

Early Detection of Hemorrhagic Stroke Using a Lightweight Deep Learning Neural Network Model



Bandi Vamsi¹, Debnath Bhattacharyya², Divya Midhunchakkkravathy¹, Jung-yoon Kim^{3*}

¹ Department of Computer Science and Multimedia, Lincoln University College, Kuala Lumpur 47301, Malaysia

² Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Greenfield, Vaddeswaram, Guntur 522502, India

³ Department of Game Media, College of Future Industry, Gachon University, Seongnam-si 13120, Gyeonggi-do, Republic of Korea

Corresponding Author Email: kjyoon@gachon.ac.kr

<https://doi.org/10.18280/ts.380616>

ABSTRACT

Received: 1 November 2021

Accepted: 6 December 2021

Keywords:

Convolution Neural Network (CNN), computed tomographic, deep learning, hemorrhagic stroke, light weight model, medical image segmentation

In present days, the major disease affecting people all across the world is “Cerebrovascular Stroke”. Computed tomographic (CT) images play a crucial role in identifying hemorrhagic strokes. It also helps in understanding the impact of damage caused in the brain cells more accurately. The existing research work is implemented on the Graphical Processing Unit (GPU’s) for stroke segmentation using heavyweight convolutions that require more processing time for diagnosis and increases the model’s cost. Deep learning techniques with VGG-16 architecture and Random Forest algorithm are implemented for detecting hemorrhagic stroke using brain CT images under segmentation. A two-step light-weighted convolution model is proposed by using the data collected from multiple- repositories to inscribe this constraint. In the first step, the input CT images are given to VGG-16 architecture and in step two, data frames are given to random forest for stroke segmentation with three levels of classes. In this paper, we explore various training time values in the detection of stroke that reduces when compared with existing heavyweight models. Experimental results have shown that when compared to other existing architectures, our hybrid model VGG-16 and random forest achieved increased results obtained are dice coefficient with 72.92 and accuracy with 97.81% which shows promising results.

1. INTRODUCTION

An unexpected disturbance occurring in the portion of the brain resulting in lack of supply of blood is an ischemic stroke [1]. When the blood veins in the brain exploded and scatter the blood to the other regions of nearby brain cells which leads to the condition [2] defined as a Hemorrhage Stroke. The hemorrhagic stroke is considered a significant stroke as its mortality rate extends from 35-52%, though this stroke case extends to only 10-30% [3]. According to the National Stroke Association (NSA), the majority 85% of the death rate [4] is caused by Hemorrhagic stroke. The examination of brain stroke including factors of broad accessibility, less expensive and vulnerability in advance detection of the stroke, Computed Tomography (CT) images are widely used [5]. In planning the clinical treatment procedures, the exact divisions of the hemorrhagic segmentation play an important role. The identification [6] of disease rather than arrangements, easy understanding and significant measures act as additional features.

1.1 Intracranial hemorrhage

Intracranial Hemorrhage (ICH) is considered the second most dangerous stroke including different etiologies with trauma, infraction, aneurysm rupture, and anticoagulant therapy from across the world. Around 37,000 to 50,000 ICH

stroke cases [7] are recorded only in the United States of America. If in case, there is any downfall in these ICH cases past in history but now these ICH stroke cases have not decreased considerably within a month it is 47% high. In 24 hours of duration, it is estimated that half of humanity is affected by ICH stroke. On the other hand results based on the treatment shows better results. The scans related to the brain called neuro-imaging is required in the detection and identifying of the ICH strokes [8], especially in meticulous areas like type, location, and size which helps in treating the patient administration. According to past studies and researches, around 18 to 68% of people with ischemic or transient stroke are affected by cerebral minor bleedings. Based on this study 44% of the patients are affected with the persistent ischemic stroke [9] while 23% are a witness in the initial stage of ischemic stroke.

1.2 Deep learning

Currently, Artificial Intelligence acts as a key technique in identifying the stroke. It is widely spread to most of the people in using this approach is beneficial in diagnosis [10] different types of strokes. However, by using the technique of heavy weights convolution, the limitation of time complexity increases and quality reduces. This also involves in high costs for radiologists in diagnosing the region of stroke. Since, out of many approaches, with the advanced technique of light

weight convolution model [11] the time complexity reduces thereby diminishing the costs. Due to this, the radiologists can easily identify the stroke region with in particular period of time accurately.

Deep learning (DL) is a division of two major fields. In clear the Artificial Intelligence (AI) acts as a major stream making machine language a subset in which deep learning acts as a specific subset. Artificial intelligence involves developing machines [12] by exhibiting intelligence. Machine learning helps in creating the algorithms using the insights of the data. Deep learning performs the actions related to specific algorithms [13] often referred to as neural networks. Deep learning is nothing but an algorithm [14] that continuously works on specified programs in speculating various works. It is interlinked with machine learning in identifying that particular algorithm.

In general, for most of the algorithms, it is very difficult for detecting the initial raw data provided as input for the system, but by insight handling techniques of the deep learning [15], it becomes easier by explaining every single layer in the module. Deep learning mainly focuses [16] on the three different layers namely the visible layer which is given as an input to the model, hidden layers that are situated in series which is invisible to the naked eye, and the output layer which produces the final result of the model [17].

Nowadays, the methods of Deep learning are widely used in several areas depending on their improvements. Mainly in terms of a medical field under image categorization and examination of computer-based methods, deep CNNs [18] are carried out flourishingly. On the other hand, the past approximations of the division of the images of CT scan in hemorrhage areas entirely depend on hand crafted mechanism whereas the deep CNNs retrieve the data involuntarily resolving difficult techniques. Hence the involuntary stroke detection can be carried out through numerous CNN techniques that include both 2D and 3D. Various parts are obtained for every patient included in their CT brain scan image. Thus, the 3D evaluation network is implemented in such CT scan divisions for spatial data drawn from various regions.

In this research work, we have used light weight neural network based on 2D transfer learning and random forest technique for extracting features over different convolutions on brain CT images for segmentation of stroke. The briefing of our work is as follows:

a) Initially, we used a two-class labeled image for stroke segmentation and this leads to an improper segmentation between skull and stroke area.

b) To overcome this limitation, we choose a three-class labeled image in which the skull and background are labeled in class-1, the gray matter is labeled in class-2 and the stroke area is labeled in class-3.

c) Different segmentation models are applied on the image data for stroke segmentation. Through the thorough literature survey, we identified few limitations such as these models require high processing power, GPU's, more training time, etc.

d) Our proposed model is a hybrid architecture which contains both deep and machine learning techniques. This model overcome some of these limitations and achieved better results.

