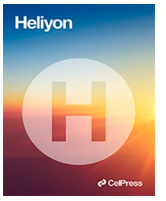
Heliyon 10 (2024) e30270

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Deep multiscale convolutional feature learning for intracranial hemorrhage classification and weakly supervised localization

Bishi He *, Zhe Xu, Dong Zhou, Lei Zhang

School of Automation (School of Artificial Intelligence), Hangzhou Dianzi University, Hangzhou, 310018, Zhejiang, China

ARTICLE INFO

Keywords:
Intracranial hemorrhage
Intraparenchymal
Subarachnoid
Intraventricular
Deep learning

ABSTRACT

Objective: This study evaluated the performance of attentional fusion model-based multiscale features in classifying intracerebral hemorrhage and the localization of bleeding focus based on weakly supervised target localization.

Methods: A publicly available dataset provided by the American College of Neuroradiology (ASNR) was used, consisting of 750,000 computed tomography (CT) scans of the brain, manually marked by radiologists for intracranial hemorrhage and five hemorrhage subtypes. A multiscale feature classification and weakly supervised localization framework based on an attentional fusion mechanism were applied, which could be annotated at the slice level and provided intracranial hemorrhage classification and hemorrhage focus localization.

Results: The designed framework achieved excellent performance for classification and localization. The area under the curve (AUC) for predicting bleeding was 0.973. High AUC values were observed for the five hemorrhage subtypes (epidural AUC = 0.891, subdural AUC = 0.991, subarachnoid AUC = 0.983, intraventricular AUC = 0.995, intraparenchymal AUC = 0.990). This model outperformed the average entry-level radiology trainee compared to previously reported data.

Conclusion: The designed method quickly and accurately detected intracerebral hemorrhage, classifying hemorrhage subtypes and locating bleeding points with image-level annotation alone. The results indicate that this framework can significantly reduce diagnostic time while improving the detection of intracerebral hemorrhage in emergencies. It can thus be integrated into the diagnostic radiology workflow in the future.

1. Introduction

Cerebral hemorrhage is a very urgent and severe disease with high mortality and disability rates. Since hematoma enlargement can lead to further deterioration of neurological deficits, irreversible damage can occur in the first few hours after the onset of intracerebral hemorrhage, making accurate and rapid diagnosis essential to reduce mortality and improve the outcome of patients [1,2]. CT is often a noninvasive and effective imaging method to detect intracerebral hemorrhage. Because the density (HounsfieldUnits, HU) of blood in brain CT images is different from other brain tissues, which is roughly within the range of 0–100 and shows the characteristics of high-density focus, radiologists can identify intracranial hemorrhage through brain CT images, understand the causes of intracranial hemorrhage, determine the subtype, location, size and severity of the bleeding focus, and the risk of imminent brain injury to judge

* Corresponding author.
E-mail addresses: hebs@hdu.edu.cn (B. He), xuzhe@hdu.edu.cn (Z. Xu), 202060291@hdu.edu.cn (D. Zhou), zz_lei@hdu.edu.cn (L. Zhang).

https://doi.org/10.1016/j.heliyon.2024.e30270
Received 10 October 2023; Received in revised form 16 April 2024; Accepted 23 April 2024
Available online 26 April 2024

whether the patient's life is endangered, and provide the central basis for clinicians to take intervention measures [3,4]. A fast and accurate automated classification and localization system for intracerebral hemorrhage can aid doctors to improve diagnostic outcomes, particularly at primary hospitals and emergency centers.

In the past decade, deep learning contributed to significant progress in medical artificial intelligence and has garnered high research interest, especially in the artificial intelligence-aided diagnosis of medical images. There is much evidence that deep learning approaches can perform well in particular medical image diagnostic tasks, such as the detection of pulmonary nodules [5,6], identification and classification of diabetic retinopathy [7], classification of lesions on chest X-ray [8–10] and benign and malignant classification of skin diseases [11], with some indicating accuracy comparable to that of specialists.

Recently, several researchers have used deep learning-based convolutional neural networks (CNNs) [12–16] to classify, detect, and segment intracranial hemorrhage in brain CT. Various medical imaging tasks, such as medical image reconstruction, enhancement, segmentation, registration, and computer-aided diagnosis and detection, have extensively used deep learning. Bar et al. [17] improved classification results by combining the previously independent segmentation and classification tasks with dependencies with relatively few datasets. Remedios et al. [18] utilized multi-instance learning to learn features from weak labels to identify massive bleeds.

Multiscale feature learning is essential when the size of lesions varies greatly, or context information is needed to identify them. Multiscale feature learning does not use the last layer of features for classification and localization but selects multiscale features for fusion before the lesion classification and localization task, to improve model performance. In image detection, detectors usually use feature pyramids to detect targets of different scales, and feature pyramid networks [19]are built by propagating semantically robust features from a high to low-level. Recent work has put forth a novel method that narrowed the differences between different scales of features through consistency supervision before feature fusion [20]. Furthermore, in CT examination datasets, three levels of labeling exist: examination level, slice-level and pixel-level. Pixel-level labels are often required to locate hemorrhagic points, which challenges dataset annotation. A weakly supervised method can be implemented when only slice level labels are used. Weakly supervised localization has been achieved by using annotations at a given image-level [21–24].

