nevus using WSIs which might function as a supplemental tool to assist pathologists by automatically screening beforehand and emphasizing the interest regions prior to review. Shujun Wang,et al., [29] performed a two-stage framework for histology gastric classification by calibrating a deep learning network with numerous instances that takes into account the various contributions made by each instance to the final prediction.

WSI of a blood clot formed in the human brain will help to identify the two major acute ischemic stroke (AIS) etiology subtypes i) Cardioembolism (CE) and ii) Large artery atherosclerosis (LAA). Convolutional Neural Networks (CNN's) [19] are the best and well-known for image datasets due to its advantages over Artificial neural networks for image data, like spatial invariance, pooling layers, filter layers and reduced computation. Anjali Gautam, et al., [11] have distinguished between ischemic stroke, normal brain and hemorrhagic stroke computed tomography scan images. The paper's authors suggest a 13-layer CNN architecture for classifying strokes, which is fed with images that have already undergone processing. The goal of the Han-Gil Jeong, et al., [15] study was to build a deep CNN that can analyse chest radiographs and accurately detect cardioembolic stroke. These radiographs were used to train the densely linked neural network, ASTROX, to recognise cardioembolic stroke and its focus regions were assessed using gradient-weighted class activation mapping. Yiheng Zhang, et al., [30] outlined a brand-new machine learning technique (extreme gradient boost) based on advanced metabolites that works in conjunction with recursive feature reduction. These metabolites help in detecting Reducing the time of acute ischemic stroke onset and acute ischemic stroke. The work by Nilanjan Dey, et al., [9] shows how a deep learning structure may be used to detect ischemic strokes in multimodality brain MRI slices. The joint segmentation and categorization approach utilized by the authors is based on convolutional neural networks (CNN) that include VGGUNet- supported segmentation. Shon Thomas, et al., [26] used random forest machine learning models using clinical and demographic variables and accurately predicted the presence of mechanical thrombectomy and major vascular occlusions in an ischemic stroke cohort.

However there are some challenges like unique data formats, image file sizes, background elimination, dataset imbalance, preprocessing large sized images and number of available pathology slides. In this work, the problem of differentiating between the AIS etiology sub-types, which are CE and LAA is addressed based on the information in the WSI ofa blood clot. This work proposes a pre-trained stacked deep learning architecture for stroke etiology identification. The individual deep learning models stacked in the proposed model are pre trained CNN architectures [19] which are trained on ImageNet [8] dataset are are used for feature extraction. The input images are preprocessed and small patches from these images are extracted to analyze the components that are present in the blood. These spatial features differentiates between the two stroke blood origin classes i.e , CE and LAA. The deep learning model's performance is measured using the weighted multi-class log loss [17] error function.

Chapter 2

Literature Review

2.1 Normalizing Histology Slides for Quantitative Analysis

Marc Macenko, et al., [21] presented a methodology for normalizing histology slides in order to perform accurate and consistent quantitative analysis to account for variations in staining, lighting, and other factors. The method involves adjusting the brightness and contrast of the slides to ensure that the images are consistent and comparable between different samples. This is important for quantitative analysis, as variations in brightness and contrast can impact the accuracy of measurements and comparisons. This ensures that the images are consistent across all slides, even if the original brightness and contrast vary between samples. The authors of the paper present several advantages to using this normalization method, including improved accuracy and reproducibility of measurements and comparisons, reduced subjectivity in data analysis, and increased efficiency in the analysis process. The algorithm presented aims to normalize the staining of histology slides for quantitative analysis. The algorithm consists of the following steps:

• Step 1: The slide's RGB data are transformed into Optical Density (OD) values, which are a more appropriate representation of the staining intensity.

• Step 2: Data that has an OD intensity below a predetermined threshold β is discarded, as it does not provide meaningful information about the staining intensity.

• Step 3: The OD tuples are subjected to Singular Value Decomposition (SVD), a matrix factorization technique that decomposes a matrix into its constituent parts.

• Step 4: The SVD directions for the two greatest singular values are combined to form a plane.

• Step 5: On the plane built in step 4, the data is projected.

• Step 6: The data is normalized to unit length so that the staining intensity is represented on a comparable scale.

• Step 7: The initial SVD direction is used to determine the angle at each location with respect to it.

• Step 8: The α^{th} and $(100-\alpha)^{th}$ percentiles of the angle serve as representations of the robust extremes.

