International Journal of Future Generation Communication and Networking
Vol. 13, No. 4, (2020), pp. 2694–2715

An Automated Classification Of Intracranial Haemorrhage Using Deep Convolution Neural Network Model

R. Aruna Kirithika[1], S. Sathiya[2], M. Balasubramanian[3],
P. Sivaraj[4]
1PhD scholar in Computer and Information Science,
2Assistant Professor in Computer Science Enginerring,
3Associate Professor in Computer Science Enginerring,
4Associate Professor in Manufacturing Engineering
1,2,3,4Annamalai University, Tamilnadu, India.
E-mail :1ashassss79@gmail.com, 2sathiya.sep05@gmail.com, 3balu.june1@gmail.com, 4cemajorsiva@gmail.com

Abstract

Traumatic brain injury might result to intracranial haemorrhage (ICH). The ICH could become a major disability or mortality when it is not precisely and timely diagnosed at the earlier stage. Due to the advanced developments in the deep learning models, automated medical diagnosis models can be developed to solve complicated decision making problem in healthcare sector. Keeping this in mind, this paper presents a new automated DL based segmentation and classification model for ICH diagnosis.The proposed method initially undergoes a set of preprocessing technique to improve the quality of the input images. Besides, instance segmentation model using DL based Depthwise Separable Network is employed to perform the segmentation process, called ISM-DL, thereby the injured regions in the brain can be identified. The proposed model also uses scale-invariant feature transform (SIFT) and residual network (ResNet) are used for feature extraction process. At last, a set of three machine learning (ML) classifiers namely logistic regression (LR), multilayer perceptron (MLP), and gradient boosting tree (GBT) are employed to determine the appropriate class labels of ICH. The performance of the proposed models are evaluated on the benchmark ICH dataset and the experimental outcome stated that the proposed model outperformed the compared methods under different aspects. Accuracy rate of 96.95% was achieved from the proposed methods using convolutional Neural Network-residual network (CNNRN).

Keywords: *Intracranial hemorrhage, Deep learning, Feature extraction, Image segmentation, Machine learning*

1. Introduction

IntracerebralHemorrhage (ICH) is a type of dangerous brain stroke that results in higher mortality and morbidity [1]. The patients affected with hemorrhage also suffer from various other diseases where the brain tissues are defected that results in edema[2]. Edema is caused because of blood vessel leakage which interrupts the nerve cell communication where the functions like speech, eyesight, memory, movement of the body are completely affected. Furthermore, the additional risk factors of ICH are head trauma, high BP, infected blood vessel walls, vein leakage, and so on. These defects are examined by various imaging modalities such as X-ray, magnetic resonance imaging (MRI), Computerized Tomography (CT), Positron Emission Tomography (PET), Single-photon emission computed tomography (SPECT) where brain hemorrhage is screened. Among them, CT scan is one of the well-known and effective methods applied for hemorrhage prediction as it is cost-effective, widely accessible, limited scanning duration for imaging. Thus, CT scan is preferred by most of the users for ICH diagnosis, surgical plans, and prominent observation of ICH patients. In past decades, radiologists view the hematoma manually by delineating the CT scan and measure the amount of tumor volume which helps in predicting the health condition of a patient. Unfortunately, it is a lengthy process correlated with inter-rater variability and the requirement for well-trained physicians has been increased in conventional system. For computing an accurate quantitative analysis of hematoma, it is essential to conduct précised and automatic segmentation.

ISSN: 2233-7857 IJFGCN

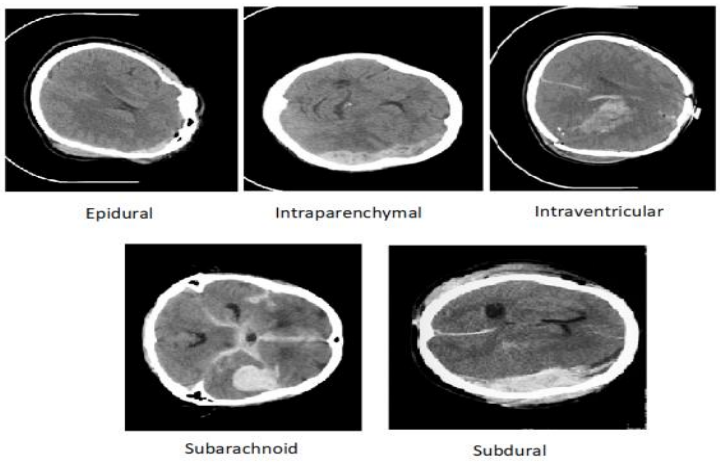


Fig.1.Images of Intracranial Haemorrhage (Epidural, Intraparenchymal.Intraventricular, Subarachnoid, Subdural)

Earlier prediction of ICH is highly important; so that appropriate treatment and sufficient medications could be offered. Hence, developers concentrated in making computer-aided diagnosis (CAD) models for ICH segmentation. Fig.1. shows images of 5 categories of ICH captured through computer tomography. In this approach, computer relied CAD mechanism used in previous ICH segmentation depends upon (a) automated segmentation where hemorrhage is predicted with no human contribution or (b) expert segmentation which requires well-trained physicians for generating best segmentation results. It is evident that massive studies were deployed in manual segmentation and only few works were projected automated ICH segmentation.Nowadays, Deep Learning (DL) related automatic segmentation models have accomplished massive attention from many researchers as it is highly applicable for computing complicated operations robustly with better accuracy. Some of the instances for BT segmentations are ischemic lesion segmentation, lung tumor segmentation, cardiac segmentation, and pancreas segmentation. Specifically, automatic hemorrhage segmentation has gained a memorandum in stroke management by dealing massive number of information which helps the physicians in decision-making process. [3] projected a collection of Deep Neural Networks (DNN) for automatic diagnosis of post-treatment ischemic stroke. In order to overcome this clinical overhead of 3D Ischemic MRI scan, [4] modeled 3D convolutionalneural network (CNN) with massive training approach of neighboring image patches as a single pass whereas automatic application of inherent class mismanagement.

[5] applied 2 CNN modules like DeconvNetsas well as multi-scale convolutional label estimation net for segmentation of acute ischemic lesion from diffusion-weighted MRI (DWI). The above-mentioned methods attempted to gain better results in state of art using various structures like 2D, 3D, and combination of both CNN with hand-engineered features and CRF is served as post-processing. Therefore, none of the above pre-defined models were employed on CT scans for ICH segmentation. Followed by, Recurrent Attention DenseNet(RADnet)applies recurrent attention DenseNet with Long short-term memory (LSTM) for segmentation and classification of brain hemorrhage from CT images in case of Traumatic Brain Injury (TBI).

Fig.2. gives a clear block diagram of the proposed work carried out on the ICH images using both conventional and deep learning models for segmentation, feature extraction and classification techniques. This paper introduces an efficient and automated DL based segmentation and classification model for ICH diagnosis. The proposed method firstly undergoes a set of preprocessing techniques to improve the quality of the input images. Next, instance segmentation model using DL based Depthwise Separable Network called ISM-DLis employed to to segment and identify the injured portions of the brain. The proposed model also uses scale-invariant feature transform (SIFT) and residual network (ResNet) are used for feature extraction process. Finally, three machine learning (ML) classifiers namely logistic regression (LR), multilayer perceptron (MLP), and gradient boosting

ISSN: 2233-7857 IJFGCN

tree (GBT) are employed to find out the proper classes of ICH. The performance of the proposed models are evaluated on the benchmark ICH dataset

2. Related works

[6]introduced 2 models according to CNN. Initially, it concentrates on predicting ICH, mass impact, and hydrocephalus in a scan level whereas alternate methods focused in predicting malicious acute infarcts. The samples are validated with consequent radiology report for equivocal findings. Hence, affected acute infarct prediction approach has provided minimum specificity and area under the curve(AUC). [7]developed 4 methods for predicting the sub-classes of ICH, calvarial fractures, midline shift, as well as mass effect. These models are verified on a large-scale dataset with numerous number of CT scans, correspondingly. Only 2 datasets have been applied for testing. Distinct deep models have been presented for 4 prediction classes. First, ResNet18 undergoes training with 5 corresponding Fully Connected (FC) layers. Finally, the simulation outcome shows that a slice has been projected into Random Forest (RF) approach for predicting the confidence of scan-level in ICH forecasting.

These 2 modules were relied on CNN along with Recurrent Neural Network (RNN) as presented for ICH prediction. [8] deployed a 40-layer CNN termed as DenseNet, with a bidirectional long short-term memory (BiLSTM) layer for computing ICH examination. Moreover, 3 auxiliary operations were performed after dense convolutional block in order to perform binary segmentation of ICH. These operations are composed of convolutional layer and deconvolution layer for up-sampling the feature maps into actual image size. Then, LSTM layer has been incorporated with inter-slice dependencies of CT scans. Therefore, training data is augmented using rotation and horizontal flipping in order to manage the count of scans for two categories. Hence, network prediction of test data is tested over the annotation of 3 expert physicians in all CT slice. The inclusion of attention layers has resulted in magnified sensitivity values.

In [9] implied a 3D joint convolutional and Recurrent Neural Network (CNN-RNN) for the purpose of prediction and classification of ICH sites. VGG-16 has been employed as CNN mechanism while bidirectional Gated Recurrent Unit (GRU) was utilized as RNN model. Thus, RNN layer performs the similar function of slice interpolation method projected by [10], but is applicable interms of neighboring slices added in classification. Consequently, it is trained and verified and sampled on massive amount of CT scans.

3. The Proposed Model

The workflow involved in the proposed model is depicted in Fig. 1. As depicted, the input images are preprocessed to raise the input quality by removing noise and increasing the contrast level. Followed by, image segmentation process takes place by a DL model. In addition, two feature extractors namely SIFT and ResNet 101 are employed to extract useful set of feature vectors. At last, three different ML models are applied to perform classification process.

3.1. Preprocessing

Initially, the input image is preprocessed in 3 levels of skull stripping, noise removal, and contrast enhancement. In general, it can be required for removing the skull region from the background region to optimal clarity. Then, bilateral filtering (BF) method is performed on the image to discard some noise exists in it. Also, contrast limited adaptive histogram equalization (CLAHE) method is utilized for increasing the contrast level of the applied image. The preprocessing has been carried out in previous module and the same has been published in [26]. A vivid explain about the preprocessing step has been done in prior module of the work done by the authors earlier.