Although many schemes have performed excellently in classifying and segmenting intracerebral hemorrhage, several constraints remain. Firstly, the classification and detection model of intracerebral hemorrhage based on CNN is a typical data-driven model. The effectiveness of CNN depends very much on the quantity and labeling quality of intracerebral hemorrhage datasets. Only a large and accurately labeled dataset can enable the model to learn useful information from medical images efficiently. For fully supervised learning, the data volume increase means that the manual annotation workload also increases. In practice, obtaining many medical image datasets is very difficult due to the patient's privacy protection concerns. In addition, image annotation requires a high-level of professional skill, which further compounds the challenges in developing classification and detection systems for cerebral hemorrhage based on CNN. Secondly, because the visual patterns extracted from different types of intracranial hemorrhage disease samples are usually different in appearance, size, and position height, most CNN networks only use the features extracted from the highest convolutional layer, resulting in unsatisfactory classification and location performance for lesion areas. Third, fully supervised localization generally requires high-quality and accurate pixel-level labeling for the localization of intracerebral hemorrhage foci. This task usually requires experienced radiologists to spend considerable time and energy on labeling, which is often unaffordable. Further, image information for Dicom format medical images is lost when the format is changed, and it can be challenging to distinguish intracranial hemorrhage from the surrounding tissues.

To overcome the abovementioned problems, we aimed to develop an end-to-end deep multiscale convolutional feature fusion framework designed to enable accurate classification and localization of intracerebral hemorrhage. Specifically, we proposed that a window adjustment optimization module can be first introduced to enhance the salience of intracranial hemorrhage in CT images and avoid the loss of image information so that the model can maximize the use of image information. Then, during model training, channel and spatial attention mechanisms could be employed to aggregate convolution features at the bottom and high levels for improved focus on lesion areas, especially small ones. This weakly supervised localization method required only section-level labeling for lesion location. Upon model training, the developed medical image intelligent diagnosis system can detect intracranial hemorrhage

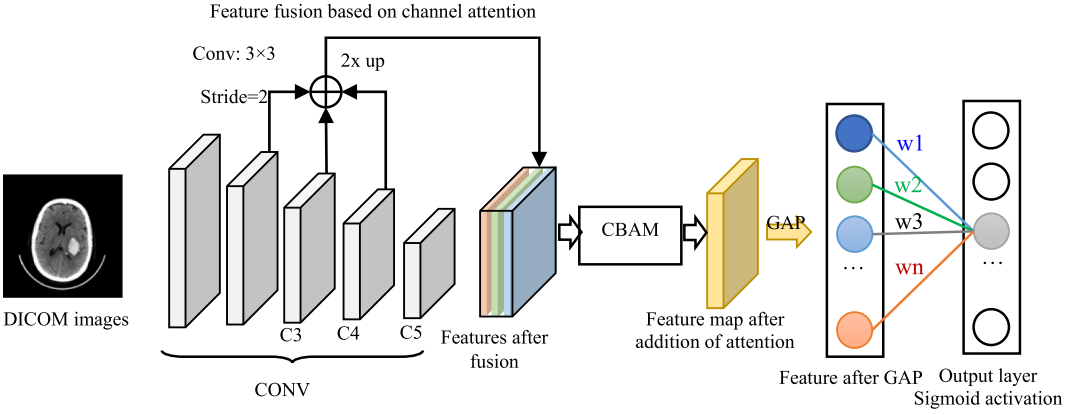


Fig. 1. Classification and weakly supervised location model of intracerebral hemorrhage.

from brain CT images and then classify and locate the subtypes of intracranial hemorrhage into epidural, subdural, subarachnoid, intraventricular, and intraparenchymal.

2. Methods

2.1. Model design

Fig. 1 represents the overall design of the model. The three critical modules in the framework were the window tuning optimization module, multiscale feature fusion module, and bleeding lesion location module (Fig. 1. Location).

A profound convolutional network for large-scale image recognition, VGG-16 [25], was selected as the backbone network, a window tuning optimization module and an attention-based multiscale feature fusion module were added to the basic intracerebral hemorrhage classification model, and the final intracerebral hemorrhage classification and weakly supervised lesion location model were obtained (Fig. 1). Specifically, feature maps of different scales were generated through multilayer convolution for the input brain image. Second, the convolution operation (stride of 2) was utilized to unify the graph size of [C3, C4, C5]. A feature fusion method based on channel attention was adopted to promote the crucial features and suppress the features that contributed little to the current task. Then, a spatial attention mechanism was adopted to fuse the spatial attention further. Finally, the prediction results of intracranial hemorrhage types were output through the GAP and output layers.

2.1.1. Window setting optimization module

The output of medical imaging equipment comprises medical images in Dicom format. Pixels often contain rich grayscale information, but the input of a deep learning model is usually an 8-bit PNG or JPG grayscale image. The information contained in such an image is far less than the 12-bit/16-bit Dicom medical image. The conversion of image format leads to the loss of the information of the original medical image. Therefore, to address this issue, a window adjustment optimization module was designed by combining convolution and activation function properties in deep learning technology, thus imitating the method used by radiologists to identify intracranial hemorrhage by adjusting the window width and window position.

