



# Detection and Classification of Brain Hemorrhage Using Hounsfield Unit and Deep Learning Techniques

Anh-Cang Phan<sup>1</sup>(✉), Hung-Phi Cao<sup>1</sup>, Thanh-Ngoan Trieu<sup>2</sup>,  
and Thuong-Cang Phan<sup>2</sup>

<sup>1</sup> Vinh Long University of Technology Education, Vinh Long, Vietnam  
{cangpa, caohungphi}@vlute.edu.vn

<sup>2</sup> Can Tho University, Can Tho, Vietnam  
{ptcang, ttngoan}@cit.ctu.edu.vn

**Abstract.** Stroke is a dangerous disease with a complex disease progression and a high mortality rate just behind cancer and cardiovascular disease. To diagnose brain hemorrhage, the doctors check CT/MRI images and rely on the Hounsfield Unit to determine the region, duration and level of bleeding. Due to the increasing number of brain haemorrhages, it will put pressure on the treating doctors. Therefore, the construction of an automatic system of segmentation and identification of brain hemorrhage with fast processing time and high accuracy is essential. In this paper, we propose a new approach based on Hounsfield Unit and deep learning techniques. It not only determines the level and duration of hemorrhage but also segments the brain hemorrhagic regions on MRI images automatically. From experiments, we compared and evaluated on three neural network models to select the most suitable model for classification. As a result, the proposed method using Hounsfield Unit and Faster RCNN Inception is time-effective and high accuracy with mean average precision (mAP) of 79%.

**Keywords:** Brain hemorrhage · CT/MRI image · Hounsfield unit · Deep learning

## 1 Introduction

Stroke is a dangerous disease progressing quickly in case of not having prompt treatment. According to statistics from the World Health Organization, stroke is the second leading cause of death and the first leading cause of disability affecting patients' daily life. In Vietnam, stroke is the seventh leading cause of death with 0.64 deaths per 100,000 cases [11]. Stroke has two main types: hemorrhagic stroke and ischemic stroke with a ratio of 85% and 15% respectively. This disease has a serious complication after recovery with 92% carrying movement sequelae, 27% severe movement sequelae, and cognitive disorders [13]. To diagnose the disease, a doctor might prescribe a CT/MRI of the cranial region. The diagnosis

is then based on the Hounsfield Unit (HU) of the cerebral hemorrhage area in the CT/MRI image [10,12]. However, with a large number of patients, doctors have more pressure that can affect the accuracy of diagnosis and treatment planning.

The advancement of computer vision has benefits in many different fields, especially medical images. The study [14] in 2017 used a convolutional neural network (CNN) with three models including LeNet, GoogLeNet, and Inception-ResNet to diagnosing brain hemorrhage. The dataset for experiments consisted of 100 cases of brain hemorrhage with CT/MRI images collected from Hospital 115 (Ho Chi Minh City, Vietnam). The research results indicated that the three models are relevant to the diagnosis of brain hemorrhage with the F1 score of LeNet, GoogLeNet, and Inception-ResNet are 0.997, 0.983, and 0.989, respectively. In 2018, experts from the University of California, Berkeley, and the University of California, San Francisco (UCSF) [8] trained a convolutional neural network called PatchFCN. This model is trained on a dataset of 4,000 CT images from hospitals affiliated with UCSF. The results show that PatchFCN has the capability of intracranial hemorrhage detection.

In general, these methods [8,14] use the CNN technique to classify medical images without considering the HU values. This can cause effects on the accuracy of the segmentation and classification of cerebral hemorrhage because HU values are often used by medical specialists to identify bleeding areas and duration of damages in reality. On the other hand, there is a lack of comparison and evaluation of neural network models on medical data. In this paper, we propose a new approach for detection and classification of brain hemorrhage based on HU values using the techniques of deep learning. Experiments were conducted to compare and evaluate the results of the four common types of cerebral hemorrhage [10,12]: epidural hematoma (EDH), subdural hematoma (SDH), subarachnoid hemorrhage (SAH), and intracerebral hemorrhage (ICH) (as in Fig. 1). Our contributions are as follows: 1) Collect medical images of cerebral hemorrhage for classification; 2) Apply HU values in automatic segmentation of cerebral hemorrhage regions to assist experts in labeling the dataset; 3) Train the multi-layer classifier of brain hemorrhage on three deep learning network models: Faster R-CNN Inception ResNet v2, and SSD MobileNet v2, and SSD Inception v2; 4) Detect, segment and quantify the time and level of cerebral hemorrhage based on HU values; 5) Compare and evaluate the classification results on these 3 models.

