



## OPEN ACCESS

## EDITED BY

Baptiste Magnier,  
Mines-Telecom Institute Alès, France

## REVIEWED BY

Zhenghua Xu,  
Hebei University of Technology, China

Dan Wu,  
Shenzhen Institutes of Advanced Technology  
(CAS), China

## \*CORRESPONDENCE

Shaohui Huang  
✉ hsh@xmu.edu.cn  
Zhanxiang Wang  
✉ wangzx@xmu.edu.cn

†These authors have contributed equally to this work

## SPECIALTY SECTION

This article was submitted to  
Brain Imaging Methods,  
a section of the journal  
Frontiers in Neuroscience

RECEIVED 07 December 2022

ACCEPTED 23 February 2023

PUBLISHED 14 March 2023

## CITATION

Gong K, Dai Q, Wang J, Zheng Y, Shi T, Yu J,  
Chen J, Huang S and Wang Z (2023) Unified  
ICH quantification and prognosis prediction in  
NCCT images using a multi-task interpretable  
network. *Front. Neurosci.* 17:1118340.  
doi: 10.3389/fnins.2023.1118340

## COPYRIGHT

© 2023 Gong, Dai, Wang, Zheng, Shi, Yu, Chen,  
Huang and Wang. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).  
The use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Unified ICH quantification and prognosis prediction in NCCT images using a multi-task interpretable network

Kai Gong<sup>1†</sup>, Qian Dai<sup>2†</sup>, Jiacheng Wang<sup>2†</sup>, Yingbin Zheng<sup>1</sup>,  
Tao Shi<sup>3</sup>, Jiaxing Yu<sup>2</sup>, Jiangwang Chen<sup>2</sup>, Shaohui Huang<sup>2\*</sup> and  
Zhanxiang Wang<sup>1\*</sup>

<sup>1</sup>The First Affiliated Hospital of Xiamen University, Xiamen University, Xiamen, Fujian, China, <sup>2</sup>Department of Computer Science, School of Informatics, Xiamen University, Xiamen, Fujian, China, <sup>3</sup>Ningbo Medical Center Lihuili Hospital, Ningbo, Zhejiang, China

With the recent development of deep learning, the regression, classification, and segmentation tasks of Computer-Aided Diagnosis (CAD) using Non-Contrast head Computed Tomography (NCCT) for spontaneous IntraCerebral Hematoma (ICH) have become popular in the field of emergency medicine. However, a few challenges such as time-consuming of ICH volume manual evaluation, excessive cost demanding patient-level predictions, and the requirement for high performance in both accuracy and interpretability remain. This paper proposes a multi-task framework consisting of upstream and downstream components to overcome these challenges. In the upstream, a weight-shared module is trained as a robust feature extractor that captures global features by performing multi-tasks (regression and classification). In the downstream, two heads are used for two different tasks (regression and classification). The final experimental results show that the multi-task framework has better performance than single-task framework. And it also reflects its good interpretability in the heatmap generated by Gradient-weighted Class Activation Mapping (Grad-CAM), which is a widely used model interpretation method, and will be presented in subsequent sections.

## KEYWORDS

intracerebral hematoma (ICH), Non-Contrast head Computed Tomography (NCCT), multi-task, ResNet, interpretability

## 1. Introduction

Spontaneous IntraCerebral Hematoma (ICH) is characterized by high incidence, high disability, and high mortality, accounting for approximately 10–20% of all strokes (Raafat et al., 2020). The initial hematoma volume is the strongest predictor of mortality and functional outcomes (Beslow et al., 2014). Fast and precise evaluation of ICH volume through Non-Contrast Computed Tomography (NCCT) images is generally considered the first and the most critical step to making further clinical decisions including medical and surgical options. As the evaluation of ICH volume is relatively a time-consuming process, head NCCT reports usually provide only a general description of ICH location and shape, whereas the ICH volume is not calculated. In most cases, the ICH volume is assessed manually by clinicians such as neurosurgeons or neurology physicians, using various methods including ABC/2, ABC/2.4, ABC/3,1/2Sh, and 2/3Sh. Among them, ABC/2, also known as the Tada formula, is the most used one. For further explanation, A is the

maximum axial hematoma diameter, B is the maximum axial diameter perpendicular to A on the same slice, and C is the vertical diameter of the hematoma (Beslow et al., 2010). The reliability of the ABC/2 formula is shown good correlation with computerized ICH volume measurements for small and uniformly shaped ICHs, but the accuracy can be affected by observer variability and imprecision (Oge et al., 2021). The method is prone to overestimate ICH volume by approximately 20% and even misestimate more for large and irregularly shaped hemorrhages (Patel et al., 2019). The 3D slicer software (<http://www.slicer.org>) provides a free open source software platform for biomedical research. It can identify hematoma pixels based on CT images and reconstruct blood clots in a three-dimensional manner, which is free from restriction by hematoma morphology and bleeding sites. Compared with manual methods, 3D slicer method is considered as a more stable and capable method of high precision for the volume calculation of most hematomas and is gradually accepted as an effective measurement method clinically (Xu et al., 2014; Chen et al., 2020; Gong et al., 2021).

Newly automatic, versatile, easily deployable, and accurate hematoma volume assessment tools are highly needed giving their clinical significance. Numerous attempts have been conducted to develop 3D slicer as well as other computer-assisted automated tools for hematoma volume evaluation (Huttner et al., 2006; Wang et al., 2009; Zhao et al., 2009; Yang et al., 2013; Xu et al., 2014), but the methods have not been fully automated yet, and still require a significant amount of decisions to be made by researchers, which remains a challenging problem.

With the recent growth of deep learning, the evaluation methods of hematomas with deep learning-based techniques have been advanced dramatically, including the ResNet model, the DenseNet model, and the H-DenseUNet model, see Zhao et al. (2021), Mantas (2020), Dawud et al. (2019), Zhou et al. (2022), and Gou and He (2021).

For the evaluation of intracranial hemorrhage volume, the deep learning method has been proved to be superior to the clinical method in stability and accuracy many times (Freeman et al., 2008; Sharrock et al., 2022). Most quantitative assessments of ICH volume are based on the segmentation or classification of ICH areas. Phaphuangwittayakul et al. (2022) propose a quantitative assessment algorithm to automatically measure both thickness and volume *via* the 3D shape mask combined with the output probabilities of the classification network. Xu et al. (2021) realize hematoma segmentation and volume evaluation by DenseUNet (Cai et al., 2020; Sharrock et al., 2021), and is based on V-Net architecture. Much literature focuses on developing better neural network architectures and training strategies to optimize ICH segmentation, on which post-processing can be performed to automatically measure the volume of intracerebral hemorrhage. In this paper, we do not have the segmentation mask of the cerebral hemorrhage area. Instead, we have the measurement results of the volume of cerebral hemorrhage by multiple doctors. On the basis of training on this, we expect the model to be able to locate the bleeding location and even mark the bleeding area, which is what we intend to continue to study in the future.

This paper describes a lightweight multi-task learning framework firstly, which is specifically designed to identify and

evaluate ICH hematoma volumes by training a large number of NCCT images collected from 258 patients with ICH. Then, based on the assessment of ICH hematoma volumes, further prognosis analysis of intracerebral patients through the multi-task framework is discussed. Finally, satisfactory results are obtained for both tasks. For the evaluation of ICH hematoma volumes, the effect of this model is even better than the assessments of some clinicians. To further improve the interpretability of the model, Gradient-weighted Class Activation Mapping (Grad-CAM) is utilized to visualize the prediction of the model, which focuses well on the bleeding area and provides a reliable basis for the predictive output of the model. This can make a lot of sense for clinical application.

