IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 May 2024, accepted 4 June 2024, date of publication 10 June 2024, date of current version 18 June 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3411932*

RESEARCH ARTICLE

# Detection and Segmentation of Subdural Hemorrhage on Head CT Images

**TUĞRUL HAKAN GENÇTÜRK[1], (Member, IEEE), FİDAN KAYA GÜLAĞIZ[1], AND İSMAİL KAYA[2]**

[1]Department of Computer Engineering, Faculty of Engineering, Kocaeli University, İzmit, 41001 Kocaeli, Turkey
[2]Department of Neurosurgery, Faculty of Medicine, Niğde Ömer Halisdemir University, 51240 Niğde, Turkey

Corresponding author: Tuğrul Hakan Gençtürk (tuhage@gmail.com)

**ABSTRACT** In today's world, there has been a significant increase in the diversity of data sources and the volume of data. This situation especially necessitates the use of technologies such as deep learning in data processing. This study thoroughly examines the processing of computed tomography (CT) images with deep learning models and their role in the diagnosis of brain hemorrhages, proposing an innovative deep learning-based model for accurately detecting and segmenting brain hemorrhages. This model combines the architectures of Mask Scoring R-CNN and EfficientNet-B2, offering an effective approach for the detection and classification of brain hemorrhages. MS R-CNN is used to detect potential hemorrhage areas in CT images, while the EfficientNet-B2 architecture serves a classification function to determine whether these areas indeed contain hemorrhages. Thus, the model offers a two-stage verification process that enhances accuracy and precision. The performance of the model has been evaluated under patient-based and random partitioning techniques using by employing two distinct datasets: an open-access and a private. In patient-based evaluation, the proposed model has an accuracy of %91.59 on open dataset and an accuracy of %90.46 on private dataset for SDH hemorrhages. In the random partitioning method, the model's accuracy rate has risen to %94.30 on open dataset and %97.33 on private dataset. Compared with similar studies in the literature, these results demonstrate that the model has a high level of accuracy and reliability. This study highlights the potential and importance of AI-supported methods in the detection of brain hemorrhages and provides a solid foundation for future work in this area. Additionally, the results obtained from an open dataset by the proposed model offer a realistic and comparable reference for future work in this field. The results obtained from a second data set also clearly demonstrate the validity of the model.

**INDEX TERMS** Deep learning, head CT scan, segmentation, subdural hemorrhage.

## I. INTRODUCTION

With the development of technology and the use of the internet, data sources have diversified and the amount of data has increased significantly. The increased size and diversity of data have made the use of new techniques and infrastructures a necessity to process and store. In addition to all these, the increase in size and type of data to be stored makes it difficult to discover meaningful information in it. At this point, new techniques and tools have been developed to process data and use different types of data as input.

Deep learning techniques have become especially popular in obtaining meaningful information from large-scale data. Another reason why these techniques have become popular is that they can take and process different types of data (audio, image, etc.) as input. In particular, the ability to process and interpret image data has enabled these techniques to be used in many areas where artificial intelligence techniques were not effective before.

One of the most critical areas where image data interpretation is essential is in healthcare. A key diagnostic technique in this field is Computerized Tomography (CT). CT scans are particularly preferred in the initial diagnosis phase due to their rapid acquisition and widespread availability. They are crucial, especially in detecting life-threatening conditions

The associate editor coordinating the review of this manuscript and approving it for publication was Gyorgy Eigner.

like internal hemorrhage, which are not visible to the naked eye or detectable through manual examination. 3D images produced by CT scans can quickly and easily identify such conditions. A prime example is brain hemorrhage, where early detection is vital for both patient survival and the preservation of their current health status. With advancements in deep learning, there has been a surge in research aimed at developing models that assist physicians in interpreting CT images of brain hemorrhages. These studies generally fall into two categories: hemorrhage detection and hemorrhage segmentation. Hemorrhage detection involves a binary classification problem, determining whether a brain hemorrhage is present or absent. In contrast, segmentation involves identifying and marking the specific area of the hemorrhage on the CT image.

Current studies on brain hemorrhage detection predominantly propose models based on deep learning; however, the use of machine learning and image processing-based methods is also common in this area. In this regard, it is possible to categorize the research conducted in this field into Convolutional Neural Networks (CNN) [1], [2], [3], [4], [5], [6], hybrid models [7], [8], [9], deep neural networks [10], [11], [12], machine learning [13], [14], and image processing [15], [16] based approaches.

When examining deep learning-based methods used for segmenting brain hemorrhages, it has been found that especially CNN [17], [18], [19] and U-Net [20], [21] architectures are commonly used in this field. While CNN-based methods are effective for the problem of segmenting brain hemorrhages, there are alternative approaches. Clustering and contour-based methods are an alternative approach for the segmentation and analysis of intracranial hemorrhages. Studies based on these methods have achieved effective results on different datasets for segmenting intracerebral hemorrhages. Examples of such studies can be found in [22], [23], [24], and [25]. There are also studies [26] that perform segmentation using classical image processing techniques. In addition to all these studies, [27] has proposed a method that, for the first time, incorporates hemorrhage expansion into the segmentation architecture and effectively transmits contextual information between slices.

When the existing studies in the literature were examined, it was determined that the studies were mostly carried out on closed datasets specific to the study, since the CT images containing brain hemorrhages were insufficient and the existing images did not contain Ground Truth data. For this reason, within the scope of the study, firstly, SDH subtype [Acute (first 3 days), subacute (3-7 days), and chronic (after 7 days) subdural hemorrhage] was used for segmentation from CT images in the CQ500 dataset, which is an open dataset, labeled by a specialist physician. Then, a model that will perform both binary classification (hemorrhage detection) and segmentation operations together is proposed. Furthermore, to ensure the proposed model works well not just in theory but in real-world scenarios, we tested its accuracy on a new, carefully compiled, and labeled private dataset aimed at identifying brain hemorrhages.

The proposed model uses Mask Scoring Region Based Convolutional Neural Network (MS R-CNN), a deep learning-based technique, in the segmentation phase. In the binary classification stage, it uses EfficientNet-B2 architecture. The MS R-CNN method identifies potential hemorrhage areas in the tomography image. The EfficientNet-B2 model, on the other hand, determines whether these regions really contain hemorrhage through the regions determined by the segmentation model. Thus, the proposed method offers a two-level verification approach.

The contributions of the proposed model can be listed as follows.

- The lack of a data set in the literature containing a sufficient number of marked brain tomographies for the purpose of segmenting subdural hemorrhage (SDH) is a major deficiency in this field. This study filled this gap by providing detailed marking of SDH-type hemorrhage in the CQ500 dataset and our private dataset by an expert.
- The results of the proposed hybrid model were obtained with both patient-based and random segmentation techniques. Thus, the generalization ability of the model and how it responds to both segmentation techniques are clearly expressed.
- When the studies carried out in this field were examined, no study was encountered that carried out the classification process over a two-stage segmented region.
- When the performance of the hybrid model was compared with similar studies in the literature, it was determined that it was more successful than studies performed on the same data set in terms of classification success.

In the second section of the article, the data set utilized and the algorithms that comprise the proposed model are discussed. The third section, dedicated to experimental studies, will provide a comprehensive presentation of the results obtained. In the fourth chapter, these results are compared with similar studies in existing literature, and the contributions of this study to science and contemporary technology are evaluated.

## II. METHOD

This section focuses on the details of the proposed model for brain hemorrhage detection. Under the Dataset subheading, the characteristics of the datasets used for analysis are presented. In the Classification and Segmentation Techniques section, the methodological approaches and techniques used for the detection of brain hemorrhage are explained.

### A. DATASET

The quality and coverage of the datasets used are vital for the detection of cerebral hemorrhage and detailed segmentation processes. The success of this process depends on the

accessibility of large and detailed datasets. Unfortunately, accessing high-quality datasets that clearly show the location and characteristics of specific medical conditions such as brain hemorrhage is a major challenge for researchers. One of the most challenging aspects of this field is the limited availability of data, especially labeled in detail. The labeling process requires deep expertise and time investment, so it can be difficult to find volunteer experts to assist in preparing such data. Additionally, privacy and ethical concerns arising from working on medical image data are also issues that need to be addressed.