The remaining of the paper is organized as follows: Sect. 2 details the related work for the detection and classification of cerebral hemorrhage. Our proposed methods are described in Sect. 3. We show and compare experimental results of our methods in Sect. 4. Finally, we draw the conclusion in Sect. 5.

## 2 Related Work

### 2.1 Hounsfield Unit (HU) in Brain Hemorrhage Segmentation

The diagnosis of cerebral hemorrhage needs a high accuracy because any error can affect the treatment regimen and the patient's recovery. In this study, we calculate the HU values, which are often used by specialists to diagnose CT/MRI

hemorrhagic images [10,12] for accurately determining the hemorrhage area, bleeding time, and the extent of bleeding. This is also a new approach proposed in our approach. We analyze input CT/MRI images in the standard DICOM format [7] without converting them to other formats, such as JPG, BMP, and PNG. The input information includes patient information, hospital information, and image data. Currently, modern CT/MRI scanners have the grayscale of -1000 to +4000 [12] while digital devices such as computer screens have the grayscale of 0 to 255. Thus, displaying the CT/MRI images on those devices will not be correct. In order to properly display the gray level on computer screens, it is necessary to convert the values by the linear transformation as in formula 1.

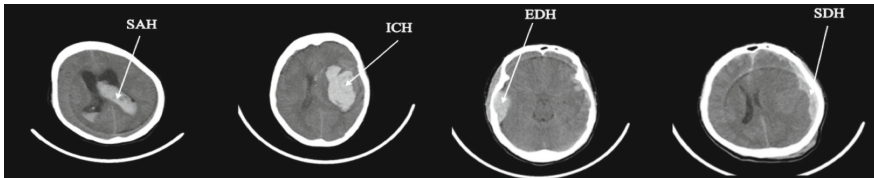


Fig. 1. Illustration of four types of brain hemorrhage on CT/MRI image [3,4]

$$HU = Pixel\_value * Rescale_{slope} * Rescale_{intercept} \tag{1}$$

where:

- *Pixel\_value*: is the value of each pixel
- *Rescale\_slope* and *Rescale\_intercept*: are the values stored in DICOM images

Specifically, a cranial DICOM image will be converted to a digital image according to the HU value. Types of tissue, water and air show different HU values, which are illustrates in Table 1. The Hounsfield scale of tissue density is based on two values: air as -1000 HU (minimum HU value) and water as 0 HU. Density of other tissues is related to this range, usually from -1000 to +1000 HU

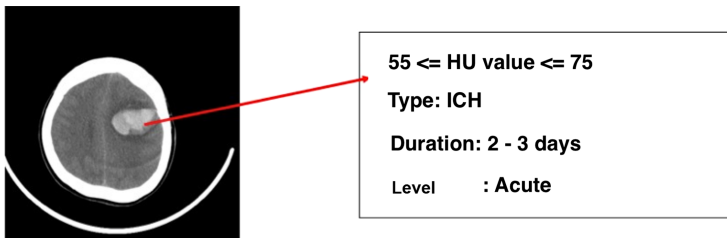


Fig. 2. Difference between bleeding and non-bleeding areas [3,4]

[1, 3]. The hemorrhagic/hematoma region has the HU values in the range from 40 to 90 (as in Fig. 2 and Table 1). We use image thresholding techniques [17] and the contours detection technique [2] to define the contours connecting all adjacent points with the same/approximately color value or contrast value. We later define a convex hull [6] surrounding the contours. The brain hemorrhagic regions are identified in this way and this is also the area for classification.