Fig. 2 depicts the window setting optimization module constructed by a $1 \times 1 \times 3$ convolution layer and a custom activation function layer. The module can input different window images by initializing different weight parameters. Additionally, the input layer of the CNN was replaced by this module so that the weight parameters of this module can be updated when training the CNN. That is, the weights were updated in a task-specific way by backpropagation to find the most appropriate weight parameters for windowing input images. At the same time, dynamic parameter updating was used to adjust the window width and position so that the model could fully use the information in medical images and learn more features of the bleeding area. The custom activation functions are shown below.

$$F_{sig}(x) = \frac{U}{1 + e^{-(Wx+b)}} \tag{3.1}$$

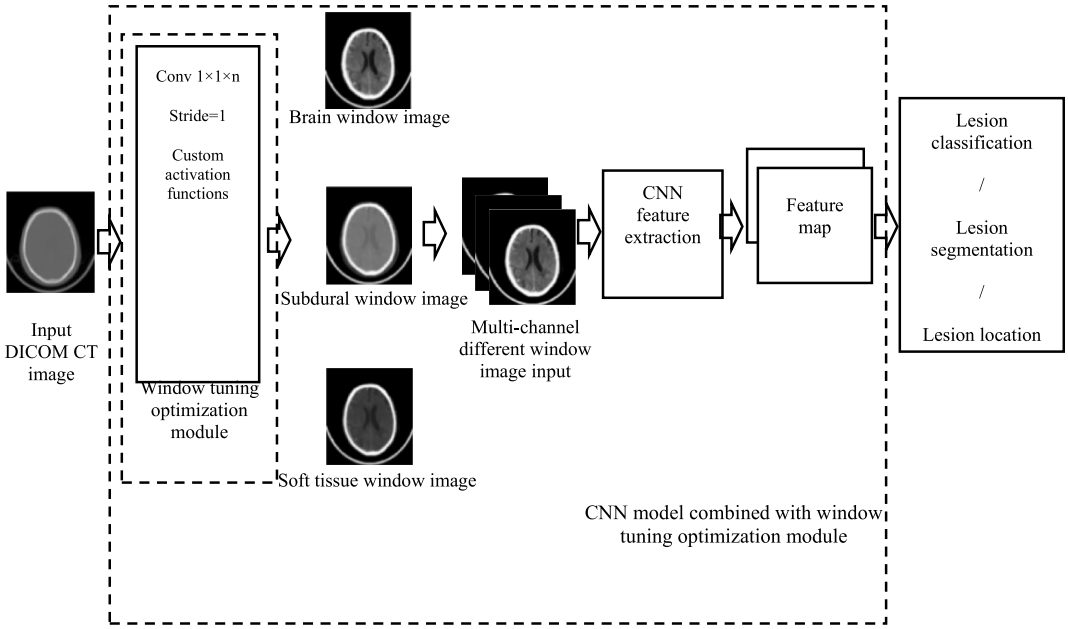


Fig. 2. Tuning window optimization module.

Where, $W = \frac{2}{WW} log\left(\frac{U}{\varepsilon} - 1\right), b = \frac{-2WL}{WW} log\ log\ \left(\frac{U}{\varepsilon} - 1\right)$, WW is the window width, WL is the window level, U is the upper limit of the gray level after activation function mapping, and ε is the margin.(Fig. 2. Location).

Using the weight parameters of the well-trained window setting optimization module of the deep learning model, the optimal window width and level could be obtained by inverse derivation from Equation (3.1) for the optimal window display of intracranial hemorrhage images. Compared with the preset window image, the contrast between the bleeding lesions and the surrounding tissues was more prominent, so the model could more intuitively distinguish whether the current image contained intracranial hemorrhage.

2.1.2. Feature fusion

Combining various scale features in convolutional networks is a crucial strategy for enhancing classification, detection, and segmentation. Various features are contained in different levels of features. For instance, low-level features acquired by shallow networks contain more position and detail information and noise. High-level features acquired by deep networks have robust semantic information but poor detail perception ability. Therefore, efficiently integrating features of different scales is crucial to enhance model performance. Traditional methods first fuse multi-level features and then use the fused features to train a predictor. Detection can only be performed after complete fusion. Fusion features usually use 'Concat' or 'Add' operations. The Concat operation directly connects two features, and its output dimension is the sum of the dimensions of input features. The add operation uses the pixel addition operation to combine multiple feature vectors into a composite vector with the same dimension. Although traditional feature fusion methods can improve recognition accuracy to some extent, it does not reflect the importance of different layer feature maps and does not consider the spatial information of lesions. Therefore, the squeeze-and-congestion (SE) module [26] was introduced into the present model to combine the different scaling features (Fig. 3). First, feature graphs C3, C4, and C5 were converted to the same size through convolution with a step size 2. Then, feature graphs were compressed to global feature vectors with a growth degree of C through average pooling and maximum pooling operations. In the excitation operation, the channel weights were obtained by feeding a global feature vector into two fully connected bottleneck structures, and then the weight of the feature vector was adjusted by learning the coefficient of channel attention. Operating with the SE module, channel interdependencies could be exploited to provide cross-channel recalibration opportunities. Then, the weighted results of the SE module output were fed into a spatial attention module (Fig. 4), which could further integrate spatial attention, and finally, the classification results of intracranial hemorrhage were obtained through GAP and the output layer (Fig. 3. Location). and (Fig. 4. Location) respectively.