Within the scope of this study, two different datasets were used to train and evaluate the proposed method. The first one is the CQ500 [28] dataset. CQ500 dataset created in 2018, at the Centre for Advanced Research in Imaging, Neurosciences and Genomics (CARING), New Delhi, India and consists of brain tomography (CT) images. In this dataset collected CT images were taken different devices as GEBrightSpeed, GE Discovery CT750 HD, GE LightSpeed, GE Optima CT660, Philips MX 16slice, Philips Access-32 CT. The images in this dataset cover various brain conditions such as cerebral hemorrhage, fracture, and midline shift. This dataset, which includes approximately 200,000 slices from 491 different CT scans, provides a comprehensive resource in the field of brain imaging. CQ500 dataset consisted of two batches and these batches including 214 (mean age 43.40; %43.92 female) and 277 (mean age 51.70; %30.31 female) scans respectively [28].

The CQ500 dataset is an openly accessible dataset designed for the detection and analysis of specific conditions, including cerebral hemorrhage. It contains a greater number of scans covering all subtypes of hemorrhages than any other open dataset in this field, giving it special significance. However, a notable challenge is that most slices in this dataset are not labeled as required for segmentation, presenting an additional problem for researchers. In this study, one of the key steps undertaken is the labeling of scans that exhibit Subdural Hematoma (SDH). These scans have been carefully marked by by Assistant Professor İsmail KAYA, a process that is crucial for the training and validation of algorithms aimed at the automatic detection and segmentation of cerebral hemorrhage. This labeling significantly enhances the dataset's value. Examples of some labeled images from the CQ500 dataset are depicted in Figure 1.

The second dataset used is a private dataset collected by our team for the detection of SDH. The images in this dataset were collected by Assistant Professor Dr. İsmail KAYA at Niğde Ömer Halisdemir University Faculty of Medicine Neurosurgery Clinic and Cumhuriyet University Faculty of Medicine Neurosurgery Clinic. The dataset includes 2434 CT images obtained from a total of 26 patients. Patients ranging from 3 to 85 years old, with a median age of 43 years for female and a mean age of 40.5 years for male. Of the patients included in the dataset, %27 were female and %73 were male (7 females and 19 males). Among the 2434 CT scans, 831 images contained Acute Subdural Hematoma (ASDH), a subtype of SDH [34]. CT images were taken with 128-slice CT devices ((Magnetom Aera; Siemens, Erlanger; Germany), (Aquilion cx edition 128 slice ct-scanner; Toshiba, Japan))in DICOM format. The collected data were initially converted to png format and anonymized. From the anonymized images, those containing ASDH were marked to obtain ground truth segmentation masks by Assistant Professor İsmail KAYA using the CVAT (Computer Vision Annotation Tool) software. During this process, the polygon marking technique was utilized [34]. Examples of some labeled images from this private dataset are depicted in Figure 2.

**TABLE 1. General layer table for VGG network architecture [30].**

| Model | Feature Extraction | | | | |
|---|---|---|---|---|---|
| | 1. Block | 2. Block | 3. Block | 4. Block | 5. Block |
| VGG-11 | Conv3-64 | Conv3-128 | Conv3-256 | Conv3-512 | Conv3-512 |
| | | | Conv3-256 | Conv3-512 | Conv3-512 |
| VGG-13 | Conv3-64 | Conv3-128 | Conv3-256 | Conv3-512 | Conv3-512 |
| | Conv3-64 | Conv3-128 | Conv3-256 | Conv3-512 | Conv3-512 |
| VGG-16 | Conv3-64 | Conv3-128 | Conv3-256 | Conv3-512 | Conv3-512 |
| | Conv3-64 | Conv3-128 | Conv3-256 | Conv3-512 | Conv3-512 |
| | | | Conv3-256 | Conv3-512 | Conv3-512 |

### B. CLASSIFICATION ALGORITHMS

In this study, three different CNN architectures commonly used in modern classification problems have been compared in terms of classification success for brain hemorrhage classification. These are VGG, ResNet, and EfficientNet architectures. These architectures have shown successful results in the field of large-scale image recognition and have formed the basis of innovative object recognition models. The most successful model as a result of this comparison has been used in the classification stage of the proposed hybrid model. The details of the compared architectures are given in the subsections.

#### 1) VGG (VISUAL GEOMETRY GROUP)

The VGG [29] architecture has been developed for effective image recognition operations especially in large-scale datasets. This architecture is a standard CNN model with many layers, and in our study, three different VGG architectures have been used: VGG-11, VGG-13, and VGG-16 [30].

VGG-11 consists of 8 convolution layers and 3 fully connected layers. It processes RGB images of 224 $\times$ 224x3 dimensions, extracting rich features with increasing numbers of parameters in each layer of the architecture. In this architecture, pooling layers are distributed to reduce the number of parameters and lighten the computational load. In the classification stage, a softmax classifier is used [30]. VGG-13 and VGG-16 architectures, while based on VGG-11, differ with changes made in the convolution layers. VGG-13 contains extra convolution layers in the first two blocks compared to VGG-11. VGG-16, starting from the third block, increases the convolution layers in each block, extracting

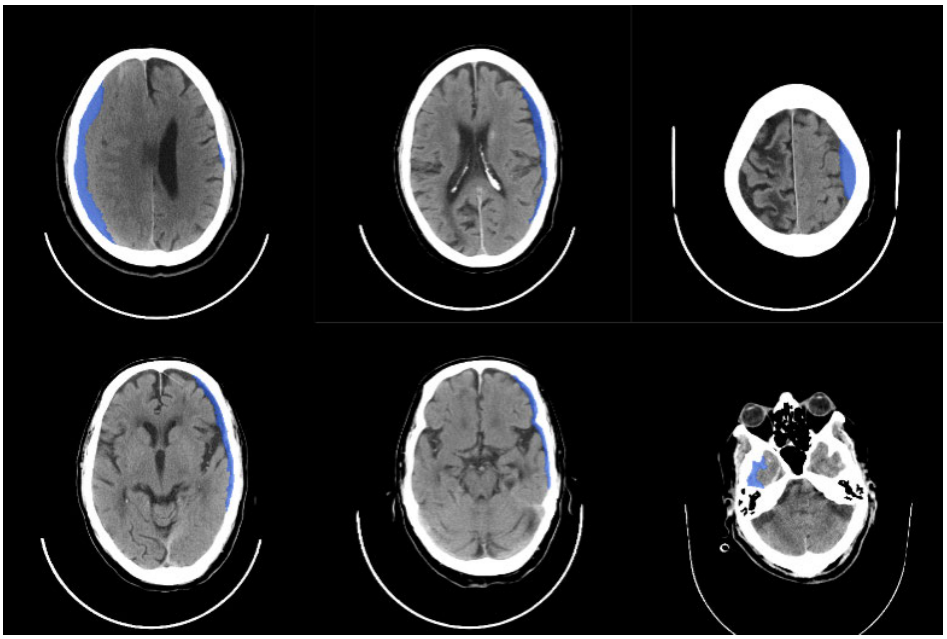

**FIGURE 1.** Labeled sample scan images from the CQ500 dataset.

**TABLE 2.** General layer table for ResNet network architecture [30].

| Model | Feature Extraction | | | | |
|---|---|---|---|---|---|
| | 1. Block | 2. Block | 3. Block | 4. Block | 5. Block |
| **Resnet-50** | 7x7, 64, Stride 2 | 3x3, MaxPool, Stride 2 $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ |
| **Resnet-101** | | 3x3, MaxPool, Stride 2 $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ | $\begin{bmatrix} 1\times1,512 \\ 3\times3,512 \\ 1\times1,2048 \end{bmatrix}\times3$ |

more detailed features. In these architectures, small $3 \times 3$ filters are used in each block, allowing for a more detailed handling of visual features. Additionally, $2 \times 2$ maximum pooling layers are added to the end of each block, intensifying the feature space. Table 1 lists the layers for different VGG models [30].