The remaining sections of this research work are arranged as follows. Section-2 discusses Related Works, Section-3 describes the DL Methods and Models, and Section-4 describes the Experimental Results of Proposed Model.

Section-5 Results and Discussions, and Section-6 discusses Conclusion and Future Scope.

2. RELATED WORK

This section discusses about existing works implemented by various researchers in the field of medical image segmentation to identify the brain stroke.

An automatic multi path convolution network model for stroke segmentation using MRI images [19]. This model takes 9 convolutions as an input in 2 dimensional with different regulations. These 9 inputs are concatenated into 3-dimensional convolutions as an output for predicting the injury anywhere in the brain. The heterogeneity dataset is having both ischemic and hemorrhagic strokes applied to this model with two-time lines: after stroke period of 1 to 5 months and more than 3 months. Out of 99 samples, 45% and 55% of subjects were used for training and testing. The image segmentation quality coefficient values among existing models like UNet, ResNet, and DeepMedic, this proposed model examined with a decent median value of 0.66 and mean value with 0.54.

A deep residual neural network for ischemic stroke segmentation using MRI images [20]. This model is inherited from the UNet convolution as a base model with high-quality features available in the input images. The loss function of image segmentation quality is applied to the proposed model to handle the imbalanced dataset. Out of 742 samples, 70% and 30% of subjects were used for training and testing. Among existing models named as RA-UNet, uResNet, and FC-ResNet, this proposed model has attained decent accuracy with 94.04%.

A neural network model for the automatic segmentation of intracerebral hemorrhagic stroke [21]. To utilize the low- and high-level semantics this model has used encoder and decoder architectures. With different loss functions, this proposed model has more attention on small regions of the stroke to overcome the imbalanced data. Out of 480 samples, 80% and 20% of subjects were used for training and testing. The performance measure of existing models used was UNet, SegNet, X-Net, and MultiResUNet, this proposed model has attained decent accuracy with 88.07% for segmentation.

A Dense convolution network model for brain CT image segmentation to overcome the limitations of UNet architecture [22]. This proposed works for image segmentation with three blocks: a) convolution with dense layer b) de-convolution with residual layer c) un-pooling block. Out of 781 samples, 64% and 36% of subjects were used for training and testing. Among existing models used in this research work named as UNet, ResUNet, FCN, and FCN dense, the proposed model performance evaluated with decent accuracy of 83.47%.

A multi-receptive network model to extract the medical information from images for segmentation [23]. This model provides receptive for each subnetwork and generates feature map information. By using the loss function, the imbalanced thinness of medical images gets eliminated. Out of 742 samples, 70% and 30% of subjects were used for training and testing. Among the existing model such as uResNet, FC-ResNet, and RA-UNet, this proposed model has attained a dice coefficient (DC) metric of 77.04%.

A multi-task convolution network for hemorrhagic stroke segmentation and classification [24]. This model uses related information about slices to combine the tasks for information sharing. Out of 1176 CT samples, 60% and 40% of subjects

were used for training, validation and testing. This proposed model attained the best accuracy of 95.3% for slice level stroke detection when compared with ResNet18 and ICHNetcls.

In the existing research works [19-24], the researchers have implemented different segmentation techniques for stroke segmentation by using heavy weight models. These models take more time for training on image data to segment the stroke area. This leads to increasing of model cost and also requires high powered GPU's to process the model. To overcome this limitation, our proposed model uses hybrid architecture which runs on low processing power and achieves good results in comparison of existing models.

3. METHODS USED

This section describes various deep learning models along with working procedure, pre-processing and augmentation techniques, dataset used in this research work.

3.1 Data pre-processing

In the selection technique, initial transformations play a key role with respect to their accuracy. The intensity scores of the tissue or region must be same for all the images. While in CT scan images, skull is always shown brighter and the rest of the image is left vice versa. In this, the configuration of the image pixel ranges from 0 to 225, based on the various features of the device being used for the study. While if the pixel unit is scattered under wider distribution the image becomes contrast and the cost also reduces which is not supported by the ICH selection. Also, the various layers with softer tissues flap in the pixel unit of the hemorrhage CT image, this disturbance shows adverse results in learning and assessment of the feature. Hence in order to obtain a clear picture with ICH selection, initially the outer portions such as the skull, skin eyes, and facial characters must be excluded. Depending on the Destruction mechanism, we can bring back the portion implementation for the initial selection of the brain CT image and later can be used for different functions to gather the intracranial brain mask area. Finally, the regions are gathered in getting the intracranial brain portions.

3.2 Data sampling

In general, the qualities are varied in ICH clinical data. Often, the regions of the brain are scattered among various places most probably around its center. In such connections most of the data become unstable and few hemorrhage regions will be moved under in small amounts leaving the entire space among the regions. In order to avoid this, some data sampling and data augmentations mechanisms came into existence. In order to catch hold of the ICH regions, a certain small shape is initially cropped by using these wide-ranging tests to address. Various features such as noise should be minimum, proper orientation, entire skull region should be enclosed and proper contrast are to be maintained. In a contrast, data sampling can easily be used in determining the unbalanced regions of hemorrhage portions considering the entire region and show cautious study, particularly to ICH regions. These sampling regions gather complex testing samples than the required data.

3.3 Data augmentation

In the naming of training sets over fitting is the major

drawback. This drawback can be resolved through data augmentation. In this theory, the exploration of augmentation methods namely traditional, spatial and intensity variations are widely used. The different forms of data augmentation are given by combining the pictures with the latest picture development methods those which rely on social networking. Given comprehensive feedback, Data augmentation is not only related to medical fields, with its evolving features it is used in both natural and medical CT scan areas. For this research work we used image flipping in both horizontal and vertical directions and different contrast levels are shown in Figure 1.

3.4 U-Net with base convolution neural network

U-Net is a Deep Convolutional neural network model which is mainly used for semantic segmentation of Image Data. It is a pre-defined architecture in which the base architecture can be any CNN model. In U-Net there are two phases which are down-sampling and up-sampling. They are also called as Encoding and Decoding. In the down sampling or Encoding phase the features in the image are extracted by using the convolutional layers. In the up sampling or the Decoding phase an image is created based on the extracted features. The detailed U-Net architecture is shown in Figure 2.