## 2. Related work

### 2.1. Deep learning in brain computerized tomography images

The application of neuroimaging technology plays an important role in the diagnosis and treatment of cerebrovascular diseases and is an indispensable auxiliary diagnostic tool. As the main imaging methods of the brain, computerized tomography (CT) and Magnetic Resonance Imaging (MRI) are widely used in clinical practice. Among them, CT examination is the first choice to find most brain diseases, including congenital brain development intracranial abnormalities, brain tumors, cerebral hemorrhage, and so on. The major advantages of CT images after reconstruction for medical image analysis are high density and clear image; it can assist clinicians to master the brain structure and abnormalities within the brain tissues easily (Padma Nanthagopal and Sukanesh Rajamony, 2013; Sachdeva et al., 2013; Vidyarthi and Mittal, 2014). Researchers utilize different automated approaches for brain disease detection and type classification through brain radio images since medical images can be scanned and uploaded to computers with a fine-resolution. For a long time in the past, Support Vector Machine (SVM) and Neural Networks (NN) techniques have been widely used due to their stable and good performance (Pan et al., 2015). But in recent years, due to the improvement of equipment computing power, Deep Learning (DL) models have created an exciting new trend in the field of machine learning, because deep architecture can efficiently represent more complex relationships without requiring a large number of nodes like traditional machine learning, such as SVM and K-Nearest Neighbor (K-NN). With the vigorous development of these technologies, they have become advanced technologies in different medical and health fields, such as bioinformatics, medical informatics, and medical image analysis. Among all kinds of deep architectures, the Convolutional Neural Network (CNN) is undoubtedly the most commonly used architecture now, which can use the convolution kernel to realize complex operations such as feature extraction (Pan et al., 2015; Ravi et al., 2016; Litjens et al., 2017). Usually, CNN is designed for image recognition tasks. The image is first processed by multiple convolutional layers (each convolutional layer is followed by a Rectified Linear Unit (ReLU) layer and a pooling layer), then

input to the fully connected layer and the ReLU layer. Finally, the output layer produces the prediction of class probabilities. Although the CNN architecture does not require manual feature extraction compared with traditional machine learning methods, it is very difficult to train a CNN model from scratch, which requires a large number of labeled data sets for adequate learning, especially for tasks such as classification and regression: the data volume requirements are relatively large. Moreover, for processing a large number of filters, such as large-scale medical 3D images such as  $256 \times 256 \times 32$ , the hardware requirements are very high (Ben Ahmed et al., 2017).

The contribution of this paper is to apply deep learning to the prediction of intracranial hemorrhage. Due to the correlation between intracranial hemorrhage and prognosis survival, this paper further proposes a lightweight multi-task learning framework based on the prediction of ICH hematoma volume. The aim is to further effectively predict the prognosis of patients according to the relevant features of the blood loss learned by the model.

## 2.2. Interpretability

Deep learning is a multi-layer representation learning method built with simple nonlinear modules that transform the previous layer representation into a higher-level, more abstract representation, so that it is able to detect increasingly abstract features. Therefore, a large number of features are also generated in the process of generating the prediction results, which makes the information very compact, especially in the deep layers. It's hard to explain this process, so we don't know why the neural network makes this prediction. This is called the "black box" problem of deep neural networks (Castelvecchi, 2016). Indeed, they are capable of producing extremely accurate predictions, but how can predictions based on features beyond comprehension be reliable? It is a challenging problem.

In medical images, the interpretability of the model is even more important: it determines whether the clinicians can trust and accept the model predictions which could be critical in clinical applications. Interpreting the predictions of deep models is hard without appropriate techniques, but a range of approaches has emerged over the past few years. One example is to show the texture of imaging features at a single layer to examine the hierarchical process of learning (Zeiler et al., 2011). In addition, numerous researches and explorations have been carried out on generating meaningful heatmaps in order to highlight the importance of individual pixel regions in the input image for the final CNN prediction. Many techniques have demonstrated the feasibility of generating heatmaps, including strategies using deconvolution (Zeiler and Fergus, 2014), layer-wise relevance propagation (Samek et al., 2016), and saliency map construction (Simonyan et al., 2013; Ghorbani et al., 2020; Thomas et al., 2020). However, these methods still have certain shortcomings, such as being susceptible to noise and artifacts, lacking sensitivity to input disturbances and qualitative criteria for evaluating the quality of back propagation (Samek et al., 2016; Smilkov et al., 2017; Ghorbani et al., 2020). Class Activation

Mapping (CAM) serves as a better alternative approach. But in its most basic implementation, CAM requires adding a pooling layer to the target model, limiting the interpretation to only one specific layer (Zhou et al., 2016). Recently, two extended versions of CAM have emerged, namely, Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017), and Grad-CAM++ (Chattopadhyay et al., 2018). Compared with CAM, both of them can interpret arbitrary layers of CNN without any architectural modification, thus increasing the flexibility of the use.

Grad-CAM is one of the most widely used techniques for prediction interpretation; it is chosen in this paper to interpret the prediction of the model and to increase the credibility of the model so that it has good interpretability in clinical application.

## 3. Method

To obtain the IntraCerebral Hemorrhage (ICH) volume and provide interpretative decision details, we propose a novel multi-task learning framework to take the complementary advantages of ICH regression and classification. Besides improving the accuracy of both tasks, the fusion is also able to enhance the interpretability which is of great significance for the clinical application.