### 2) ResNet

One of the biggest challenges in deep learning models today is the difficulty in training and saturation as the depth of the model increases. ResNet [31] is a specialized CNN architecture designed to overcome this problem. ResNet adopts the "residual learning" approach to optimize the training process of the model. This approach tries to model a part or "residue" of the target function that the model needs to learn.

The fundamental idea of ResNet is to facilitate the learning process by adding the input of layers to their output. This is referred to as a "shortcut" or "residual connection." These connections enable the model to train rapidly and efficiently, even as it becomes deeper. Residual connections particularly help reduce the problem of vanishing gradients as the depth of the network increases.

The ResNet architecture has many variations with different depths, and in this study, the ResNet-50 and ResNet-101 architectures are considered. These architectures are known for their ability to maintain the model's performance even in very deep networks. As the feature maps' resolution and depth decrease within the network, they have an average pooling layer after the convolution filters and a fully connected layer at the output of the network for making predictions. Also, outputs outside of the model's residual

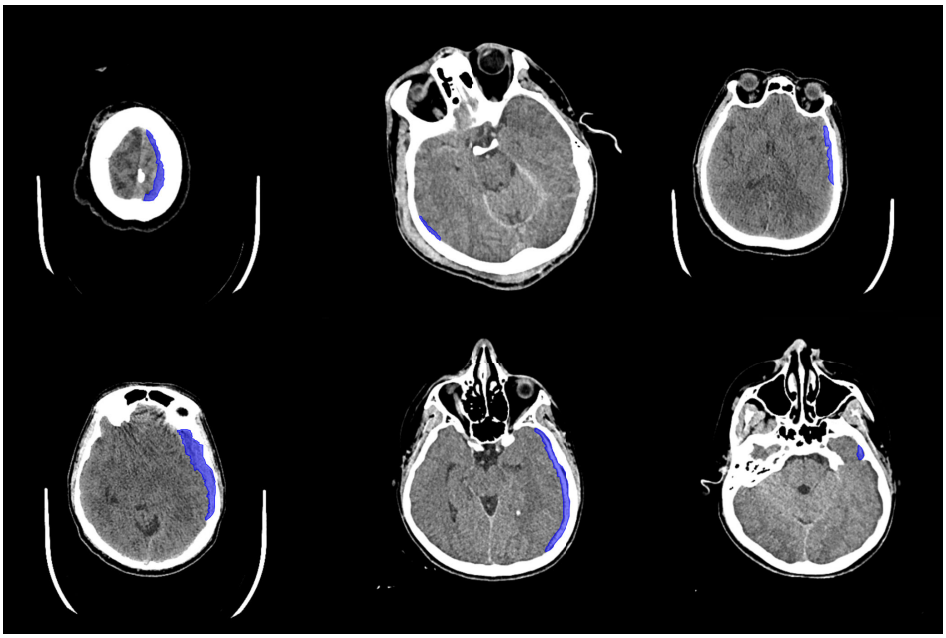

**FIGURE 2.** Labeled sample scan images from the private dataset.

networks and inputs of the residual convolution layers are combined with a simple summation process.

The ResNet-50 and ResNet-101 architectures are specifically designed for deeper and more complex tasks. Each has a different number of convolution, normalization, and activation layers. The detailed layer structure of the ResNet-50 and ResNet-101 models is shown in Table 2.

#### 3) EfficientNet

EfficientNet [32] is an innovative CNN architecture in the field of deep learning that aims to establish a balanced relationship between model complexity and computational efficiency. The key feature of this architecture is its simultaneous optimization of model width, depth, and resolution using a compound scaling method. Employing Mobile Inverted Bottleneck Convolutions (MBConv), this architecture significantly reduces computational load compared to traditional layer structures. MBConv blocks use channel expansion and squeezing processes together to minimize the model's computational cost while maximizing performance.

The EfficientNet series consists of seven different models ranging from B0 to B7. Each model differs in increasing complexity and capacity. In our study, the first four models of this series (B0-B4) have been evaluated. The model structures include various MBConv blocks, and the number of channels used in each block is increased with different numbers of filters. The advantage of this structure is that it expands the model's feature extraction capability while reducing computational costs.

The structural and performance differences between the different models of the EfficientNet series have been examined in detail in Table 3. In this table, significant features such as the number of layers, input resolution, and feature vector size of each model are presented comparatively. The model depth increases from B0 to B7, with EfficientNet-B7 having approximately 3.5 times more depth than the B0 model. Due to the observed decrease in accuracy rates as the complexity of the models used in our study increased, models beyond B4 have been excluded from evaluation.

**TABLE 3.** General feature table of EfficientNet models [30].

| Model | Layer Count | Input Resolution | Feature Vector Size |
|---|---|---|---|
| **EfficientNet-B0** | 240 | 224x224 | 1280 |
| **EfficientNet-B1** | 342 | 240x240 | 1280 |
| **EfficientNet-B2** | 342 | 260x260 | 1408 |
| **EfficientNet-B3** | 387 | 300x300 | 1536 |
| **EfficientNet-B4** | 477 | 380x380 | 1792 |
| **EfficientNet-B5** | 579 | 456x456 | 2048 |
| **EfficientNet-B6** | 669 | 528x528 | 2304 |
| **EfficientNet-B7** | 816 | 600x600 | 2560 |

### C. SEGMENTATION ALGORITHM

Image segmentation is a technique used to distinguish specific targets or areas in digital images. Especially in medical imaging, segmentation is critically important for determining the precise boundaries of pathological changes or specific structures. Segmentation techniques used for the detection of brain hemorrhages are employed to highlight and identify hemorrhage regions in the image. At this stage, it is important to choose the most appropriate method specific to the problem. Within the scope of the study, the Mask Scoring R-CNN method, which has previously been proven [34] successful in the segmentation of brain hemorrhages,

was preferred. Below, the segmentation technique used in this study is discussed.

#### 1) MASK SCORING R-CNN

Mask Scoring R-CNN [33] is an advanced version of the existing Mask R-CNN model and includes innovative features to enhance segmentation performance. This model is specifically designed to improve the accuracy of masking operations. As an innovative approach, this method adds a Mask IoU (Intersection Over Union) component to the output of the network. This addition enables more accurate scoring of the generated masks, thereby significantly improving the quality of segmentation results.

The operation process of Mask Scoring R-CNN consists of two main stages. The first stage focuses on creating masks for accurate object instances. The second stage involves generating MaskIoU scores for background objects. This model has a structure consisting of four main stages, and the first three are the same as the original Mask R-CNN model. Initially, an advanced feature extraction process is carried out through a backbone architecture built using ResNet-101+FPN. In the second stage, candidate Region Proposals (ROIs) are determined as outputs of the RPN (Region Proposal Network). In the third step, features are extracted for each candidate ROI using RoIAlign, and a segmentation mask is created. In the fourth and final stage, a MaskIoU score is calculated between the predicted mask and the actual mask.

Figure 3 visually explains the general architecture of Mask Scoring R-CNN. In this architecture, the added MaskIoU header, which improves the accuracy of segmentation masks, functions to assign a score to each mask during segmentation. This scoring assesses the congruence between the mask and the actual target, measuring the accuracy of the segmentation. This additional evaluation stage is a significant factor in increasing the accuracy of the model, especially for applications such as medical imaging and complex object recognition. The innovative structure of Mask Scoring R-CNN sets new standards in advanced image analysis applications by increasing the precision and reliability of segmentation algorithms.

### D. PROPOSED HYBRID MODEL

Hybrid methodologies, which integrate different approaches, are techniques that often lead to more comprehensive and accurate results. Preferring these methods while working with complex situations and large datasets generally provides more efficient and successful outputs. Our study's proposed hybrid approach to brain hemorrhage detection fundamentally consists of two main components: segmentation and classification. This dual approach encompasses both the detection and classification of brain hemorrhages. The developed methodology is visually presented in Figure 4.