Figure 2 shows an input image is given into the model. The dimensions of the input image are 160 x 160 x 3. In the down sampling process the first layer of the model has 32, 64, 64 convolutions each of the size 3 x 3 respectively with a stride of 2. Max-pool is applied to the input image in the first layer which converts it into an image of dimensions 40 x 40 x 64. The similar procedure is applied to the input image until it reaches the dimensions of 10 x 10 x 256 with an intermediate reduction of 40 x 40 x 128 and 20 x 20 x 256 respectively. The output of the down sampling method is sent into the first layer of up sampling where the image of the dimensions 10 x 10 x 256 is decoded into 20 x 20 x 256. The similar procedure is applied until the image has the dimensions of 160 x 160 x 3. The final image of the dimensions 160 x 160 x 3 is the desired filter which is used to segment the input image.

3.5 U-Net with base as ResNet

In Figure 3, the U-Net is implemented using two ResNet's by connecting the up-sampling and down-sampling networks. By making this as a base architecture, for down-sampling a total of 3 skip connections and for up-sampling a total of 4 skip connections are used. An input image is induced into the model. The dimensions of the input image are 128 x 128 x 3. The first layer of the model consists of a filter of a size 3 x 3 and a stride of 2. There are 32, 64, 64 such convolutions persisting in the first layer of the down sampling process. In addition to the convolutions there is a skip connection from the input image to the desired output of the dimensions 128 x 128 x 128. The image is converted into the dimensions 64 x 64 x 128 by encoding it with max-pool. This process of encoding is done until the resultant feature vector is of size 16 x 16 x 1024.

After achieving the final feature vector, then it will be decoded to form the mask for the input image. The steps involved in this decoding process are first the feature vector is up sampled to 32 x 32 x 512 then using the filters of size 3 x 3 with stride of 2 the feature vector is decoded after that the vector will get up sampled to a size of 64 x 64 x 256. This is repeatedly done until the feature vector will be of size 128 x 128 x 3. This 128x 128 x 3 is the desired segmented output image.

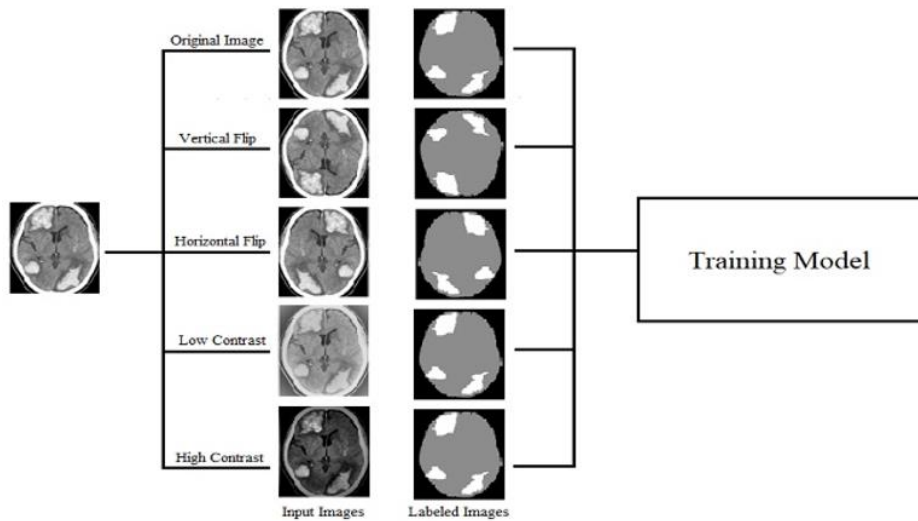


Figure 1. Different forms of data augmentation on brain images

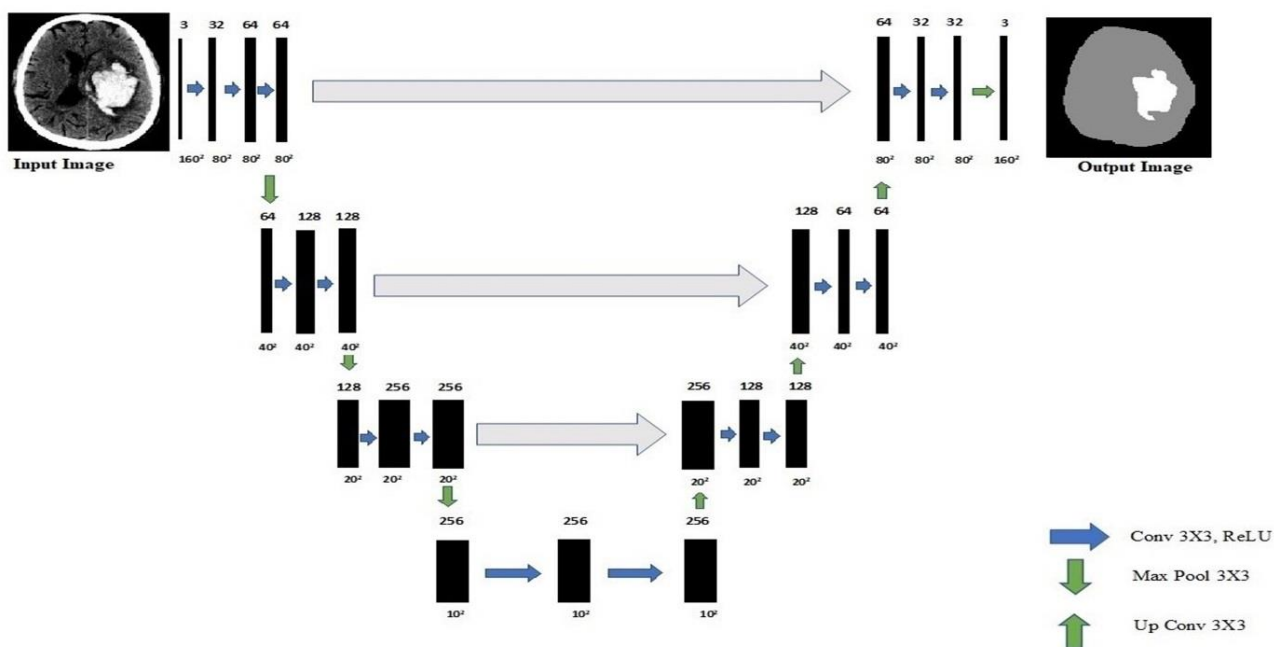


Figure 2. Block of down and up sampling convolutions involved in U-Net architecture with