Before recalibration, global statistics must be performed for each channel. The global compression across the entire spatial domain is obtained through global maximum and average pooling. Compared with the original SE module, this model uses an additional pooling layer to extract richer high-level features. $U \in R^{h \times w \times c}$ Represents the feature graph after transformation, where $h \times w \times c$ is the dimension. The compression operation aggregates the features of $h \times w$ to obtain a $1 \times 1 \times c$ feature map to describe the channels. If z is a compressed vector, the cth element of z is computed as Equation (3.2).

$$z^c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u^c(i,j) \tag{3.2}$$

The model needs to learn the weights for each channel to recalibrate the feature diagram for the channel field. Based on nonlinear channel interdependence, the automatic gate mechanism was adopted to output channel attention, and the importance coefficient of the channel is as Equation (3.3):

$$s = \sigma(W_2 \times ReLU(W_1 \times z)) \tag{3.3}$$

where σ is the sigmoid operation, $W_1 \in R^{\frac{C}{r} \times C}$, $W_2 \in R^{C \times \frac{C}{r}}$. $W1$ and $W2$ parameterize the two fully connected layers with different compression rates to form a bottleneck structure so that the model can adaptively adjust the channel weights according to learning objectives. After channel c passes through the channel attention module, the final output $\widetilde{x}_c$ is as Equation (3.4):

$$\widetilde{x}_c = s^c . u^c \ c \in \{0, 1, ..., C-1\} \tag{3.4}$$

Where s is the importance coefficient of the channel, and u is the feature graph. The attention channel module has relevance for the classification and localization of intracerebral hemorrhage as it is difficult to distinguish the diseased area from the normal area of different hemorrhage types, and the use of a single feature map or independent processing of multiple feature maps cannot provide sufficient feature information for the classification of intracerebral hemorrhage.

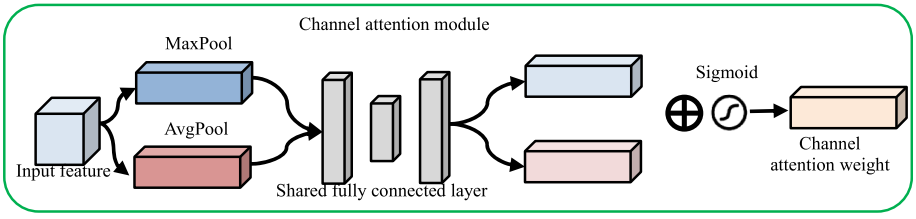


Fig. 3. Channel attention module.

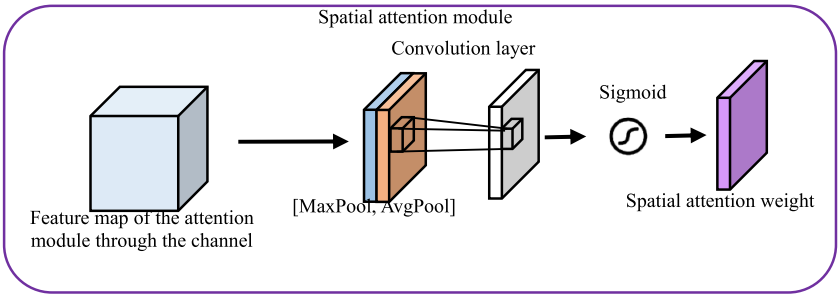


Fig. 4. Spatial attention module.

As channel attention learns the "what", with spatial attention, the model learns the "where". The spatial attention module differs from channel attention in emphasizing "where" is the most critical characteristic information. These are first connected along the channel axis using average and maximum pooling operations to obtain valid feature descriptors before calculating spatial attention. Then, a convolution layer is used to generate spatial attention as Equation (3.5),

$$M_s(F) = \sigma\left(f^{7\times7}\left(\left[F^{avg_s}_{avg}, F^{\max_s}_{\max}\right]\right)\right) \tag{3.5}$$

where σ represents sigmoid operation, and $f^{7\times7}$ is the convolution function of kernel 7×7.

2.1.3. *Weakly supervised localization*

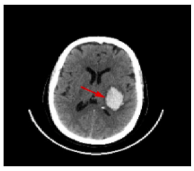
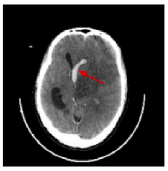
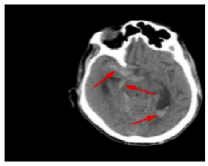
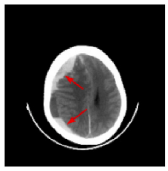
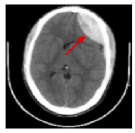
For weakly supervised localization, class-activated mapping (CAM) [27] was used to find the most differentiated part in the image to achieve lesion location. To improve the accuracy of lesion location, multiscale features were fused with channel attention, and then features were further extracted with spatial attention. Finally, a heatmap was obtained based on CAM. Twenty percent of the maximum activation value in the heatmap was selected as the binarization threshold, and the heatmap generated by CAM took the threshold. After that, the minimum enclosing rectangle with the maximum connected domain was found, and its boundary coordinates were extracted as the output of the model to generate BBox.

2.2. *Experimental protocol*

2.2.1. *Dataset used*

A public dataset was used for our research, provided by the American College of Neuroradiology (ASNR) through the Kaggle platform for phase II training. It includes a total of 750,000 CT scans of the brain corresponding to 6 types of labels prelabeled by radiologists, namely, whether there was intracranial hemorrhage (any), epidural, intraparenchymal, intraventricular, subarachnoid, and subdural. When the sections were labeled intracerebral hemorrhage, at least one of the five hemorrhage subtypes was labeled. Intraparenchymal hemorrhage (IPH) refers to the hemorrhage wholly located inside the brain; Intraventricular (IVH) or subarachnoid hemorrhage (SAH) refers to hemorrhage that infiltrates into the intracerebral space containing cerebrospinal fluid (ventricle or subarachnoid cistern); Extra-axial hemorrhage, including subdural hemorrhage (SDH) and epidural hemorrhage (EDH), refers to the hemorrhage gathered in the tissue covering around the brain. In addition, because of multiple causes, several cerebral hemorrhage subtypes may occur in a patient at once, manifesting as various hemorrhage subtypes on the same CT scan slice. The specific bleeding

Table 1
Imaging characteristics of various subtypes of hemorrhage.

