Received: 24 September 2021 | Revised: 15 February 2022 | Accepted: 9 March 2022

DOI: 10.1002/hbm.25845

TECHNICAL REPORT

WILEY

Deep learning derived automated ASPECTS on non-contrast CT scans of acute ischemic stroke patients

Zehong Cao[1,2] | Jiaona Xu[3,4] | Bin Song[5] | Lizhou Chen[5] | Tianyang Sun[2] | Yichu He[2] | Ying Wei[2] | Guozhong Niu[4] | Yu Zhang[1] | Qianjin Feng[1] | Zhongxiang Ding[6] | Feng Shi[2] | Dinggang Shen[2,7]

[1]School of Biomedical Engineering Southern Medical University, Guangzhou, China

[2]Department of Research and Development, Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China

[3]The Fourth School of Clinical Medicine, Zhejiang Chinese Medicine University, Hangzhou, China

[4]Department of Neurology, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China

[5]Department of Radiology, West China Hospital, Sichuan University, Chengdu, China

[6]Department of Radiology, Key Laboratory of Clinical Cancer Pharmacology and Toxicology Research of Zhejiang Province, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China

[7]School of Biomedical Engineering, ShanghaiTech University, Shanghai, China

Correspondence
Zhongxiang Ding, Department of Radiology, Key Laboratory of Clinical Cancer Pharmacology and Toxicology Research of Zhejiang Province, Affiliated Hangzhou First People's Hospital, Zhejiang University School of Medicine, Hangzhou, China.
Email: hangzhoudzx73@126.com

Feng Shi, Department of Research and Development, Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China.
Email: feng.shi@uii-ai.com

Dinggang Shen, School of Biomedical Engineering, ShanghaiTech University, Shanghai, China.
Email: dinggang.shen@gmail.com

Funding information
National Key Research and Development Program of China, Grant/Award Number: 2018YFC0116400; National Science Foundation of China, Grant/Award Numbers: 62131015, 81871337; Science and Technology Commission of Shanghai Municipality (STCSM), Grant/Award Number: 21010502600

Abstract

Ischemic stroke is the most common type of stroke, ranked as the second leading cause of death worldwide. The Alberta Stroke Program Early CT Score (ASPECTS) is considered as a systematic method of assessing ischemic change on non-contrast CT scans (NCCT) of acute ischemic stroke (AIS) patients, while still suffering from the requirement of experts' experience and also the inconsistent results between readers. In this study, we proposed an automated ASPECTS method to utilize the powerful learning ability of neural networks for objectively scoring CT scans of AIS patients. First, we proposed to use the CT perfusion (CTP) from one-stop stroke imaging to provide the golden standard of ischemic regions for ASPECTS scoring. Second, we designed an asymmetry network to capture features when comparing the left and right sides for each ASPECTS region to estimate its ischemic status. Third, we performed experiments in a large main dataset of 870 patients, as well as an independent testing dataset consisting of 207 patients with radiologists' scorings. Experimental results show that our network achieved remarkable performance, as sensitivity and accuracy of 93.7 and 92.4% in the main dataset, and 95.5 and 91.3% in the independent testing dataset, respectively. In the latter dataset, our analysis revealed a high positive correlation between the ASPECTS score and the prognosis of patients in 90DmRs. Also, we found ASPECTS score is a good indicator of the size of

Zehong Cao, Jiaona Xu, and Bin Song contributed equally to this work.

CTP core volume of an infraction. The proposed method shows its potential for automated ASPECTS scoring on NCCT images.

KEYWORDS
ASPECTS, asymmetry, deep learning, ischemic, stroke

1 | INTRODUCTION

Stroke is the leading cause of long-term disability and the second most common incidence of death worldwide (Kim et al., 2020). One in six people worldwide would have a stroke in their lifetime. More than 13.7 million people have a stroke each year, of which approximately 80% are ischemic strokes and 5.8 million lead to death (Murphy & Werring, 2020). Clinical manifestations of stroke generally include weakness or paralysis or loss of sensation in any part of the body, even hemiplegia, loss of balance and coordination, speech and visual impairment, severe headaches, memory decline, difficulty swallowing, and involuntary eye movements (Jones, O'Connell, David, & Chalder, 2020; Ojaghihaghighi, Vahdati, Mikaeilpour, & Ramouz, 2017; Runchey & McGee, 2010). Non-contrast Computed Tomography (NCCT) is generally used as first-line imaging to determine the site of brain hemorrhage and infarction, which leads to specific treatment processes (Powers et al., 2019; Yew & Cheng, 2009).

Previously, the diagnosis of ischemic stroke patients was based on the 1/3 rule, where ischemic lesions less than 1/3 of the area supplied by the middle cerebral artery (MCA) were considered suitable for thrombolytic therapy with NCCT images (Van Seeters et al., 2013). While nowadays, Alberta Stroke Program Early Computed Tomography Score (ASPECTS; Schroder & Thomalla, 2016) is widely used as a systematic method for quantifying early ischemic changes (EIC) with NCCT images. Early studies have shown that ASPECTS is more reliable than the 1/3 rule (Barber, Demchuk, Zhang, & Buchan, 2000). ASPECTS scoring is defined as follows. First, the MCA vascular territory was separated into 10 regions of each hemisphere. There are six cortical regions, for example, M1 corresponding to the frontal operculum, M2 to the anterior temporal lobe, M3 to the posterior temporal lobe, with M4, M5, M6 being superior to M1, M2, M3, respectively. The four subcortical regions include caudate (C), lentiform (L), internal capsule (IC), and insular (I). Then, if the left and right sides of one region show signs of ischemic stroke, such as hypodensity, gray-white distinction, and focal swelling, 1 point is deducted from the initial score of 10 points (Neuhaus et al., 2020). ASPECTS scores were found correlated with clinical outcomes, where the American Heart Association/American Stroke Association (AHA/ASA) guidelines recommend selecting patients with ASPECTS > 6 for intravascular thrombectomy (Powers et al., 2019). Although most randomized trials excluded patients with an ASPECTS ≤6 or 7, recent studies suggest that patients with low ASPECTS may also benefit from thrombectomy (Kaesmacher et al., 2019; Yoo et al., 2016).