3.2. DL based Segmentation Process

Once the images are preprocessed, DL based segmentation process takes place to determine the affection regions in the image. This model [11] is composed of the features of low attributes and

ISSN: 2233-7857 IJFGCN

processing cost when compared with remarkable convolution process. The key objective of depth-wise separable convolution to degrade the better convolution integral as depth-wise convolution as well as point-wise convolution.

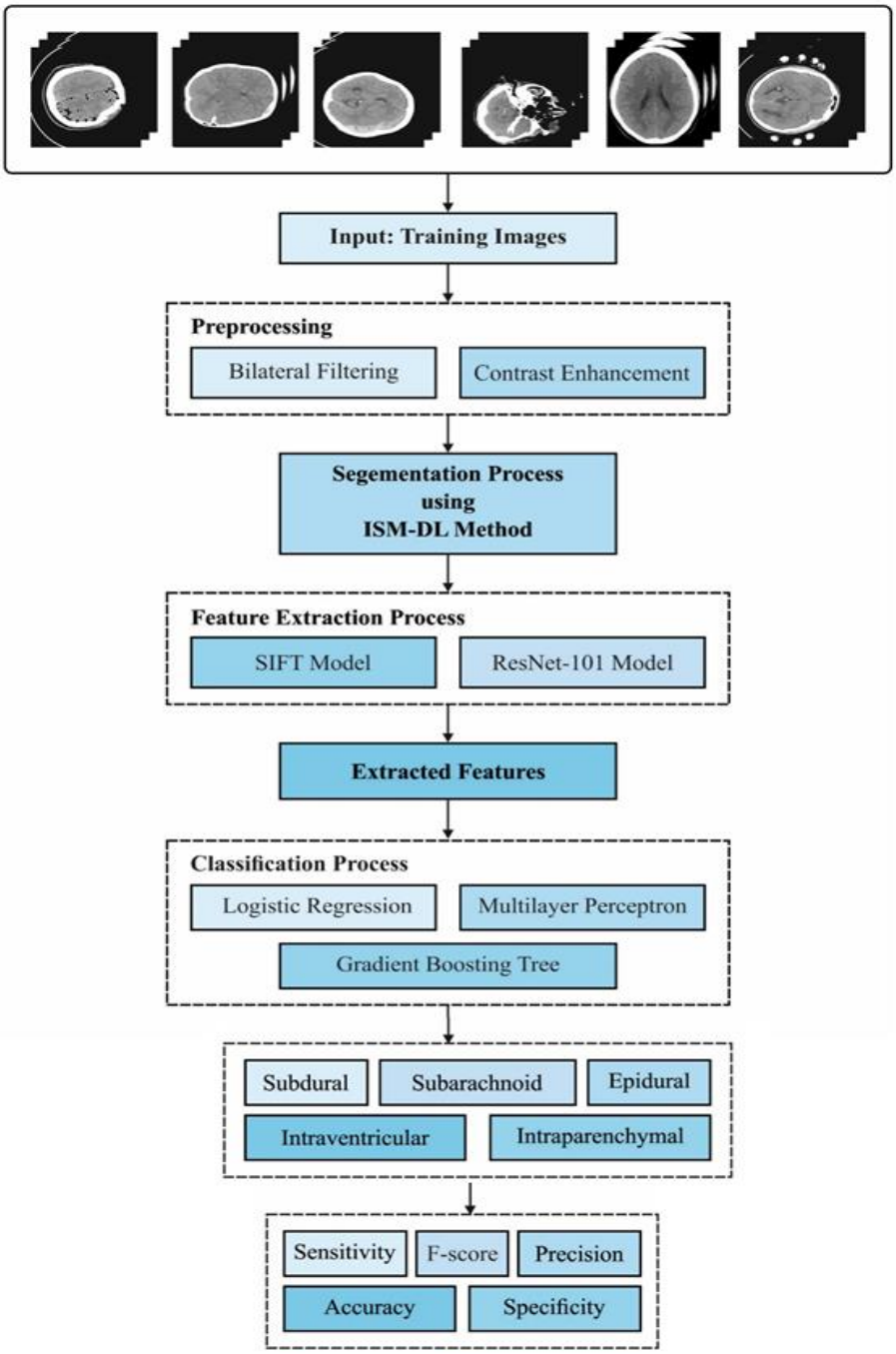


Fig. 2. Overall process of proposed method

Hence, comparison is performed among depth-wise separation convolution and standard convolution.Assume the input volume with width and height D_f, and count of incoming channels M. When a color image is considered as an input, then$M$ is symmetrical to 3 RGB channels. For standard

ISSN: 2233-7857 IJFGCN

convolution, usage of filters over input channels and integration of these measures are performed in single step. Additionally, N convolution kernels of structure $D_k * D_k * M$ which is employed as input in standard CNN, the final volume might be $D_g * D_g * N$. The cost of this convolution mechanism is $N * D_k^2 * D_g^2 * M$. Considering the similar input volume for comparing depth-wise separable convolution divides the convolution into 2 portions depth-wise convolution and point-wise convolution. Initially, depthwise convolution uses convolution for a single input channel simultaneously. Hence, convolution kernel of structure $D_k * D_k * 1$ has been utilized for a single input channel in depthwise convolution where M convolution kernels are essential for input volume. Under the application of M outputs from these convolutions jointly, a resultant volume with shape of $D_g * D_g * M$ has been considered [12]. Finally, depth-wise convolution would be effective in point-wise convolution that contributes in linear integration of every layer. Basically, a layer $1 * 1$ convolution operation is performed in M layers. Moreover, N filters have the output similarly as standard convolution $D_g * D_g * N$. The overall cost of these 2 stages might be $M * D_k^2 * D_g^2 + M * D_g^2 * N$ that is $(D_k^2 + N) * M * D_g^2$. The impact of depth-wise separable convolution is expressed in the following:

$$\frac{\left(D_k^2 + N\right) * M * D_g^2.}{D_k^2 * M * N * D_g^2.} = \frac{1}{N} + \frac{1}{D_k^2}. \quad (1)$$

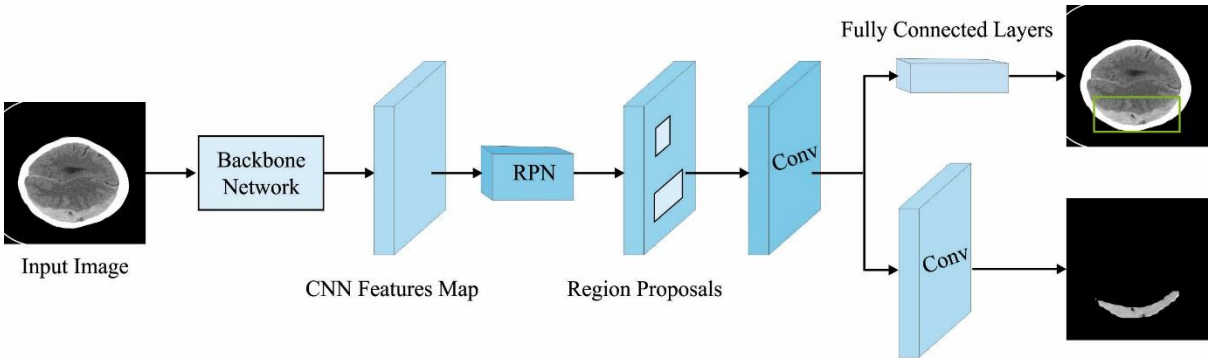


Fig. 3. Structure of ISM-DL

For sample, assume the final feature volume N as well as kernel of size, and accomplish better ratio. Thus, the processing resources are essential for depth-wise separable convolution which is minimum than standard convolution. Fig. 3 shows the structure of ISM-DL model.

3.3. Feature extraction Process

For the extraction of useful sets of feature vectors from the segmented image, SIFT and ResNet101 models are employed.The SIFT detector is used for extracting the useful descriptive image features such as invariant for scaling, rotation, and illumination. Also, these points are projected in the high-contrast region, especially on the object edges. One of the essential attributes regarding these features are that therelevant positions among them should remain the same without any modification.

3.3.1 Scale Invariant Feature Transform

The major phases in SIFT feature extraction is defined in the following:

- The initial phase is scale-space extrema extraction: here, the interest points are scale as well as rotation invariant have been explored. The difference of Gaussian (DoG) function has been employed.
- Followed by, a key point localization as well as filtering is carried out. Here, position and scale for output interest points are identified. Key points are decided as it is fast in image distortion.

ISSN: 2233-7857 IJFGCN

- Then, Orientation Assignment is performed where maximum orientation is allocated for every key point position according to the local image-gradient directions.
- Finally, feature description has been performed. Local image gradients are estimated at decided scale in a neighborhood of a key point. Also, 128D feature descriptor has been attained.

In past decades, CNN has gained maximum concentration from most of the developers due to its effective performance on image classification. It is combined with Transfer Learning (TL) and hyper-parameter tuning [13]. Moreover, AlexNet, VggNet, GoogleNet, ResNet have been employed in this literature for advanced Deep CNNs (DCNN), and TL is performed to show the classification of MRI datasets. The studies based on ICH prediction is provided in this literature. Also, TL is a technique used in FC layer of existing CNN trained are deleted and viewed as a feature extractor. The attributes of hyper-parameter-tuning scheme is not initialized by a system which is essential for tuning and optimizing the attributes on the basis of simulation outcome attained from MRI computation and make the performance more viable. Fig.4 shows the keypoints that has been extracted by the SIFT feature extraction algorithm used on various ICH images.