	IPH	IVH	SAH	SDH	EDH
Position	Inside the brain	Intraventricular	Between arachnoid and pia mater	Between dura and arachnoid	Between dura and skull
CT images					
Pathogenesis	Hypertensio, trauma, arteriovenous malformation, tumor	Related to intraparenchymal and subarachnoid hemorrhage	Rupture of aneurysm, arteriovenous malformation, trauma	Trauma	After surgery, trauma
Shape	Usually, round	Similar to ventricular shape	Along sulci and fissure	Crescent	Fusiform

in each subtype is shown in Table 1. (Table 1 location).

2.2.2. Training process and evaluation indicators

Our method was implemented based on the PyTorch framework and accelerated by an NVIDIA GTX 1080Ti 12G × 4 graphics card. The training rounds included 20 epochs. The optimizer adopts the Adam method, batch_size is 8. The initial learning rate was 0.001, which was decreased by 10 times every 5 epochs. The model with the highest score in the validation set was finally chosen for testing. The initialization parameters of the window tuning optimization module are presented in Table 2 (Table 2 location).

The receiver operating curve (ROC) analysis and AUC values were used to evaluate the performance. The ROC curve was drawn with FPR as the x-axis, TPR as the y-axis, and AUC as the region under ROC curves. The average precision (AP) was used to assess disease localization. AP is a comprehensive index of accuracy and recall rate, and its value is the region under the PR curve drawn with recall as x-axis and precision as y-axis (TP is taken as a prediction box with IoU greater than 0.5).

$$AP = \sum_{k=1}^{n} P(k)\Delta R(k) \tag{4.1}$$

In Equation (4.1), $\Delta R(k)$ is the evenly divided interval between [0,1] in the abscissa. After the average accuracy AP of each classification was obtained, the mAP was obtained by taking the mean of the average accuracy to evaluate the overall positioning accuracy.

2.2.3. Ablation study

Ablation experiments were conducted on the proposed network framework to prove the effectiveness of window setting optimization module and multiscale feature fusion. Each module was removed sequentially to verify the model's performance differences. Only the network structure was altered in the following three trials to ensure a fair comparison. All other experimental settings remained the same.

3. Results

Intracerebral hemorrhage and subtypes of intracerebral hemorrhage were analyzed in the whole dataset, of which 644,870 sections did not contain intracerebral hemorrhage, 75,859 sections contained two types of hemorrhage, 24,826 sections contained three types of hemorrhage, and 6217 sections contained four types of hemorrhage. A total of 1 % (~7500 sections) of the data were randomly selected as the test set according to the proportion of bleeding subtypes, and the rest were used as training sets. Fig. 5 shows the proportion of intracranial hemorrhage subtypes. Intracranial hemorrhage accounted for approximately 14.34 % (14.41 %), but each subtype of hemorrhage was uneven. Epidural hemorrhage was the lowest, accounting for 0.42 % (0.45 %). Subdural hemorrhage was the highest, accounting for 6.26 % (6.15 %). Parenchymal hemorrhage, ventricular hemorrhage and subarachnoid hemorrhage accounted for 4.80 % (4.90 %), 3.48 % (3.44 %), and 4.74 % (5.09 %), respectively. (Fig. 5 location).

Based on the trained model, the classification and localization performance of the test dataset were evaluated. The ROC curves and AUC scores of each category are displayed in Fig. 6, where 'any' indicates detecting intracranial hemorrhage, and the rest are the classification results (Fig. 6 location).

To evaluate and compare the localization precision of weakly supervised lesions proposed in this research, mAP = 0.2 was calculated. Specifically, the AP values of each hemorrhage subtype are $AP_{IVH} = 0.07$, $AP_{IPH} = 0.27$, $AP_{SAH} = 0.18$, and $AP_{SDH} = 0.48$. Among these, the number of EDH samples was too small, so the localization accuracy was not calculated temporarily. SDH and IPH achieved better performance.

Results of the ablation experiment are presented in Table 3; the average AUC of our method was 0.973. Compared with the sequential removal of the other two modules, the average AUC was improved. Only the network structure was altered in the following three trials. All other experimental parameters remained the same (Table 3 location).

For the window setting optimization module, different inputs were used for training and evaluation under the same dataset, and the same parameter settings and the classification results for intracranial hemorrhage are presented in Table 4. (Table 4 Location).

In Table 4, initialization refers to initializing the input HU value into different window images. The three windows are the cerebral, subdural, and soft tissue windows. The cerebral window's width (WW) and WL were set to 80, 40, the subdural window's WW and WL were set to 200 and 80, and the soft tissue window's WWWL was set to 380 and 40, respectively. The absence of weights in training means that the window tuning optimization module only completes mapping the HU value to fixed window images, and its weight parameters were not updated, equivalent to the original VGG-16 model. The brain window's WL and width (WW) were set to 40. The subdural window's WL and width (WW) were set to 80.

From the visualization standpoint, the single-window initialization model and the three-window initialization model were selected, and the weight parameters W and B of their window tuning optimization module were extracted. Based on Equation (3.1), the optimal

Table 2
Initialization weight parameter values of the window tuning optimization module.