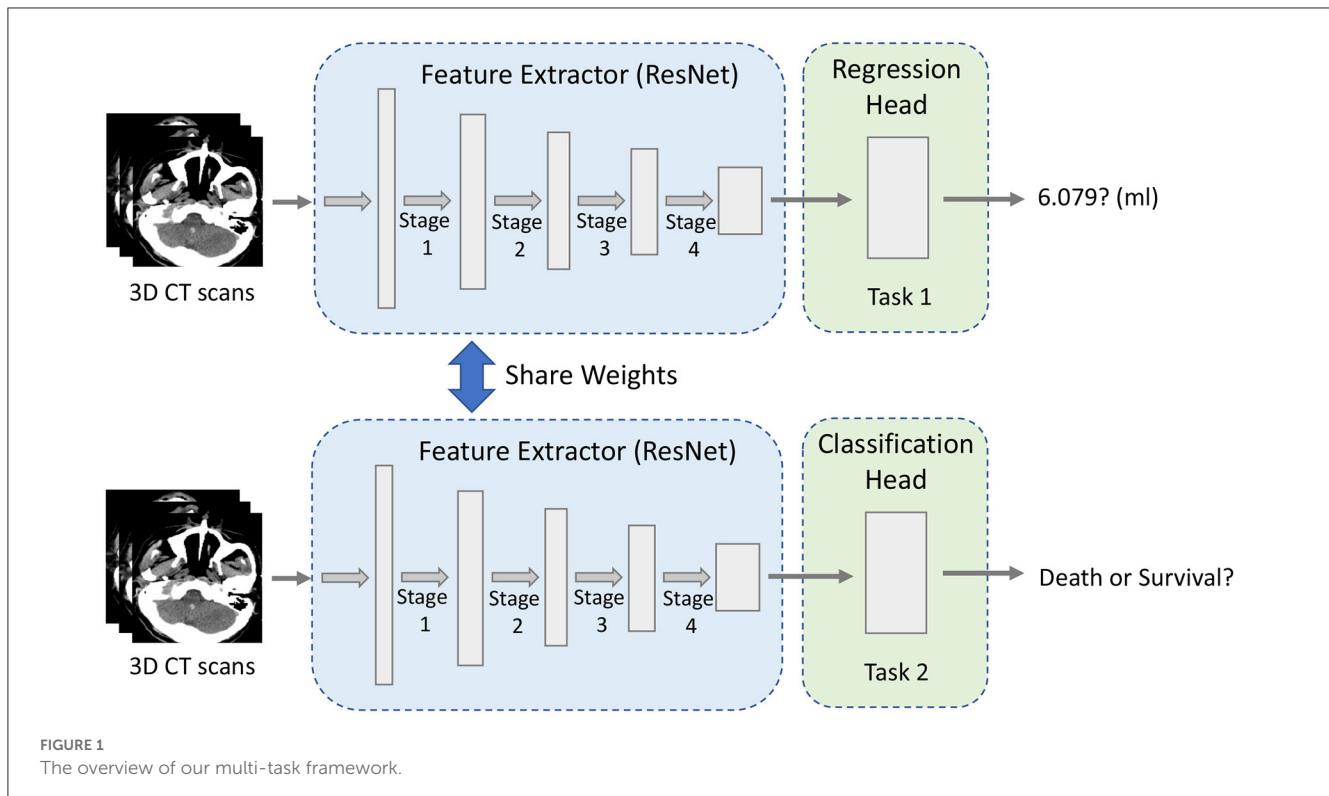
### 3.1. Overall architecture

As described above, the proposed network adopts the multi-task learning manner allowing the model to learn from different annotations. It can not only boost knowledge extraction by sharing the convolutional parameters but also can accelerate the learning program by calculating gradients from multiple branches. The motivation is the correlation between ICH regression and classification tasks. The research in Brott et al. (1997) shows that the increase of bleeding volume in the early stage of ICH is an important factor determining the prognosis of ICH patients. Unfortunately, the mortality rate of ICH is about 40% per month, with 61–88% of survivors having degrees of residual disability. In this context, the hematoma size is a key character for prognosis predictions. Therefore, it is more accurate and acceptable to predict the prognosis of patients based on a certain accurate assessment of the bleeding amount.

Our multi-task CNN architecture is proposed to jointly learn from both tasks, whose details are shown in Figure 1. The major components of the proposed network include a shared module for brain image feature extraction and two heads corresponding to different tasks. The algorithms for regression and classification tasks are built on ResNet deep neural network architecture (He et al., 2016), which includes the residual information extracted from the previous layer and mitigated the adverse performance by using a large number of layers.

### 3.2. Efficient feature extraction

Experience suggests that the multi-task model often achieves better results than the single-task model. Compared to single-task learning, multi-task learning has the following advantages: shared



encoder can reduce parameter amounts and save computation, associated tasks can improve each other's performance by sharing information and complementing each other. Also, it is helpful for the bottom-sharing encoder to learn the common feature representation.

The essence of multi-task learning lies in sharing the presentation layer and making the tasks interact with each other when the correlation between the predicted goals is relatively high.<sup>1</sup> Consequently, the parameter sharing layer will not bring too much loss, the parameter sharing layer can strengthen the parameter sharing, multiple target models can be trained jointly, reduce the parameter scale of the model and prevent the model from overfitting. Our final experimental results also fully demonstrate the effectiveness of applying multi-task learning to our tasks.

For multi-task learning, learning a general feature representation is very important for the model. As shown in Figure 1, the robust feature extraction module is designed to extract features of a 3D image and obtain a high-quality feature vector, which is used for downstream tasks. The module is based on 3D ResNet34 (He et al., 2016) because the residual structure effectively solves the problem of model degradation due to its depth, skip connection and also enhances the information transfer between the upper and lower layers.

Given a 3D CT image of  $224 \times 224 \times 32$  pixel size, the features are extracted by one stem block and four stages of ResNet (see Figure 1). Specifically, the first layer is a convolutional layer with

a  $7 \times 7 \times 7$  kernel size and 64 filters. The pooling layer is set after that to decrease the image's size and save the computation cost. The max pooling layer with a kernel size of  $3 \times 3 \times 3$  and stride 2 is used here empirically. The second, third, fourth, and fifth phases are made up of multiple convolutional operations, corresponding to the four stages in Figure 1. The first is a convolution layer with a filter size of  $3 \times 3 \times 3$  and the rest of the three modules are stacked on the top of each layer. The second is a convolution layer with the filter size of  $3 \times 3 \times 3$  and there are four residual blocks. Each convolutional layer in the network applies zero-padding, after each convolutional layer, a batch normalization layer (Ioffe and Szegedy, 2015) is applied to speed up model training and convergence. To allow the application to images of arbitrary size, 3D feature maps are not flattened but dense layers are implemented as convolutions with the size of  $1 \times 1 \times 1$  (Long et al., 2015).

After the entire feature extraction process, the final extracted feature map is obtained. The network has two output heads, one for ICH volume regression and another one concerns the classification of the prognosis. Finally, the features extracted by the last layer of the module are fed to each task-specific head.

### 3.3. Regression head

The first task-specific head for regression only has an average pooling layer and one fully connected layer. The final output is a scalar referring to the amount of ICH volume predicted by the model. This regression task is performed to predict the IntraCerebral Hemorrhage (ICH) volume in each 3D CT scan, task-specific head module to find the most intensive crux features of

<sup>1</sup> In our task, ICH volume regression and patient prognosis classification based on brain CT images. The underlying features of both should be similar, both at the bleeding site in the brain.

ICH. The regression loss is defined as the Mean Square Error Loss (MSE Loss) defined by  $L_{reg}$ :

$$L_{reg} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_{reg_i} - y_{reg_i})^2, \quad (1)$$

Where  $\hat{y}_{reg_i}$  is the predicted output of the regression head, and  $N$  is the number of CT images.

### 3.4. Classification head

The second task-specific head for classification has a simple linear classifier which consists of a global average pooling, a linear layer, and a ReLU activation function. Finally, the softmax function is applied to the output to obtain the probability of the model prediction, which is between 0 and 1. Generally, we use 0.5 as the classification decision threshold: a probability greater than 0.5 is fixed to 1, which means that the class belongs to, whereas a probability less than 0.5 is fixed to 0, meaning that the class does not belong to. The upstream framework is the same and shares the weight of the feature extraction module of the regression task. The classification loss is defined as the Binary Cross-Entropy Loss (BCE Loss) called  $L_{cls}$ :

$$L_{cls} = \frac{1}{N} \cdot \sum_{i=1}^N [y_{cls_i} \cdot \log(\hat{y}_{cls_i}) + (1 - y_{cls_i}) \cdot \log(1 - \hat{y}_{cls_i})], \quad (2)$$

Where  $\hat{y}_{cls_i}$  is the predicted output of the classification head, and  $N$  represents the number of CT images.

For our multi-task framework, the feature extractor is shared and the two heads fit the outputs of the two tasks separately. Therefore, the loss of multi-task framework is a combination of the losses of two tasks. Through the simple weighting of loss, the network weights are updated to optimize.

### 3.5. Interpretive design

Traditional deep learning applications lack interpretability and thus faces limitation in clinical practice. We propose to utilize the advanced Grad-CAM to explore how the results are obtained. Specifically, class weights derived in Grad-CAM used the equation:

$$w_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (3)$$

Where  $Z$  denotes the total number of elements in a feature map  $A$ ,  $A_{ij}^k$  represents the data of the feature map in channel  $k$ , and the coordinate is  $ij$ ,  $y^c$  denotes the score predicted by the network for class  $c$ .

This equation is backpropagated through the prediction score of class  $c$  to obtain the gradient information that is backpropagated back to feature layer  $A$ , namely  $w_k^c$ , which represents the importance of each channel of feature map  $A$ , here for class  $c$ . The higher the value, the greater the contribution, and the more important it is considered by the model. Finally, weighted summation is performed and the final Grad-CAM heatmap is

obtained by ReLU activation function. Consequently, the heatmap generation uses the formula:

$$L_{Grad-CAM}^c = ReLU \left( \sum_k w_k^c A^k \right), \quad (4)$$

Where the ReLU (Rectified Linear Unit) function allows to evaluate features with positive impact only. In this paper, we use the last convolutional layer in computing the weights as suggested previously.