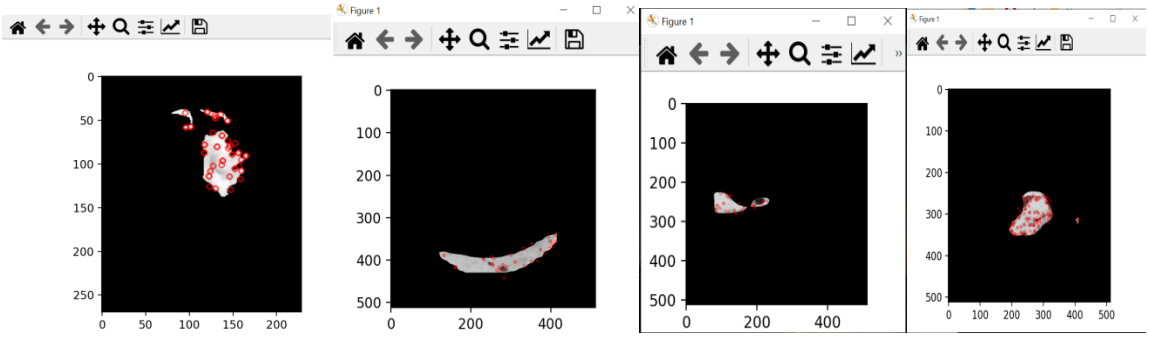


Fig 4 SIFT keypoints

3.3.2 ResNet101

TheResNet depends upon deep structures with optimal convergence behaviors and better accuracy as introduced by [14]. ResNet is manufactured by numerous stacked residual units and massive layers. Therefore, the count of operations differs on the basis of various structures. From the above, residual units are comprised of convolutional, pooling layers. With respect to Residual Unit, it trains numerous DNN for accomplishing the ILSVRC 2015 championship and gain limited error rate classification for first 5 categories, that is frequent in count of variables which is minimum than VGGNet. The main aim of ResNet, HighWay Nets, applies the skip connection for inducing few input as layer indiscriminately for combining the data flow to eliminate the data loss to convert the layer and gradient diminishing issues. Moreover, reducing noise means averaging and it manages both training accuracy as well as generalization.

One of the significant ways is to enhance the label data and to accomplish maximum training accuracy as well as précised level of traversal. The ResNet architecture is significant to stimulate the training for ultra DNN and to maximize the method's accuracy. The accuracy is considered and leads in massive inclination whereas the depth is increased and limits the accuracy. It also tends to make over-fitting issues as the error is maximized in test samples; however, the training examples are also incorporated. When a shallow network satisfies the accuracy of saturation and congruent mapping layers, then the error might be reduced and deeper system tends to make training examples. The principle behind the existing output is to use congruent mapping from ResNet.

ResNet applies a residual block for resolving decomposition as well as gradient diminishing issues generally occurs in typical CNN. Next, residual block deepens the network and improvises the function of newly developed approach. ResNet networks are flexible, effective in the ImageNet [15] classification process. The function of residual function is depicted as given below:

ISSN: 2233-7857 IJFGCN

$$y = F(x, W) + x \quad (2)$$

where$x$ denotes the input of residual block; W represents the weight of residual block; y means the simulation outcome of residual block. The architecture of a block is projected in Fig.4.

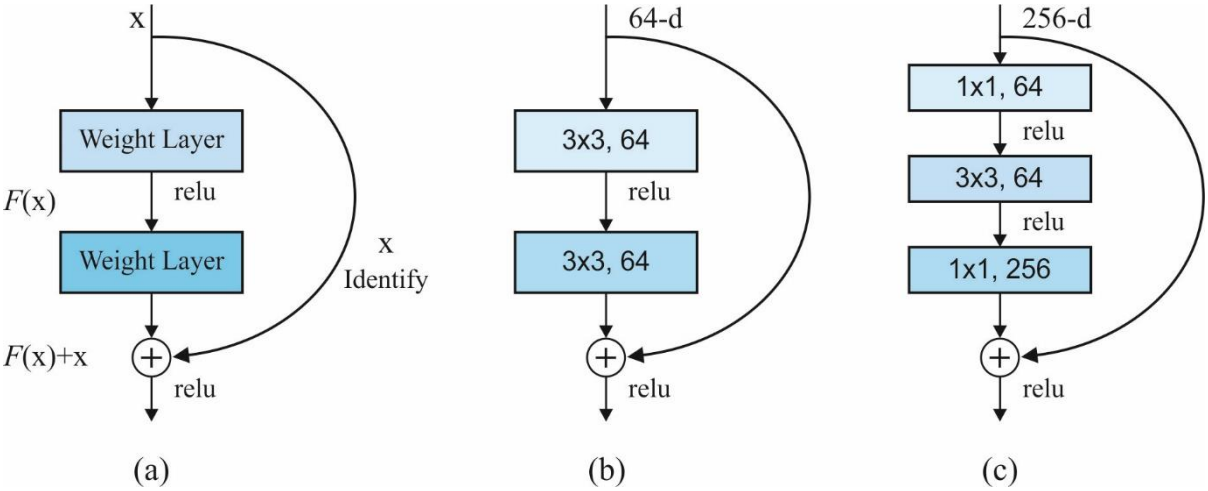


Fig.5.a) Residual Block b) Two Layer Deep c) Three Layer Deep

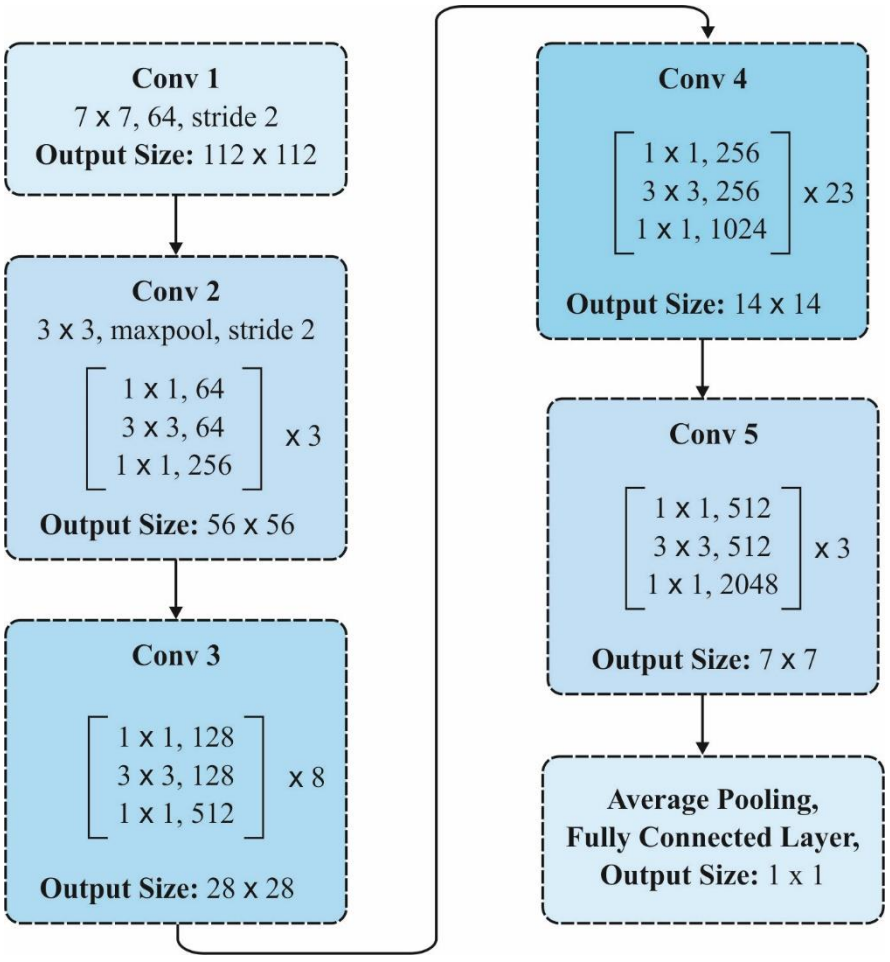


Fig.6. Architecture of ResNet-101

ResNet network is comprised of various residual blocks where convolution kernel size of convolution layer differs. The traditional architecture of ResNet is RetNet18, RestNet50, and RestNet101. A

ISSN: 2233-7857 IJFGCN

feature obtained by ResNet is suited in FC layer for the purpose of computing image classification. Fig. 6 illustrates the architecture of ResNet-101 model. In general, softmax classifier is applied for classification. In this case, three models namely MLP, LR, and GBT are used for classification process.

3.4. Image Classification

During the image classification process, three ML models namely LR, MLP, and GBT models are employed to identify the distinct class labels of the input image provided.

3.4.1. LR model

LR is a commonly employed binary classification model. In LR model, a single outcome parameter, $y_i(i = 1, \dots, n)$, is coded as 1 with probability p_i and 0 in conjunction with probability $1 - p_i$. Followed by, p_i differs in a function of explanatory parameters like x_i which is numerically represented as given below as [16]:

$$E[y_i|x_i, \beta] = p_i = \frac{e^{x_i\beta}}{1 + e^{x_i\beta'}}, \tag{3}$$

whereβ denotes a vector of attributes with consideration of $X_{i0} = 1$; thus, the intercept β_0 shows a constant term. Therefore, an assumption with an intercept as added from vector β is repeated until reaching a termination condition. Next, logistic (logit) transformation is defined as a logarithm of positive response which is illustrated as [17],

$$\mathfrak{y}_i = \ln\left(\frac{p_i}{1 - p_i}\right) = x_i\beta. \tag{4}$$

The logit function is depicted as matrix form as given below,

$$\mathfrak{y} = X\beta. \tag{5}$$

The regularized log likelihood is demonstrated as

$$\ln L(\beta) = \sum_{i=1}^{N} (y_i \ln p_i + (1 - y_i) \ln (1 - p_i)) - \frac{\lambda}{2}\|\beta\|^2, \tag{6}$$

where regularization (penalty) term $(\lambda/2)\|\beta\|^2$ is included for accomplishing better generalization.