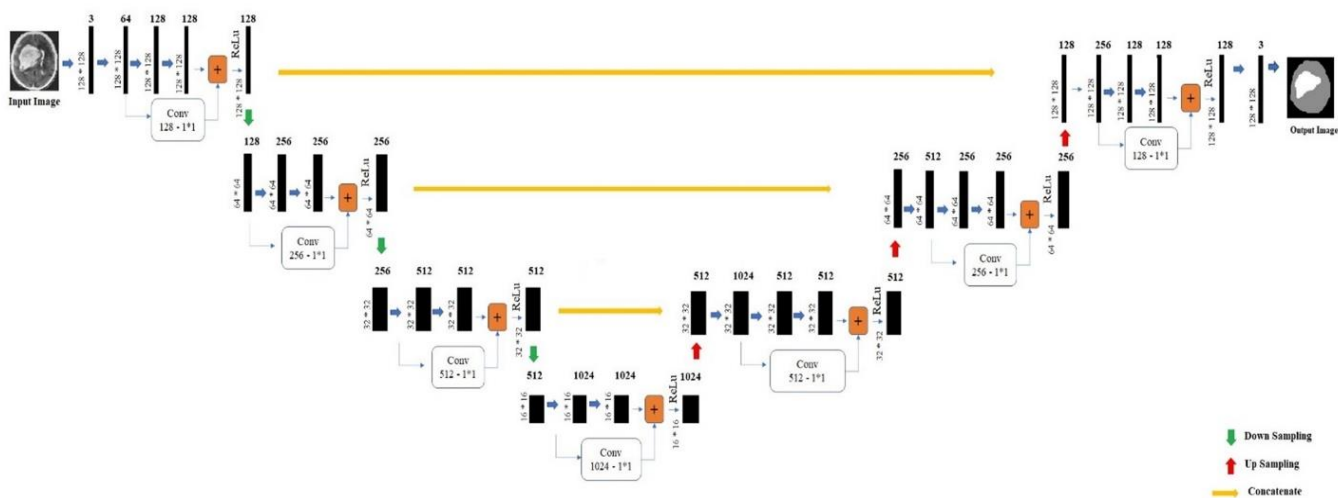


Figure 3. Encoder and decoder convolutions involved in U-Net architecture with base ResNet

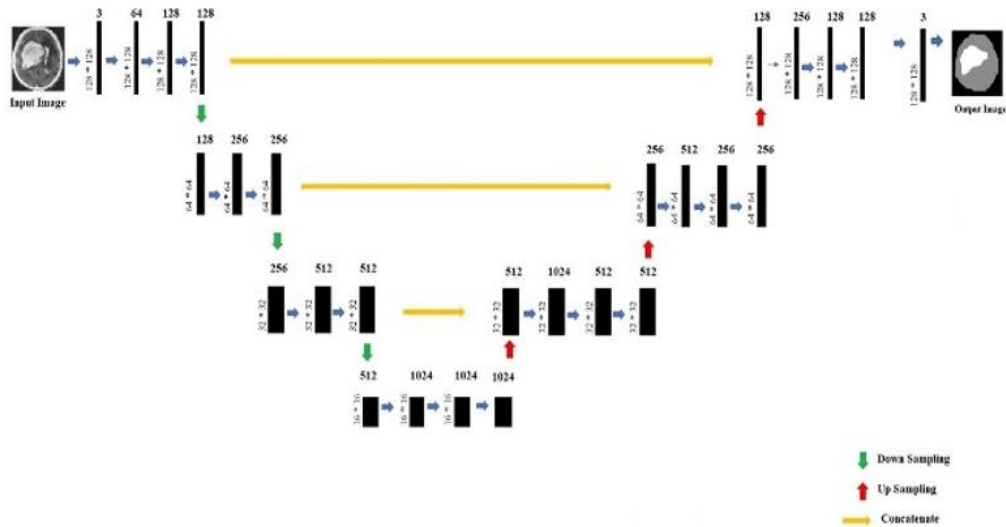


Figure 4. Convolution (Conv) blocks involved in U-Net architecture with base VGG-16

3.6 U-Net with base as VGG-16

Visual Geometric Group (VGG), VGG-16 is the base architecture used in Figure 4. The model is fed with the input given by us. The input images are of the dimensions 128 x 128 x 3. The first layer of the model consists of a filter of a size 3 x 3 and a stride of 2. There are 32, 64, 64 such convolutions persisting in the first layer of the down sampling process. Max Pool is used for down-sampling of the image into the dimensions 64 x 64 x 128. The same process is iteratively done until a feature vector of dimensions 16 x 16 x 1024 is obtained. Decoding of the resultant final vector is iteratively done in the next steps until we get a feature vector of the size 128 x 28 x 3, which is the required output image.

3.7 VGG-16 architecture

VGG is a network with 16 layers proposed by the Visual Geometric Group as shown in Figure 5. It consists of trainable parameters in all the 16 layers and layers like Max pool are also present in between them. VGG-16 has a total of 5 blocks in it. It has max pooling layers preceded by two continuous 2 convolution layers. They are succeeded by 3 blocks of 3 convolution layers tagged with max pooling layers after them. The detailed summary of VGG-16 architecture is available under Table 1.

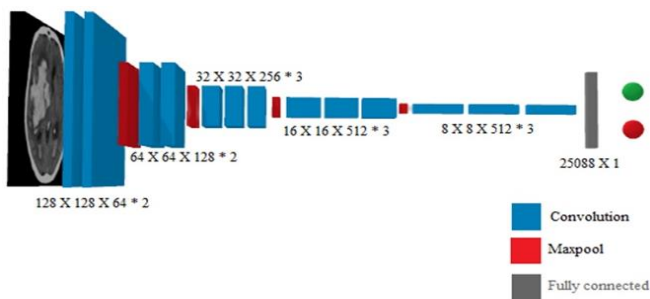


Figure 5. Convolution blocks involved in VGG-16 architecture

VGG-16 takes an input image of the size 128x 128 x 3. There are miniscule receptive convolution kernels of the dimensions 3 x 3 with a stride of 1. Each convolution layer

consists of a stack of these kernels and the images are passed through them to train the parameters. It also consists of Max pool layers of the size 2 x 2 and a stride value of 2. The activation function used in this model is ReLU and it also has an optimizer called Adam.

Table 1. Summary of VGG-16 architecture

Layer Type	Activation Shape	Parameters	Connected to
Input	(128,128,3)	0	-
Conv-1	(128,128,64)	1792	Input
Conv-2	(128,128,64)	36928	Conv-1
Max Pooling-1	(64,64,64)	0	Conv-2
Conv-3	(64,64,128)	73856	Max Pooling-1
Conv-4	(64,64,64)	147584	Conv-3
Max Pooling-2	(32,32,128)	0	Conv-4
Conv-5	(32,32,256)	295168	Max Pooling-2
Conv-6	(32,32,256)	590080	Conv-5
Conv-7	(32,32,256)	590080	Conv-6
Max Pooling-3	(16,16,256)	0	Conv-7
Conv-8	(16,16,512)	1180160	Max Pooling-3
Conv-9	(16,16,512)	2359808	Conv-8
Conv-10	(16,16,512)	2359808	Conv-9
Max Pooling-4	(8,8,512)	0	Conv-10
Conv-11	(8,8,512)	2359808	Max Pooling-4
Conv-12	(8,8,512)	2359808	Conv-11
Conv-13	(8,8,512)	2359808	Conv-12
Max Pooling-5	(4,4,512)	0	Conv-13
Flatten	25088	0	Max Pooling-5
Dense	2	50178	Flatten

3.8 Dataset

The dataset used in this research work is collected from multiple image repositories [25-28]. It includes total of 578 brain CT subjects which contains 463 are stroke images and 115 are non-stroke images.