In the first stage, brain tomography scans stored in DICOM format were converted to PNG format. Then, these scans were divided into two main categories: 'healthy' and 'hemorrhagic'. Scans containing hemorrhages were split into %80 for training and %20 for testing. The MS R-CNN segmentation model was trained using the designated training data. This model identifies images in the test dataset that contain potential signs of brain hemorrhages and marks these areas with a mask and bounding box. If the model does not detect signs of hemorrhage in a scan, that scan is classified as 'healthy'. However, the segmentation model can sometimes mistakenly mark healthy areas as 'hemorrhagic'. To reduce such false positives, the EfficientNet-B2 classification model is implemented as a second layer of security. This model forms the second phase of the hybrid methodology, evaluating the areas segmented in the first phase to determine whether they truly contain brain hemorrhages.

During the training process of the EfficientNet-B2 model, data from two different sources were used. The first input is the bounding box information of ''healthy'' brain scans that have been incorrectly marked due to misleading symptoms by our segmentation model. The second input is the part of brain scans determined as ''hemorrhagic'' used in training. These scans are also processed through the segmentation model. The model analyzes the scans and marks hemorrhagic regions with bounding boxes. These boxes represent areas with actual hemorrhages, and this information is also included in the training of the EfficientNet-B2 model. Both types of data are used in the training process of the EfficientNet-B2 model. False positives, where the model mistakenly marks healthy brain areas as hemorrhagic, help prevent misclassification, while bounding boxes indicating real hemorrhages improve the model's ability to correctly identify hemorrhages. This dual approach enhances both the precision and accuracy of the model.

## III. EXPERIMENTAL RESULTS AND ANALYSIS

### A. SETUP AND PERFORMANCE METRICS

During the study, a server equipped with an Intel Xeon 2.2GHz processor, 28 GB RAM and 16 GB NVIDIA Tesla P100 GPU was used. In the software development process, Python programming language was preferred.

The performance evaluation of the model was made using accuracy, precision, recall, F1 score and AUC (Area Under the Curve) metrics. Accuracy refers to the ratio of how many of all the predictions the model makes are actually correct. Precision indicates the ratio of correctly identified hemorrhage images among all those predicted to be hemorrhage.

Recall refers to the proportion of correctly detected hemorrhage images out of all images that are actually hemorrhage. The F1 score is the harmonic mean of precision and precision and reflects the balanced performance of the model. AUC is an indicator of how well the classifier can distinguish classes; This is based on the relationship between the false positive rate and the true positive rate. A high AUC value indicates that the model can distinguish classes more effectively. The formulas used to calculate these metrics are

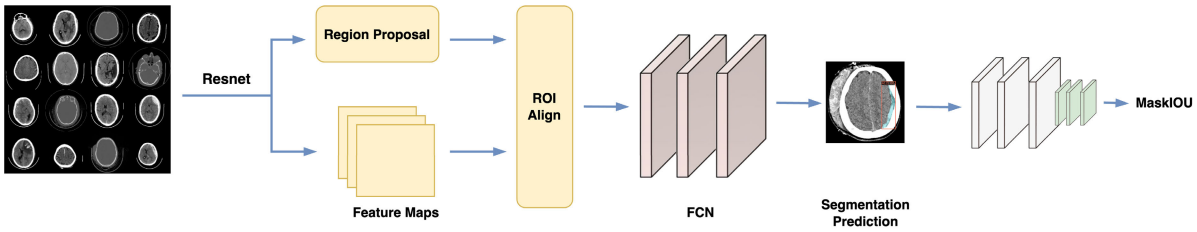


**FIGURE 3.** General architecture of mask scoring R-CNN [34].

included in equations 1, 2, 3 and 4 respectively.

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \quad (1)$$

$$\text{Precision} = TP/(TP + FP) \quad (2)$$

$$\text{Recall} = TP/(TP + FN) \quad (3)$$

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

The terms used in these formulas are as follows: True Positive (TP), which denotes the number of images that the model correctly predicts as hemorrhagic; True Negative (TN), which denotes the number of images that the model correctly predicts as non-hemorrhagic; False Positive (FP), which denotes the number of images that the model predicts as hemorrhagic but are actually non-hemorrhagic; and False Negative (FN), which denotes the number of images that the model predicts as non-hemorrhagic but are actually hemorrhagic.

In determining the accuracy of segmentation methods, the overlap ratio (IoU - Intersection over Union) value was used as a threshold. IoU measures how much the predicted boundary (the object boundary determined by the algorithm) overlaps the actual object boundary. This measurement is calculated with equations 5. A and B in the formula represent the outputs of the model's prediction and actual data, respectively. To evaluate the accuracy of an image segmentation, the IoU intersection rate must be above a specified threshold value. Typically, an IoU threshold of 0.5 or higher is used to determine whether a prediction is a true positive (true) or a false positive (false). In this study, the IOU value was determined as 0.5 and detection and segmentation accuracy rates were calculated based on this value.

$$\text{IOU} = \frac{A \cap B}{A \cup B} \quad (5)$$

The mAP (mean average precision) criterion was used to evaluate detection and segmentation accuracies. mAP is the average of the precision and recall metrics calculated for detected bounding boxes. First, the average precision (AP) is calculated for each class (such as subdural hemorrhage and background).Then, these average precisions are averaged across classes to obtain mAP. In order for a detection to be considered a true positive, the IoU of that detection must be above a certain threshold. Once all images are examined against this threshold value, precision and recall values can be calculated. Precision represents the total number of correct examples produced by the model and is calculated by the following formula.

$$P = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \quad (6)$$

The total number of positive predictions the model can produce is measured by a recall calculated as follows:

$$R = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \quad (7)$$

The area under the PR curve is used to determine average accuracy. Equations 8 shows the calculation of AP:

$$\text{AP} = \sum_{n=1}^{N} [R(n) - R(n-1)] \cdot \max P(n) \quad (8)$$

N is the total number of PR points identified. mAP is calculated using the following equation:

$$mAP = \frac{1}{n} \sum_{n=1}^{n} AP_i \quad (9)$$

In addition to all these metrics, to evaluate the proposed model Gradient-weighted Class Activation Mapping (Grad-CAM) [44] visualization was used. Grad-CAM shows which regions of the image the CNN-based methods focus on during the decision-making phase. In order to obtain the Grad-CAM map of a class c, the importance of each feature map for class c must first be obtained. This importance value is given in equation 10.

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (10)$$

In equation 10, $\alpha_k^c$ express the 'importance' (contribution of the k-th feature map to the gradients) of feature map k for a class c. $A_{ij}^k$ is the value of the k-th feature map in i,j coordinates. The y symbol represents the output class and the $y^c$ value is the $c^{\text{th}}$ value of the output class and Z is the term used to normalize the gradient sum.

$$L_{\text{Grad-CAM}}^c = \text{Re}\, LU\left(\sum_k \alpha_k^c A^k\right) \quad (11)$$

In equation 11, $L_{\text{Grad-CAM}}^c$ represents the Grad-CAM map of class c. The $A^k$ value in the formula refers the feature