The weight parameters	Brain window	Subdural window	Soft tissue window
W	0.13843335	0.05537334	0.02914386
b	−5.5373344	−4.4298673	−1.1657546

Fig. 5. Proportion of intracerebral hemorrhage subtypes in the training and test sets.

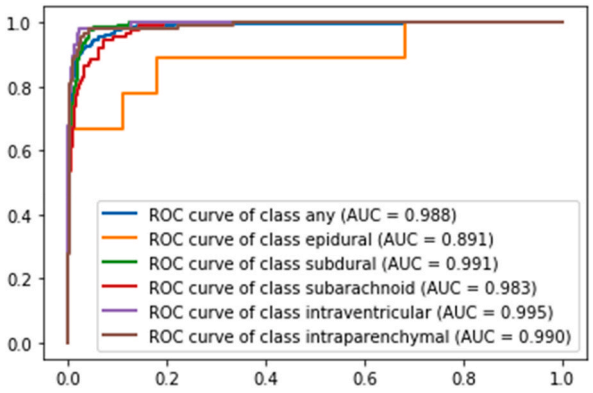


Fig. 6. ROC curves and AUC scores for each category.

Table 3
Effectiveness of the two modules.

	Our Method	w/o module window	w/o multiple feature fusion	Junior radiologist[Ye,2019]	Senior radiologist[Ye,2019]
Any	0.988	0.943	0.91	0.97	1.0
Epidural	0.891	0.845	0.82	0.85	0.96
Subdural	0.991	0.965	0.913	0.79	0.97
Subarachnoid	0.983	0.9	0.86	0.66	0.96
Intraventricular	0.995	0.98	0.871	0.84	0.97
intraparenchymal	0.990	0.91	0.864	0.89	0.92
Average	0.973	0.923	0.873		

Table 4
AUCs input from different models.

Input	Window tuning optimization module		Prediction of intracranial hemorrhage	Note
	Initialization	Weight training	AUC	
HU	–	–	0.891	VGG-16
HU	Brain window	–	0.937	VGG-16
HU	Three windows	–	0.961	VGG-16
HU	–	√	0.924	VGG-16+ window tuning optimization module
HU	Brain window	√	0.959	VGG-16+ window tuning optimization module
HU	Three windows	√	**0.973**	VGG-16+ window tuning optimization module

window width WW and *WL* could be deduced backward to visualize the brain image (Fig. 7). The intracranial images visualized with the optimal window width significantly differed from those visualized with the fixed window width (Fig. 7 location).

The three-window optimization was also used as input, and a heatmap based on CAM was generated to locate the lesions in the test set, depicted in Fig. 8. Here, the first column is a manually labeled BBox, the second column is a heatmap with feature fusion removed from the model, and the third column is a heatmap with multiscale feature fusion. Multiple feature fusion resulted in more complete localization of the bleeding lesion (lines 1 and 4), more accurate localization (line 2), and more obvious localization of the lesion (lines 3 and 4). Multiple feature fusion could locate more bleeding lesions for multiple types of intracerebral hemorrhage in a single section (line 5). (Fig. 8 location).

4. Discussion

As many patients with intracerebral hemorrhage often receive CT examination in the emergency department, a rapid and highly accurate diagnostic protocol is warranted. Furthermore, in many clinical centers, junior radiologists or emergency physicians typically provide the initial diagnosis for head CT. Unfortunately, their expertise might be lower in comparison to more experienced radiologists, and time limitations may be high in such settings, which may result in missed diagnosis or misdiagnosis, highlighting a pressing need for intelligent diagnostic aids in this context.

Compared with the existing method proposed by Chen et al. [28] for weakly supervised localization of intracranial hemorrhage, our network localization performance mAP was improved from 0.133 to 0.2, an increase of 5 %. Although there is still a large gap between mAP and strong supervision, based on weak supervision, only section-level annotations are needed, and pixel-level annotations are not needed, which provides a feasible solution for the utilization of medical images lacking pixel-level annotations. Based on the experimental results, we showed that using three-window input can achieve higher performance than single-window input. Compared with the fixed window image model, the present model can achieve better results after using the window adjustment optimization model. Furthermore, multiscale feature fusion can not only improve classification AUC but also improve lesion localization performance.

At the same time, the weakly supervised location method and experimental result data in the present study also provide an analytical and comparative path for future research. The present study leveraged process radiologists use in identifying lesion types and the window adjustment optimization algorithm enhanced the detection of intracerebral hemorrhage by obtaining an optimal window width and window position value and optimizing network input. This method can be widely applied to X-ray, CT, magnetic resonance, and other medical image processing.

The study also indicates that integrating low-level and high-level multiscale features through an attentional mechanism, classification and localization ability of an image analysis model can be improved. A major strength of the weakly supervised object localization method is that only image-level annotation is required to locate lesions, which provides a feasible reference scheme for the location of lesions in other medical image diagnostic systems. The limitation of the present report includes the limited samples for epidural type of intracranial hemorrhage, therefore the localization performance for epidural lesions was not high. In the future, small-sample learning methods will be further studied to improve the localization performance to help develop a more accurate intelligent diagnosis system for intracranial hemorrhage.

5. Conclusion

Based on deep learning methods, a novel framework of intracranial hemorrhage classification and weakly supervised location is proposed, which can quickly realize the preliminary screening of intracranial hemorrhage and realize the classification and localization of intracranial hemorrhage subtypes. In addition, the performance of intracerebral hemorrhage classification and weakly supervised location is further improved through window setting optimization and a multiscale feature fusion mechanism. The effectiveness of the overall framework and each optimization module was verified through extensive experiments and evaluations on open data sets.