• Step 9: The extreme values are converted back to OD space.

• Step 10: The optimal stain vectors are output, which can be used to normalize the staining intensity of the histology slides for quantitative analysis.

2.2 VGG19

VGG19 is a deep CNN architecture [19] that was proposed by the Visual Geometry Group of Oxford University in 2014 which performed classification accurately on the ImageNet dataset. It is 19 layers deep and has 16 convolutional layers, 5 same pooling layers and 3 fully connected layers. The default input size for the VGG19 model is 224*224. The size of the kernels is 3*3 in all the convolutional layers. In the convolutional operations, to preserve the spatial resolution of the image, spatial padding was used and outputs from each layer are passed through ReLU activation function. The output layer has 1000 classes which represents the probabilities of the class labels belonging to ImageNet dataset and uses softmax activation for the classification purpose. Using transfer learning [18], it is possible to fine tune this by using the base feature extractors to produce spatial output and use these feature maps for further downstream classification for the custom problem statement

2.3 ResNet-152

ResNet [13] became the winner of ILSVRC held in 2015, and also winner of MS C0C0 2015 competition. Unlike VGG19, ResNet-152 is a very deep CNN architecture with 152 layers but still less complex than VGG19. This architecture has residual skip connections which helps it to carry information through the deep layers and avoid problems like vanishing and exploding gradient descents. The shortcut/skip connections pass the output from one layer to another far way layer directly without any modifications, which helps the old information to sustain longer. The entire architecture can be divided into 5 convolutional blocks followed by a fully connected layer. The kernel sizes in the first block and rest of the blocks are 7*7 and 3*3 respectively. For each block except the first, 3 layers are stacked one over the other, which are

1*1, 3*3, 1*1 convolutions respectively. The 1*1 convolutions are responsible for reducing and restoring the dimensions.

2.4 EfficientNet-B0

Instead of designing models which are too deep, wide or having high resolution to increase the models performance, EfficientNets [25] allow these upgrades in a more principled way. Every EfficientNet architecture contains 7 blocks excluding the stem and final layers, which also contain different numbers of sub-blocks, whose total number rises from EfficientNet-B0 to EfficientNet-B7. The stem contains preprocessing steps like rescaling, padding, normalization etc. Each block uses multiple MBConv (an inverted residual block) layers and the outputs of the last block are connected to a 1*1 convolution followed by pooling and a fully connected layer. The number of layers in EfficientNet-B0 is 237, and increases as we move on to EfficientNet-B7 which has 813 layers. Better performance may result from properly balancing the network depth, width, and resolution. Using compound scaling method, which scales the networks over width, depth, and resolution, equilibrium between all dimensions can be achieved. The scaling can be performed this way:

$$\text{depth} : d = \alpha^{\theta}$$

$$\text{width} : w = \beta^{\theta}$$

$$\text{resolution} : r = \gamma^{\theta}$$

$$\text{s.t } \alpha\beta^2\gamma^2 \approx 2, \text{ and } \alpha \geq 1, \beta \geq 1, \gamma \geq 1$$

where d, r, we are the coefficients for scaling the depth, resolution and width respectively, θ is the compound coefficient and α, β, γ are the scaling coefficients which control the depth, width and resolution of the network to achieve the trade-off between model size and accuracy.

2.5 Ensembling EfficientNets for the Classification and Interpretation of Histopathology Images

The paper "Ensembling EfficientNets for the Classification and Interpretation of Histopathology Images"[32] proposes a deep learning-based method to classify histopathology images into different categories such as benign or malignant tumors. The authors use a state-of-the-art architecture called EfficientNet as a base model and propose an ensemble

strategy that combines the outputs of multiple EfficientNets trained on different subsets of the data to improve the overall classification accuracy.

In addition to the classification task, the authors also propose a novel interpretation method based on attention maps, which highlight the regions of the image that the model uses to make its prediction. They show that this method can provide insights into the decision-making process of the model and can help identify regions of interest in the image that are most relevant for the classification task.