In case of binary results, the loss function or deviance DEV is a negative log −likelihood which is represented as,

$$DEV\ (\hat{\beta}) = -2 \ln L(\beta). \tag{7}$$

3.4.2. MLP model

In general, MLPsystem is composed of input, hidden, and output layers. The MLPcontains distinct hidden layers which enable the system with processing capabilities for generating the network outputs [18]. Fig. 6 implies the MLP network with hidden layer. It is comprised of weights connecting among layers. Hence, output values are estimated by the given procedures,

ISSN: 2233-7857 IJFGCN

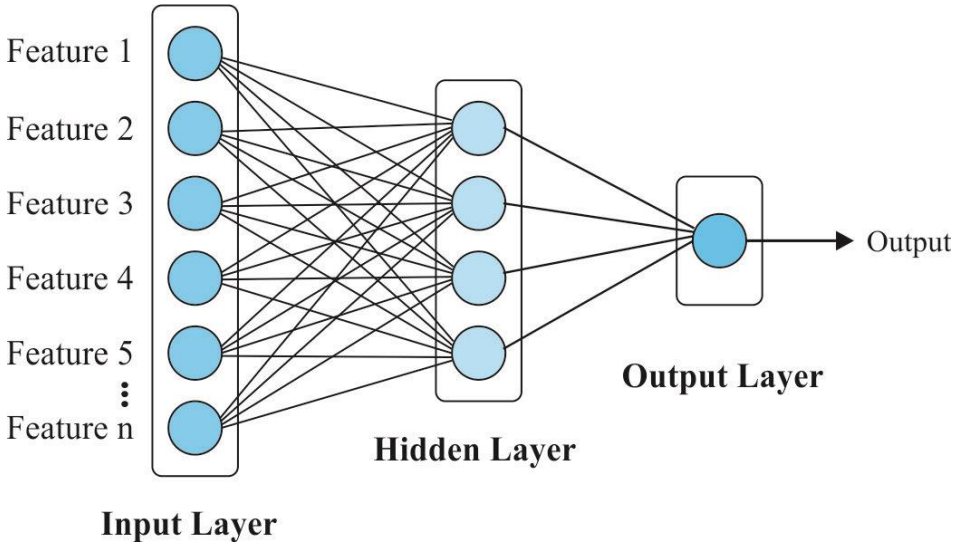


Fig.7. Structure of MLP

Initially, the sum of weights is estimated as given in the following:

$$S_j = \sum_{i=1}^{n} w_{ij} x_i + \beta_i, \tag{8}$$

where$x_i$implies input parameter, w_{ij}defines the weight from input variable x_i and neuron j, as well as $\beta_i$refers to input parameters' bias term.

Secondly, the resultant values of the neurons in the hidden layers are produced from derived values of weighted summation (Eq. (8)) under the application of activation function. The well-known function is defined as a sigmoid function which is expressed as given below:

$$f_j(x) = \frac{1}{1 + e^{-S_j}}, \tag{9}$$

where$f_j$denotes the sigmoid function for neuron j and $S_j$represents the summation of weights.

As a result, the resultant of neuron j is measured in the following:

$$y_j = \sum_{i=1}^{k} w_{ij} f_j + \beta_j, \tag{10}$$

where$y_j$signifies the final outcome of neuron j, w_{ij}defines the weight from final parameter y_i and neuron j, $f_j$defines the activation function for neuron j, and $\beta_i$demonstrates the last variable's bias term.

3.4.3. GBT model

The norm boosting means a collection of models which transform the weak learners into effective learners, knowing that weak learner is moderate over a random selection whereas a strong learner is highly more effective [19]. Thus, GTB is evolved from ML approach which is applied in regression as well as classification issues. Moreover, it generates a predictive mechanism as ensemble from vulnerable predictive models like Decision Trees (DT) as specific approach employed in GBT. It develops various phases like alternate boosting methods;however, it generalizes by optimizing random differentiable loss function.

ISSN: 2233-7857 IJFGCN

GBTis an ML approach which is applied in regression and classification issues. Henceforth, the gradient boosting applies a collection of weak moduleswhere it forms a robust method. Consequently, the final model considers a vector of attributes $x \in R^n$for gaining a measure $F(x) \in R$thus,$F_i(x) = F_{i-1}(x) + \gamma_i h_i(x)$, where $h_i$denotes a function which develops a tree and $\gamma_i \in R$refers the weight correlated with ith tree, thus the 2 terms are learned in a training phase. Followed by, the root causes for applying GBT in contrast to alternate predictive schemes is that it provides best *out-of-the-box*outcomes. Additionally, by assuming the base approach to understand the issues in traditional DT which is assumed as ensemble method.

4. Experimental Evaluation

For experimentation,ICH dataset [20] is employed. The classical dataset is composed of 171 images under epidural class, 24 images under intra-ventricular, 72 images under intra-parenchymal, 56 images under subdural as well as 18 images under subarachnoid classFollowed by, the hemorrhage image size is 512*512 pixels. The sample images from datasets are depicted in Fig. 8.

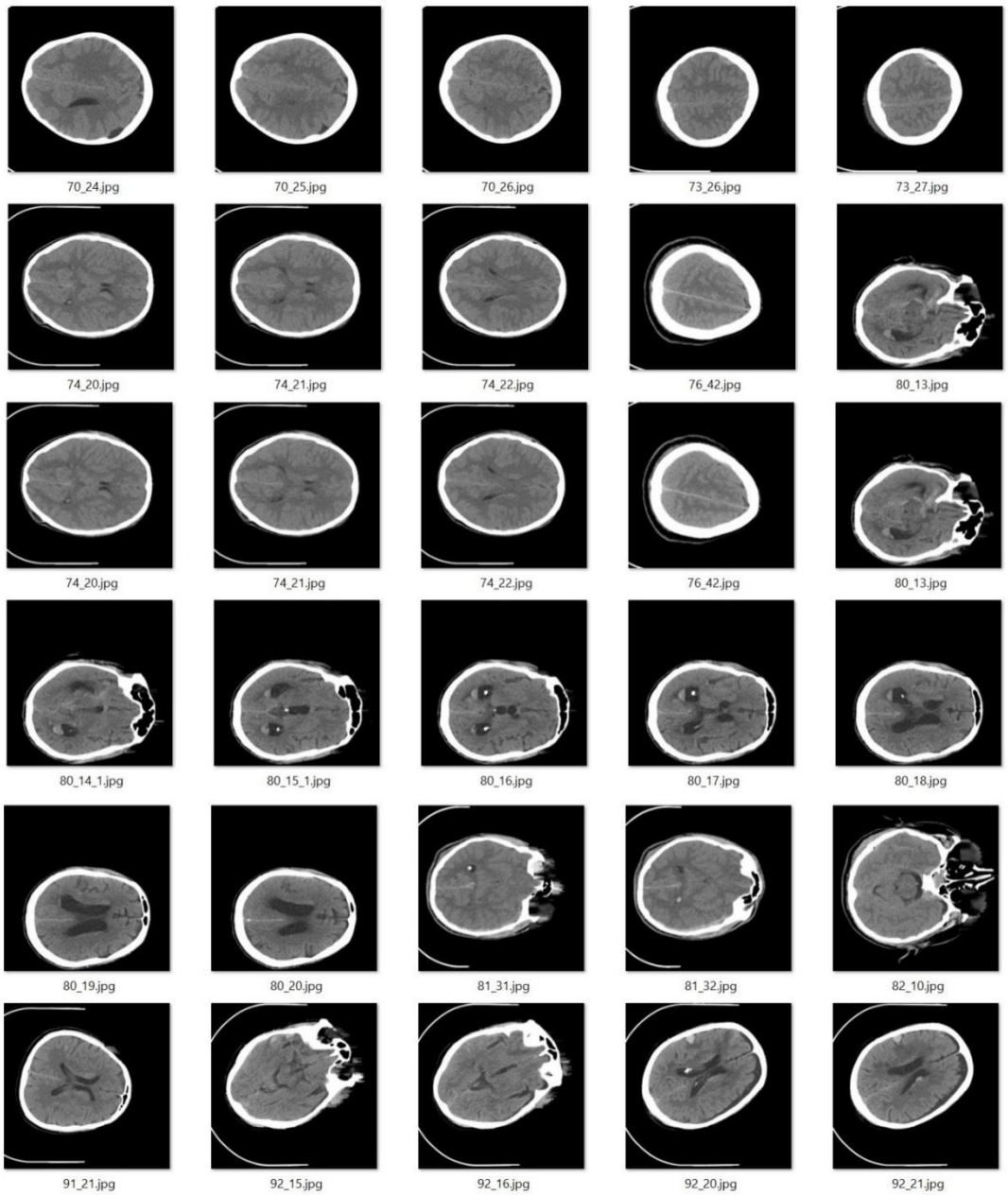
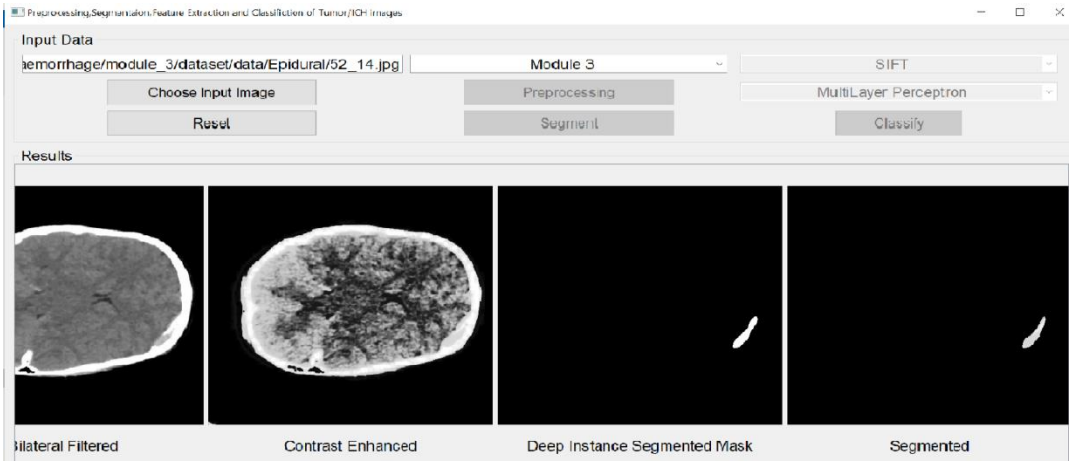


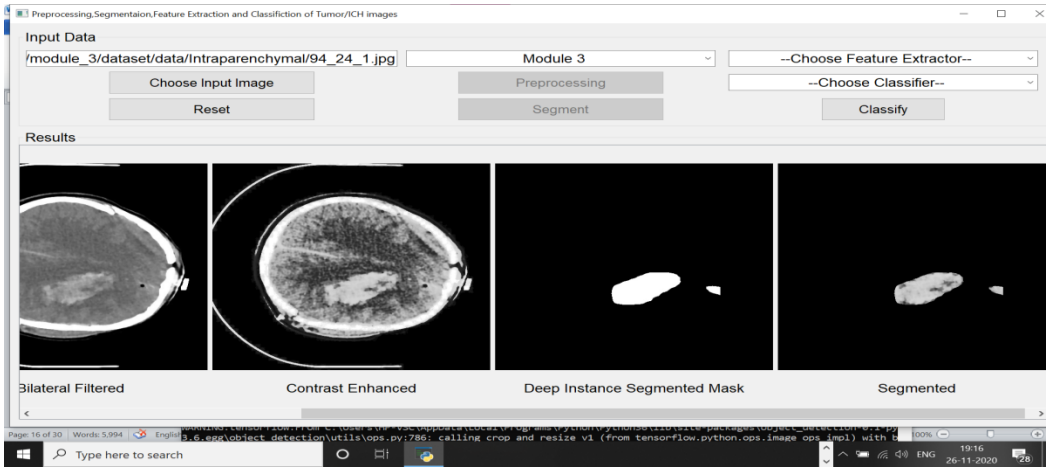
Fig. 8. Dataset Sample Images

ISSN: 2233-7857 IJFGCN

Visualization of simulation outcome accomplished by a proposed method is demonstrated in Fig. 8. The input actual image is gone through series of processes which is shown in fig. 9a 9bpreprocessed, segmented images and Deep instance mask are showcased correspondingly. The figures implied that the projected approach effectively preprocesses and found the tumor regions properly.