3.9 Random Forest (RF)

Random Forest is an Ensemble and it consists of a plethora of decision trees. It is a supervised machine learning model, mainly focuses on its accuracy and steady prediction by assembling various decision trees and by combining them simultaneously. The decision trees involved in this Random

Forest [29, 30] give out class predictions and the class with the highest number of votes will be categorized as the model's prediction. The added advantage is that we can eliminate individual errors since there is an involvement of multiple decision trees. The feature importance of every sub tree is calculated by Eq. (1) as follows:

$$RF \text{ for every feature } f_i = \frac{\sum_{\text{subtree } j \in \text{sum of all features of sub tree } i \text{ in every tree } j}^i \text{ every feature of sub tree } i}{\text{Total number of trees}} \quad (1)$$

4. EXPERIMENTAL RESULTS OF PROPOSED MODEL

This section describes experimental results of proposed hybrid architecture, and work flow of VGG-16 architecture with RF model implemented for segmentation of brain stroke.

Figure 6 shows the proposed segmentation model takes brain CT images as input from the available data set. The pre-processing technique is applied on all input images for noise removal, contrast balancing, dimensionality adjustment, etc. The pre-processed data is segregated, out of which 80% is used for training and 20% is used for testing. In the training phase, the convolutions of the VGG-16 model are trained to extract the features to detect the stroke in the brain. The convolutions of dimensions 128x128 each containing 64 filters are extracted from the first layers of the VGG-16 model. These convolutions are applied to the input images to get a feature set of size 64. Using this feature set, the random forest is trained to segment the stroke area. For VGG-16 model we used classification loss function which is categorical cross entropy. The loss function for soft-max is calculated by Eq. (2) as follows:

$$\text{Loss Function} = \left(\frac{e^{sp}}{\sum_i^c e^{si}} \right) \quad (2)$$

where, 'c' is number of classes, 'sp' is score of positive class, and 'si' is score of individual class.

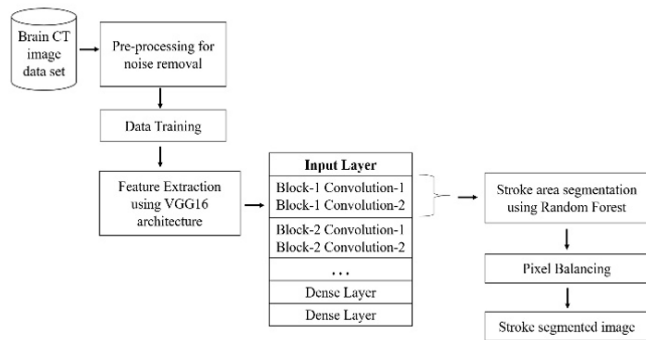


Figure 6. Interconnected layers of convolutional blocks using Random Forest

4.1 Data flow

Figure 7 describes the workflow of the proposed model to detect the brain stroke by using CT images as input. The input image size 128x128x3 is undergone a feature extraction process by the convolutions of sizes 128x128 with 64 filters each. These convolutions are extracted from the first layers of the VGG-16 model which is trained for stroke detection. The output of feature extraction is a feature map containing a set of

64 images. Each image in the feature map is flattened and all these flattened images are combined to create a data frame with 64 columns. The extra label column is added to the data frame by flattening the labeled image. This data frame containing the pixel values of each filter is given to a random forest as an input. The random model is trained based on the pixel values available in data frames to predict the filtered image. The predicted filter is used for segmenting the brain stroke images.

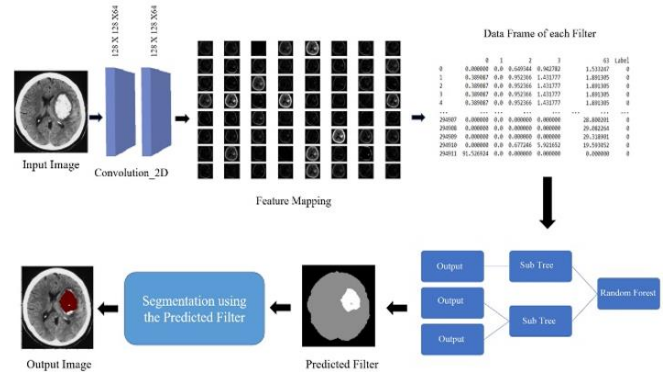


Figure 7. Dataflow of proposed model

4.2 Segmentation using predicted filter

Figure 8 describes the segmentation process using predicted filters. The filtered image consists of three classes. The three-class labeled image is an intermediary output of the network. By using this we are constructing the output image. The class containing the white pixels represents the stroke area, the class containing gray pixels represents the gray matter and the class containing black pixels represents the remaining area of the image. Here the necessity of the three-class labeled image is to differentiate the pixel density of the stroke-affected area and bone area. The class white area of the predicted label is mapped with the input image to generate a stroke segmented image in which the stroke affected is highlighted in red color.

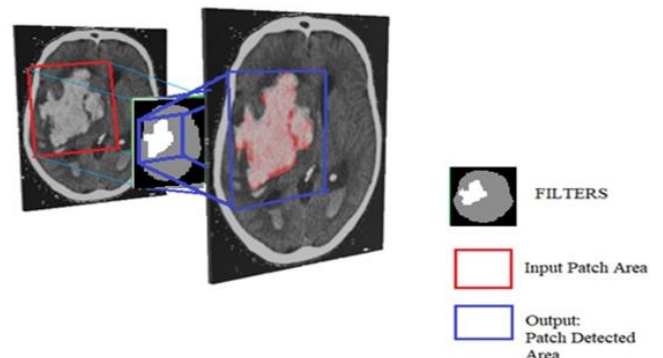


Figure 8. Detected patch area

4.3 Proposed Algorithm

Algorithm 1: Hybrid Segmentation Model

- Step 1: Apply data pre-processing on brain CT images
- Step 2: Segregate stroke and non-stroke images
- Step 3: Train the VGG-16 model on brain CT images
- Step 4: Extract the feature set from the Conv_2D_2 layer of the trained model.