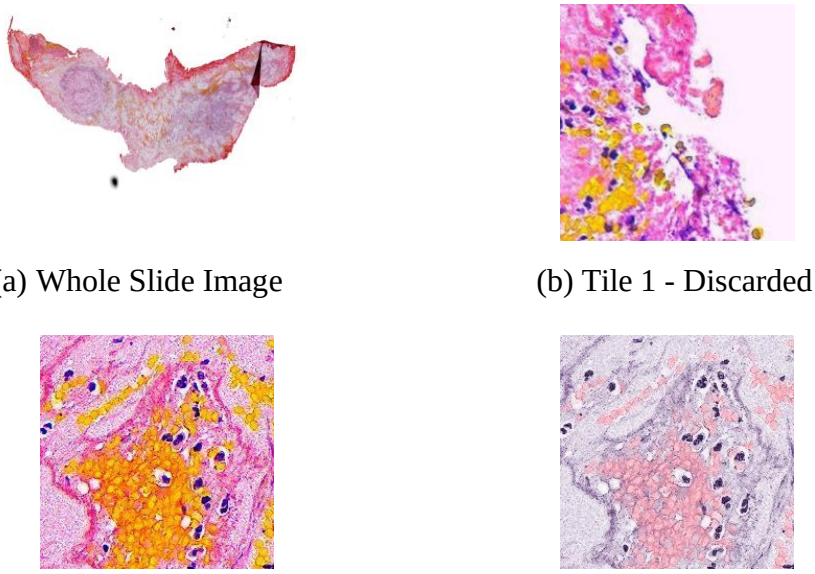
The authors evaluate their method on a large dataset of histopathology images and demonstrate that their ensembling strategy improves the classification accuracy compared to a single model. They also show that their interpretation method can provide meaningful insights into the model's decision-making process and can help identify regions of interest in the image. The proposed method has the potential to improve the accuracy and interpretability of histopathology image analysis, which is important for the diagnosis and treatment of various diseases.

Chapter 3

PROPOSED METHODOLOGY

3.1 DATASET

In this work, experimentation analysis is done on Mayo Clinic-STRIP AI [2] dataset. This dataset consists of 754 high resolution WSI images in TIFF format, each belonging to a blood clot origin. This is a class imbalance dataset as around 551 images belong to the class CE and 203 images belong to class LAA. Figure 2(a) shows a sample WSI from the dataset.



(a) Whole Slide Image (b) Tile 1 - Discarded

(c) Tile 2 - Preserved (d) Normalized tile

Figure 3.1: Whole slide image tiles

3.2 DATA PREPROCESSING

These high resolution WSI images need to be preprocessed before passing for training.
The entire preprocessing can be divided into the following sections.

3.2.1 Tiling and filtering

Each WSI is tiled [6] into multiple small images using the openslide deep zoom generator [12]. The sample tiles are as shown in Figure 2(b) and Figure 2(c). The tiles with white or mostly white backgrounds are discarded and also the tiles that do not carry important information are ignored. The tiles that satisfy the threshold conditions of mean and standard deviations are filtered and added to the final training set. Tiles which are dark and have higher standard

deviation are considered into the training set. For this dataset, the mean threshold is set to 180 and standard deviation threshold is set to 50, i.e tiles with only mean value less than 180 and standard deviation value more than 50 will be taken into consideration. The tile in Figure 2(b) is eliminated as the mean and standard deviation values are 208 and 59 respectively. The tile in Figure 2(c) is included in the training process as it has met the threshold conditions of mean and standard deviation.

3.2.2 Color Normalization

The images are then color normalized [23]. To ensure that differences in color intensity and shading, which can be introduced due to various factors such as slide preparation, staining methods and imaging device variability, do not affect the analysis and interpretation of the images. This also eliminates the tiles that do not carry useful or relevant information. When the algorithm proposed by Marc Macenko et al., [21] is applied on the images, there is a possibility for the covariance matrix to have rank less than the number of channels, which means that the data is linearly dependent or there are not enough unique patterns in the data. This is possible with images that are extremely noisy, have very low contrast, or have no tissue present in the image. In these cases, the normalization process will fail and images of these kinds are eliminated.

Layer	**Output shape**	**Parameters**
Input Layer	512*512*3	0
EfficientNet-B0 feature extractor	16*16*1280	4,049,571
ResNet-152 feature extractor	16*16*1280	58,370,944
VGG19 feature extractor	16*16*512	20,024,384
Concatenation	1*3048	0
Batch Normalization	1*3048	15,360
Dense	1*128	491,648
Dense	1*2	258

Table 3.1: Output shapes and parameters of the proposed architecture

3.2.3 Data augmentation and rescaling

Data augmentation techniques like rotating, increasing the brightness, zooming, vertical and horizontal flips, etc are applied to the tiles. Also the images are rescaled by a factor of 255 before passing into the model. Figure 2(d) shows the fully preprocessed image after applying data cleaning and color normalization techniques. To handle the class imbalance, at most 6 tiles from an imagethat belong to class CE and at most 22 tiles that belong to class LAA have been selected.