However, in clinical practice, ASPECTS is a subjective scoring system and thus presents poor inter-rater reliability (Farzin et al., 2016). In recent years, automated ASPECTS methods based on machine learning have been proposed to provide objective and fast recommendations of scores based on the NCCT data. Their performances were also reported to be comparable to experienced neuroradiologists (Hoelter et al., 2020; Naganuma et al., 2021). Hoelter et al. concluded that among the three most common scoring software on the market (Syngo.via Frontier ASPECT Score Prototype V2, Brainomix e-ASPECTS and RAPID ASPECTS), the highest correlation ($r = 0.871$, $p < .001$) was found between expert reading and Brainomix (e-ASPECTS software), followed by Frontier V2, then RAPID (Hoelter et al., 2020). Neuhaus et al. compared the results of the automated software e-ASPECTS with ground truth acquired from experts and obtained fair agreement in 178 baseline NCCT scans (Neuhaus et al., 2020). Brinjikji et al. separately compare the inter-class concordance of ASPECTS scores given by 16 experts for 60 patients in the condition with and without the e-ASPECTS aid in NCCT and follow-up CT, and they found the e-ASPECTS software could improve the accuracy of ischemic prediction (Brinjikji et al., 2021). However, it was reported that the performance of e-ASPECTS software would be impacted by the slice thickness. Neuberger et al. found that the ASPECTS scores acquired from e-ASPECTS software have no significant difference between thin slice thickness and ground truth. While when it applies to thick slice thickness images, the performance would have a substantial decrease, which has significant difference with its ground truth (Neuberger et al., 2020). Besides using expert scoring from NCCT as golden standard, recent studies employed Diffusion-Weighted Imaging (DWI) (Yoshimoto et al., 2019). Briefly, DWI can detect ischemic tissue within a few minutes, and has become a sensitive and specific imaging technique for acute ischemia. For example, Qiu et al. proposed an automatic ASPECTS using random forest by training on radiomics features of each partition with CT images of 157 patients followed by DWI within 1 hr as ground truth. They achieved an agreement with radiologists' manual scorings where accuracy was up to 84.9% for dichotomous results on a testing set of 100 subjects (Kuang et al., 2019). DWI is sensitive in the detection of small and early infarcts in acute ischemic stroke. However, DWI is not used as a first-line imaging modality in most of stoke centers. Despite the MR availability in these facilities, one-stop-shop stroke imaging with CT was widely adopted and DWI might be used for the follow-up scans. Regarding this, the approach of using DWI as reference standard has a great challenge of collecting sufficient data, and so in this study we used CTP as the reference.

The human brain is largely symmetrical, and comparing the left and right sides of a brain region is wildly used in the clinical evaluation of medical imaging (Kalavathi, Senthamilselvi, & Prasath, 2017; Shi

et al., 2012). Similarly, in the image analysis field (Chen et al., 2017; Liu, Zhang, Adeli, & Shen, 2019; Munsell et al., 2015; Shi et al., 2012), Herzog et al. performed an early diagnosis of Alzheimer's disease by comparing the differences between the left and right hemispheres of the brain (Herzog & Magoulas, 2021). Li et al. utilized the property of symmetry of brain structures by taking the differential features of the hippocampus on both sides as an input channel (Li et al., 2021). In the case of ischemic stroke, the CT appearance of affected regions may also show asymmetry that could be utilized by the machine learning model and thus contribute to the automated ASPECTS scoring.

In this article, we propose a novel method, named deep asymmetry network (DA-Net), to score ASPECTS from NCCT images. (1) We propose to utilize the infarct region in the CTP as the ground truth of NCCT to train models. As general interval time of DWI and clinical CT is too long, that is, almost beyond 2 hr, that image features are inconsistent. In contrast, CT perfusion and CT scans are generally separated by a few minutes in a one-stop scanning scenario, which is relatively reliable and will provide a large amount of data for training better performance models. Among CTP parameter maps, the cerebral blood flow (CBF) represents the rate of delivery of arterial blood to a capillary bed in the brain tissue, measured in units of ml of blood per 100 g of tissue per minute. The smaller the CBF value, the more severe the ischemia of the tissue. When the CBF value in a region decreases to 30% of the normal brain tissue in the opposite brain side, the cerebral infarction in that region is considered as infarct core area. The infract core is initially invisible on NCCT at the hyperacute ischemic stage especially when the onset to imaging time is less than 3 hr. As the disease progress, the ischemic signs as mentioned above gradually become visible (Rudkin, Cerejo, Tayal, & Goldberg, 2018). According to Voleti et al., there is a moderate correlation between NCCT ASPECTS and CTP core volume (defined by CBF < 30%) in acute ischemic strokes ($r = -0.55$, $p < .0001$) (Voleti et al., 2021). In this study, we chose CTP as the reference standard to evaluate the ischemic signs in NCCT for patients in acute stage. (2) Here we propose a novel deep asymmetry network that compares left and right differences in CT for ASPECTS scoring. (3) Besides the main dataset, we also introduce an independent testing dataset to evaluate the performance of our method, with manual results from radiologists as ground truth.