## 4. Experiments

### 4.1. Ethics approval and consent to participate

The study is undertaken in compliance with the principles of the Declaration of Helsinki and is approved by the ethics committees of the First Affiliated Hospital of Xiamen University (Approved ID: SL-2020KY034). All individual patient data used for the analysis are collected by providers after obtaining appropriate consent and agreements.

### 4.2. Data collection

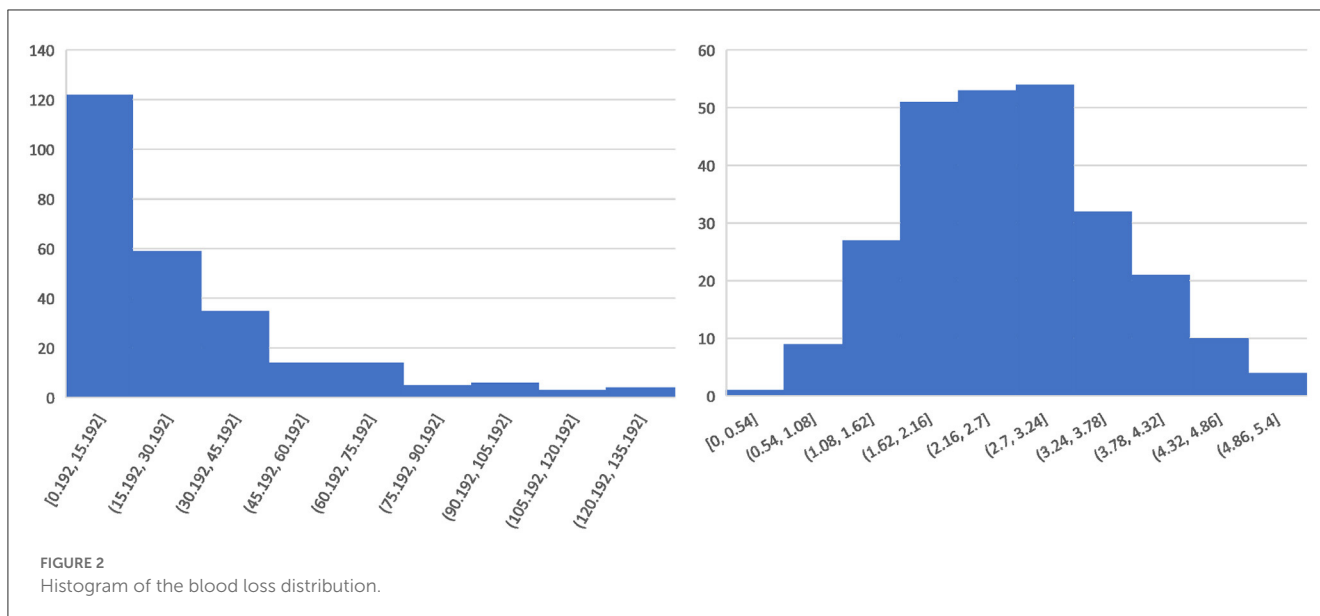
This study involves 258 patients with spontaneous intracerebral hematoma (ICH) admitted to the First Affiliated Hospital of Xiamen University between April 2017 and February 2019 to develop the prognosis model. These axial brain Computed Tomography (CT) scans, taken on admission using a Philips Brilliance 64-row spiral CT scanner are exported from the Department of Neurosurgery through Picture Archiving and Communication Systems (PACS) and stored in Digital Imaging and Communications in Medicine (DICOM) format.

### 4.3. Data preprocessing

To evaluate the effectiveness of the developed network, we collect a new ICH dataset with 258 patients, of which 227 patients survived and 31 patients dead. All patients are given a table with the patient's name and measurements of the ICH volume (milliliter, ml) from eight professional neurosurgeons (with 2, 3, 3, 4, 4, 8, 15, and 20 years of experience) and the 3D slicer software, which is considered the gold standard.

For the regression labels, because of the uneven distribution of ICH volume (shown in Figure 2 left), we calculate the third root of the ICH volume and convert it to the ICH size (millimeter, mm), which is more balanced in distribution (shown in Figure 2 right) and more beneficial to model training. For the final prediction results of the regression model, we perform post-processing and then convert them into ICH volumes for further analysis and comparisons.

Considering the intensity range of brain characteristics (Fosbinder and Orth, 2011), we apply a brain window that clips [40, 110] Hounsfield units (HU) and normalized



the input to  $[0, 1]$ . After cutting out the blank part of the brain in the image, the image size is resized down to  $32 \times 224 \times 224$  by linear interpolation. Since ICH can occur anywhere in the brain with multiple subtypes simultaneously, we do not crop the images and don't use patch images for training, which may lead to unstable results and false positives (Hu et al., 2011).

#### 4.4. Evaluation metrics

For the regression task, we choose the following three indicators to evaluate, which include Mean Square Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). And MAPE is the relative error, which is the percentage of absolute error and truth. For classification task, three widely-used metrics are used for quantitative analysis, which are accuracy, ROC curve, and area under the curve(AUC).

#### 4.5. Implementation details

Our network is implemented on Pytorch (Paszke et al., 2019) and trains using SGD (Robbins and Monro, 1951) with 300 epochs, an initial learning rate of  $1 \times 10^{-4}$ , a momentum of 0.9, and a weight decay of  $5 \times 10^{-4}$ . The whole architecture is trained on one GeForce RTX 3080 Ti GPU, and each GPU has a batch size of 8. For a fair comparison, our method and all models follow the same training settings.

#### 4.6. Results

The performance of our proposed approach for regression and classification tasks is reported in Tables 1, 2, respectively.

To verify the effectiveness and feasibility of our proposed framework, Table 1 selects five existing methods widely used

TABLE 1 Quantitative comparisons for the effectiveness of regression model.

Methods	MAE	MSE	MAPE(%)
Clinician 1	5.21	65.93	28.78
Clinician 2	6.14	84.10	38.50
clinician 3	4.46	39.15	38.68
Clinician 4	4.77	50.05	34.13
Clinician 5	4.88	72.04	29.60
Clinician 6	5.30	56.08	44.65
Clinician 7	3.86	38.84	20.14
Clinician 8	4.02	40.98	24.42
X3D (Feichtenhofer, 2020)	18.28	490.19	99.98
R(2+1)D (Tran et al., 2018)	7.09	80.02	45.59
P3D (Qiu et al., 2017)	10.24	143.54	171.18
TSM (Lin et al., 2019)	9.51	130.69	160.93
Times Former (Bertasius et al., 2021)	10.49	158.13	171.09
TIN-ResNet-18 (Shao et al., 2020)	7.90	107.52	110.68
TIN-ResNet-34 (Shao et al., 2020)	9.71	134.73	159.54
TIN-ResNet-50 (Shao et al., 2020)	10.87	170.85	184.58
Uniformizing-3D (Zunair et al., 2020)	7.87	118.06	102.38
Multi-task (ours)	4.91	52.65	43.00

The first eight lines denote the measurements of eight clinicians.

for 3D image classification tasks for comparison, including X3D (Feichtenhofer, 2020), R(2+1)D (Tran et al., 2018), P3D (Qiu et al., 2017), TSM (Lin et al., 2019), Times Former (Bertasius et al., 2021), TIN (Shao et al., 2020) with backbone of ResNet-18, 34, 50 and Uniformizing-3D (Zunair et al., 2020), which is a 3D CNN for CT scans. For providing a fair comparison, we obtain the

final classification results on the official implementations of these compared methods. We train these models on our dataset and set the same experimental parameters as ours. As shown in Table 1, the model trained by deep learning performs well. The method we choose based on ResNet architecture achieves the best results and is even more suitable for the tasks of regression and classification of 3D brain CT images, with MAE reaching 4.91, the MSE reaching 51.92, and MAPE reaching 43.00%. Compared to the evaluation results of eight clinicians (first eight rows in Table 1), our method also has competitive performance, outperforming the evaluation results of three clinicians (clinicians 1, 2, 6), while also having better interpretability.