Finally, the training set consists of 2569 images belonging to class CE and 2302 images belonging to class LAA. These images are obtained after filtering the tiles which have white backgrounds that are not meeting the thresholds of mean and standard deviation. All these images are color normalized.

3.3 EFFICIENTNET-B0, VGG19 AND RESNET-152 STACKED MODEL

This subsection describes the architecture of the stacked model for classification of blood clot stroke origin. The preprocessed images are given as input to the proposed model shown in Figure 1. The model is a stacked architecture [22] of EfficientNet-B0 [25], VGG19 [24] and ResNet-152 [13] where the backbones of the CNN architectures are pre-trained on ImageNet dataset and fine-tuned. Stacking models is a kind of ensemble [10] technique where there is no weightage given to the models based on their performance. The first layer of the model accepts RGB images of shape $512 * 512$. The feature extractor outputs are of shapes $16 * 16 * 1280$, $16 * 16 * 2048$, $16 * 16 * 512$ from the three backbone models EfficientNet-B0, ResNet-152 and VGG19 respectively are passed through a global average pooling [20] layer. These outputs are individually flattened and obtained outputs are of shapes $1 * 1280$, $1 * 2048$ and $1 * 512$. These outputs are concatenated to obtain a single output of shape $1 * 3840$.

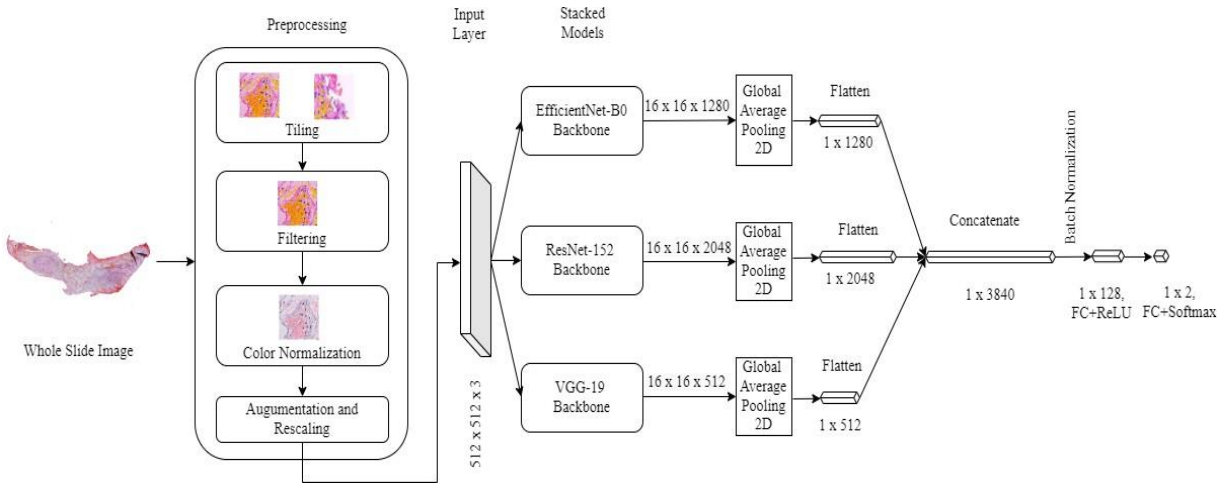


Figure 3.2: Architecture of stacked EfficientNet-B0, VGG19 and ResNet-152 model for classification of stroke blood clot origin

A batch normalization layer [14] is used after flattening and concatenating the feature extraction outputs. This is followed by a hidden dense layer having 128 neurons with ReLU [1] activation function. The output layer of the model has two neurons with softmax activation [7] function and provides the probabilities of each class label. The total parameters of this implementation is 82, 952, 165, out of which 499, 586 are trainable. The complete architecture details of the model are shown in Table 1. The proposed architecture is fine-tuned on the Mayo Clinic-STRIP AI dataset and optimal set of parameters are found.

Chapter 4

EXPERIMENTAL SETUP AND IMPLEMENTATION DETAILS

In this work, experiments are performed on the Kaggle Platform. The process of image preprocessing and model training are executed using GPU. Kaggle provides NVIDIA TESLA P100 GPU with 15.9 GB memory and a disk space of 73.1 GB.