(a)



(b)

Fig.9. Experimental results for ICH images a) Epidural b) Intraventricular

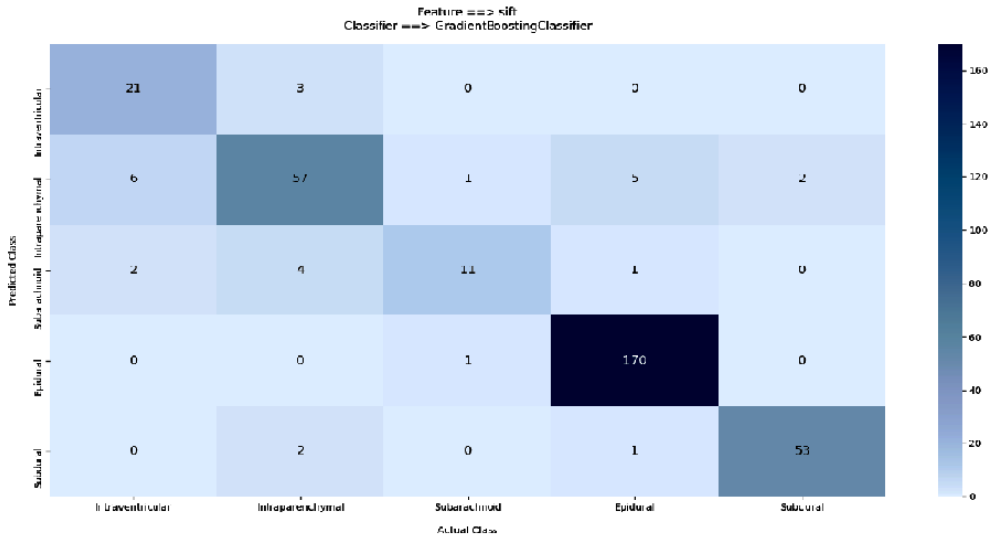
Table 1 and Fig. 10illustrate the confusion matrices produced by the diverse presented techniques on classification of ICH. Fig. 10a signifies that the SIFT-GBT approach has accurately categorized a collection of 21 images under intra-ventricular class, 59 images under intraparenchymal class, 12 images under subarachnoid class, 169 images under epidural class, and 54 images from subdural class.

Measures	**SIFT-GBT**				
	Intraventricular	**Intraparenchymal**	**Subarachnoid**	**Epidural**	**Subdural**
TP	21	59	12	169	54
FN	3	13	6	2	2
FP	8	8	4	4	2
TN	309	261	319	166	283
Measures	**SIFT-LR**				

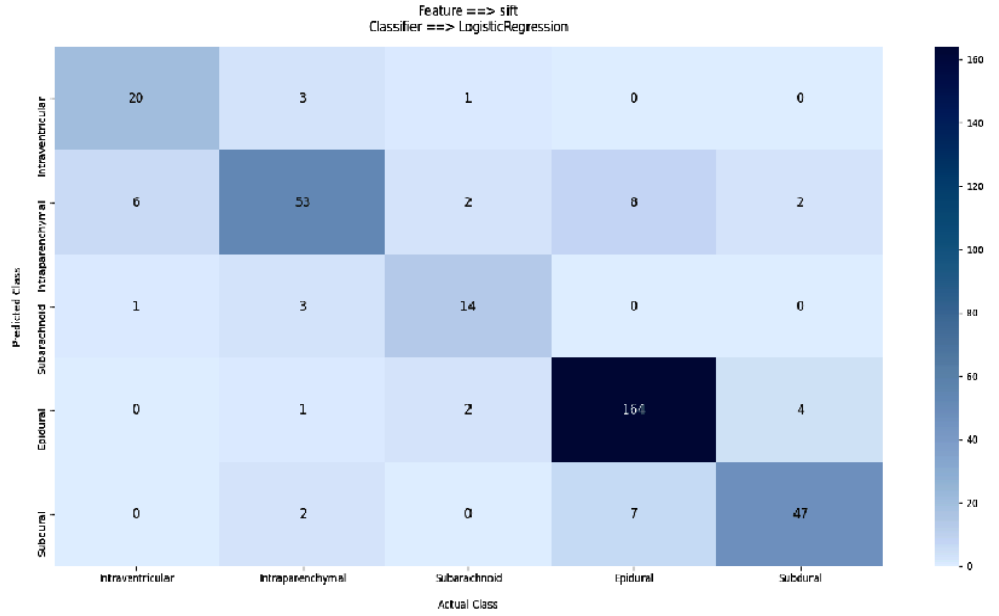
ISSN: 2233-7857 IJFGCN

	Intraventricular	**Intraparenchymal**	**Subarachnoid**	**Epidural**	**Subdural**
TP	20	56	13	159	48
FN	4	16	5	12	8
FP	7	14	11	19	12
TN	310	255	312	151	273
Measures	**SIFT-MLP**				
	Intraventricular	**Intraparenchymal**	**Subarachnoid**	**Epidural**	**Subdural**
TP	22	55	12	168	51
FN	2	17	6	3	5
FP	12	5	3	9	4
TN	305	264	320	161	281

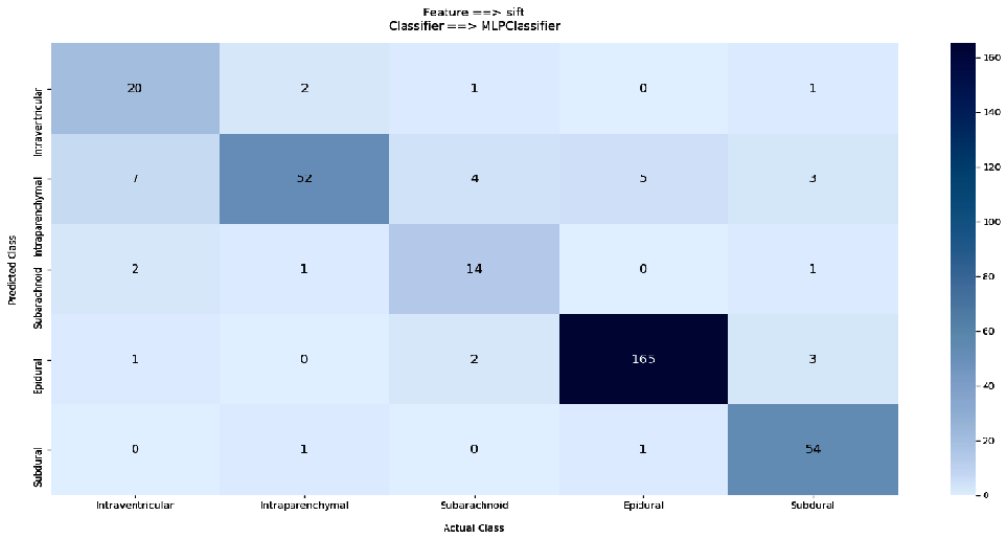
Table 1 Manipulation from Confusion Matrix for SIFT-GBT, SIFT-LR and SIFT-MLP



(a)



(b)



(c)

Fig. 10. Confusion Matrix a) SIFT-GBT b) SIFT-LR c) SIFT-MLP

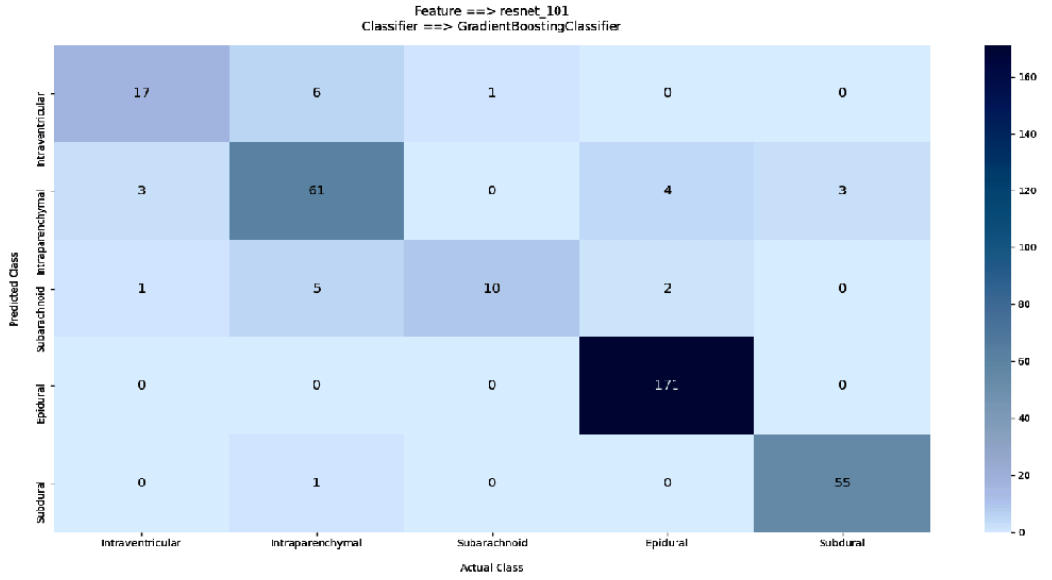
In addition, the SIFT-LR method has provided a significant classification with massive number of 20 images under intraventricular class, 56 images under intraparenchymal class, 13 images under subarachnoid class, 159 images from epidural class, and 48 images from subdural class.Furthermore, the SIFT-MLP scheme has effectively divided 22 images under intraventricular class, 55 images from intraparenchymal class, 12 images from subarachnoid class, 168 images under epidural class, and 51 images under subdural class.