To further show the stability and effectiveness of our model training, we plot the curve of the indicator change during the model training and validation process, including MAE and MSE, as shown in Figure 3. From the figure, we can see that whether in the training or validation phase, the MAE and MSE loss functions of our model gradually decrease and converge to be stable in 300 epochs, which indicates that the model obtained by our training is stable and reliable. The training of other models also shows a similar trend.

The results of the classification task shown in Table 2 demonstrate the effectiveness of the multi-task network. Compared with the single-task model, that is, trained from scratch, the performance of the multi-task model is also greatly improved due to the shared encoder parameters, with accuracy improved by 5.66% and AUC improved by 5.68%, and both reached more than 90%.

TABLE 2 Quantitative comparisons of the effectiveness of multi-task framework.

Methods	ACC (Accuracy)	AUC
ResNet (single-task)	88.68	87.94
ResNet (multi-task, ours)	<b>94.34</b>	<b>93.62</b>

“Single-task” denotes the classification model train from scratch. Bold values indicate the optimal result.

More visually, Figure 4 shows the ROC curves of our multi-task model and single-task model. The performance of the multi-task model is better than others, and the area of AUC remains over 90%, which fully demonstrates its effectiveness.

### 4.7. Interpretability

To enhance the interpretability of the model for the prediction of ICH volume, we visualize the output of the last convolutional layer of the regression model. As mentioned earlier, we introduce Grad-CAM (Gradient-weighted Class Activation Mapping) (Selvaraju et al., 2017) to generate heatmaps, which show the region of interest of the model.

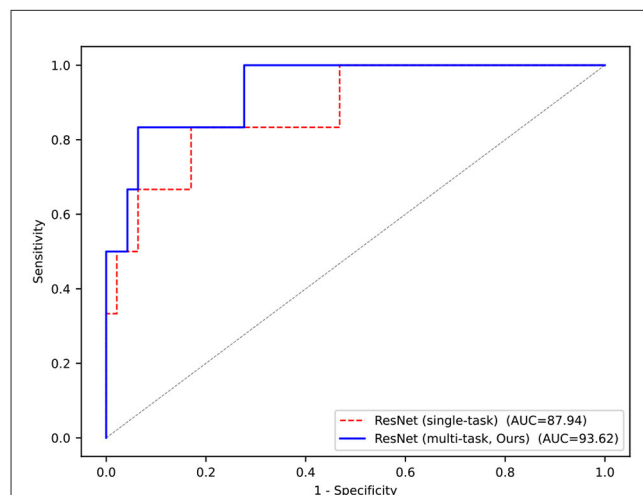


FIGURE 4 ROC curves of our multi-task framework and compared method.

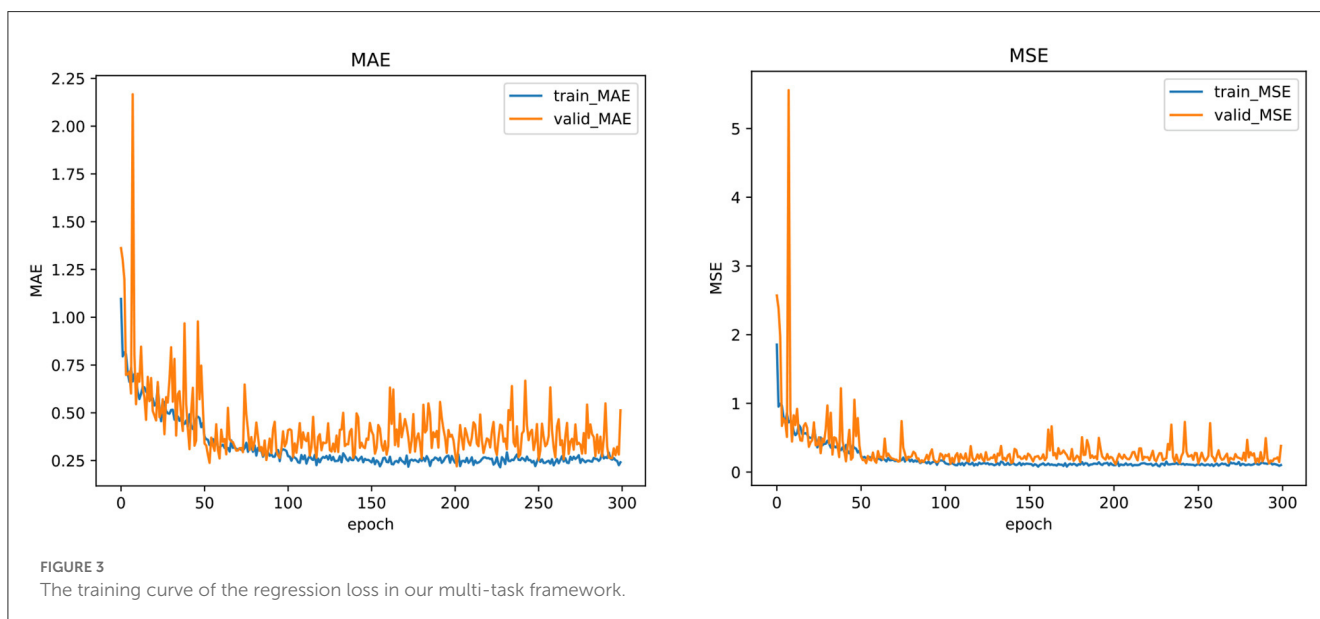
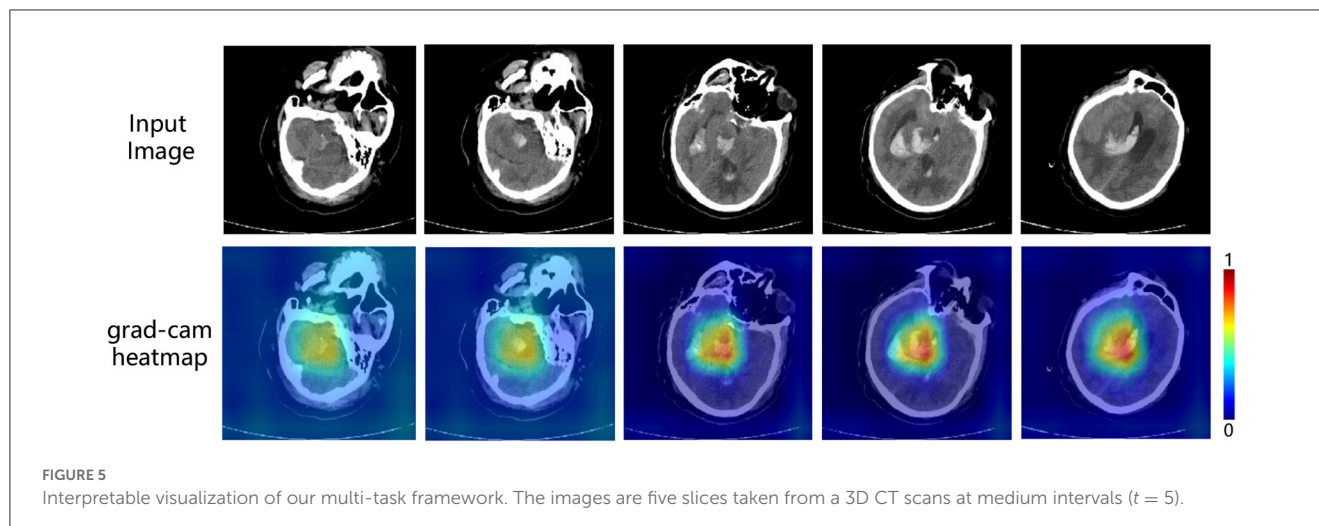


FIGURE 3 The training curve of the regression loss in our multi-task framework.