2 | METHOD

2.1 | Overview

In this study, we propose a framework for automatically generating ASPECTS scores from NCCT images. Briefly, in the preprocessing stage, we adjust the orientation of NCCT to a standard space to facilitate the comparison of the left and right sides of brain regions. Then, we adopt a segmentation neural network, namely VB-Net (Han, 2019; Hua et al., 2020), to remove skulls and obtain ASPECTS regions, respectively. Finally, a deep asymmetry network (DA-Net) is designed to automatically score ASPECTS for each pair of brain regions as Figure 1a.

2.2 | Preprocessing

Many stroke patients have decreased consciousness to cooperate with the image scan, and thus the brain orientation in CT images varies largely. To this end, we first align the routine CT scan to a standard CT template to ensure image symmetry, that is, the left and right sides show the same structural regions. Given that, in clinics, the identification of local ischemia largely depends on the contralateral information. Skull stripping is then performed to reduce the impact of non-brain tissues, which is a standard process in brain image analysis. Finally, images are normalized to intensity range -1 to 1 and in-plane resolution of $0.5 \times 0.5\,\text{mm}^2$.

2.3 | Acquiring ASPECTS regions via VB-Net

In this step, we segment 10 paired ASPECTS regions from preprocessed CT images. Based on the widely used V-net (Milletari, Navab, & Ahmadi, 2016) in segmentation tasks, our previous work proposed VB-Net which further combines bottleneck layers for fewer network parameters and thus faster speed. The VB-Net has been successfully used in organ and lesion segmentations (Han, 2019; Shi et al., 2021). We use a cascade structure to segment 20 regions in a coarse-to-fine manner with the ground truth outlined by radiologists according to the ASPECTS regions. In the first stage, we train a coarse segmentation network on downsampled images to roughly locate the region of interest to be scored. Then the original resolution images were used to train a another finely segmentation network within the partition detected in first stage to accurately segment the scored regions. In testing, a cascade of above two network (VB-Net) is employed for inference. Considering the image slice thickness (3–6 mm) is usually much larger than that of in-plane resolution, the first two down blocks are performed in the x and y directions, and the last two down blocks are performed in the x, y, and z directions simultaneously. The architecture of VB-Net is shown in Figure 1b.

2.4 | Proposed deep asymmetry network (DA-Net) for ASPECTS scoring

In this article, we propose a novel asymmetry network for comparing the left and right regions and predicting a score. The framework consists of feature extraction and contralateral comparison, as neurologists usually evaluate ASPECTS scores by comparing bilateral asymmetric vascular territory. For each of 10 brain regions, the left and right sides from the segmented image are used as two 3D patches for network input. The patches have a size of [80, 160, 120], window level of 30, and window width of 50. For the network to better learn the asym-