Step 5: Covert the feature set into a Data Frame.

Step 6: Create a Label attribute in Data Frame by Flattening the Labeled Image.

Step 7: Train Random Forest Algorithm with Decision Tree as a base estimator on Data Frame is calculated by Eq. (3) as follows:

$$Gini\ Index\ (G) = \sum_{i=1}^k \hat{P}_{mk} (1 - \hat{P}_{mk}) \quad (3)$$

where, ‘m’ is number of leaf nodes, ‘k’ defines classes, and the node which is having more ‘G’ value become the root node.

Step 8: Segment the brain CT image using the predicted filter.

5. RESULTS AND DISCUSSIONS

This section describes the evaluation metrics to measure and compare the performance of the proposed model with existing models. The results of stroke and non-stroke patients to detect the brain stroke using segmentation. Our research work is implemented on a system with configuration includes: ‘Intel i3 9th generation processor’ with ‘8 GB RAM’.

5.1 Performance evaluation metrics

5.1.1 Accuracy

Accuracy is used for evaluating the number of true predictions to its total number of predictions is calculated by Eq. (4) as follows:

$$Accuracy = \frac{Sum\ of\ true\ predictions}{Total\ number\ of\ predictions} \quad (4)$$

5.1.2 Dice Coefficient (DC)

The dice coefficient is an evaluation metric for image segmentation. It is measured by using Intersection over Union (IoU) of each class in the predicted image with labeled image and is calculated by Eq. (5) as follows:

$$DC(X, Y) = \frac{\sum_{i=1}^N 2 * \frac{|X_i \cap Y_i|}{|X_i| + |Y_i|}}{N} \quad (5)$$

where, ‘X’ defines the predicted image and ‘Y’ defines the labeled image and ‘N’ represents number of classes. ‘Xi’ represents the pixels of ‘class-i’ in predicted image and ‘Yi’ represents the pixels of ‘class-i’ in labeled image. |Xi| represents number of pixels of ‘class-i’ in predicted image and |Yi| represents number of pixels of ‘class-i’ in labeled image.

5.2 Performance of proposed model with existing models

Figure 9 shows the pre-processed brain CT images of 8 patients, in which 6 patients are affected with stroke and remaining 2 are non-stroke patients. Figure 10 shows the labeled brain CT images of 8 patients. These labeled images contain three classes represented with white, gray, and black colors. Patients 4&8 are not affected with stroke, so these images are labeled with only two classes. The images shown in Figure 11 are predicted by our proposed model. Figure 12 shows output images of our proposed stroke segmentation on brain CT images.

The gradient in Receiver Operating Characteristic (ROC) curve determines the true positive rate there by diminishing false positive rate which are used mostly in decimal categorization to rely classifier outcome. On the other hand, Area under the Curve (AUC) is computed in determining capability of a classifier to differentiate among classes and can be utilized as briefing of ROC curve. When AUC is at its peak value, the performance among positive and negative classes of the model is differentiated. The AUC and ROC curve between existing and proposed model is shown in Figure 14.

Figure 13 shows an amalgam of multiple predicted labels alongside their patient numbers. It is clearly evident from the figure that U-Net with base ResNet gives highly accurate segmentation filters but takes the highest time for the whole prediction. The time taken by these models is shown in Table 2. In order to reduce the training time, we use U-Net with base VGG-16 and simple Convolutional neural network model. But this model also requires extremely high computational power, which is practically impossible for regular computers. Our model can approximately achieve the same results as these segmentation models and the dice coefficient values shown in Table 3. But with exponentially less computation power than U-Net that is, it can process and predict with a regular CPU also.

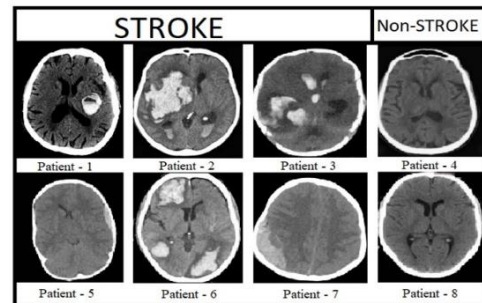


Figure 9. Original images of stroke and non-stroke patients

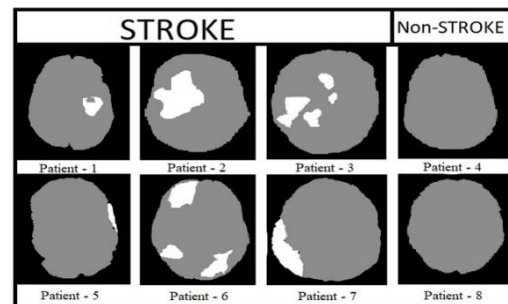


Figure 10. Labeled images of stroke and non-stroke patients

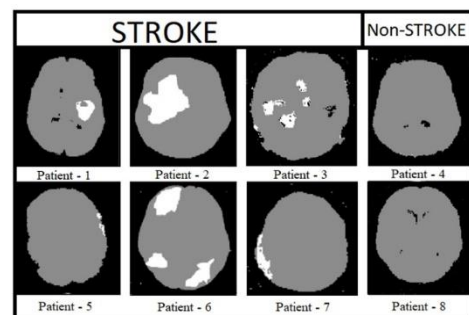


Figure 11. Predicted filter images of stroke and non-stroke patients

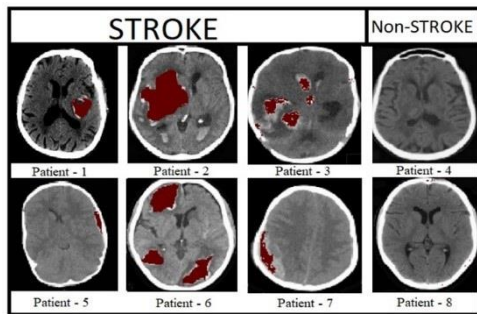


Figure 12. Output images of stroke and non-stroke patients

Table 2 shows the time taken for each algorithm to train is shown. So while observing U-Net with the base as ResNet it took around 36 hours of training time as it contains 50 convolutions for Down-sampling and 50 convolutions for Up-sampling and 4 fully connected network of dimensions 1024*1. In U-Net with base as VGG-16 it took 22 hours with its 16 Down-sampling Convolutions and 16 Up-sampling convolutions and 4 fully connected network with the same 1024*1 dimensions. In U-Net with base as normal CNN it took less time of 14 hours for training as it has least number of convolutions (9 for Down-sampling and 9 for Up-sampling and 3 fully connected network of dimensions 256*1) when compared to ResNet and VGG-16. But with our proposed model, there are only 16 Convolutions and a fully connected layer of 1024*2 along with a Random Forest model. As random forest takes negligible amount of training time when compared to deep learning algorithms the total training time of our proposed model is only 4hrs.