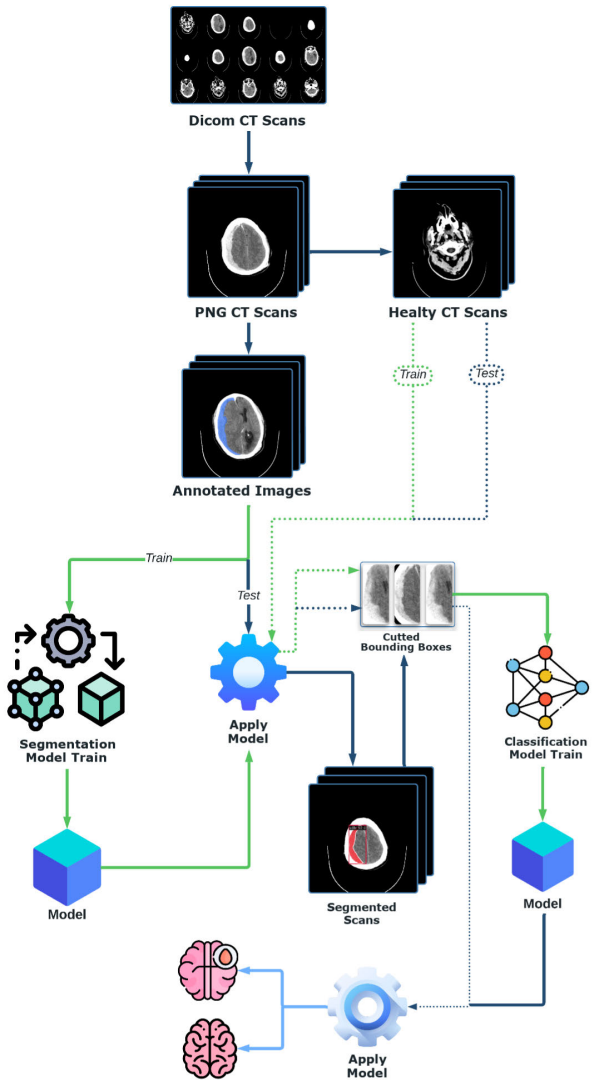


**FIGURE 4.** General architecture of the proposed system.

map of the $k^{\text{th}}$ layer. The first stage of the proposed model is the detection of the hemorragic area. Therefore, when the class label of the hemorragic area is considered as c and the Grad-CAM map is obtained, it will be visualized which features are focused on to detect this hemorragic area.

### B. RESULTS

Within the scope of the study, a two-stage hybrid model has been developed. To select the model to be used in the classification stage, a comparison of VGG, ResNet, and EfficientNet architectures was first conducted on the CQ500 dataset. The obtained results are comparatively presented in Figure 5 and Figure 6. Evaluating these figures, it appears that the EfficientNet-B2 model is the most suitable among the evaluated models for the classification of brain hemorrhages using CT images. Therefore, the EfficientNet-B2 model has been used as the classifier in the hybrid model.

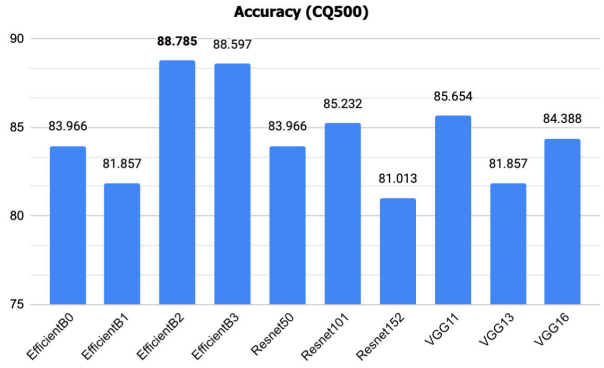


**FIGURE 5.** Accuracies of classification algorithms.

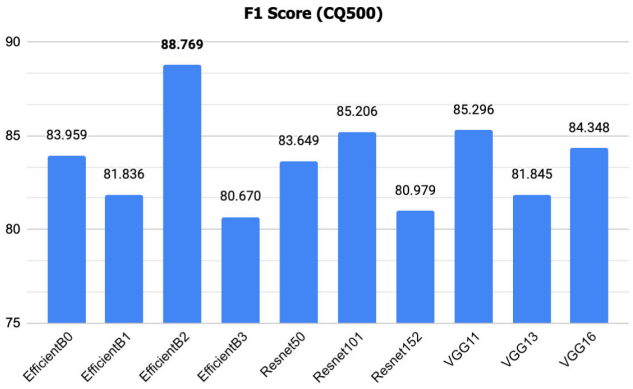


**FIGURE 6.** F1 score results of classification algorithms.

To obtain the final results through the hybrid model, the dataset must first be correctly partitioned. At this stage, two different techniques, patient-based and random partitioning, were used. In patient-based partitioning, all images of a patient are included either in the training set or the test set, completing the model's training and testing process. The main purpose of this approach is to prevent the model from misleadingly showing high performance by seeing different images of the same patient. Patient-based partitioning helps us better understand how well the model can generalize in real-world situations. In random partitioning, images in the dataset are distributed randomly into training and test sets. In this case, images of the same patient can be found in both training and test sets. This method is generally preferred as a faster way to assess the overall performance of the model. However, the disadvantage of this approach is that the model's performance can be misleading due to images of the same patient being in both training and test sets. Studies [35], [36], [37], [38], [39], [40] that have performed data partitioning with both approaches are present in the literature.

Figure 7 shows the results obtained when only the classifier model is used, i.e., when images are subjected to the classification algorithm without going through the segmentation process on CQ500 dataset. In patient-based partitioning, the model's accuracy rate is measured at %88.79, precision at %84.91, recall at %91.84, and F1 score at %88.24. These

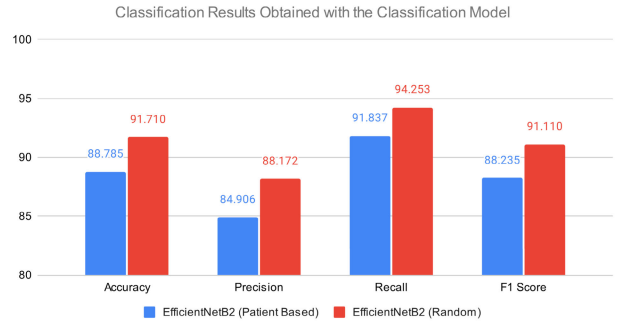


**FIGURE 7.** Results obtained with only classifier model (CQ500 Dataset).

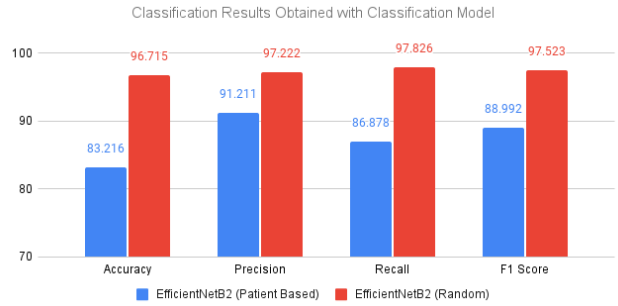


**FIGURE 8.** Results obtained with only classifier model (Private Dataset).

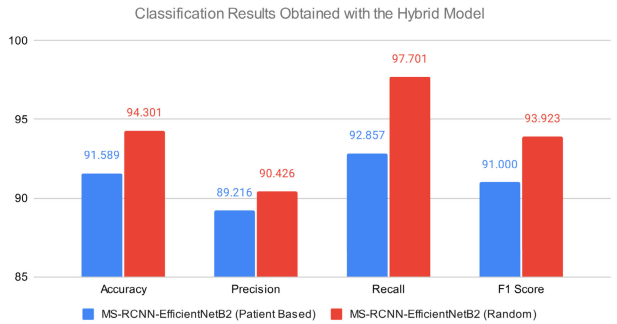


**FIGURE 9.** Classification results obtained with hybrid model (CQ500 Dataset).

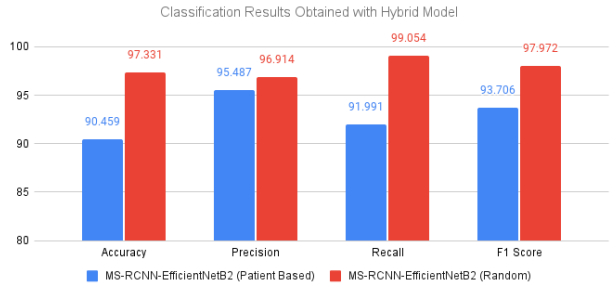


**FIGURE 10.** Classification results obtained with hybrid model (Private Dataset).

results indicate that the model generally performs well and can make accurate predictions. In random partitioning, the model's performance further increases. The accuracy rate is %91.71, precision is %88.17, recall is %94.25, and F1 score is %91.11. Also results obtained from private dataset shown in Figure 8. In patient-based partitioning, the model's accuracy rate is measured at %83.22, precision at %91.21, recall at %86.88, and F1 score at %88.99. In random partitioning, the model's performance further increases as in the CQ500 dataset. The accuracy rate is %96.72, precision is %97.22, recall is %97.83, and F1 score is %97.52.