**Table 1.** X-ray absorption was measured on a CT/MRI image by hounsfield unit (HU) [1, 3, 12]

Matter	Density (HU)
Water	0
Bone	1,000
Air	-1,000
Gray matter	35-40
White matter	20
Hematoma	40-90

## 2.2 Convolutional Neural Network (CNN) for Training and Classifying

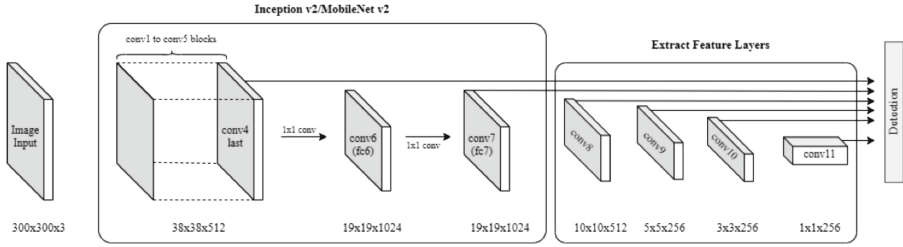
The choice of the appropriate deep learning models and the quality dataset plays an important role in improving classification accuracy. With many powerful deep learning models for images, we decide to take the advantage of two effective deep learning models, SSD [9] and Faster R-CNN [15]. After preparing the training dataset, we perform a training phase on these two models including SSD and Faster R-CNN with three neural networks consisting of Inception v2, MobileNet v2, Inception ResNet v2 for feature extraction. Below is a brief description of the models we applied in the proposed methods.

### 2.2.1 Single Shot MultiBox Detector Architecture (SSD)

The SSD model [9] comes up with the idea of bounding boxes to create boxes in multiple locations of the image. It performs calculations and evaluations on each box for segmentation and object classification. The SSD architecture shown in Fig. 3 is designed to optimize object detection time as quickly as possible. The SSD model consists of two main phases.

*Phase 1 - Convolutional predictors for detection:* SSD uses basic networks such as VGG16, Inception, MobileNet, and ResNet to extract image features. The filters used have the size of  $3 \times 3 \times p$  over the layers for prediction instead of using the fully connected layer as in other network models. This is greatly reducing computation costs.

*Phase 2 - Multi-scale feature maps for detection:* The size of feature maps will decrease along with the depth of the network when using filters. This helps



**Fig. 3.** Single shot MultiBox detector architecture [9]

in detecting objects at different sizes and different scales. In this study, we conduct experiments of the SSD model with two basic networks, Inception and MobileNet. The architecture of the SSD model is presented in Fig. 3.

SSD has the objective loss function with localization loss and confidence loss. The localization loss is used to evaluate the detection task and the confidence loss is used to evaluate the classification task. The localization loss is a Smooth L1 loss [5] between the predicted box ( $l$ ) and the ground truth box ( $g$ ). Let  $x_{ij}^p = \{1, 0\}$  corresponds to the  $i$ -th box that matches the ground truth box  $j$ -th of class  $p$ . The Loss function is a weighted sum of the localization loss ( $loc$ ) and the confidence loss ( $conf$ ), as in Eq. 2.

$$L_{loc}(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \quad (2)$$

where:

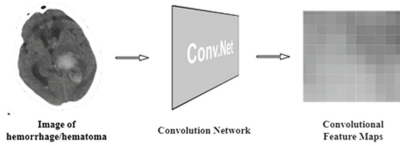
$$- L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ik}^m \text{smooth}_{L1}(l_i^m - g_j^m)$$

$$- L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0)$$

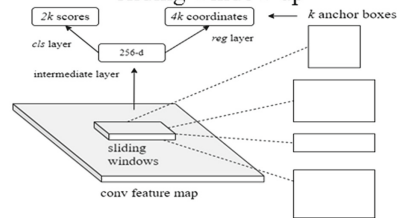
**2.2.2 Faster R-CNN Architecture**

The Faster Region-based Convolutional Network model (Faster R-CNN) [15] is improved from the Fast R-CNN model [5] with the replacement of the selective

**RPN Phase 1:** Feed-forward the input image through DNN to obtain convolutional features



**RPN Phase 2:** Use a convolutional sliding window up



**Fig. 4.** Two main phases of RPN [15]

search algorithm to the region proposal network (RPN). Faster R-CNN works with two main phases including the first phase using the RPN to create the proposed zones and the second phase segmenting and classifying objects on the proposed zones. The RPN accepts an input image of any size and outputs region proposals with the probability containing objects.