Table 3 shows the Dice coefficients of each algorithm with different input images are shown. Similarly the time taken by the algorithm is proportional to the number of convolutions. The dice coefficients of the algorithms are also proportional to the number of Convolutions. Since with increase in convolutions also increases the feature extraction ability. But in our proposed model the use of Random Forest algorithm to support the VGG-16 model is to increase the feature extraction ability without increase in the number of Convolutions. So the U-Net with base as ResNet has highest average of Dice Coefficients as it has highest number of convolutions, followed by our proposed model. Both U-Net with base VGG and U-Net with base CNN have less average dice coefficient values as they both contains less number of convolution.

Table 2. Comparison of training time

Parameter	U-Net (base=VGG-16)	U-Net (base=ResNet)	U-Net (base=CNN)	Proposed Model (VGG-16+RF)
Training time in hours	22	36	14	4

Table 3. Comparison of dice coefficient values in percentages

Patient	U-Net (base=VGG-16)	U-Net (base=ResNet)	U-Net (base=CNN)	Proposed Model (VGG-16+RF)
A	65.07	68.32	61.30	67.47
B	71.30	70.99	70.01	71.78
C	72.83	74.89	71.16	72.92
D	65.18	69.31	61.39	67.28

Table 4. Accuracy of proposed model with existing models

Model	AUC	Accuracy
U-Net (base = VGG-16)	0.91	88.93
U-Net (base = ResNet)	0.87	94.36
U-Net (base = CNN)	0.93	83.27
Proposed Model (VGG-16 + RF)	0.96	97.81

The performance comparison between existing models with proposed model is shown in Table 4. From this, it is observed that when compared with various existing U-Net models metrics, our proposed model shows AUC with 0.96 and accuracy with 97.81%. Figure 14 shows AUC metrics of the proposed model with existing models.

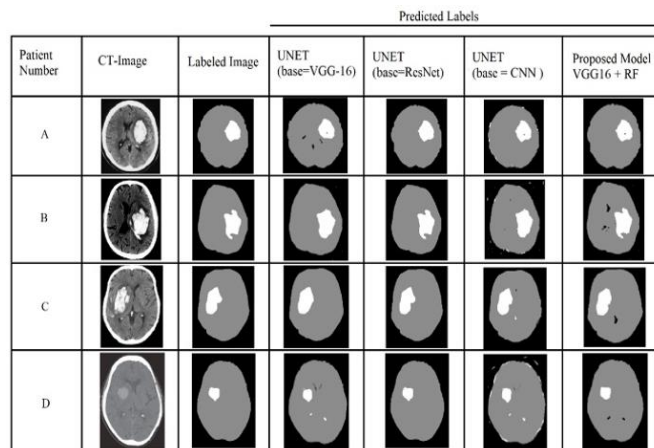


Figure 13. Comparison results of existing models with proposed model

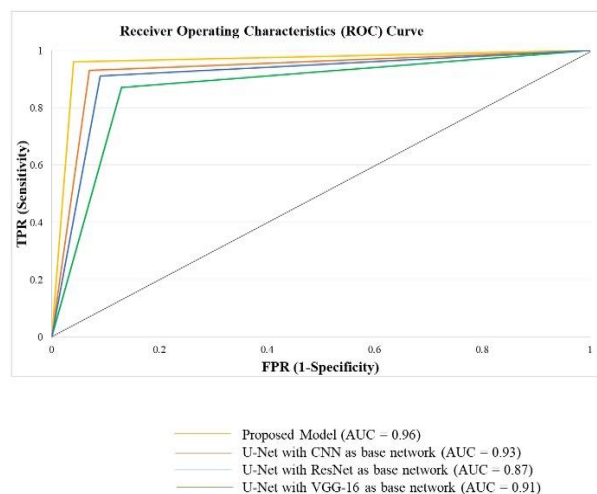


Figure 14. AUC and ROC performance metrics

6. CONCLUSION AND FUTURE SCOPE

In this work, it is observed that the training time in detecting the brain stroke requires high computational power for existing deep learning models. This leads to the major limitation for patients those are affected with a hemorrhage

stroke. Framing this as a main feature, our proposed model is implemented by VGG-16 Architecture combined with Random Forest algorithm. In general, random forest plays a major role in acquiring accurate and balanced outcomes. By using this, our proposed light weight model reduces training time to get accurate results when compared with the existing models. Since the convolution neural networks are prone to translate invariance among the samples requires more time for encoding and decoding. In future work, we aim to choose transfer learning from 2D to 3D light weight Convolutional Neural Network with various Machine Learning models to detect the brain stroke with increased accuracy.

ACKNOWLEDGMENT

This research work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1G1A1100341).