In general, the model shows high performance in both partitioning methods and effectively classifies brain hemorrhages. However, the model shows higher performance in the randomly partitioned dataset. This suggests that the model's ability to generalize information learned from specific patients may be somewhat limited in patient-based partitioning.

Figure 9 shows the results of hybrid method using the CQ500 dataset. In patient-based partitioning, the model's accuracy rate is %91.59, precision rate is %89.22, recall is %92.86, and F1 score is %91.00. In random partitioning, the model's performance has further increased. On CQ500 dataset, the accuracy rate is %94.30, precision rate is %90.43, recall is %97.70, and F1 score is %93.92. Figure 10 shows the results obtained from the developed hybrid method using the private dataset. In patient-based partitioning, the model's accuracy rate is %90.46, precision rate is %95.49, recall is %91.99, and F1 score is %93.70. On private dataset, the accuracy rate is %97.33, precision rate is %96.91, recall is %99.05, and F1 score is %97.97. These results show that the model displays a higher overall performance in a randomly partitioned dataset. The closeness of the model's recall and accuracy rates indicates that it can accurately predict both positive (hemorrhage) and negative (no hemorrhage) examples.

**TABLE 4.** Classification results obtained with the hybrid model.

| Method | Year | Accuracy | AUC | Split Method |
|---|---|---|---|---|
| Hybrid Model (CQ500 Dataset) | 2023 | 91.5888 | 92.9200 | Patient Based |
| Hybrid Model (CQ500 Dataset) | 2023 | 94.3005 | 96.9200 | Random |
| [42] | 2020 | - | 96.5400 | - |
| [43] | 2022 | 71.7000 | 90.9000 | Patient Based |
| [28] | 2018 | - | 94.0000 | - |
| [45] | 2023 | - | 90.0-92.0 | - |
| [3]* | 2023 | 96.2000 | - | - |
| [1]* | 2021 | 99.8600 | - | Random |
| [41]* | 2021 | - | 93.9000 | Random |

In the proposed method, one of the most basic factors affecting the final classification result is the segmentation process performed in the first stage. Therefore, providing segmentation results will also be useful in evaluating the general model. When the IoU value is taken as 0.5, with the MS R-CNN method, the mAP value was obtained as 0.8780 on the CQ500 dataset [34] and as 0.8980 on the private dataset.These results demonstrate that the segmentation process in the proposed method achieves high accuracy, providing a solid foundation for the subsequent classification stages.

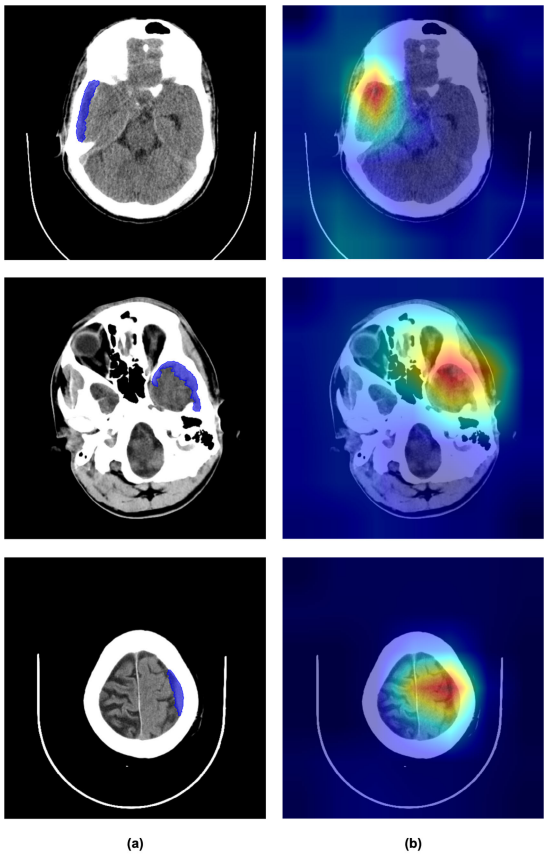


**FIGURE 11.** Local explainability of proposed hybrid with Grad-CAM (CQ500 Dataset).

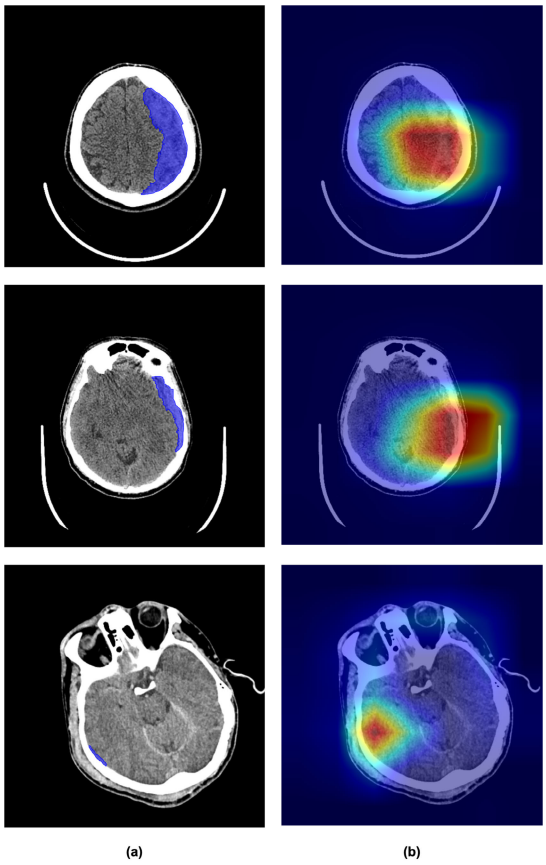


**FIGURE 12.** Local explainability of proposed hybrid with Grad-CAM (Private Dataset).

In addition to model accuracy evaluation, Grad-CAM images were also obtained showing the areas on which the proposed model focuses for the detection of SDH. With the obtained Grad-CAMs, it is clearly shown where the important features of the model are located. Figure 11 shows sample Grad-CAM visualisations obtained from the CQ500 dataset, and Figure 12 shows sample Grad-CAM visualisations obtained from the private dataset.Visualizations obtained from both datasets clearly show the success of the model. When Figures 11 and 12 are compared, it is seen that the model focuses better on the SDH area in the Grad-CAM images obtained from the private dataset. The main factor causing this situation is the difference in the devices used to obtain CT images. Images with higher resolution and more detail were obtained when shooting with devices with higher slice values. This enabled the model to focus on more accurate points.

Table 4 presents a comparison of the proposed method with studies conducted on the CQ500 dataset in the literature. There is no study in the literature performing the partitioning of SDH hemorrhages on the CQ500 dataset. Therefore, the comparison results in the table include the final classification success of the proposed model. Looking at similar studies in the literature, the study conducted by [1] achieved the highest accuracy rate (%99.86). However, the dataset was reduced for use in the study, which would influence the accuracy value, making the results not entirely comparable with the proposed model in our study. The studies by [3] and [41] used the CQ500 dataset only for validation and used different datasets in the training process. Despite being trained with datasets containing more comprehensive and higher-resolution samples, the study by [3] achieved a success rate similar to our model (%96.2), while the study by [1] showed lower performance (AUC value of 86) than our model.

Some studies have evaluated their proposed models with the help of the AUC metric. When compared with studies that base their evaluations on the AUC metric, our model appears to demonstrate superior performance.

## IV. CONCLUSION

With the development of technology, applications supported by artificial intelligence are being utilized in many different areas. One of the most significant fields where these applications have gained popularity is medicine. In this field, where early diagnosis is crucial, the goal is to expedite the diagnostic process using AI-supported applications and to improve quality of life through early diagnosis and treatment.