Figure 4 illustrates the operation of the RPN algorithm, in which phase 1 uses a convolution layer with a size of  $3 \times 3$  and 5 layers of max pooling size  $2 \times 2$  to create feature maps. The second phase of RPN uses a sliding window ( $n \times n$ ) on the feature maps. The results achieved are the positions of the objects and the probability containing the objects on the feature maps. The Loss function is measure by formula 3 and 4 [15].

$$Loss(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i * L_{reg}(t_i, t_i^*) \quad (3)$$

$$SmoothL1Loss(x, y) = \begin{cases} 0.5(x_i - y_i)^2 & \text{if } |x_i - y_i| < 1 \\ |x_i - y_i| - 0.5 & \text{otherwise} \end{cases} \quad (4)$$

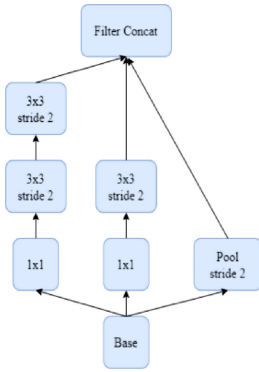
where:

- $i$ : is the index of anchor in mini-batch
- $p_i$ : is the predicted probability of anchor  $i$  is an object
- $p_i^*$ : is the ground-truth label value, 1: if anchor is positive and 0: if anchor is negative
- $t_i$ : is a 4-dimensional vector representing the coordinate values of the predicted bounding box
- $t_i^*$ : is a 4-dimensional vector representing the coordinate values of the ground-truth box corresponding to the positive anchor.
- $L_{cls}$ : is the log loss of 2 classes (object and non-object)
- $L_{reg}$ : is the SmoothL1Loss

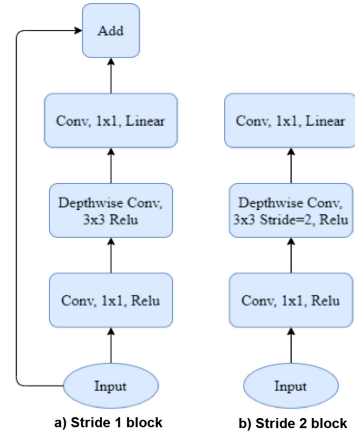
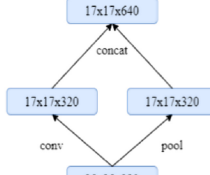
## 2.3 CNN Models for Feature Extraction

### 2.3.1 Inception V2

The Inception v2 network [19] is an artificial neural network made up of multiple layers of CNN networks with the architecture described in Fig. 5 . The Inception v2 network is trained on more than 10 million images of 1,000 different classes of objects (ImageNet database version 2012). This is a network model with a low error rate (3.46%). In this research, we retrain this model with SSD architecture on the brain hemorrhagic image dataset [19].



**Fig. 5.** Architecture of the inception v2 [19]



**Fig. 6.** Architecture of the MobileNet v2 [16]

### 2.3.2 MobileNet V2

MobileNet v2 network [16] is an artificial neural network designed to be optimized for low configuration devices with limited computing power but still ensuring accuracy (as shown in Fig. 6). Comparing with the VGG network model, the MobileNet v2 network reduces up to 75% of the parameters with a total of 30 computational layers. In this study, we retrain the MobileNet v2 network model together with the Inception v2 network for comparison and evaluation.

### 2.3.3 Inception Resnet V2

Inception Resnet v2 [19] is a network model built on a combination of the Inception and Residual Network architectures. Inception Resnets overall architecture is evaluated as a deep learning architecture network and trained with more than 1000 layers on ImageNet-2012 image set. The input of the model is  $299 \times 299$  images and the output is a list of the predicted results.