Funding

No funds, grants, or other support was received for this study.

Information sharing statement

Datasets used in the study will be provided upon request to the authors.

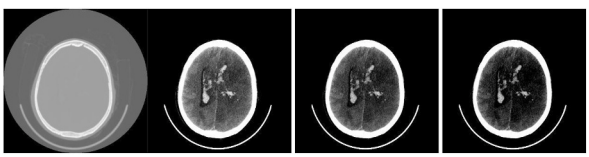


Fig. 7. Visualization of optimal window width and window position.

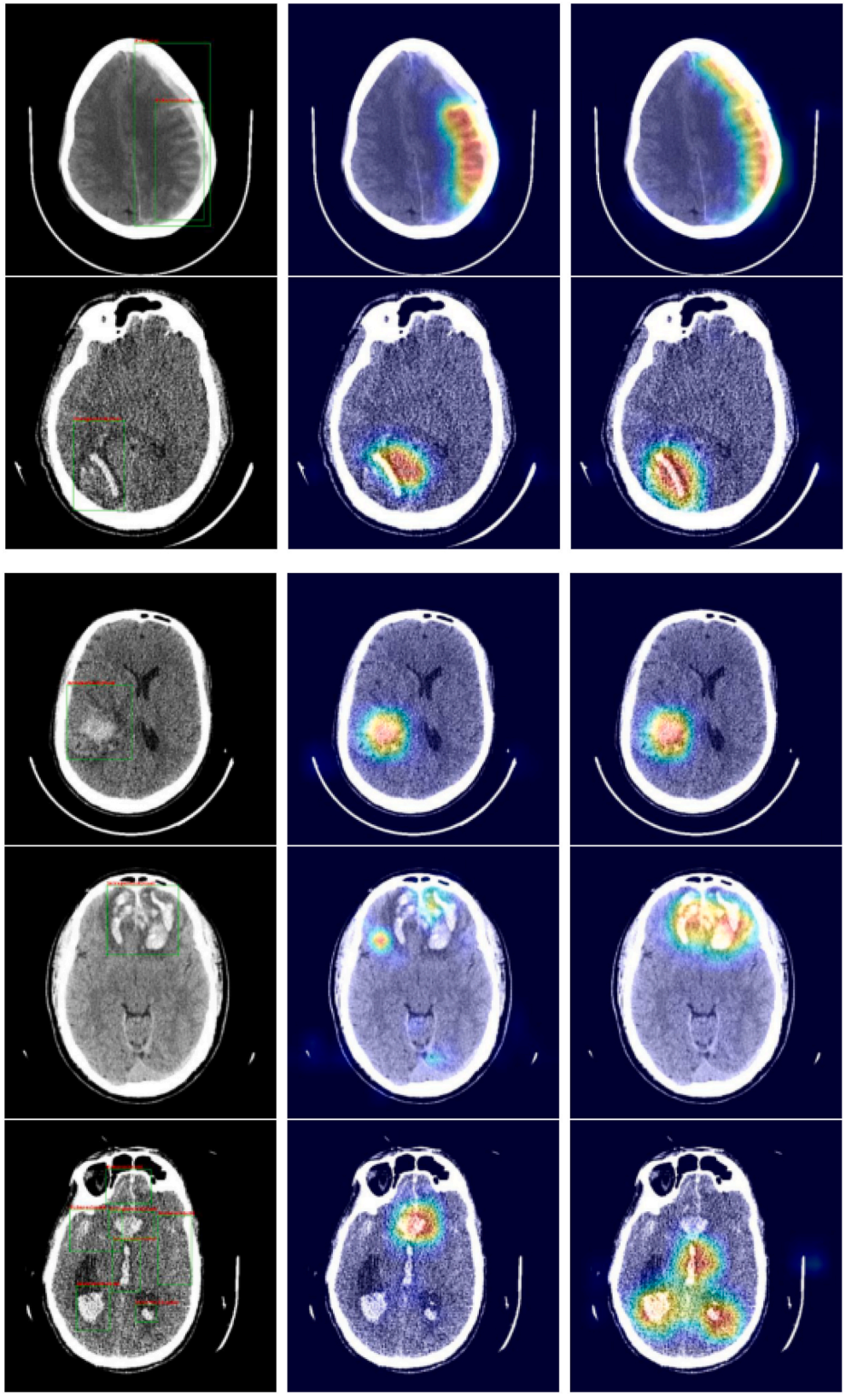


Fig. 8. Focal location based on CAM heat map.

Data availability statement

The datasets generated and/or analyzed during the current study are available in the public repositories: https://www.kaggle.com/competitions/rsna-intracranial-hemorrhage-detection/data.

CRediT authorship contribution statement

Bishi He: Supervision, Project administration, Funding acquisition, Data curation, Conceptualization. **Zhe Xu:** Writing – original draft, Visualization, Resources, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Dong Zhou:** Writing – review & editing, Writing – original draft, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Lei Zhang:** Writing – review & editing, Visualization, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:The authors declare no competing interests. The authors have no relevant financial or non-financial interests to disclose.

Acknowledgement

None.