In this direction, within the scope of the study, an artificial intelligence-based two-level hybrid model was developed

to support doctors in the detection and classification of brain hemorrhages. At the first level, the MS R-CNN model separates scan images into ''hemorrhage'' and ''healthy''. Then, it marks the areas where images contain hemorrhage. At this stage, some images that do not contain hemorrhage can be labeled as ''hemorrhagic'' by the model. In order to correct the error of the model in the first stage, two-level verification is provided with the EfficientNet-B2 model in the second stage. The proposed model has an accuracy value of %91.59 on a patient basis and %94.30 on a random basis for SDH on CQ500 dataset and has an accuracy value of %90.46 on a patient basis and %97.33 on a random basis for SDH on private dataset. Although the model was trained with a limited amount of data, it has higher accuracy and AUC value than studies in the literature that trained using the CQ500 dataset. The fact that the results obtained from the private dataset used are compatible with the results obtained from the open dataset showed the generalizability of the proposed model.The proposed model clearly demonstrates the contribution of artificial intelligence-based applications to be developed in this field in supporting doctors.The fact that the model obtained results on an open data set offers the opportunity to make more realistic comparisons for studies in this field.

In the study, model training was carried out by converting DICOM data into images with. png extension. In future studies, the model can be improved and its accuracy can be increased by using metadata in DICOM. At the same time, training can be performed on more comprehensive datasets with data diversity to increase the overall performance and applicability of the model. This will have a significant impact on the generalization and success of the model.

## REFERENCES


[1] P. Kumaravel, S. Mohan, J. Arivudaiyanambi, N. Shajil, and H. N. Venkatakrishnan, ''A simplified framework for the detection of intracranial hemorrhage in CT brain images using deep learning,'' *Current Med. Imag. Formerly Current Med. Imag. Rev.*, vol. 17, no. 10, pp. 1226–1236, Oct. 2021, doi: 10.2174/1573405617666210218100641.

[2] S. P. Singh, L. Wang, S. Gupta, B. Gulyás, and P. Padmanabhan, ''Shallow 3D CNN for detecting acute brain hemorrhage from medical imaging sensors,'' *IEEE Sensors J.*, vol. 21, no. 13, pp. 14290–14299, Jul. 2021, doi: 10.1109/JSEN.2020.3023471.

[3] M. Asif, M. A. Shah, H. A. Khattak, S. Mussadiq, E. Ahmed, E. A. Nasr, and H. T. Rauf, ''Intracranial hemorrhage detection using parallel deep convolutional models and boosting mechanism,'' *Diagnostics*, vol. 13, no. 4, p. 652, Feb. 2023, doi: 10.3390/diagnostics13040652.

[4] L. Cortés-Ferre, M. A. Gutiérrez-Naranjo, J. J. Egea-Guerrero, S. Pérez-Sánchez, and M. Balcerzyk, ''Deep learning applied to intracranial hemorrhage detection,'' *J. Imag.*, vol. 9, no. 2, p. 37, Feb. 2023, doi: 10.3390/jimaging9020037.

[5] M. Burduja, R. T. Ionescu, and N. Verga, ''Accurate and efficient intracranial hemorrhage detection and subtype classification in 3D CT scans with convolutional and long short-term memory neural networks,'' *Sensors*, vol. 20, no. 19, p. 5611, Oct. 2020, doi: 10.3390/s20195611.

[6] M. Karki, J. Cho, E. Lee, M. H. Hahm, S. Y. Yoon, M. Kim, S. Park, ''CT window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings,'' *Artif. Intell. Med.*, vol. 106, no. 101850, pp. 1–8 , 2020, doi: 10.1016/j.artmed.2020.101850.

[7] O. Ozaltin, O. Coskun, O. Yeniay, and A. Subasi, ''Classification of brain hemorrhage computed tomography images using OzNet hybrid algorithm,'' *Int. J. Imaging Syst. Technol.*, vol. 33, no. 1, pp. 69–91, 2023, doi: 10.1002/ima.22806.

[8] D. Wang, C. Wang, L. Masters, and M. Barnett, ''Masked multi-task network for case-level intracranial hemorrhage classification in brain CT volumes,'' presented at the Med. Image Comput. Comput. Assist. Intervent., Lima, Peru, Oct. 2020.

[9] S. Santhoshkumar, V. Varadarajan, S. Gavaskar, J. J. Amalraj, and A. Sumathi, ''Machine learning model for intracranial hemorrhage diagnosis and classification,'' *Electronics*, vol. 10, no. 21, p. 2574, Oct. 2021, doi: 10.3390/electronics10212574.

[10] L. Zhang, W. Miao, C. Zhu, Y. Wang, Y. Luo, R. Song, L. Liu, and J. Yang, ''A weakly supervised-guided soft attention network for classification of intracranial hemorrhage,'' *IEEE Trans. Cogn. Develop. Syst.*, vol. 15, no. 1, pp. 42–53, Mar. 2023, doi: 10.1109/TCDS.2022.3141591.

[11] A. C. Phan, H. D. Tran, and T. C. Phan, ''Efficient brain hemorrhage detection on 3D CT scans with deep neural network,'' presented at the 8th Int. Conf. Future Data Security Eng. (FDSE), Nov. 2021.

[12] C. S. S. Anupama, M. Sivaram, E. L. Lydia, D. Gupta, and K. Shankar, ''Synergic deep learning model-based automated detection and classification of brain intracranial hemorrhage images in wearable networks,'' *Pers. Ubiquitous Comput.*, vol. 26, no. 1, pp. 1–10, Feb. 2022, doi: 10.1007/s00779-020-01492-2.

[13] J.-L. Solorio-Ramírez, M. Saldana-Perez, M. D. Lytras, M.-A. Moreno-Ibarra, and C. Yáñez-Márquez, ''Brain hemorrhage classification in CT scan images using minimalist machine learning,'' *Diagnostics*, vol. 11, no. 8, p. 1449, Aug. 2021, doi: 10.3390/diagnostics11081449.

[14] H. Chen, S. Khan, B. Kou, S. Nazir, W. Liu, and A. Hussain, ''A smart machine learning model for the detection of brain hemorrhage diagnosis based Internet of Things in smart cities,'' *Complexity*, vol. 2020, pp. 1–10, Sep. 2020, doi: 10.1155/2020/3047869.

[15] G. Zhang, K. Chen, S. Xu, P. C. Cho, Y. Nan, X. Zhou, C. Lv, C. Li, and G. Xie, ''Lesion synthesis to improve intracranial hemorrhage detection and classification for CT images,'' *Computerized Med. Imag. Graph.*, vol. 90, Jun. 2021, Art. no. 101929, doi: 10.1016/j.compmedimag.2021.101929.

[16] K. M. Pai, ''Region detection and segmentation of brain hemorrhage using algorithmic approach of image processing,'' *J. Phys., Conf. Ser.*, vol. 2161, no. 2022, pp. 1–8, 2022, Art. no. 012070, doi: 10.1088/1742-6596/2161/1/012070.

[17] P. Inkeaw, S. Angkurawaranon, P. Khumrin, N. Inmutto, P. Traisathit, J. Chaijaruwanich, C. Angkurawaranon, and I. Chitapanarux, ''Automatic hemorrhage segmentation on head CT scan for traumatic brain injury using 3D deep learning model,'' *Comput. Biol. Med.*, vol. 146, no. 2022, pp. 1–11, Art. no. 105530, doi: 10.1016/j.compbiomed.2022.105530.

[18] A. C. Phan, H. P. Cao, T. N. Trieu, and T. C. Phan, ''Detection and classification of brain hemorrhage using Hounsfield unit and deep learning techniques,'' presented at the 7th Int. Conf. Future Data Security Eng. (FDSE), Nov. 2020.

[19] H. Yao, C. Williamson, J. Gryak, and K. Najarian, ''Automated hematoma segmentation and outcome prediction for patients with traumatic brain injury,'' *Artif. Intell. Med.*, vol. 107, Jul. 2020, Art. no. 101910, doi: 10.1016/j.artmed.2020.101910.

[20] M. D. Hssayeni, M. S. Croock, A. D. Salman, H. F. Al-khafaji, Z. A. Yahya, and B. Ghoraani, ''Intracranial hemorrhage segmentation using a deep convolutional model,'' *Data*, vol. 5, no. 1, p. 14, Feb. 2020, doi: 10.3390/data5010014.