In the heatmap, the brighter the color is, the more attention the model pays to. As shown in Figure 5, the highlighted areas are concentrated in certain parts of the brain with some white spots. And according to our confirmation with a professional clinician, this is exactly the location of the cerebral hemorrhage in the image. Therefore, it is reasonable to say that in the prediction of ICH volume, our model accurately locates the location of bleeding, to make the prediction, and the result is acceptable to clinicians. At the same time, it also provides a basis for subsequent prognostic analysis tasks, as the ICH volume is a key factor affecting patient mortality.

## 5. Discussion

### 5.1. Clinical significance for proposed algorithm

Stroke is the second-leading cause of death, and the third-leading cause of death and disability combined in the world (Feigin et al., 2021). Intracerebral hematoma (ICH) is a common and fatal subtype of stroke, which is characterized by rupture of arterial blood vessels and formation of hematoma within the brain parenchyma, such as the basal ganglia (Zhang et al., 2020; Kinoshita et al., 2021). The formation of a large hematoma compresses the surrounding brain tissue, resulting in neurological deterioration, intracranial hypertension, global cerebral ischemia, and brain herniation, which can affect vital autonomic structures (Yu et al., 2020). The mortality rate of ICH is about 40% per month, with 61–88% of survivors having high degrees of residual disability (Van Asch et al., 2010). Management of increased intracranial pressure is the mainstay of ICH treatment in the acute phase (Kase and Hanley, 2021). Patients with excessive bleeding and coma often require surgical treatment to relieve the hematoma, brain tissue compression, and secondary brain injury (Lin et al., 2014). However, as an invasive process, iatrogenic injury is inevitable during hematoma evacuation, which means the accurate assessments of the patients' condition and appropriate selections for surgical interventions are critically important. The

hematoma size is a key character for prognosis predictions and clinical decision making. Supratentorial hematomas larger than 30 ml (Kerebel et al., 2013; Huang et al., 2021) and infratentorial hematomas larger than 15 ml (de Oliveira Manoel, 2020) are often considered to result in poor prognosis and tend to require surgical evacuation. Thus, the accuracy of volume evaluation for ICH patients has significant clinical implications.

There are mainly four kinds of methods to evaluate ICH volume, including the mathematical formula method, tool measurement method, CT machine measurement method, and software method (Chen et al., 2020). Among them, Tada formula, a kind of formula method, is the most commonly used one. However, the accuracy of Tada formula can be affected by both hematoma shape and volume, resulting in imprecise disease evaluation and potential improper treatments (Gong et al., 2021). 3D Slicer is shown to be more precise than most methods (Chen et al., 2020), but the modeling process is relatively time-consuming for clinicians. With the development of deep learning techniques, artificial intelligence has provided promising results in the medical field and may help with medical image interpretation and prognosis prediction, offering more reliable results than manual decisions by clinicians. Several previous studies have used deep learning-based techniques to evaluate ICH volumes (Roh et al., 2019), but there are still some shortcomings. For instance, some studies use the manually estimated hematoma volume as the label to train the model, which is not accurate enough to evaluate the performance of the model because the relatively large error of manual estimation leads to the inaccuracy of the label. In addition, it does not take the problem of model interpretation into account.

For this reason, this study uses the results of 3D slicer as labels to improve the accuracy of model training and evaluation. At the same time, in addition to achieving competitive accuracy with clinicians, our model has good interpretability and is easier to be understood and accepted by clinicians. Intuitively, we attribute the accuracy of model prediction to the fact that the heatmap shows that it focuses on the information related to the bleeding area, which is in turn the target of the model being explained. This demonstrates the accuracy of the effect of our model, and its concerns are consistent with those of professional physicians.



Furthermore, the evaluation of hematoma volume to predict patient prognosis is not accurate enough because the ICH density, shape, location, the middle line deviation, the effacement of Sylvian fissure and perimesencephalic cisterns (Kim, 2012) are all related to the prognosis of ICH. Therefore, this study uses the complete CT images as the training objects and does not mark a specific region of interest, so that the results are more reasonable and robust.

## 5.2. Limitations and further directions

Our experimental results prove the effectiveness of our proposed framework well, but to better apply it clinically, we should test it on more data. This is also the current limitation of our method. First, our data are collected from a single center and the result should be further validated by data from multiple centers in terms of predicting prognosis due to differences in treatment levels. Second, given the variable performance of ICH, the sample size is still needed to be expanded, including hematomas of different sizes, locations, shapes, and densities. Third, cerebral hemorrhage is a dynamic process, and a single NCCT can only be used to evaluate static intracranial conditions. Therefore, subsequent studies can use different concurrent images as time series for prognostic evaluation. Fourth, besides the radiological features, the clinical features including age, the severity of other underlying diseases, the GCS coma score, etc. are all associated with prognosis (Wang et al., 2014). Using multimodal data may further advance the performance of such clinical prediction models.

## 6. Conclusion

We introduce a multi-task framework for ICH volume (regression) and patient prognosis (classification) prediction in NCCT. First, we design a robust feature extractor through the ICH volume prediction task, which effectively locates and focuses on ICH regions. This is effectively demonstrated in the Grad-CAM. By exploring the relationship between ICH volume and patient prognosis, we find that the effectiveness of feature extraction on the

former task affected the performance on the latter task. A feature extractor that can well locate the hemorrhage area and pay attention to the hematoma, which is very beneficial for the subsequent task of prognostic analysis. The experimental results also prove it well. Compared with existing methods, our framework achieves impressive results on both tasks, and its reliable interpretability also makes it possible for clinical application.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Author contributions

KG, ZW, and SH contributed to the study concept and design. KG and YZ collected the data. KG, QD, and JW performed the statistical analysis and wrote the manuscript. YZ, TS, JY, and JC revised the manuscript. All authors contributed to the article and agreed to the submitted version.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Ben Ahmed, K., Hall, L., Goldgof, D., Liu, R., and Gatenby, R. (2017). Fine-tuning convolutional deep features for MRI based brain tumor classification. *Med. Imaging Comput. Aided Diagn.* 10134, 613–619. doi: 10.1117/12.2253982
- Bertasius, G., Wang, H., and Torresani, L. (2021). "Is space-time attention all you need for video understanding?" in *International Conference on Machine Learning (Honolulu)*, 2, 4.
- Beslow, L. A., Ichord, R. N., Gindville, M. C., Kleinman, J. T., Bastian, R. A., Smith, S. E., et al. (2014). Frequency of hematoma expansion after spontaneous intracerebral hemorrhage in children. *JAMA Neurol.* 71, 165–171. doi: 10.1001/jamaneurol.2013.4672
- Beslow, L. A., Ichord, R. N., Kasner, S. E., Mullen, M. T., Licht, D. J., Smith, S. E., et al. (2010). ABC/XYZ estimates intracerebral hemorrhage volume as a percent of total brain volume in children. *Stroke* 41, 691–694. doi: 10.1161/STROKEAHA.109.566430
- Brott, T., Broderick, J., Kothari, R., Barsan, W., Tomsick, T., Sauerbeck, L., et al. (1997). Early hemorrhage growth in patients with intracerebral hemorrhage. *Stroke* 28, 1–5. doi: 10.1161/01.STR.28.1.1
- Cai, S., Tian, Y., Lui, H., Zeng, H., Wu, Y., and Chen, G. (2020). Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. *Quant. Imaging Med. Surg.* 10, 1275. doi: 10.21037/qims-19-1090
- Castelvecchi, D. (2016). Can we open the black box of AI? *Nat. News* 538, 20. doi: 10.1038/538020a
- Chattopadhyay, A., Sarkar, A., Howlader, P., and Balasubramanian, V. N. (2018). "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," in *IEEE Winter Conference on Applications of Computer Vision (Lake Tahoe, NV: IEEE)*, 839–847.
- Chen, M., Li, Z., Ding, J., Lu, X., Cheng, Y., and Lin, J. (2020). Comparison of common methods for precision volume measurement of hematoma. *Comput. Math. Methods Med.* 2020, 1–11. doi: 10.1155/2020/6930836
- Dawud, A. M., Yurtkan, K., and Oztoprak, H. (2019). Application of deep learning in neuroradiology: brain haemorrhage classification using transfer learning. *Comput. Intell. Neurosci.* 2019, 1–12. doi: 10.1155/2019/4629859
- de Oliveira Manoel, A. L. (2020). Surgery for spontaneous intracerebral hemorrhage. *Crit. Care* 24, 1–19. doi: 10.1186/s13054-020-2749-2