Table of Abbreviations

AP	Average precision
AUC	Area under the curve
CAM	Class-activated mapping
CNN	Convolutional neural networks
CT	Computed tomography
ROC	Receiver operating curve
WL	Window level
WW	Window's width

References

[1] J.J. Heit, M. Iv, M. Wintermark, Imaging of intracranial hemorrhage, Journal of stroke 19 (1) (2017) 11.
[2] C.J. Van Asch, M.J. Luitse, G.J. Rinkel, I. van der Tweel, A. Algra, C.J. Klijn, 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 (2) (2010) 167–176.
[3] D.B. Larson, L.W. Johnson, B.M. Schnell, S.R. Salisbury, H.P. Forman, National trends in CT use in the emergency department: 1995–2007, Radiology 258 (1) (2011) 164–173.
[4] L. Papa, I.G. Stiell, C.M. Clement, A. Pawlowicz, A. Wolfram, C. Braga, S. Draviam, G.A. Wells, Performance of the Canadian CT Head Rule and the New Orleans Criteria for predicting any traumatic intracranial injury on computed tomography in a United States Level I trauma center, Acad. Emerg. Med. 19 (1) (2012) 2–10.
[5] P. Sermanet, S. Chintala, Y. LeCun, Convolutional neural networks applied to house numbers digit classification, in: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), IEEE, 2012, pp. 3288–3291.
[6] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2097–2106.
[7] R. Gargeya, T. Leng, Automated identification of diabetic retinopathy using deep learning, Ophthalmology 124 (7) (2017) 962–969.
[8] F. Ciompi, K. Chung, S.J. Van Riel, A.A.A. Setio, P.K. Gerke, C. Jacobs, E.T. Scholten, C. Schaefer-Prokop, M.M. Wille, A. Marchiano, Towards automatic pulmonary nodule management in lung cancer screening with deep learning, Sci. Rep. 7 (1) (2017) 46479.
[9] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison, Proc. AAAI Conf. Artif. Intell. (2019) 590–597.
[10] B. Liu, W. Chi, X. Li, P. Li, W. Liang, H. Liu, W. Wang, J. He, Evolving the pulmonary nodules diagnosis from classical approaches to deep learning-aided decision support: three decades' development course and future prospect, J. Cancer Res. Clin. Oncol. 146 (2020) 153–185.
[11] A. Hekler, J.S. Utikal, A.H. Enk, C. Berking, J. Klode, D. Schadendorf, P. Jansen, C. Franklin, T. Holland-Letz, D. Krahl, Pathologist-level classification of histopathological melanoma images with deep neural networks, Eur. J. Cancer 115 (2019) 79–83.
[12] P.D. Chang, E. Kuoy, J. Grinband, B.D. Weinberg, M. Thompson, R. Homo, J. Chen, H. Abcede, M. Shafie, L. Sugrue, Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT, Am. J. Neuroradiol. 39 (9) (2018) 1609–1616.
[13] J. Cho, K.-S. Park, M. Karki, E. Lee, S. Ko, J.K. Kim, D. Lee, J. Choe, J. Son, M. Kim, Improving sensitivity on identification and delineation of intracranial hemorrhage lesion using cascaded deep learning models, J. Digit. Imag. 32 (2019) 450–461.
[14] M.D. Hssayeni, M.S. Croock, A.D. Salman, H.F. Al-Khafaji, Z.A. Yahya, B. Ghoraani, Intracranial hemorrhage segmentation using a deep convolutional model, Data 5 (1) (2020) 14.
[15] R. Kirithika, S. Sathiya, M. Balasubramanian, P. Sivaraj, Brain tumor and Intracranial haemorrhage feature extraction and classification using conventional and deep learning methods, Eur. J. Mol. Clin. Med 7 (7) (2020) 237–258.
[16] H. Ye, F. Gao, Y. Yin, D. Guo, P. Zhao, Y. Lu, X. Wang, J. Bai, K. Cao, Q. Song, Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network, Eur. Radiol. 29 (2019) 6191–6201.

[17] A. Bar, M.M. Havakuk, Y. Turner, M. Safadi, E. Elnekave, Improved ich classification using task-dependent learning, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE, 2019, pp. 1567–1571.
[18] S. Remedios, Z. Wu, C. Bermudez, C.I. Kerley, S. Roy, M.B. Patel, J.A. Butman, B.A. Landman, D.L. Pham, Extracting 2D weak labels from volume labels using multiple instance learning in CT hemorrhage detection, in: Medical Imaging 2020: Image Processing, SPIE, 2020, pp. 66–75.
[19] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, Feature pyramid networks for object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2117–2125.
[20] C. Guo, B. Fan, Q. Zhang, S. Xiang, C. Pan, Augfpn: improving multi-scale feature learning for object detection, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 12595–12604.
[21] J. Choe, H. Shim, Attention-based dropout layer for weakly supervised object localization, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 2219–2228.
[22] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: visual explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 618–626.
[23] X. Zhang, Y. Wei, J. Feng, Y. Yang, T.S. Huang, Adversarial complementary learning for weakly supervised object localization, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 1325–1334.
[24] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, Learning deep features for discriminative localization, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2921–2929.
[25] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv Preprint arXiv:1409.1556, 2014.
[26] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.
[27] T. Tagaris, M. Sdraka, A. Stafylopatis, High-resolution class activation mapping, in: 2019 IEEE International Conference on Image Processing (ICIP), IEEE, 2019, pp. 4514–4518.
[28] Z. Chen, R. Ji, J. Wu, Y. Shen, Multi-scale features for weakly supervised lesion detection of cerebral hemorrhage with collaborative learning, in: Proceedings of the 1st ACM International Conference on Multimedia in Asia, 2019, pp. 1–7.