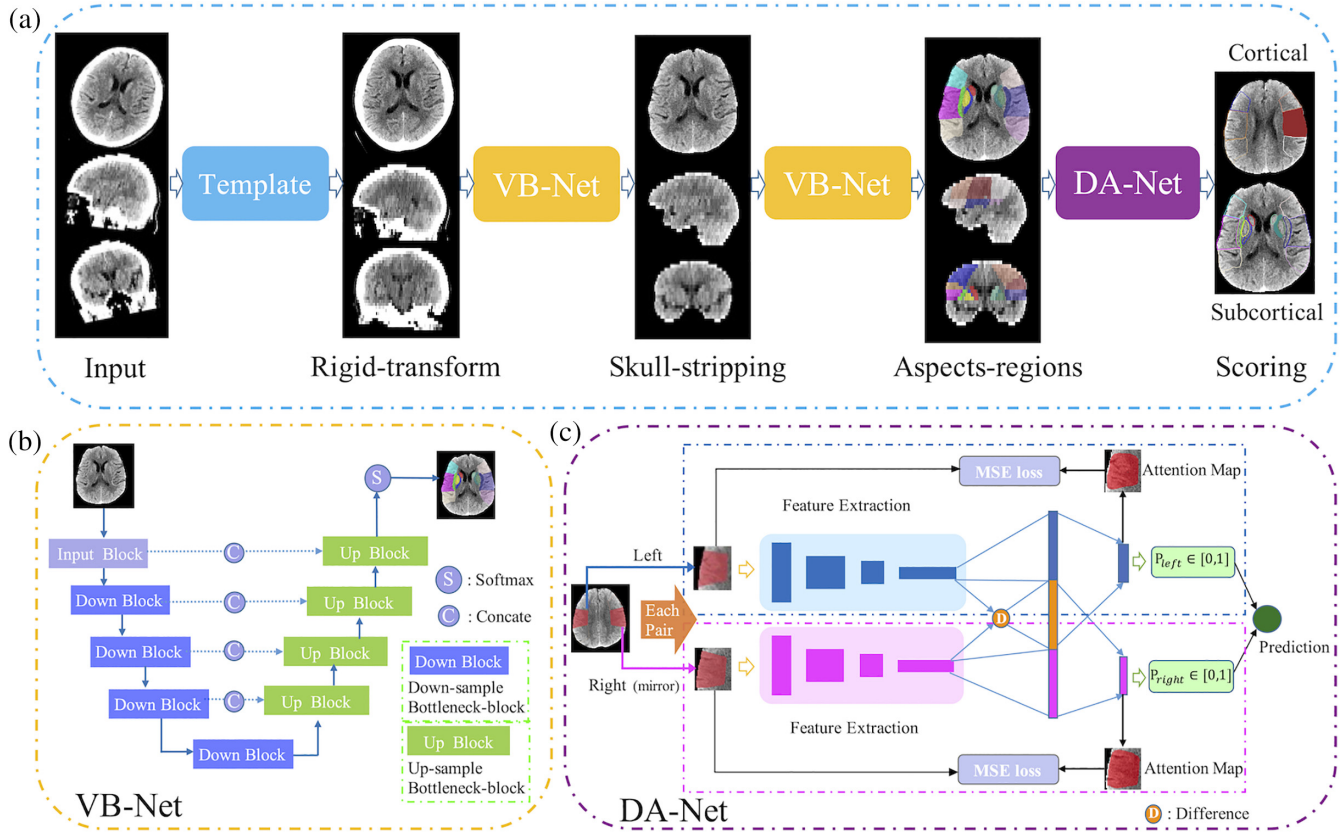

FIGURE 1 Framework of the proposed method. (a) The pipeline where operations include aligning, skull-stripping, region segmentation, and ASPECTS scoring. (b and c) The details about VB-Net and DA-Net, respectively

metric information by comparison, the right-side region is flipped. Then, these two regions go through convolutional layers independently.

We adopt DenseNet (Huang, Liu, Van Der Maaten, & Weinberger, 2017) as the backbone of the feature-extraction module with four dense blocks, and each dense block consists of Batch Norm, ReLu, and convolution layers. The transition module serves as the connection between dense blocks, and the size of feature maps is reduced by the pooling layer. However, due to individual anisotropy of CT brain images, with an intra-layer resolution of $0.5 \times 0.5\,\text{mm}^2$ and 5 mm thickness, we select $1 \times 3 \times 3$ convolution and $3 \times 1 \times 1$ convolution, instead of $3 \times 3 \times 3$ convolution in each dense block to enhance the feature map fidelity. After these layers, the high-level parameter maps of left and right sides are compared and their difference map is computed. Briefly, we calculate the differences between two sides in the feature domain, as shown as operation *D* in Figure 1c. The diagnostic results of left regions are acquired using a one-dimension feature vector composed of left features and *D* which is left map minus right map is used as input of a classifier to predict ischemic stroke. One point is deducted from the corresponding hemisphere if the output of DA-Net is true. The right side employs the same operation except *D* is replaced by a vector of right feature maps minus the left feature maps, and the final score takes the minimum between two hemispheres.

2.5 | Enhanced loss function for prediction regularization

In order to promote the accuracy of automatic scoring to the ground truth, we design a weighted BCE loss to balance normal regions and ischemic regions. And then, we constrain the network attention to focus on the ROI by computing the MSE loss between the attention map and ASPECTS regions.

Specifically, we utilize the binary cross-entropy (BCE) loss (Ruby, Theerthagiri, Jacob, & Vamsidhar, 2020) that is commonly used for classification. BCE loss compares the predicted probabilities of each sample with the actual class output it belongs to (0 or 1) and then penalizes it according to the error between probabilities and the expected value. If the error is larger, then the magnitude of parameter adjustments of the network in the backpropagation training process should be larger, thus rendering the predicted value of the model to be closer to the true value. BCE loss is defined as follows:

$$Log_{loss} = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)),$$

y_i is the *i*th sample label (1 for positive, and 0 for negative), $p(y_i)$ is the *i*th sample's predicted probability of being positive.

Due to the large amount of data containing normal regions instead of lesions, a weighting factor w_i^{class} inversely proportional to the effective number of samples is introduced for BCE loss to balance the ratio between positive and negative samples in the training set:

$$w_i^{class} = \sum_{k=1}^{N} \frac{1_{\{y_k \neq y_i\}}}{N},$$

where $class \in \{0, 1\}$, N is the total number of all samples, and y_i is the ith label.

Since the input patches of the network are 3D and contain ASPECTS regions as well as their surrounding regions, we apply an online attention module to focus on the network's attention on the ASPECTS regions (e.g., M1-M6, I, C, L, IC). This is to learn all the important features for classification efficiently and also generate the corresponding attention map. As shown in Figure 1c, the class activation mapping (CAM; Zhou, Khosla, Lapedriza, Oliva, & Torralba, 2016) is computed in the fully connected part of the network, and then the MSE is calculated between the ASPECTS region and the attention map in each classification branch, which is used to constrain the network attention.

3 | EXPERIMENT

3.1 | Data

In this study, we worked on two datasets, (1) main dataset, and (2) independent testing set. In the *main dataset*, there are paired NCCT and CTP images from 870 subjects within 8 hr from symptom onset, among which 694 subjects were randomly selected as the training set and the remaining 176 patients were used as the testing set. These data were collected in Huaxi Hospital of Sichuan University using the CT scanner of SOMATOM Definition Flash with 70 kV from Siemens. We applied post-processing software in the workstation to acquire the CTP parameter maps, including CBF. In CBF map, the ischemic core was obtained with the threshold of relative CBF (rCBF) <30%, as this region has relatively 70% or more decreased cerebral blood flow when compared to its opposite side (Mokin et al., 2017). Then, we computed the volume ratio of the core occlusion region to each ASPECTS region, and the region was considered to belong to the ischemic region if the ratio exceeded 10%. This was taken as the ground truth for the main dataset. As shown in Figure 2, the core occlusion mask could be obtained from CBF, and then, after comparing the volume ratio, it is used for defining the ground-truth affected ASPECTS region (Jia, Wu, Wang, & Shen, 2010).