- Feichtenhofer, C. (2020). "X3D: expanding architectures for efficient video recognition," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (Seattle, WA: IEEE), 203–213.
- Feigin, V. L., Stark, B. A., Johnson, C. O., Roth, G. A., Bisignano, C., Abady, G. G., et al. (2021). Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol.* 20, 795–820. doi: 10.1016/S1474-4422(21)00252-0
- Fosbinder, R., and Orth, D. (2011). *Essentials of Radiologic Science*. Philadelphia, PA: Lippincott Williams & Wilkins.
- Freeman, W. D., Barrett, K. M., Bestic, J. M., Meschia, J. F., Broderick, D. F., and Brott, T. G. (2008). Computer-assisted volumetric analysis compared with ABC/2 method for assessing warfarin-related intracranial hemorrhage volumes. *Neurocrit. Care* 9, 307–312. doi: 10.1007/s12028-008-9089-4
- Ghorbani, A., Ouyang, D., Abid, A., He, B., Chen, J. H., Harrington, R. A., et al. (2020). Deep learning interpretation of echocardiograms. *NPJ Digital Med.* 3, 1–10. doi: 10.1038/s41746-019-0216-8
- Gong, K., Shi, T., Zhao, L., Xu, Z., and Wang, Z. (2021). Comparing the inter-observer reliability of the Tada formula among neurosurgeons while estimating the intracerebral haematoma volume. *Clin. Neurol Neurosurg.* 205, 106668. doi: 10.1016/j.clineuro.2021.106668
- Gou, X., and He, X. (2021). Deep learning-based detection and diagnosis of subarachnoid hemorrhage. *J. Healthc Eng.* 2021. doi: 10.1155/2021/9639419
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV: IEEE), 770–778.
- Hu, S., Coupé, P., Pruessner, J. C., and Collins, L. D. (2011). "Validation of appearance-model based segmentation with patch-based refinement on medial temporal lobe structures," in *MICCAI Workshop on Multi-Atlas Labeling and Statistical Fusion* (Toronto, ON), 28–37.
- Huang, X., Jiang, L., Chen, S., Li, G., Pan, W., Peng, L., et al. (2021). Comparison of the curative effect and prognosis of stereotactic drainage and conservative treatment for moderate and small basal ganglia haemorrhage. *BMC Neurol.* 21, 1–8. doi: 10.1186/s12883-021-02293-7
- Huttner, H. B., Steiner, T., Hartmann, M., Köhrmann, M., Juettler, E., Mueller, S., et al. (2006). Comparison of ABC/2 estimation technique to computer-assisted planimetric analysis in warfarin-related intracerebral parenchymal hemorrhage. *Stroke* 37, 404–408. doi: 10.1161/01.STR.0000198806.67472.5c
- Ioffe, S., and Szegedy, C. (2015). "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *International Conference on Machine Learning* (Lille), 448–456.
- Kase, C. S., and Hanley, D. F. (2021). Intracerebral hemorrhage: advances in emergency care. *Neurol Clin.* 39, 405–418. doi: 10.1016/j.ncl.2021.02.002
- Kerebel, D., Joly, L.-M., Honnart, D., Schmidt, J., Galanaud, D., Negrier, C., et al. (2013). A french multicenter randomised trial comparing two dose-regimens of prothrombin complex concentrates in urgent anticoagulation reversal. *Crit. Care* 17, 1–11. doi: 10.1186/cc11923
- Kim, H. J. (2012). The prognostic factors related to traumatic brain stem injury. *J. Korean Neurosurg. Soc.* 51, 24–30. doi: 10.3340/jkns.2012.51.1.24
- Kinoshita, K., Hamanaka, G., Ohtomo, R., Takase, H., Chung, K. K., Lok, J., et al. (2021). Mature adult mice with exercise-preconditioning show better recovery after intracerebral hemorrhage. *Stroke* 52, 1861–1865. doi: 10.1161/STROKEAHA.120.032201
- Lin, B., Yang, H., Cui, M., Li, Y., and Yu, J. (2014). Surgicel™ application in intracranial hemorrhage surgery contributed to giant-cell granuloma in a patient with hypertension: case report and review of the literature. *World J. Surg. Oncol.* 12, 1–5. doi: 10.1186/1477-7819-12-101
- Lin, J., Gan, C., and Han, S. (2019). "TSM: temporal shift module for efficient video understanding," in *Proceedings of the IEEE International Conference on Computer Vision* (Seoul: IEEE), 7083–7093.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88. doi: 10.1016/j.media.2017.07.005
- Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully convolutional networks for semantic segmentation," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (Boston, MA: IEEE), 3431–3440.
- Mantas, J. (2020). "Classification of intracranial hemorrhage subtypes using deep learning on CT scans," in *The Importance of Health Informatics in Public Health during a Pandemic, Vol. 272* (Clifton), 370.
- Oge, D. D., Arsava, E. M., Pektezel, M. Y., Gocmen, R., and Topcuoglu, M. A. (2021). Intracerebral hemorrhage volume estimation: Is modification of the ABC/2 formula necessary according to the hematoma shape? *Clin. Neurol Neurosurg.* 207, 106779. doi: 10.1016/j.clineuro.2021.106779
- Padma Nanthagopal, A., and Sukanesh Rajamony, R. (2013). Classification of benign and malignant brain tumor CT images using wavelet texture parameters and neural network classifier. *J. Visualizat.* 16, 19–28. doi: 10.1007/s12650-012-0153-y
- Pan, Y., Huang, W., Lin, Z., Zhu, W., Zhou, J., Wong, J., et al. (2015). "Brain tumor grading based on neural networks and convolutional neural networks," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Milan: IEEE), 699–702.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: an imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* 32, 8024–8035. doi: 10.48550/arXiv.1912.01703
- Patel, A., Schreuder, F. H., Klijn, C. J., Prokop, M., Ginneken, B., v., et al. (2019). Intracerebral haemorrhage segmentation in non-contrast CT. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-54491-6
- Phaphuangwittayakul, A., Guo, Y., Ying, F., Dawod, A. Y., Angkurawaranon, S., and Angkurawaranon, C. (2022). An optimal deep learning framework for multi-type hemorrhagic lesions detection and quantification in head CT images for traumatic brain injury. *Appl. Intell.* 52, 7320–7338. doi: 10.1007/s10489-021-02782-9
- Qiu, Z., Yao, T., and Mei, T. (2017). "Learning spatio-temporal representation with pseudo-3d residual networks," in *Proceedings of the IEEE International Conference on Computer Vision* (Venice: IEEE), 5533–5541.
- Raafat, M., Ragab, O. A., Abdelwahab, O. M., Salama, M. M., and Hafez, M. A. (2020). Early versus delayed surgical evacuation of spontaneous supratentorial intracerebral hematoma: a prospective cohort study. *Surg. Neurol Int.* 11, 103. doi: 10.25259/SNI\_103\_2020
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., et al. (2016). Deep learning for health informatics. *IEEE J. Biomed. Health Inform.* 21, 4–21. doi: 10.1109/JBHI.2016.2636665
- Robbins, H., and Monro, S. (1951). "A stochastic approximation method. *Ann. Math. Stat.* 22, 400–407. doi: 10.1214/aoms/117729586
- Roh, D., Sun, C.-H., Murthy, S., Elkind, M. S., Bruce, S. S., Melmed, K., et al. (2019). Hematoma expansion differences in lobar and deep primary intracerebral hemorrhage. *Neurocrit. Care* 31, 40–45. doi: 10.1007/s12028-018-00668-2
- Sachdeva, J., Kumar, V., Gupta, I., Khandelwal, N., and Ahuja, C. K. (2013). Segmentation, feature extraction, and multiclass brain tumor classification. *J. Digit. Imaging* 26, 1141–1150. doi: 10.1007/s10278-013-9600-0
- Samek, W., Binder, A., Montavon, G., Lapuschkin, S., and Müller, K.-R. (2016). Evaluating the visualization of what a deep neural network has learned. *IEEE Trans. Neural Networks Learn. Syst.* 28, 2660–2673. doi: 10.1109/TNNLS.2016.2599820
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). "Grad-cam: visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision* (Venice: IEEE), 618–626.
- Shao, H., Qian, S., and Liu, Y. (2020). Temporal interlacing network. *Proc. AAAI Conf. Artif. Intell.* 34, 11966–11973. doi: 10.1609/aaai.v34i07.6872
- Sharrock, M. F., Mould, W. A., Ali, H., Hildreth, M., Awad, I. A., Hanley, D. F., et al. (2021). 3D deep neural network segmentation of intracerebral hemorrhage: development and validation for clinical trials. *Neuroinformatics* 19, 403–415. doi: 10.1007/s12021-020-09493-5
- Sharrock, M. F., Mould, W. A., Hildreth, M., Ryu, E. P., Walborn, N., Awad, I. A., et al. (2022). Bayesian deep learning outperforms clinical trial estimators of intracerebral and intraventricular hemorrhage volume. *J. Neuroimaging* 32, 968–976. doi: 10.1111/jon.12997
- Simonyan, K., Vedaldi, A., and Zisserman, A. (2013). Deep inside convolutional networks: visualising image classification models and saliency maps. *arXiv preprint arXiv:1312.6034*. doi: 10.48550/arXiv.1312.6034
- Smilkov, D., Thorat, N., Kim, B., Viégas, F., and Wattenberg, M. (2017). Smoothgrad: removing noise by adding noise. *arXiv preprint arXiv:1706.03825*. doi: 10.48550/arXiv.1706.03825
- Thomas, K. A., Kidziński, L., Halilaj, E., Fleming, S. L., Venkataraman, G. R., Oei, E. H., et al. (2020). Automated classification of radiographic knee osteoarthritis severity using deep neural networks. *Radiol. Artif. Intell.* 2, e190065. doi: 10.1148/ryai.2020190065
- Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., and Paluri, M. (2018). "A closer look at spatiotemporal convolutions for action recognition," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT: IEEE), 6450–6459.
- Van Asch, C. J., Luitse, M. J., Rinkel, G. J., van der Tweel, I., Algra, A., and Klijn, C. J. (2010). Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *Lancet Neurol.* 9, 167–176. doi: 10.1016/S1474-4422(09)70340-0
- Vidyarthi, A., and Mittal, N. (2014). "Comparative study for brain tumor classification on MR/CT images," in *Proceedings of the International Conference on Soft Computing for Problem Solving* (New Delhi), 889–897.
- Wang, C.-W., Juan, C.-J., Liu, Y.-J., Hsu, H.-H., Liu, H.-S., Chen, C.-Y., et al. (2009). Volume-dependent overestimation of spontaneous intracerebral hematoma volume by the ABC/2 formula. *Acta Radiol.* 50, 306–311. doi: 10.1080/02841850802647039
- Wang, C.-W., Liu, Y.-J., Lee, Y.-H., Hueng, D.-Y., Fan, H.-C., Yang, F.-C., et al. (2014). Hematoma shape, hematoma size, Glasgow coma scale score and ICH score:

which predicts the 30-day mortality better for intracerebral hematoma? *PLoS ONE* 9, e102326. doi: 10.1371/journal.pone.0102326

Xu, J., Zhang, R., Zhou, Z., Wu, C., Gong, Q., Zhang, H., et al. (2021). Deep network for the automatic segmentation and quantification of intracranial hemorrhage on CT. *Front. Neurosci.* 14, 541817. doi: 10.3389/fnins.2020.541817

Xu, X., Chen, X., Zhang, J., Zheng, Y., Sun, G., Yu, X., et al. (2014). Comparison of the Tada formula with software slicer: precise and low-cost method for volume assessment of intracerebral hematoma. *Stroke* 45, 3433–3435. doi: 10.1161/STROKEAHA.114.007095

Yang, W., Feng, Y., Zhang, Y., Yan, J., Fu, Y., and Chen, S. (2013). Volume quantification of acute infratentorial hemorrhage with computed tomography: validation of the formula  $1/2ABC$  and  $2/3Sh$ . *PLoS ONE* 8, e62286. doi: 10.1371/journal.pone.0062286

Yu, H., Cao, X., Li, W., Liu, P., Zhao, Y., Song, L., et al. (2020). Targeting connexin 43 provides anti-inflammatory effects after intracerebral hemorrhage injury by regulating yap signaling. *J. Neuroinflamm.* 17, 1–19. doi: 10.1186/s12974-020-01978-z

Zeiler, M. D., and Fergus, R. (2014). “Visualizing and understanding convolutional networks,” in *European Conference on Computer Vision* (Cham), 818–833.

Zeiler, M. D., Taylor, G. W., and Fergus, R. (2011). “Adaptive deconvolutional networks for mid and high level feature learning,” in *Proceedings of the IEEE International Conference on Computer Vision* (Barcelona: IEEE), 2018–2025.

Zhang, J., Cheng, H., Zhou, S., Huang, L., Lv, J., Wang, P., et al. (2020). 3D-printed model-guided endoscopic evacuation for basal ganglia hemorrhage. *Sci. Rep.* 10, 1–7. doi: 10.1038/s41598-020-62232-3

Zhao, K.-J., Liu, Y., Zhang, R.-Y., Wang, X.-Q., Gao, C., and Shen, J.-K. (2009). A precise, simple, convenient and new method for estimation of intracranial hematoma volume—the formula  $2/3 Sh$ . *Neurol. Res.* 31, 1031–1036. doi: 10.1179/174313209X385662

Zhao, X., Chen, K., Wu, G., Zhang, G., Zhou, X., Lv, C., et al. (2021). Deep learning shows good reliability for automatic segmentation and volume measurement of brain hemorrhage, intraventricular extension, and peripheral edema. *Eur. Radiol.* 31, 5012–5020. doi: 10.1007/s00330-020-07558-2

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A. (2016). “Learning deep features for discriminative localization,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV: IEEE), 2921–2929.

Zhou, Q., Zhu, W., Li, F., Yuan, M., Zheng, L., and Liu, X. (2022). Transfer learning of the ResNet-18 and DenseNet-121 model used to diagnose intracranial hemorrhage in CT scanning. *Curr. Pharm. Des.* 28, 287–295. doi: 10.2174/1381612827666211213143357

Zunair, H., Rahman, A., Mohammed, N., and Cohen, J. P. (2020). “Uniformizing techniques to process CT scans with 3D CNNs for tuberculosis prediction,” in *International Workshop on Predictive Intelligence in Medicine* (Cham), 156–168.