Table 2 and Fig. 11 depict the confusion matrices produced by the distinct projected techniques on the classification of ICH. Fig. 9a signifies the CNNRN101-GBT approach has accurately categorized a set of 20 images from intraventricular class, 59 images under intraparenchymal class, 9 images under subarachnoid class, 171 images under epidural class, and 56 images under subdural class.

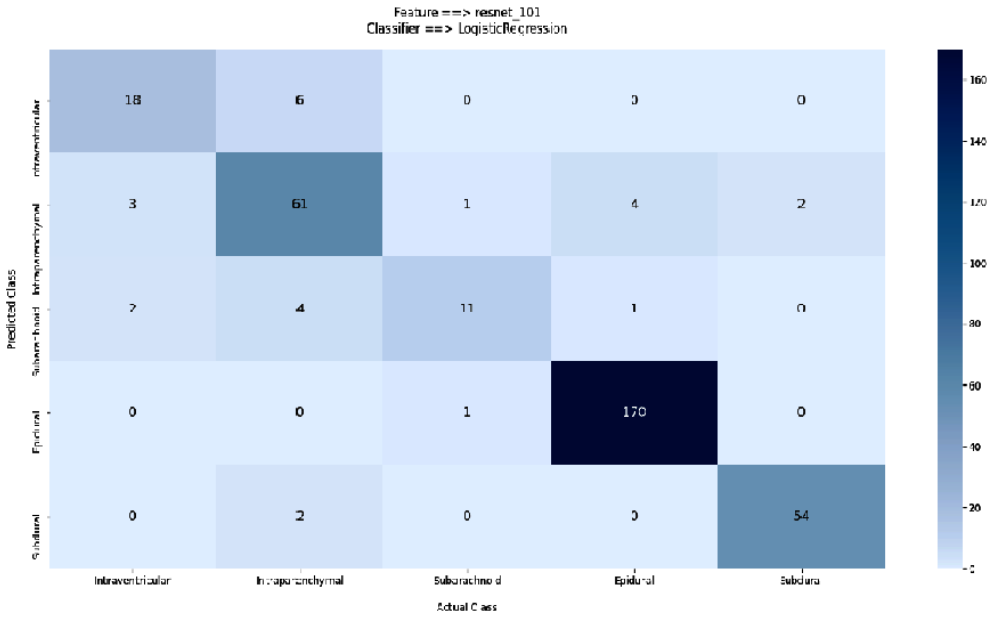
ISSN: 2233-7857 IJFGCN

Measures	CNNRN101-GBT				
	Intraventricular	**Intraparenchymal**	**Subarachnoid**	**Epidural**	**Subdural**
TP	20	59	9	171	56
FN	4	13	9	0	0
FP	7	9	0	6	4
TN	310	260	323	164	281
Measures	**CNNRN101-LR**				
	Intraventricular	**Intraparenchymal**	**Subarachnoid**	**Epidural**	**Subdural**
TP	20	60	11	170	54
FN	4	12	7	1	2
FP	7	10	2	5	2
TN	310	259	321	165	283
Measures	**CNNRN101-MLP**				
	Intraventricular	**Intraparenchymal**	**Subarachnoid**	**Epidural**	**Subdural**
TP	15	72	9	166	52
FN	9	0	9	5	4
FP	2	23	1	1	0
TN	315	246	322	169	285

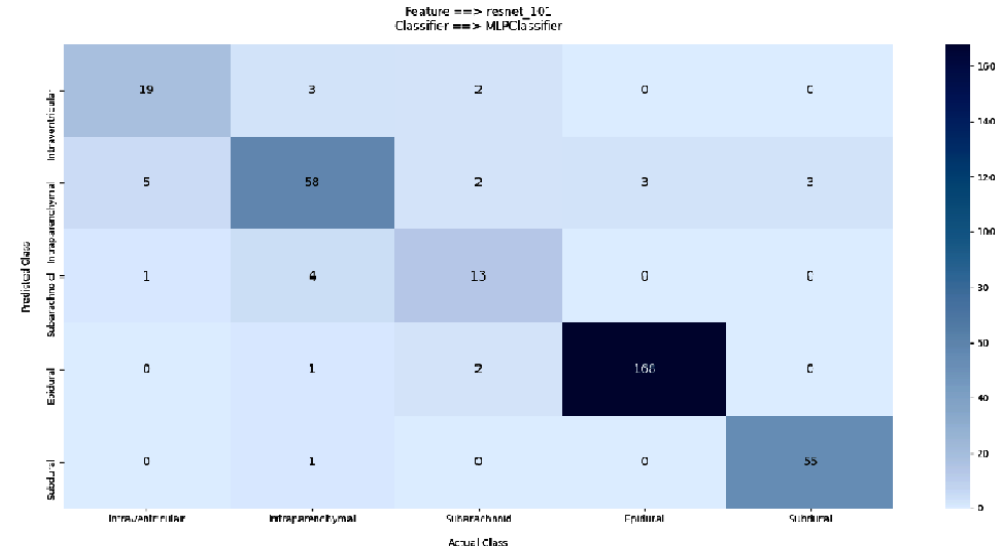
Table 2 Manipulation from Confusion Matrix for CNNRN101-GBT, CNNRN101-LR, and CNNRN101-MLP



(a)



(b)



(c)

Fig. 11. Confusion Matrix a) CNNRN101-GBT b) CNNRN101-LR c) CNNRN101-MLP

Additionally, the CNNRN101-LR framework has provided a significant classification with the massive count of 20 images under intraventricular class, 60 images under intraparenchymal class, 11 images under subarachnoid class, 170 images under epidural class, and 54 images under subdural class. Also, the CNNRN101-MLP scheme has significantly divided a total number of 15 images under intraventricular class, 72 images under intraparenchymal class, 9 images under subarachnoid class, 166 images under epidural class, and 52 images under subdural class.

Table 3 and Figs. 12-13 demonstrates the classification result of the presented approaches by means of distinct values. From the gained values, it is clear that the SIFT-LR scheme has attained lower sensitivity of 83.82%, specificity of 96.56%, accuracy of 93.16%, precision of 85.76%, and F-score of 84.69%. Simultaneously, the SIFT-MLP method has concluded with moderate performance over the SIFT-LR approach with the sensitivity of 84.81%, specificity of 97.35%, accuracy of 96.13%, precision of 84.8%, and F-score of 84.07%. Besides, competing function is exhibited by the SIFT-

ISSN: 2233-7857 IJFGCN

GBT scheme with the sensitivity of 86.27%, specificity of 98.04%, accuracy of 96.93%, precision of 85.92%, and F-score of 85.88%. Similarly, the CNNRN101-MLP technology has illustrated considerable outcome with sensitivity of 80.49%, specificity of 97.98%, accuracy of 96.83%, precision of 90.69%, and F-score of 83.64%. In addition, the CNNRN101-GBT approach has implemented moderate results with a sensitivity of 83.06%, specificity of 97.9%, accuracy of 96.95%, precision of 90.16%, and F-score of 84.84%.

Methods	**Sensitivity (%)**	**Specificity(%)**	**Accuracy(%)**	**Precision(%)**	**F-score(%)**
CNNRN101-GBT	83.06	97.90	96.95	90.16	84.84
CNNRN101-LR	84.72	97.96	96.95	87.60	85.72
CNNRN101-MLP	80.49	97.98	96.83	90.69	83.64
SIFT-GBT	86.27	98.04	96.93	85.92	85.88
SIFT-LR	83.82	96.56	93.16	85.76	84.69
SIFT -MLP	84.81	97.35	96.13	84.80	84.07

Table 3 Result Analysis of Proposed Methods interms of Sensitivity, Specificity, Accuracy, Precision, and F-score

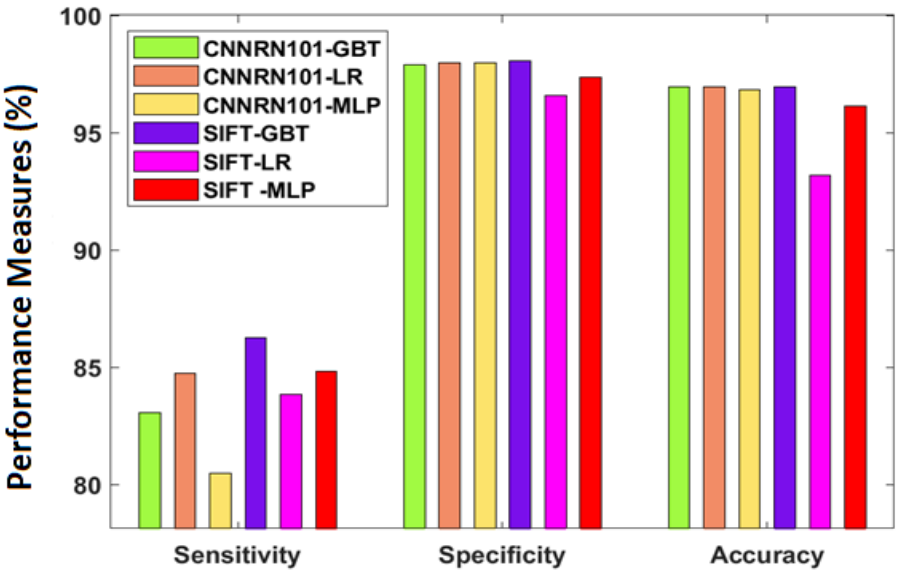


Fig. 12. Result analysis of proposed model

Therefore, the CNNRN101-LR scheme has surpassed the traditional proposed models with a higher sensitivity of 84.72%, specificity of 97.96%, accuracy of 96.95%, precision of 87.6%, and F-score of 85.72%.