4.1 Hyper parameter settings

This section discusses the hyper parameters used during the implementation of the proposed methodology.

4.1.1 Adam Optimizer:

The proposed architecture uses the Adam optimization function [16]. It is employed because it takes minimal running time and training settings and changes the learning rate for each weight in the neural network. Initial learning rate is set as 0.001. The weights are updated as follows:

$$w_{t+1} = w_t + \alpha g_t$$

where,

$$g_t = \beta g_{t-1} + (1 - \beta)\left(\frac{\delta L}{\delta w_t}\right)$$

w_{t+1} denotes the weights of the neural network at time t +1. w_t denotes the weights at time t, g_t is the aggregate of current gradients, α_t denotes the learning rate, δw_t denotes the derivative of weights at time t. δL is the derivative of loss function, β is the average moving parameter, g_{t-1} is the aggregate of previous gradients.

4.1.2 Categorical Cross Entropy

Here, the loss function used for calculating the error during back propagation is categorical cross entropy [27]. Categorical loss function is given as :

$$-\frac{1}{r}\sum_{i=1}^{r} a_i log(\widehat{a}_i) + (1 - a_i) log(1 - \widehat{a}_i)$$

where, r represents the number of images in the dataset, a_i is the ground truth probability and $\widehat{a}_i$ is the predicted probability.

4.2 Model evaluation

This section describes the process of evaluating the proposed architecture. Weighted multi-class logarithmic loss function is used for the evaluation of the model after training. This loss function is very similar to the categorical cross entropy with a difference that the classes are assigned weights during the calculation of the loss. Weighted multi-class logarithmic loss is given as:

$$\frac{-\sum_{i=1}^{c} w_i \sum_{j=1}^{r_i} \frac{y_{ij}}{r_i} ln(\widehat{y}_{ij})}{\sum_{i=1}^{c} w_i}$$

where c, r represents the number of classes, and number of images respectively, y_{ij} denotes whether image i belongs to class j in the training set, $\widehat{y}_{ij}$ is the predicted probability of image i belonging to class j, and w_i denotes the weight of class i.

4.3 Model Prediction

Three tiles are randomly fine-tuned from the test image and preprocessed before passing into the stacked model and predictions are made. The prediction values for the three tiles are averaged and considered as the final prediction for the test image.

Chapter 5

Results and analysis

Experiment	Model	Input size	Epochs	Batch size	Weighted multi-class logarithmic loss
1	CNN	224	30	32	0.74833
2	EfficientNet-B0	224	30	32	0.73246
3	EfficientNet-B0	512	60	16	0.69366
4	EfficientNet-B0	512	40	32	0.69332
5	ResNet-152	224	30	32	0.73490
6	ResNet-152	512	60	16	0.69597
7	ResNet-152	512	40	32	0.69854
8	VGG19	224	30	32	0.82692
9	VGG19	512	40	32	0.70645
10	VGG19	512	60	16	0.81135
11	Stacked EfficientNet-B0, ResNet-152	224	30	32	0.74695
12	Stacked EfficientNet-B0, ResNet-152	512	40	32	0.69484
13	Stacked EfficientNet-B0, ResNet-152	512	60	16	0.71081
14	Stacked EfficientNet-B0, VGG19	224	30	32	0.78061

15	Stacked EfficientNet-B0, VGG19	512	40	32	0.75251
16	Stacked ResNet-152,VGG19	512	40	32	0.69792
17	Stacked ResNet-152, VGG19	512	60	16	0.69943
18	Stacked EfficientNet-B0, VGG19, ResNet-152	512	60	16	0.71019
19	Stacked EfficientNet-B0 , VGG19, ResNet-152	512	40	32	0.69312

Table 5.1: Experimental results

Table II shows experimental results obtained for different combinations of stacked EfficientNet-B0, ResNet-152 and VGG19 models with various values of input image size, number of epochs and batch size on Mayo Clinic - Strip AI dataset. Total 20 experiments were conducted. Experiments from 1 to 10 are done using individual models whereas experiments from 11 to 20 are stacked models with various combinations of EfficientNet-B0, ResNet-152 and VGG19. The experiments were performed by changing the input size to the models and also varying the hyper parameters like epochs and batch sizes during the training process.