Besides, we also collected an *independent testing set*, which included 207 patients with anterior circulation large vessel occlusion treated with thrombectomy. This data was acquired between January 2018 and May 2020 from Hangzhou First People's Hospital of Zhejiang University. All patients had baseline CT. These data were acquired using the CT scanner of SOMATOM Definition Flash with 120 kV from Siemens. For each subject, an experienced radiologist (J.X.) evaluated the corresponding ASPECTS score, which is used as the ground truth for this dataset. This study was approved by the Institutional Review Boards of participating institutes.

3.2 | Model evaluation metrics

We compared the experimental results of the proposed method with the popular Radiomics-based method. The radiomics-based method is

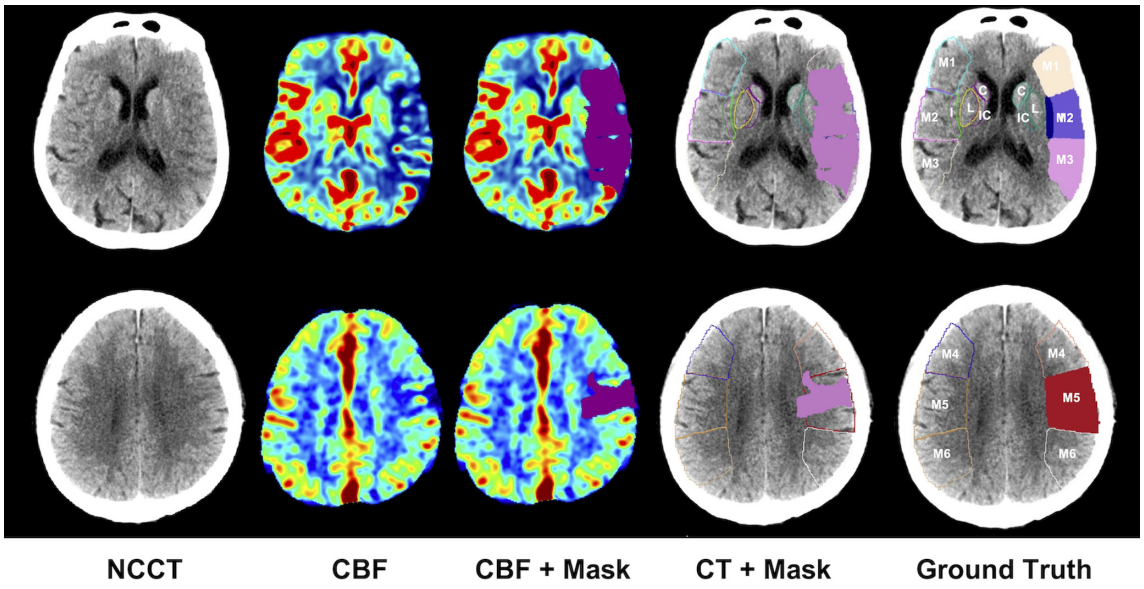

FIGURE 2 Procedure to extract the ischemic cores from CBF and determine the affected ASPECTS regions in CT. The rows are for two image slices of a representative patient. The first column shows the input NCCT image slices. The second column shows the related CBF images. The third column shows the ischemic core defined on CBF with the rule of rCBF <30%. The fourth column matches the core region on NCCT. In the last column, an ASPECTS region is defined as affected if the ratio of the core volume in this region over the region volume

widely used in many clinical studies, and a similar method was used in Qiu's work for ASPECTS scoring (Kuang et al., 2019). Briefly, the comparison method first extracted radiomics features using pyradiomics (van Griethuysen et al., 2017) for each region, and then used these features to train a model for each region with random forest to determine whether this region is an affected ischemic region.

We present boxplots, Bland–Altman plots, and histograms to reflect the variability between automated ASPECTS and ground truth (CTP or expert labeling). Boxplots with scatter points represent the distribution of the generated scores over the ground truth, where the closer the distribution to the line $y = x$, the better match both scores. Bland–Altman plots show the agreement between a total automated score and the ground truth. Its horizontal axis indicates the difference between two methods, and the blue lines refer to an interval of 1.96 standard deviations (SD). The more data that fall within the 95% confidence interval (CI) (equal to [−1.96SD, 1.96SD]) and the closer mean value is to 0, the better agreement between the automatic scores and the ground truth. Histograms illustrate the frequency of variance between automated scores and the ground truth, where the higher the frequency around 0, the smaller the difference between the two methods. Besides, we evaluated the model performance from the metrics commonly used in classification tasks, for example, sensitivity, specificity, accuracy, and area under the curve (AUC). We quantified the model performance separately for cortical and subcortical regions. There is experimental evidence that ASPECTS > 6 versus ≤ 6 is a threshold to distinguish whether a large ischemic stroke occurs, which we analyzed using dichotomized ASPECTS.