REFERENCES

- [1] Soltanpour, M., Greiner, R., Boulanger, P., Buck, B. (2019). Ischemic stroke lesion prediction in CT perfusion scans using multiple parallel U-nets following by a pixel-level classifier. In 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), pp. 957-963. <https://doi.org/10.1109/BIBE.2019.00179>
- [2] Liu, L., Wu, F.X., Wang, J. (2019). Efficient multi-kernel DCNN with pixel dropout for stroke MRI segmentation. *Neurocomputing*, 350: 117-127. <https://doi.org/10.1016/j.neucom.2019.03.049>
- [3] Park, A., Chute, C., Rajpurkar, P., Lou, J., Ball, R.L., Shpanskaya, K. (2019). Deep learning-assisted diagnosis of cerebral aneurysms using the HeadXNet model. *JAMA Network Open*, 2(6): e195600-e195600. <https://doi.org/10.1001/jamanetworkopen.2019.5600>
- [4] Ginat, D.T. (2020). Analysis of head CT scans flagged by deep learning software for acute intracranial hemorrhage. *Neuroradiology*, 62(3): 335-340. <https://doi.org/10.1007/s00234-019-02330-w>
- [5] Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J.N., Wu, Z., Ding, X. (2020). Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Medical Image Analysis*, 63: 101693. <https://doi.org/10.1016/j.media.2020.101693>
- [6] Bandi, V., Bhattacharyya, D., Midhunchakkravarthy, D. (2020). Prediction of brain stroke severity using machine learning. *Revue d'Intelligence Artificielle*, 34(6): 753-761. <https://doi.org/10.18280/ria.340609>
- [7] <https://algorithmia.com/blog/introduction-to-deep-learning>, accessed on 9th February 2021.
- [8] <https://towardsdatascience.com/a-gentle-introduction-to-deep-learning-part-1-introduction-43eb199b0b9>, accessed on 9th February 2021.
- [9] Dhar, R., Falcone, G.J., Chen, Y., Hamzehloo, A., Kirsch, E.P., Noche, R.B. (2020). Deep learning for automated measurement of hemorrhage and perihematomal edema in supratentorial intracerebral hemorrhage. *Stroke*, 51(2): 648-651. <https://doi.org/10.1161/STROKEAHA.119.027657>
- [10] Schultheiss, M., Noël, P.B., Riederer, I., Thiele, F., Kopp, F.K., Renger, B. (2020). Towards subject-level cerebral infarction classification of CT scans using convolutional networks. *PloS One*, 15(7): e0235765. <https://doi.org/10.1371/journal.pone.0235765>
- [11] Almaslukh, B. (2021). A lightweight deep learning-based pneumonia detection approach for energy-efficient medical systems. *Wireless Communications and Mobile Computing*, 2021: 1-14. <https://doi.org/10.1155/2021/5556635>
- [12] Li, L., Wei, M., Liu, B., Atchaneeyasakul, K., Zhou, F., Pan, Z. (2020). Deep learning for hemorrhagic lesion detection and segmentation on brain CT images. *IEEE Journal of Biomedical and Health Informatics*, 25(5): 1646-1659. <https://doi.org/10.1109/jbhi.2020.3028243>
- [13] Ibtehaz, N., Rahman, M.S. (2020). MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*, 121: 74-87. <https://doi.org/10.1016/j.neunet.2019.08.025>
- [14] Kuang, Z., Deng, X., Yu, L., Wang, H., Li, T., Wang, S. (2020). Ψ -net: Focusing on the border areas of intracerebral hemorrhage on CT images. *Computer Methods and Programs in Biomedicine*, 194: 105546. <https://doi.org/10.1016/j.cmpb.2020.105546>
- [15] Shi, W., Liu, H. (2019). Modified U-net architecture for ischemic stroke lesion segmentation and detection. In 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 1: 1068-1071. <https://doi.org/10.1109/IAEAC47372.2019.8997642>
- [16] Naseer, A., Yasir, T., Azhar, A., Shakeel, T., Zafar, K. (2021). Computer-aided brain tumor diagnosis: Performance evaluation of deep learner CNN using augmented brain MRI. *International Journal of Biomedical Imaging*, 2021: 1-11. <https://doi.org/10.1155/2021/5513500>
- [17] Xu, Y., Holanda, G., Fabrício, L., de F. Souza, L., Silva, H., Gomes, A., Silva, I., Ferreira, M., Jia, C., Han, T., de Albuquerque, V.H.C., Rebouças Filho, P.P. (2020). Deep learning-enhanced Internet of medical things to analyze brain CT scans of hemorrhagic stroke patients: A new approach. *IEEE Sensors Journal*. <https://doi.org/10.1109/jsen.2020.3032897>
- [18] Cui, K., Fu, P., Li, Y., Lin, Y. (2021). Bayesian fully convolutional networks for brain image registration. *Journal of Healthcare Engineering*, 2021: 1-12. <https://doi.org/10.1155/2021/5528160>
- [19] Xue, Y., Farhat, F.G., Boukrina, O., Barrett, A.M., Binder, J.R., Roshan, U.W., Graves, W.W. (2020). A multi-path 2.5 dimensional convolutional neural network system for segmenting stroke lesions in brain MRI images. *NeuroImage: Clinical*, 25: 102118. <https://doi.org/10.1016/j.nicl.2019.102118>
- [20] Liu, L., Kurgan, L., Wu, F.X., Wang, J. (2020). Attention convolutional neural network for accurate segmentation and quantification of lesions in ischemic stroke disease. *Medical Image Analysis*, 65: 101791. <https://doi.org/10.1016/j.media.2020.101791>
- [21] Hu, K., Chen, K., He, X., Zhang, Y., Chen, Z., Li, X., Gao, X. (2020). Automatic segmentation of intracerebral hemorrhage in CT images using encoder-decoder convolutional neural network. *Information Processing & Management*, 57(6): 102352. <https://doi.org/10.1016/j.ipm.2020.102352>
- [22] Chen, L., Bentley, P., Mori, K., Misawa, K., Fujiwara,

- M., Rueckert, D. (2018). DRINet for medical image segmentation. *IEEE Transactions on Medical Imaging*, 37(11): 2453-2462. <https://doi.org/10.1109/TMI.2018.2835303>
- [23] Liu, L., Wu, F.X., Wang, Y.P., Wang, J. (2020). Multi-receptive-field CNN for semantic segmentation of medical images. *IEEE Journal of Biomedical and Health Informatics*, 24(11): 3215-3225. <https://doi.org/10.1109/JBHI.2020.3016306>
- [24] Guo, D., Wei, H., Zhao, P., Pan, Y., Yang, H.Y., Wang, X. (2020). Simultaneous classification and segmentation of intracranial hemorrhage using a fully convolutional neural network. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 118-121. <https://doi.org/10.1109/ISBI45749.2020.9098596>
- [25] Kuang, H., Najm, M., Chakraborty, D., Maraj, N., Sohn, S.I., Goyal, M. (2019). Automated ASPECTS on noncontrast CT scans in patients with acute ischemic stroke using machine learning. *American Journal of Neuroradiology*, 40(1): 33-38. <https://doi.org/10.3174/ajnr.A5889>
- [26] Hssayeni, M. (2020). Computed tomography images for intracranial hemorrhage detection and segmentation. *PhysioNet*. <https://doi.org/10.13026/4nae-zg36>
- [27] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23): e215-e220. <https://doi.org/10.1161/01.CIR.101.23.e215>
- [28] Vamsi, B., Debnath, B., Kiran, V. (2021). *Brain_Stroke CT-Images*. Mendeley Data. <https://doi.org/10.17632/363csnhzmd.1>
- [29] Krishnaveni, P.R., Kishore, G.N. (2020). Image based group classifier for brain tumor detection using machine learning technique. *Traitement du Signal*, 37(5): 865-871. <https://doi.org/10.18280/ts.370520>
- [30] Vamsi, B., Bhattacharyya, D. (2021). Detection of brain stroke based on the family history using machine learning techniques. In: Saha S.K., Pang P.S., Bhattacharyya D. (eds) *Smart Technologies in Data Science and Communication. Lecture Notes in Networks and Systems*, vol 210. Springer, Singapore. https://doi.org/10.1007/978-981-16-1773-7_2