[21] J. Xu, R. Zhang, Z. Zhou, C. Wu, Q. Gong, H. Zhang, S. Wu, G. Wu, Y. Deng, C. Xia, and J. Ma, ''Deep network for the automatic segmentation and quantification of intracranial hemorrhage on CT,'' *Frontiers Neurosci.*, vol. 14, pp. 1–7, Jan. 2021, doi: 10.3389/fnins.2020.541817.

[22] I. Kumar, C. Bhatt, and K. U. Singh, ''Entropy based automatic unsupervised brain intracranial hemorrhage segmentation using CT images,'' *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 6, pp. 2589–2600, Jun. 2022, doi: 10.1016/j.jksuci.2020.01.003.

[23] H. Yao, C. Williamson, J. Gryak, and K. Najarian, ''Brain hematoma segmentation using active learning and an active contour model,'' presented at the 7th Int. Work Conf. Bioinf. Biomed. Eng. (IWBBIO), Granada, Spain, May 2019.

[24] W. Tu, L. Kong, R. Karunamuni, K. Butcher, L. Zheng, and R. McCourt, ''Nonlocal spatial clustering in automated brain hematoma and edema segmentation,'' *Appl. Stochastic Models Bus. Ind.*, vol. 35, no. 2, pp. 321–329, Mar. 2019, doi: 10.1002/asmb.2431.

[25] N. S. Bhadauria, I. Kumar, H. S. Bhadauria, and R. B. Patel, "Hemorrhage detection using edge-based contour with fuzzy clustering from brain computed tomography images," *Int. J. Syst. Assurance Eng. Manage.*, vol. 12, no. 6, pp. 1296–1307, Dec. 2021, doi: 10.1007/s13198-021-01269-7.

[26] N. Farzaneh, C. A. Williamson, C. Jiang, A. Srinivasan, J. R. Bapuraj, J. Gryak, K. Najarian, and S. M. R. Soroushmehr, "Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries," *Diagnostics*, vol. 10, no. 10, p. 773, Sep. 2020, doi: 10.3390/diagnostics10100773.

[27] X. Li, G. Luo, W. Wang, K. Wang, Y. Gao, and S. Li, "Hematoma expansion context guided intracranial hemorrhage segmentation and uncertainty estimation," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 3, pp. 1140–1151, Mar. 2022, doi: 10.1109/JBHI.2021.3103850.

[28] S. Chilamkurthy, R. Ghosh, S. Tanamala, M. Biviji, N. G. Campeau, V. K. Venugopal, V. Mahajan, P. Rao, and P. Warier, "Deep learning algorithms for detection of critical findings in head CT scans: A retrospective study," *Lancet*, vol. 392, no. 10162, pp. 2388–2396, Dec. 2018, doi: 10.1016/s0140-6736(18)31645-3.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," presented at the 3rd Int. Conf. Learn. Represent. (ICLR), San Diego, CA, USA, May 2015.

[30] T. H. Gençtürk, F. K. Gülağız, and I. Kaya, "Derin Öğrenme Yöntemleri kullanılarak BT taramalarında beyin kanaması teşhisinin karışlaştırmalı bir analizi," *J. Intell. Syst., Theory Appl.*, vol. 6, no. 1, pp. 75–84, Mar. 2023, doi: 10.38016/jista.1215025.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[32] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," presented at the 36th Int. Conf. Mach. Learn. (ICML), Long Beach, CA, USA, Jun. 2019.

[33] Z. Huang, L. Huang, Y. Gong, C. Huang, and X. Wang, "Mask scoring R-CNN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seoul, South Korea, Jun. 2019, pp. 6402–6411.

[34] T. H. Gençtürk, I. Kaya, and F. Kaya Gülagiz, "A comparative study on subdural brain hemorrhage segmentation," presented at the Int. Conf. Comput., Intell. Data Anal., Kocaeli, Turkey, Sep. 2022.

[35] M. Maftouni, A. C. C. Law, B. Shen, Z. J. K. Grado, Y. Zhou, and N. A. Yazdi, "A robust ensemble-deep learning model for COVID-19 diagnosis based on an integrated CT scan images database," presented at the IISE Annu. Conf. Expo, May 2021.

[36] A. P. Umejiaku, P. Dhakal, and V. S. Sheng, "Detecting COVID-19 effectively with transformers and CNN-based deep learning mechanisms," *Appl. Sci.*, vol. 13, no. 6, p. 4050, Mar. 2023, doi: 10.3390/app13064050.

[37] J. Yang, R. Shi, D. Wei, Z. Liu, L. Zhao, B. Ke, H. Pfister, and B. Ni, "MedMNIST V2—A large-scale lightweight benchmark for 2D and 3D biomedical image classification," *Sci. Data*, vol. 10, no. 1, pp. 1–10, Jan. 2023, doi: 10.1038/s41597-022-01721-8.

[38] P. Silva, E. Luz, G. Silva, G. Moreira, R. Silva, D. Lucio, and D. Menotti, "COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis," *Informat. Med. Unlocked*, vol. 20, Jan. 2020, Art. no. 100427, doi: 10.1016/j.imu.2020.100427.

[39] M. Abd Elaziz, A. Dahou, N. A. Alsaleh, A. H. Elsheikh, A. I. Saba, and M. Ahmadein, "Boosting COVID-19 image classification using MobileNetV3 and Aquila optimizer algorithm," *Entropy*, vol. 23, no. 11, p. 1383, Oct. 2021, doi: 10.3390/e23111383.

[40] P. Huang, S. Park, R. Yan, J. Lee, L. C. Chu, C. T. Lin, A. Hussien, J. Rathmell, B. Thomas, C. Chen, R. Hales, D. S. Ettinger, M. Brock, P. Hu, E. K. Fishman, E. Gabrielson, and S. Lam, "Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: A matched case-control study," *Radiology*, vol. 286, no. 1, pp. 286–295, Jan. 2018, doi: 10.1148/radiol.2017162725.

[41] E. Lin and E. Yuh, "Semi-supervised learning for generalizable intracranial hemorrhage detection and segmentation," 2021, *arXiv:2105.00582*.

[42] N. T. Nguyen, D. Q. Tran, N. T. Nguyen, and H. Q. Nguyen, "A CNN-LSTM architecture for detection of intracranial hemorrhage on CT scans," 2020, *arXiv:2005.10992*.

[43] M. López-Pérez, A. Schmidt, Y. Wu, R. Molina, and A. K. Katsaggelos, "Deep Gaussian processes for multiple instance learning: Application to CT intracranial hemorrhage detection," *Comput. Methods Programs Biomed.*, vol. 219, Jun. 2022, Art. no. 106783, doi: 10.1016/j.cmpb.2022.106783.

[44] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.

[45] J. Teneggi, P. H. Yi, and J. Sulam, "Examination-level supervision for deep learning-based intracranial hemorrhage detection on head CT scans," *Radiol., Artif. Intell.*, vol. 6, no. 1, pp. 1–10, Jan. 2024, doi: 10.1148/ryai.230159.

**TUĞRUL HAKAN GENÇTÜRK** (Member, IEEE) received the B.S. and M.S. degrees in computer engineering from Kocaeli University, Kocaeli, in 2021 and 2023, respectively, where he is currently pursuing the Ph.D. degree. Since 2018, he has been a Researcher with the Embedded and Sensor Systems Laboratory, Kocaeli University. In 2022, he began serving as a Research Assistant with the Computer Engineering Department, Kocaeli University.

**FİDAN KAYA GÜLAĞIZ** received the B.Sc., M.S., and Ph.D. degrees in computer engineering from Kocaeli University, Turkey, in 2010, 2012, and 2018, respectively. She is currently an Assistant Professor with the Department of Computer Engineering, Kocaeli University. Her main research interests include artificial intelligence, machine learning, and distributed systems.

**İSMAİL KAYA** received the Medical degree from Eskişehir Osmangazi University, following initial studies at Karadeniz Technical University. He specialized in neurosurgery with Cumhuriyet University, where he researched the correlation between spinal canal types and transpedicular screw placement. He is currently continuing his research as an Assistant Professor at the Faculty of Medicine, Department of Neurosurgery, Niğde Ömer Halisdemir University.