## 2.4 Measurements for Evaluating the Accuracy of Classification

It is necessary to have a suitable method to evaluate and compare the use of network models. The commonly used method for multi-class classification problems is Precision-Recall [18], as in formula 5. High precision relates to the low False Positive rate, i.e., the proportion of the negatives having positive outcomes. High recall means a high True Positive rate, i.e., the rate of omitting objects that are actually positive is low.

$$precision = \frac{TP}{TP + FP} \qquad recall = \frac{TP}{TP + FN} \qquad (5)$$

F1 is a combination of Precision and Recall, which is the harmonic mean value of the Precision and Recall. F1 tends to get the value, which is closer to the lower

value between Precision and Recall. F1 is high if both Precision and Recall are high [18].

$$F1 = \frac{1}{\frac{1}{precision} + \frac{1}{recall}} \tag{6}$$

In addition to the two methods mentioned above, the mean average precision (mAP) measurement is considered a popular method to evaluate model accuracy when performing Faster R-CNN and SSD network models. It has the values in the range between 0 and 1. The mAP measure is determined in Eq. 7.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \tag{7}$$

in which  $AP_i$  is the average precision for  $i^{th}$  class in  $N$  class.

### 3 Proposed Method

In this paper, we propose the implementation consists of two processing phases: the training phase of the network model and the testing phase. We perform these two phases on the three network models (Faster R-CNN Inception ResNet v2, and SSD MobileNet v2, and SSD Inception v2) so that we can evaluate and choose a suitable model for applying in brain hemorrhage detection and classification. Specifically, the training phase consists of 5 stages described in Fig. 7.

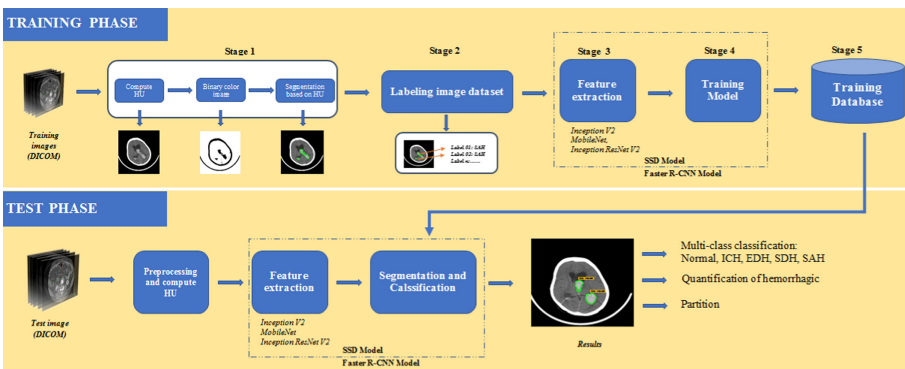


Fig. 7. General model of the proposed method

**Stage 1- Data Preprocessing and Automatic Hemorrhage Segmentation Based on HU:** A DICOM file captured from the CT/MRI scanner will be converted to a digital image (.jpg). The extent, duration, and location of cerebral hemorrhage were determined based on the HU values calculated by formula 1 and Table 1 as shown in Sect. 2.1. The results of stage 1 is a digital image dataset with the brain hemorrhage regions highlighted.



**Stage 2- Labeling Image Dataset With the Support of Experts:** Each digital image will be automatically labeled based on the location of the brain hemorrhage areas under the supportive supervision of specialists. At the end of stage 2, the outcome is the labeled brain hemorrhage image dataset.

**Stage 3 - Feature Extraction:** To reduce the computation time and classify hemorrhage quickly, this stage is to perform feature extraction. We extract the features using the three network models including Inception v2, MobileNet, and Inception ResNet v2 (presented in Sect. 2.3).

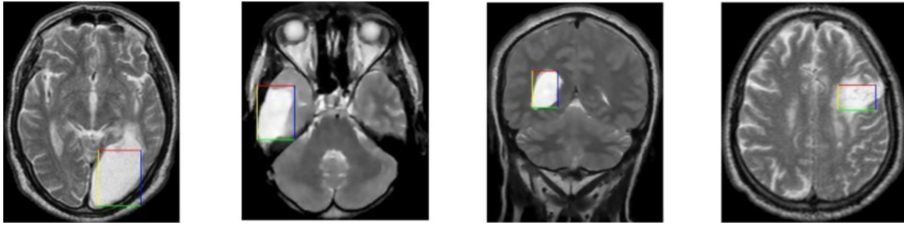
**Stage 4- Training:** The extracted features in Stage 3 are used to train in three network models SSD Inception v2, SSD MobileNet v2, and Faster R-CNN Inception ResNet v2. We monitor the training process based on the Loss value of each network architecture presented in Sect. 2.2. This process will be done until the Loss value does not improve (not decrease) after a certain number of iterations. We stop training the model and move to the testing phase to compare and evaluate the model.