Experiment 20, which is a stacked model of EfficientNet-B0, ResNet-152 and VGG19 , using the input size of 512, number of epochs as 40 and batch size as 32, has achieved the minimum weighted multi-class logarithmic loss value of 0.69312 among all the experiments conducted. Experimental results show that the proposed architecture achieves the objective of classification of ischemic stroke blood clot origin between LAA and CE using WSI images.

Chapter 6

Conclusion

Classification of ischemic blood stroke origin is done using whole slide images of blood clots. The proposed stacked model of EfficientNet-B0, ResNet-152 and VGG19 performs the task ofclassifying the brain stroke etiology subtypes i.e., Cardioembolism and Large artery atherosclerosis. This stacked model is evaluated using the metric weighted multiclass logarithmic loss for the task of classification of ischemic blood stroke origin on Mayo Clinic - Strip AI dataset. The performance of the proposed stacked EfficientNet-B0, ResNet- 152 and VGG19 model is compared with CNN and other individual and stacked models of EfficientNet-B0, ResNet-152 and VGG19 models. Comparison results on weighted multiclass logarithmic loss shows that the proposed model achieves minimum loss value, as this architecture has the advantages of all three individual models. This allows us to do away with the requirement for human aid and the uncertainties associated with it. With this work, healthcare providers will have the ability to more accurately determine the source of blood clots that cause fatal strokes, allowing physicians to provide more effective post-stroke treatment and decreasing the chance of a recurrent stroke.

The proposed stacked model is initialized with the weights obtained from the models VGG19, ResNet-152 and EfficientNetB0 that are pre trained on ImageNet dataset. As a future work, theusage of models that are pre trained on digital pathology whole slide images needs to be explored as it improves the performance.

References

[1] Abien Fred Agarap. Deep Learning using Rectified Linear Units (ReLU). arXiv preprint arXiv:1803.08375, 2018.

[2] Ryan Holbrook Sobhi Jabal VikashGupta Ashley Chow, Barbaros. Mayo Clinic - STRIP AI. Kaggle, 2022.

[3] Wei Ba, Rui Wang, Guang Yin, Zhigang Song, Jinyi Zou, Cheng Zhong, Jingrun Yang, Guanzhen Yu, Hongyu Yang, Litao Zhang, and Chengxin Li. Diagnostic assessment of deep learning for melanocytic lesions using whole-slide pathological images. Translational Oncology, 14(9):101161, 2021.

[4] Wei Ba, Shuhao Wang, Cancheng Liu, Yuefeng Wang, Huaiyin Shi, and Zhigang Song. Histopathological Diagnosis System for Gastritis Using Deep Learning Algorithm. Chinese Medical Sciences Journal, 36(3):204–209, 2021.

[5] Jon Braatz, Pranav Rajpurkar, Stephanie Zhang, Andrew Y. Ng, and Jeanne Shen. Deep Learning-Based Sparse Whole-Slide Image Analysis for the Diagnosis of Gastric Intestinal Metaplasia. arXiv preprint arXiv:2201.01449, 2022.

[6] Erich Bremer, Joel Saltz, and Jonas S Almeida. ImageBox 2 – Efficient and Rapid Access of Image Tiles from Whole-Slide Images Using Serverless HTTP Range Requests. Journal of Pathology Informatics, 11(1):29, 2020.

[7] Binghui Chen, Weihong Deng, and Junping Du. Noisy Softmax: Improving the Generalization Ability of DCNN via Postponing the Early Softmax Saturation. 07 2017.

[8] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pages 248– 255, 2009.

[9] Nilanjan Dey and Rajinikanth V. 7 - Automated detection of ischemic stroke with brain MRI using machine learning and deep learning features. In Nilanjan Dey and Rajinikanth V.,

editors, Magnetic Resonance Imaging, Primers in Biomedical Imaging Devices and Systems, pages 147–174. 2022.

[10] M. A. Ganaie, Minghui Hu, Muhammad Tanveer, and Ponnuthurai Nagaratnam Suganthan. Ensemble deep learning: A review. Engineering Applications of Artificial Intelligence, 115:105151, 2021.

[11] Anjali Gautam and Balasubramanian Raman. Towards effective classification of brain hemorrhagic and ischemic stroke using CNN. Biomedical Signal Processing and Control, 63:102178, 2021.