3.3 | Results

3.3.1 | Evaluation of ASPECTS region segmentation

We trained a specialized VB-Net with 626 labeled cases and tested it with 86 cases. These data were randomly selected from the main dataset and manually labeled by an experienced rater for 20 ASPECTS

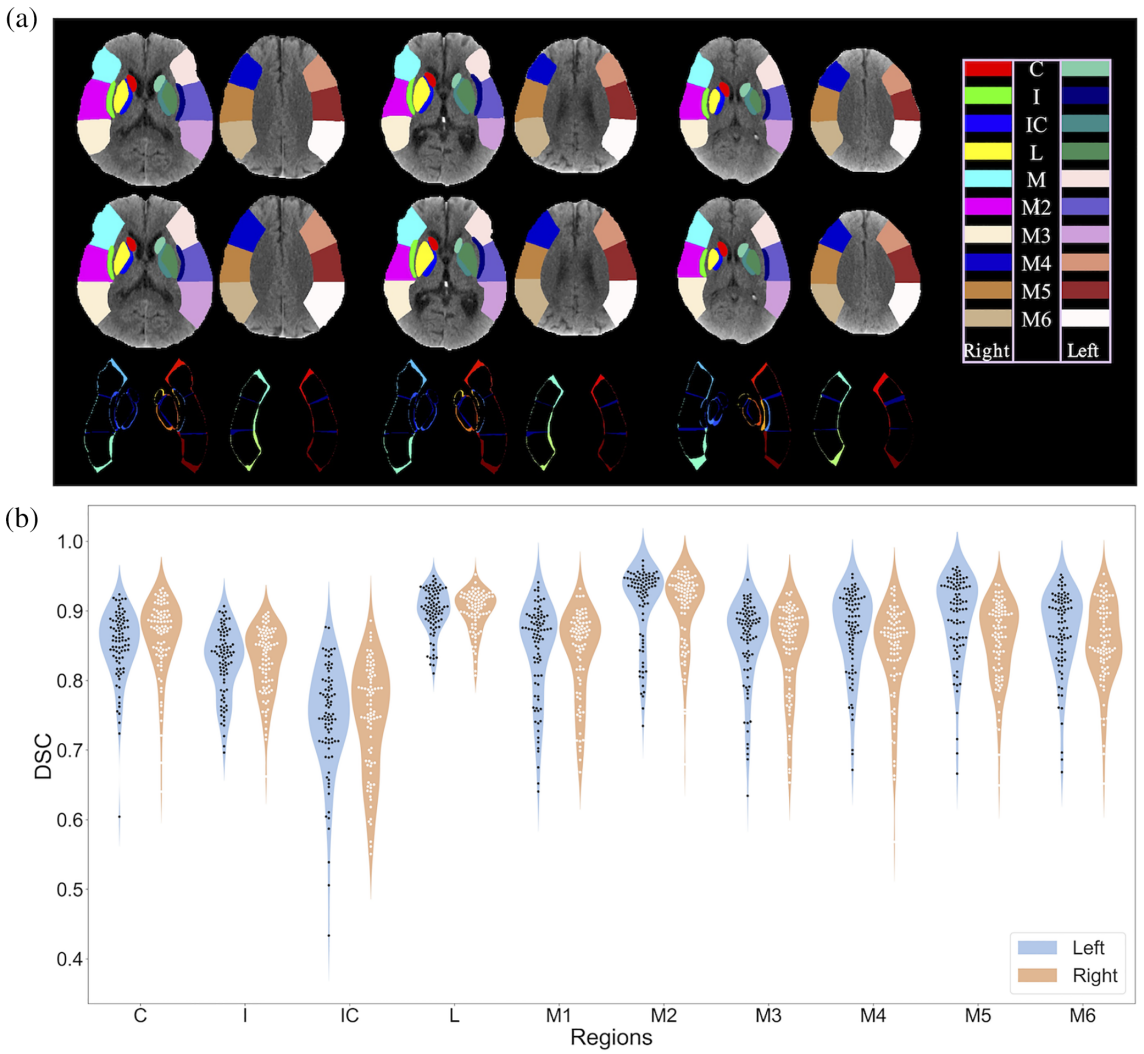

FIGURE 3 (a) Visualization of three subjects from segmentation test from left to right. The first row is the output of network, the second is the ground truth, and the third is difference between output and corresponding labels. (b) The Dice results for segmenting 10 ASPECTS regions in hemispheres of 86 testing images. C, caudate; I, insula; IC, internal capsule; L, lentiform; M, MCA. Each color block represents a scoring region

regions on NCCT images. Figure 3a shows the results of our VB-Net for ASPECTS regions segmentation. As can be seen from the third row, the segmentation result of our network output is basically the same as the labels, except for the edges, where we add post-processing to smooth the boundaries in order to be more realistic. Since the input of our scoring network is a patch containing the corresponding ASPECTS partition, it does *not only* have a strict ASPECTS partition *but also* embraces part of the surrounding area.