**Stage 5 - Storing Features and Training Parameters:** At the end of stage 4, we receive the features, labels, and training parameter values. Those will be stored in the database for classification of cerebral hemorrhage. Next, we compare and evaluate the network models through the testing dataset to evaluate the accuracy with mAP, presented in Sect. 2.4.

## 4 Experimental Results

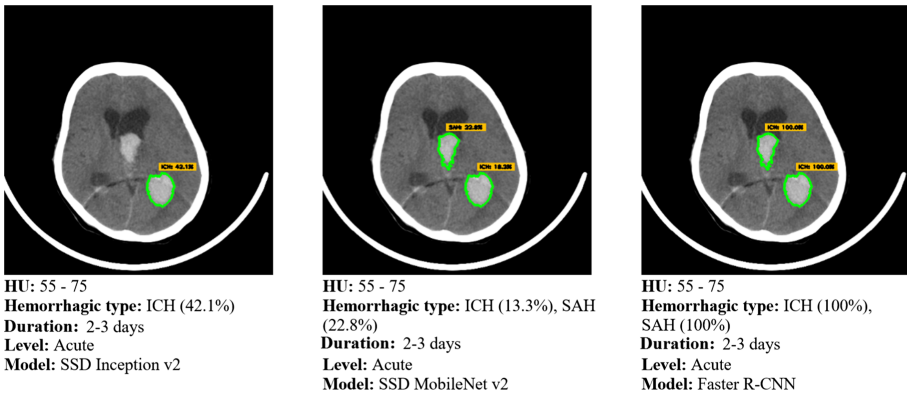
The experiments were conducted on the three models SSD Inception v2, SSD MobileNet v2, Faster R-CNN Inception ResNet v2 in the same Google Colab environment with Ubuntu 18.04. The configuration of the computer is 32 GB RAM and Nvidia Tesla P100 GPU. The library used to support the training of the network model is TensorFlow GPU version 1.5. Since there are no public datasets for labeled cerebral hemorrhage images, we collected a dataset from several hospitals with 479 MRI images with size  $512 \times 512$ , including 79 images of EDH type, 54 images of SDH type, 90 images of SAH type and 256 images of ICH type. Based on HU values, the dataset is automatically partitioned and labeled under the support of experts to contribute to the high reliability training dataset. The dataset will be randomly divided into a training dataset and a testing dataset at the rate of 80% (382 images) and 20% (97 images) respectively. Since pretrained weights for medical images are not available, we do not use transfer learning and train all the layers. The weights of these models are initialized using “COCO dataset” and a sigmoid output layer has been added to get our final output labels. In addition, during training, we tweak the parameters of these models to obtain high accuracy.

Figure 9 shows the results after performing detection and segmentation of the brain hemorrhage. Compared with the results of the FBB method [4] as shown in Fig. 8, the proposed method has accuracy 100% for detecting the contours of the entire hemorrhage region. Especially, our proposed method can detect all bleeding types on the same CT/MRI image.



**Fig. 8.** Hemorrhage segmentation results using FBB algorithm [4]

We evaluate the accuracy and identification time of the three network models. First, the three models need to be assessed on the accuracy of cerebral hemorrhage classification. We determine the Loss value of the three network models to decide to stop training the model and move to the testing phase. Figure 10 shows the results comparing the Loss value of the three network models after a number of iterations with the Learning\_rate of 0.0003. Figure 10.c shows the Loss value of the Faster R-CNN is very low below 10% (Loss\_value < 0.01) compared to the other models (Fig. 10.a and Fig. 10.b) after 60,000 training times. This means that the error rate in predicting hemorrhage types and training frequency of Faster R-CNN model is lowest compared to the other models. Figure 11 shows the classification accuracy of the three network models by two measurements Average Precision (AP) and mean Average Precision (mAP). In terms AP, the Faster R-CNN Inception ResNet v2 model has the most stable classification results in comparison with the other two models (Fig. 11.a). Similarly, when evaluating the models with mAP, Faster R-CNN Inception ResNet v2 gives the highest result with mAP = 0.79 for all 4 classes of brain hemorrhage (Fig. 11.b). SSD Inception v2 and SSD MobileNet v2 have the lower mAP results of 0.72 and 0.75, respectively.