ISSN: 2233-7857 IJFGCN

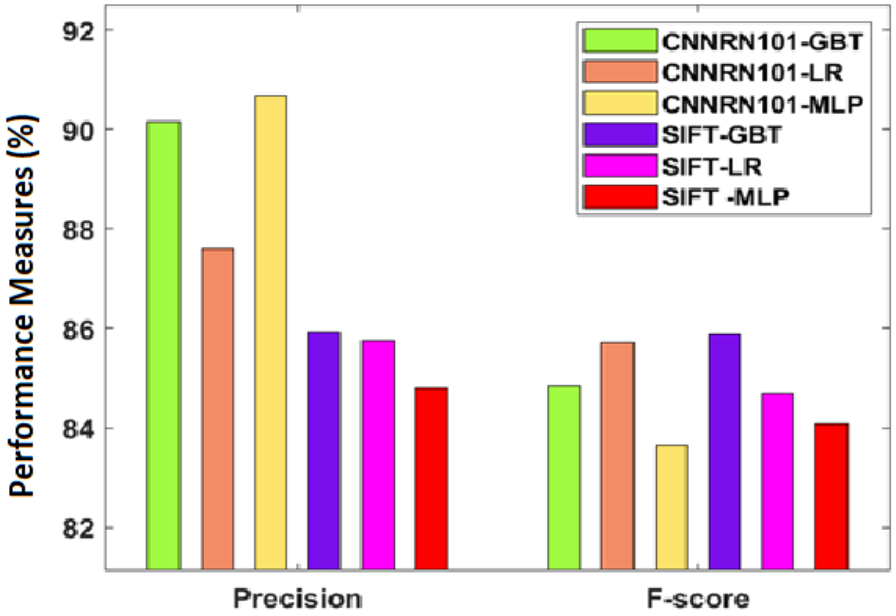


Fig. 13. Precision and F-score analysis of proposed method

Table 4 and Figs. 13-15 implies the comparative results examination of the presented scheme with classical approaches [21-25] with respect to diverse metrics. Fig. 12 examines the classifier outcomes analysis of the projected method with respect to sensitivity, specificity. The experimental outcomes pointed that the CNN technology has represented competing results with the sensitivity of 87% and specificity of 88.1%,

Methods	**Sensitivity(%)**	**Specificity(%)**	**Accuracy(%)**	**Precision(%)**	**F-score(%)**
Proposed Methods					
CNNRN101-GBT	83.06	97.90	96.95	90.16	84.84
CNNRN101-LR	84.72	97.96	96.95	87.60	85.72
CNNRN101-MLP	80.49	97.98	96.83	90.69	83.64
SIFT-GBT	86.27	98.04	96.93	85.92	85.88
SIFT-LR	83.82	96.56	93.16	85.76	84.69
SIFT -MLP	84.81	97.35	96.13	84.80	84.07
Existing Methods					
HIS-DCM	63.10	88.60	87.00	88.19	-
ANN	60.18	70.13	69.78	70.08	-
HIS-DL	88.75	90.42	89.30	95.20	-
WTNN-IIHD	83.33	97.48	88.35	89.90	-
CNN	87.00	88.10	87.50	87.90	-

ISSN: 2233-7857 IJFGCN

SVM	76.30	79.43	77.31	77.50	-
MLP	-	-	93.30	-	-
KNN	-	-	60.00	-	-

Table 4 Result Analysis of Existing with Proposed Methods interms of Sensitivity, Specificity, Accuracy, Precision, and F-score

Fig. 14 examined the classifier results examination of the newly presented framework in light of accuracy. The experimental outcomes have showcased that the KNN scheme accomplished poor performance by achieving minimum accuracy of 60%. Additionally, the ANN technology has obtained moderate accuracy of 69.78%. Similarly, the SVM model has gained acceptable accuracy of 77.31%. Meantime, the HIS-DCM and CNN frameworks have managed to obtain reasonable similar outcome with an accuracy of 87% and 87.5% respectively. Subsequently, the linear WTNN-IIHD and HIS-DL have expressed some satisfactory outcomes with an accuracy of 88.35% and 89.3% correspondingly. At the same time, the SIFT-LR and MLP technologies have obtained acceptable and closer results with an accuracy of 93.16% and 93.3% respectively. Simultaneously, the SIFT-MLP and CNNRN101-MLP methodologies have obtained satisfactory results with an accuracy of 96.13% and 96.83%. Even though the SIFT-GBT scheme has referred competing results with the accuracy of 96.93%, it failed to surpass the CNNRN101-GBT and CNNRN101-LR models which have achieved higher and identical accuracy of 96.95%.

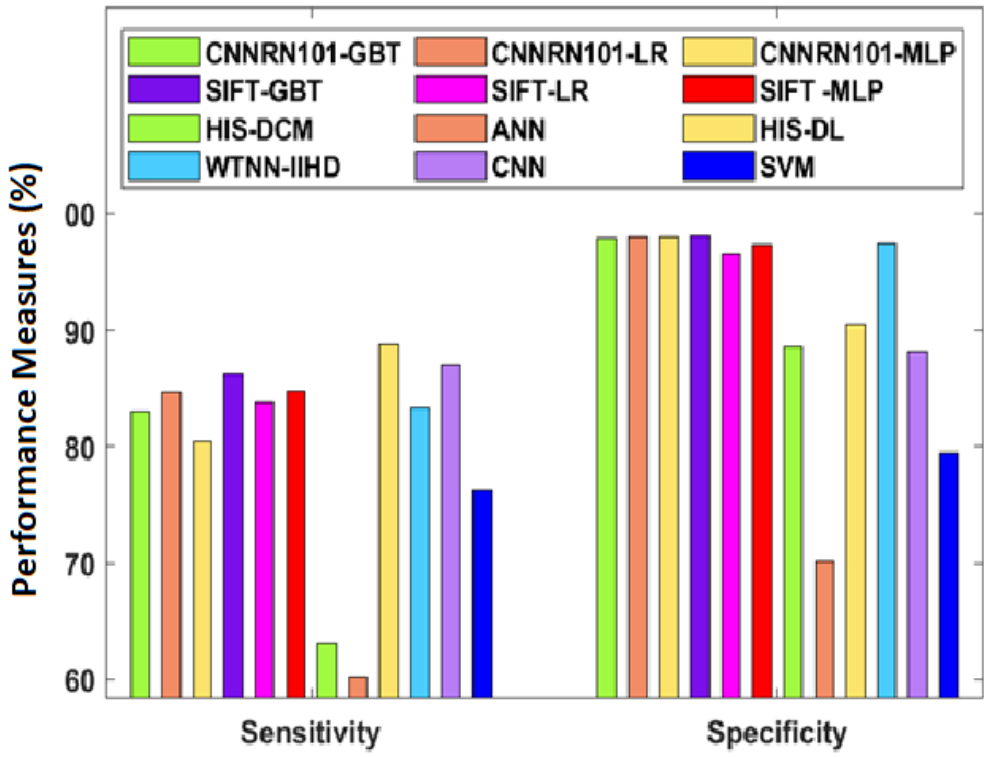


Fig. 14. Comparative analysis of proposed method interms of Sensitivity and specificity

ISSN: 2233-7857 IJFGCN

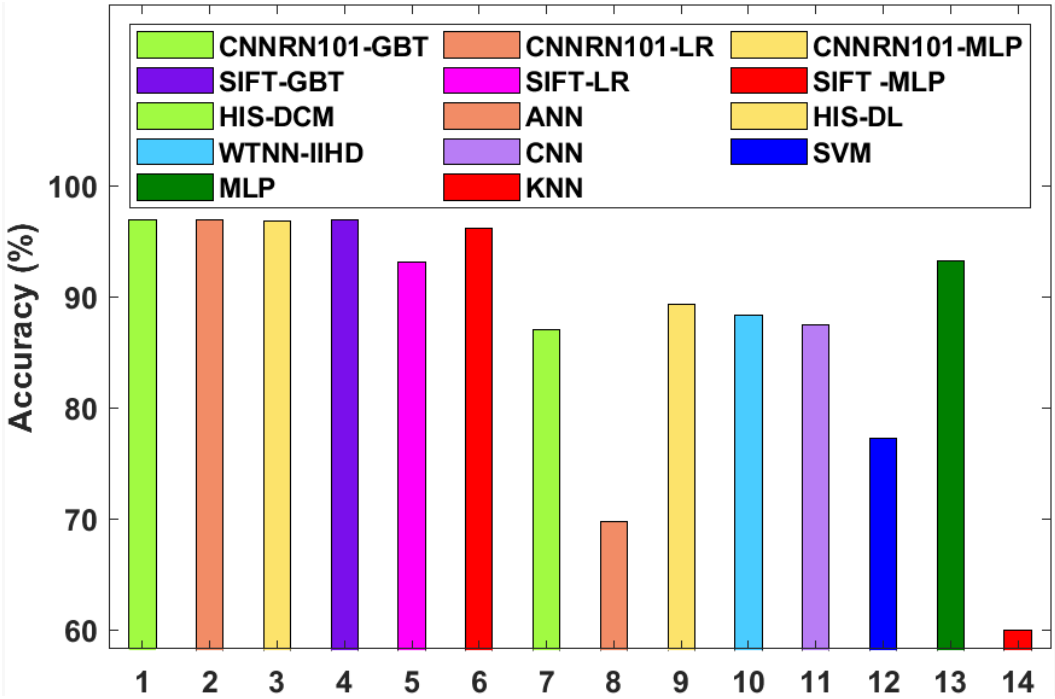
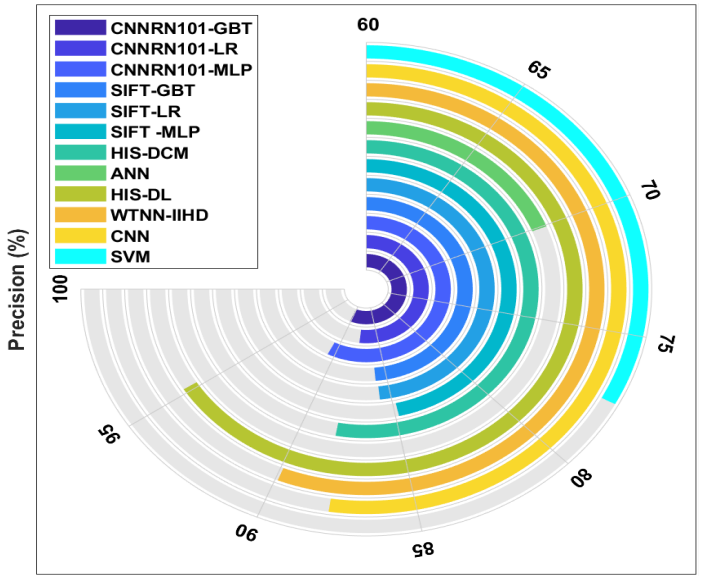


Fig. 15. Comparative analysis of proposed method interms of Accuracy

Figs. 16-17 is the Radial chart that examine the classifier results analysis of the projected method with respect to precision, F-score. The experimental results implied that the ANN method accomplishes inferior function by gaining lower precision of 70.08%. Moreover, the SVM model has attained considerable precision of 77.5%. Likewise, the SIFT-MLP framework has provided moderate precision of 84.8% and F-score of 84.07%. Concurrently, the SIFT-LR and SIFT-GBT methodologies have managed to obtain acceptable results with the precision of 85.76%, 85.92%, and F-score of 84.69%, 85.88% respectively. Consecutively, the linear CNN and HIS-DCM have showcased satisfactory outcomes with the precision of 87.9% and 88.19% respectively. Meantime, the CNNRN101-LR scheme has accomplished considerable results with a precision of 87.6% and F-score of 85.72%. At the same time, the WTNN-IIHD models have gained considerable results with the precision of 89.9%. Though the CNNRN101-GBT and CNNRN101-MLP techniques have displayed competing outcomes with the precision of 90.16%, 90.69, and F-score of 84.84%, 83.64, it failed to perform well the HIS-DL model which has obtained an optimal precision of 95.2%.