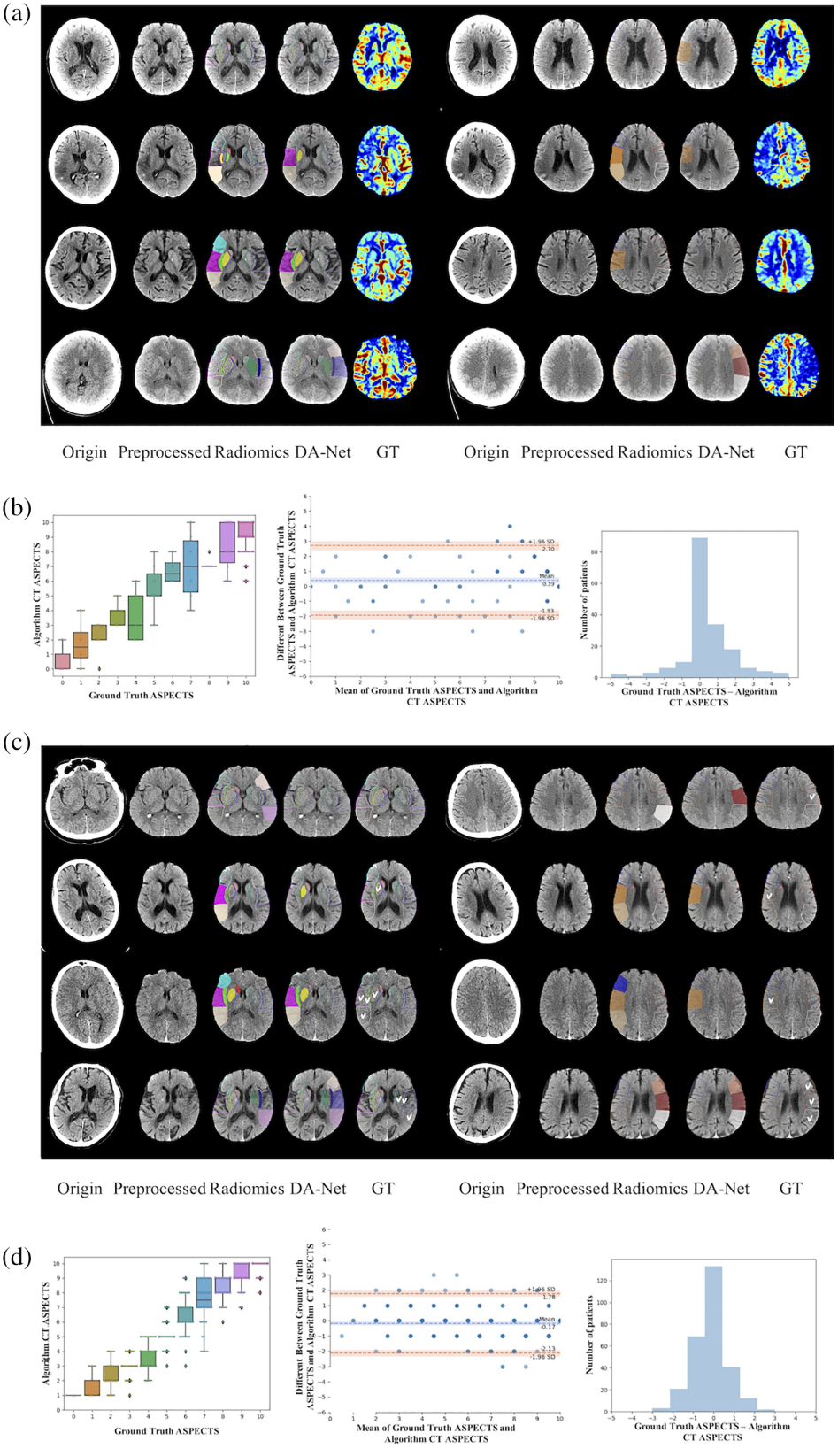

FIGURE 4 Results of two datasets in line (a)–(d). (a and c) The visualization of four subjects from the main dataset and the independent dataset. Corresponding original CT, preprocessed CT, radiomics result, DA-Net result (by our proposed method), and ground truth (GT) of (a) from CBF, of (c) from labeling of radiologist (J.X). For radiomics and DA-Net results, those solid parts represented the automatically detected ischemic regions. Radiomics results without side information referred to the ground-truth ischemic orientation for generating the visualization results. (b and d) Statistical charts using the testing set in the main dataset and independent dataset. In the left of (b) and (d) are box plots between ground-truth and automatically-estimated CT ASPECTS. In the middle of (b) and (d) are Bland–Altman plots. In the right of (b) and (d) are histograms of difference between ground-truth and automatically-estimated CT ASPECTS

From this point of perspective, our segmentation results achieve the expected goal. Quantitative results in Figure 3b show the Dice of 20 regions on the testing set and the average Dice is 0.852.

3.3.2 | Evaluation of proposed method on the main dataset with 176 cases

In Figure 4a, we show the visualization results of four subjects from the main dataset. We compared the ASPECTS score results of the radiomics method and our proposed method concerning the ground truth. Regarding the accuracy of scores, the results of our proposed method, that is, the solid regions were more consistent with the dark blue areas of ground truth (GT), while the radiomics method is slightly less accurate, and the information on the location of the lesion (left or right brain) cannot be given automatically. Considering the old lesion, the radiomics method and our proposed method both identified them successfully. However, those inconspicuous lesions that are difficult to observe in NCCT confused the radiomics method but were mostly recognized successively by our proposed method.

Figure 4b shows the results of the main dataset. The boxplot also illustrates that the two methods have a high agreement, that is, boxes are approximately distributed on the diagonal, indicating that the number of cases falling in the same fraction was remarkably close. Darker points in the Bland–Altman plot indicated that the amount of data aggregated at this point is larger, and the difference between the two scoring methods concentrated around 0. This is considered as a small difference in the final scoring results, and contains almost all the data at a confidence level of 95%, indicating that the scoring results of our algorithm had an excellent concordance with the ground truth. From the histogram, we can also observe that the ultimate scores of most patients, that is, the ground-truth scores and the algorithmic automatic scores were consistent, with a smaller portion having a slight difference.