**Fig. 9.** Hemorrhage segmentation results of the proposed method

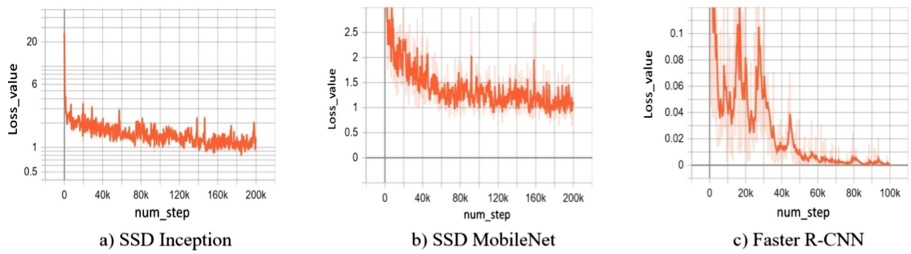


Fig. 10. Comparison of the loss values of three network models

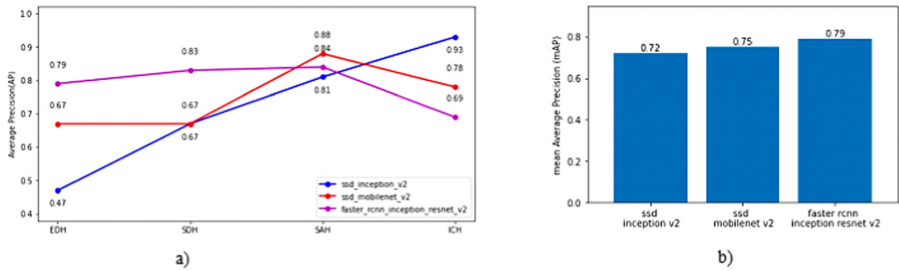


Fig. 11. Comparing the metrics of AP and mAP of the three network models for 4 classes

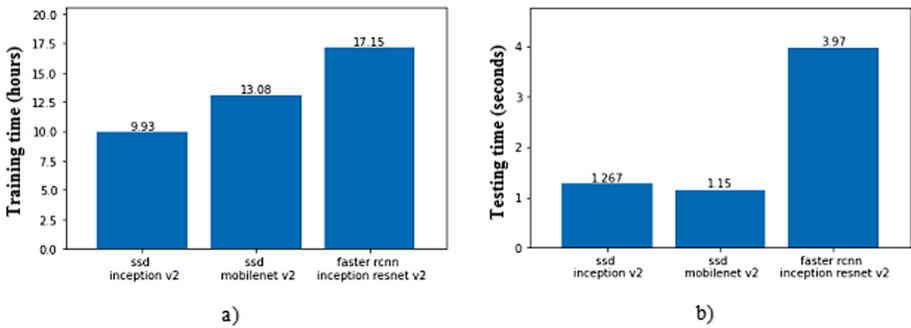


Fig. 12. Comparing training time (a) and testing time (b) of the three network models

In the same environment, the training time is quite different with 17 h 15 min of the Faster R-CNN Inception ResNet v2, 13 h 08 min of the SSD MobileNet v2, and 9 h 93 min of the SSD Inception v2. We can see that the SSD Inception v2 has the fastest training and reasonable classification time compared to other models but the mAP measurement is the lowest (see Fig. 11.b and Fig. 12). The Faster R-CNN network model has longer training time but mAP is the highest at 79%.