[12] Adam Goode, Benjamin Gilbert, Jan Harkes, Drazen Jukic, and Mahadev Satyanarayanan. OpenSlide: A vendor-neutral software foundation for digital pathology. Journal of Pathology Informatics, 4(1):27, 2013.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. 2015.

[14] Sergey Ioffe and Christian Szegedy. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv preprint arXiv:1502.03167, 2015.

[15] Han-Gil Jeong, Beom Joon Kim, Tackeun Kim, Jihoon Kang, Jun Yup Kim, Joonghee Kim, Joon-Tae Kim, Jong-Moo Park, Jae Guk Kim, Jeong-Ho Hong, Kyung Bok Lee, Tai Hwan Park, Dae-Hyun Kim, Chang Wan Oh, Moon-Ku Han, and Hee-Joon Bae. Classification of cardioembolic stroke based on a deep neural network using chest radiographs. eBioMedicine, 69:103466, 2021.

[16] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980, 2014.

[17] Maksim Lapin, Matthias Hein, and Bernt Schiele. Analysis and Optimization of Loss Functions for Multiclass, Top-k, and Multilabel Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40:1533–1554, 2016.

[18] Diane Larsen-Freeman. Transfer of Learning Transformed. Language Learning, 63, 2013.

[19] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, 1998.

[20] Min Lin, Qiang Chen, and Shuicheng Yan. Network In Network. arXiv preprint arXiv:1312.4400, 2013.

[21] Marc Macenko, Marc Niethammer, J. Marron, David Borland, John Woosley, Xiaojun Guan, Charles Schmitt, and Nancy Thomas. A Method for Normalizing Histology Slides for Quantitative Analysis. Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009, 9:1107–1110, 2009.

[22] Bohdan M. Pavlyshenko. Using Stacking Approaches for Machine Learning Models. IEEE Second International Conference on Data Stream Mining & Processing (DSMP), pages 255–258, 2018.

[23] Santanu Roy, Shyam Lal, and Jyoti R. Kini. Novel Color Normalization Method for Hematoxylin Eosin Stained Histopathology Images. IEEE Access, 7:28982–28998, 2019.

[24] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556, 2014.

[25] Mingxing Tan and Quoc V. Le. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 2019.

[26] Shon Thomas, Paula de la Pena, Liam Butler, Oguz Akbilgic, Daniel M. Heiferman, Ravi Garg, Rick Gill, and Joseph C. Serrone. Machine learning models improve prediction of large vessel occlusion and mechanical thrombectomy candidacy in acute ischemic stroke. Journal of Clinical Neuroscience, 91:383–390, 2021.

[27] Vladimir Vovk. In Fields of Logic and Computation II. 2015.

[28] Ching-Wei Wang, Sheng-Chuan Huang, Yu-Ching Lee, Yu-Jie Shen, Shwu-Ing Meng, and Jeff L. Gaol. Deep learning for bone marrow cell detection and classification on whole-slide images. Medical Image Analysis, 75:102270, 2022.

[29] Shujun Wang, Yaxi Zhu, Lequan Yu, Hao Chen, Huangjing Lin, Xiangbo Wan, Xinjuan Fan, and Pheng-Ann Heng. RMDL: Recalibrated multi instance deep learning for whole slide gastric image classification. Medical Image Analysis, 58:101549, 2019.

[30] Yiheng Zhang, Dayu Zhu, Tao Li, Xiaoya Wang, Lili Zhao, Xiaofei Yang, Meijuan Dang, Ye Li, Yulun Wu, Ziwei Lu, Jialiang Lu, Yating Jian, Heying Wang, Lei Zhang, Xiaoyun Lu, Ziyu Shen, Hong Fan, Wenshan Cai, and Guilian Zhang. Detection of acute ischemic stroke and backtracking stroke onset time via machine learning analysis of metabolomics. Biomedicine Pharmacotherapy, 155:113641, 2022.

[31] Changjiang Zhou, Yi Jin, Yuzong Chen, Shan Huang, Rengpeng Huang, Yuhong Wang, Youcai Zhao, Yao Chen, Lingchuan Guo, and Jun Liao. Histopathology classification and localization of colorectal cancer using global labels by weakly supervised deep learning. Computerized Medical Imaging and Graphics, 88:101861, 2021.

[32] Kallipolitis, A. & Revelos, Kyriakos & Maglogiannis, Ilias. (2021). Ensembling EfficientNets for the Classification and Interpretation of Histopathology Images. Algorithms. 14. 278.