ISSN: 2233-7857 IJFGCN

Fig. 16. Comparative analysis of proposed method interms of Precision

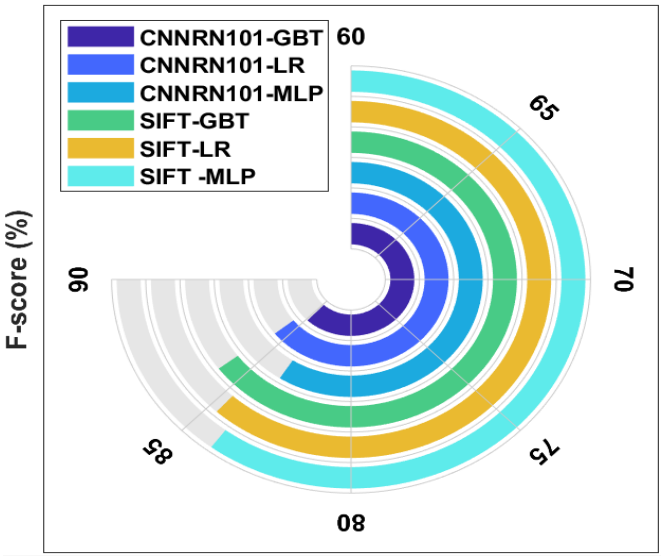


Fig. 17. Comparative analysis of proposed method interms of F-score

5. Conclusion

This paper has developed an automated DL based segmentation and classification model for ICH diagnosis. Primarily, the input images are preprocessed to raise the input quality by removing noise and increasing the contrast level. Followed by, image segmentation process takes place by a DL model. In addition, two feature extractors namely SIFT and ResNet 101 are employed to extract useful set of feature vectors. At last, three different ML models are applied to perform classification process. The application of DL models for both segmentation and feature extraction helps to achieve improved classifier outcome. For experimental validation, a set of simulation were carried out on benchmark ICH dataset. The obtained simulation outcome pointed out that the proposed model has given accuracy level of 96.95% which is comparatively higher than the conventional model.

References

[1] van Asch, C.J., Luitse, M.J., Rinkel, G.J., van der Tweel, I., Algra, A., Klijn, C.J.: Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and metaanalysis. Lancet Neurol. 9(2), 167–176 (2010)

[2] Saulle, M.F., Schambra, H.M.: Recovery and rehabilitation after intracerebralhemorrhage. In: Seminars in Neurology, vol. 36, p. 306. NIH Public Access (2016)

[3] Choi, Y., Kwon, Y., Lee, H., Kim, B.J., Paik, M.C., Won, J.H.: Ensemble of deep convolutional neural networks for prognosis of ischemic stroke. In: Crimi, A., Menze, B., Maier, O., Reyes, M., Winzeck, S., Handels, H. (eds.) BrainLes 2016. LNCS, vol. 10154, pp. 231–243. Springer, Cham (2016).

[4] Kamnitsas, K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 36, 61–78 (2017)

[5] Loganathan, J., Janakiraman, S., & Latchoumi, T. P. A Novel Architecture for Next Generation Cellular Network Using Opportunistic Spectrum Access Scheme. Journal of Advanced Research in Dynamical and Control Systems,(12), 1388-1400.

[6] Prevedello, L.M.; Erdal, B.S.; Ryu, J.L.; Little, K.J.; Demirer, M.; Qian, S.; White, R.D. Automated critical test findings identification and online notification system using artificial intelligence in imaging. Radiology 2017, 285, 923–931.

[7] Chilamkurthy, S.; Ghosh, R.; Tanamala, S.; Biviji, M.; Campeau, N.G.; Venugopal, V.K.; Mahajan, V.; Rao, P.; Warier, P. Deep learning algorithms for detection of critical findings in head CT scans: A retrospective study. Lancet 2018, 392, 2388–2396.

ISSN: 2233-7857 IJFGCN

[8] Grewal, M.; Srivastava, M.M.; Kumar, P.; Varadarajan, S. RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 281–284.
[9] Ranjeeth, S., Latchoumi, T. P., & Victer Paul, P. (2019). Optimal stochastic gradient descent with multilayer perceptron based student's academic performance prediction model. Recent Advances in Computer Science and Communications. https://doi. org/10.2174/2666255813666191116150319.
[10] Lee, H.; Yune, S.; Mansouri, M.; Kim, M.; Tajmir, S.H.; Guerrier, C.E.; Ebert, S.A.; Pomerantz, S.R.; Romero, J.M.; Kamalian, S.; et al. An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. Nat. Biomed. Eng. 2019, 3, 173.
[11] Ranjeeth, S., Latchoumi, T. P., & Paul, P. V. (2020). A Survey on Predictive Models of Learning Analytics. Procedia Computer Science, 167, 37-46.
[12] Li, Y., Xu, X. and Yuan, C., 2020. Enhanced Mask R-CNN for Chinese Food Image Detection. *Mathematical Problems in Engineering*, *2020*.
[13] Gupta, S., Thakur, K. and Kumar, M., 2020. 2D-human face recognition using SIFT and SURF descriptors of face's feature regions. *The Visual Computer*, pp.1-10.
[14] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
[15] Maeda-Gutiérrez, V., Galván-Tejada, C.E., Zanella-Calzada, L.A., Celaya-Padilla, J.M., Galván-Tejada, J.I., Gamboa-Rosales, H., Luna-García, H., Magallanes-Quintanar, R., Guerrero Méndez, C.A. and Olvera-Olvera, C.A., 2020. Comparison of Convolutional Neural Network Architectures for Classification of Tomato Plant Diseases. *Applied Sciences*, *10*(4), p.1245.
[16] M. Maalouf, T. B. Trafalis, and I. Adrianto, "Kernel logistic regression using truncated Newton method," Computational Management Science, vol. 8, no. 4, pp. 415–428, 2011.
[17] Hamed, A.A., Li, R., Xiaoming, Z. and Xu, C., 2013. Video genre classification using weighted kernel logistic regression. *Advances in Multimedia*, *2013*.
[18] Tarkhaneh, O. and Shen, H., 2019. Training of feedforward neural networks for data classification using hybrid particle swarm optimization, Mantegna Lévy flight and neighborhood search. *Heliyon*, *5*(4), p.e01275.
[19] Flores, V. and Keith, B., 2019. Gradient boosted trees predictive models for surface roughness in high-speed milling in the steel and aluminum metalworking industry. *Complexity*, *2019*.
[20] Loganathan, J., Latchoumi, T. P., Janakiraman, S., & parthiban, L. (2016, August). A novel multi-criteria channel decision in co-operative cognitive radio network using E-TOPSIS. In Proceedings of the International Conference on Informatics and Analytics (pp. 1-6).
[21] Hssayeni, M.D., Croock, M.S., Salman, A.D., Al-khafaji, H.F., Yahya, Z.A. and Ghoraani, B., 2020. Intracranial Hemorrhage Segmentation Using A Deep Convolutional Model. *Data*, *5*(1), p.14.
[22] Davis, V. and Devane, S., 2017, December. Diagnosis & classification of brain hemorrhage. In *2017 International Conference on Advances in Computing, Communication and Control (ICAC3)* (pp. 1-6). IEEE.
[23] Danilov, G., Kotik, K., Negreeva, A., Tsukanova, T., Shifrin, M., Zakharova, N., Batalov, A., Pronin, I. and Potapov, A., 2020. Classification of Intracranial Hemorrhage Subtypes Using Deep Learning on CT Scans. *Studies in Health Technology and Informatics*, *272*, pp.370-373.
[24] Karki, M., Cho, J., Lee, E., Hahm, M.H., Yoon, S.Y., Kim, M., Ahn, J.Y., Son, J., Park, S.H., Kim, K.H. and Park, S., 2020. CT window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings. *Artificial Intelligence in Medicine*, p.101850.
[25] Shahangian, B. and Pourghassem, H., 2013, September. Automatic brain hemorrhage segmentation and classification in CT scan images. *8th Iranian Conference on Machine Vision and Image Processing (MVIP)* (pp. 467-471). IEEE.

ISSN: 2233-7857 IJFGCN

[26] ArunaKirithika.R, Sathiya.S et. al.,2020. Brain Tumor and Intracranial Haemorrhage Feature Extraction and Classification using Conventional and Deep Learning methods. *European Journal of Molecular and Clinical Medicine.*

[27] Sathiya.S, Balasubramanian.M, Palanivel.S., 2013, Detection and Recognition of Characters in Place Name Board for Driving. *International Journal of Computer Applications (0975 8887) Volume 80 - No. 7*

[28] S. Sathiya et al. 2014,Pattern Recognition Based Detection Recognition of Traffic Sign Using SVM, International Journal of Engineering and Technology (IJET)Vol 6 No 2 Apr-May 2014

ISSN: 2233-7857 IJFGCN