## 5 Conclusion

In this paper, we propose a new approach based on Hounsfield Unit and deep learning techniques. Based on Hounsfield Unit, our method not only determines the levels and duration of brain hemorrhage but also supports in automatic hemorrhage segmentation on CT/MRI images. The proposed method applies deep learning techniques with 3 network models SSD Inception v2, SSD MobileNet v2, and Faster R-CNN Inception ResNet v2. Our method can detect and classify multiple types of brain hemorrhages that appear on the same CT/MRI image of the patient. The experimental results show that the proposed method can detect multiple types of brain hemorrhage on the same CT/MRI image and the automatic hemorrhage segmentation achieves 100% accuracy. In which, the training model with Faster R-CNN Inception ResNet v2 achieves the mean average precision of 79% for the four types of hemorrhage. This research can be extended to the problem of collecting cranial CT/MRI images from the hospitals for segmentation and classification of brain hemorrhages automatically based on HU. Specialists and doctors can be supported to accurately diagnose cerebral hemorrhage and offer appropriate treatment regimens.

## References

1. <http://radclass.mudr.org/content/hounsfield-units-scale-hu-ct-numbers>
2. Brummer, M.E., Mersereau, R.M., Eisner, R.L., Lewine, R.R.: Automatic detection of brain contours in MRI data sets. *IEEE Trans. Med. Imaging* **12**(2), 153–166 (1993)
3. Buzug, T.M.: *Computed Tomography from Photon Statistics to Modern Cone-Beam CT*. Springer, Berlin (2008). <https://doi.org/10.1007/978-3-540-39408-2>. <http://www.amazon.de/Computed-Tomography-Photon-Statistics-Cone-Beam/dp/3540394079>
4. Fazli, S., Nadirkhanlou, P.: A novel method for automatic segmentation of brain tumors in MRI images. *CoRR abs/1312.7573* (2013). <http://arxiv.org/abs/1312.7573>
5. Girshick, R.B.: Fast R-CNN. *CoRR abs/1504.08083* (2015). <http://arxiv.org/abs/1504.08083>
6. Graham, R.L., Yao, F.F.: Finding the convex hull of a simple polygon. *J. Algorithms* **4**(4), 324–331 (1983)
7. ISO12052:2017: Health informatics Digital imaging and communication in medicine (DICOM) including workflow and data management. International Organization for Standardization (ISO) (2017)
8. Kuo, W., Häne, C., Yuh, E.L., Mukherjee, P., Malik, J.: Patchfcn for intracranial hemorrhage detection. *CoRR abs/1806.03265* (2018). <http://arxiv.org/abs/1806.03265>
9. Liu, Wei., et al.: SSD: single shot MultiBox detector. In: Leibe, Bastian, Matas, Jiri, Sebe, Nicu, Welling, Max (eds.) *ECCV 2016*. LNCS, vol. 9905, pp. 21–37. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2)
10. Ly, N.L., Dong, V.H.: *Brain Injury*. Medical Publishing House (2013)
11. Pham, L.T., Nguyen, N.L., Phan, L.T.H.: *Health Statistics Yearbook*. Medical Publishing House (2015)

12. Pham, N.H., Le, V.P.: CT Head Injury. Medical Publishing House (2011)
13. Phan, A.C., Phan, T.C., Vo, V.Q., Le, T.H.Y.: Automatic detection and classification of brain hemorrhage on CT/MRI images. In: Twentieth National Conference: Selected Issues of Information and Communication Technology, pp. 246–252 (2017)
14. Phong, T.D., et al.: Brain hemorrhage diagnosis by using deep learning. In: Proceedings of the 2017 International Conference on Machine Learning and Soft Computing, pp. 34–39. ACM (2017)
15. Ren, S., He, K., Girshick, R.B., Sun, J.: Faster r-CNN: Towards real-time object detection with region proposal networks. In: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R. (eds.) In: NIPS, pp. 91–99 (2015)
16. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv 2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520 (2018)
17. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. *J. Electron. Imaging* **13**(1), 146–166 (2004)
18. Simon, L., Webster, R., Rabin, J.: Revisiting precision and recall definition for generative model evaluation. CoRR abs/1905.05441 (2019), <http://arxiv.org/abs/1905.05441>
19. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-first AAAI Conference on Artificial Intelligence, AAAI (2017)