Table 1 for the main dataset part shows the evaluation of the prediction results for each region, given by our scoring model. It can be seen that, when using all regions as a criterion, all metrics (sensitivity, specificity, accuracy, and AUC) of our model outperformed the classifier trained on radiomics features, by achieving 90.6% in accuracy and also superior sensitivity (which is desired by clinicians in clinic). For ASPECTS >6 versus ASPECTS ≤6, an important threshold for the severity of stroke, the scoring results of our method showed an extremely high accuracy of 92.4% and sensitivity of 93.7%, well above the performance of the random forest with radiomics features. Besides, the performance of our method was superior in subcortex than in cortex.

3.3.3 | Evaluation of our proposed method on the independent dataset with 207 cases

From Figure 4c, our proposed method shows coherence with the expert in cortical and subcortical regions for ASPECTS scores. But the radiomics method tends to more easily include old lesions.

In the left of Figure 4d, the boxplot reports the distribution of the automatically-estimated NCCT ASPECTS at each individual ASPECTS, and the intra-class correlation coefficient (ICC) between scores by our method and experts was 0.924 based on all cases. In the middle of Figure 4d, Bland–Altman agreement plot is shown for total ASPECTS between results by our proposed method and experts. The mean difference in total ASPECTS between scores of our proposed method and experts was minimal (i.e., −0.17; 95% CI, −2.13 to 1.78). And the Spearman correlation between our method and expert reading is 0.9195 (p-value <.05), between radiomics and expert reading is 0.2443 (p-value <.05). A histogram is provided in the right of Figure 4d showing the scoring differences between algorithm and ground truth.

From Table 1 for the independent dataset part, it can be observed that our proposed method demonstrates better robustness in all regions compared to the main dataset. For the regions with ASPECTS >6, the sensitivity (95.5%), specificity (86.6%), accuracy (91.3%) and AUC (0.911) all reach a high level. To demonstrate the credibility of our model, we tested the difference between scores of the proposed method and ground truth through the paired T-test. And the results show that the difference is no significant (p-value is .28). However, the radiomics method seems not suitable for the independent dataset, with the sensitivity, specificity, accuracy, and AUC all being inferior to results obtained in the main dataset. The p-value of paired T-test between radiomics scores and radiologist's scores reached significance ($p < .05$).

3.3.4 | ASPECTS-based clinical index analysis on the independent dataset with 207 cases

Table 2 shows demographic and clinical information of 207 patients with acute ischemic stroke undergoing thrombectomy. The median age is 72 years (interquartile range [IQR]: 63–81), including 81 females (39.1%). The most common risk factors include diabetes, hyperlipidemia, coronary heart disease, atrial fibrillation, hypertension, tobacco, and stroke, corresponding to 16.9, 23.7, 15, 60.9, 64.3, 27.1, and 10.6% of 207 subjects, respectively. A total of 23.7% of patients received intravenous antiplatelet therapy. The median of 90DmRS is 3, with range of 1–5. The median CTP core volume is 22, with range of 9.2–44 among 144 patients, which is defined as the area with rCBF <30% and acquired by MIStar software (Apollo Medical Imaging Technology, Melbourne, Australia). At the same time, we calculated the significant difference between ASPECTS score >6 and ASPECTS score ≤6 for each index and found the p-values of CTP core volume and preoperative NIHSS were statistically significant.

We also employed the Spearman correlation to examine the relationship between each variable (with ground truth, our proposed method, and radiomics) and the critical clinical indexes. Table 3 illustrates that ASPECTS scores from all methods have significant relevance with 90DmRS, and a higher ASPECTS score is associated with lower 90DmRS, which is consistent with previous clinical research. The correlation of CTP core volume and ASPECTS also show that smaller infraction area has higher ASPECTS score. In addition, for the

TABLE 1 The performance metrics of our method and the comparison method in each ASPECTS region, cortex, subcortex, and region with ASPECTS >6 on the testing set of the main dataset and independent dataset

	Ours				Radiomics features			
Region	**Sensitivity**	**Specificity**	**Accuracy**	**AUC**	**Sensitivity**	**Specificity**	**Accuracy**	**AUC**
Main dataset								
C	0.400	0.976	0.959	0.688	0.800	0.940	0.936	0.870
I	0.849	0.928	0.913	0.888	0.516	0.743	0.702	0.630
IC	0.636	0.963	0.942	0.800	0.083	0.887	0.830	0.485
L	1.000	0.885	0.901	0.943	0.696	0.730	0.725	0.713
M1	0.731	0.966	0.930	0.848	0.231	0.883	0.784	0.557
M2	0.909	0.899	0.901	0.904	0.303	0.820	0.719	0.561
M3	0.682	0.920	0.890	0.801	0.360	0.904	0.825	0.632
M4	0.708	0.905	0.878	0.807	0.080	0.925	0.801	0.502
M5	0.849	0.871	0.866	0.860	0.294	0.745	0.655	0.519
M6	0.773	0.893	0.878	0.833	0.522	0.824	0.784	0.673
All	0.807	0.922	0.906	0.864	0.389	0.840	0.776	0.614
Cortical	0.788	0.909	0.891	0.849	0.298	0.850	0.761	0.574
Subcortical	0.836	0.940	0.930	0.888	0.524	0.825	0.798	0.674
>6	**0.937**	**0.862**	**0.924**	**0.900**	0.851	0.433	0.778	0.642
Independent dataset								
C	0.241	0.996	0.923	0.619	0.621	0.865	0.840	0.743
I	0.843	0.861	0.853	0.852	0.481	0.826	0.667	0.654
IC	0.750	0.946	0.933	0.848	0.250	0.795	0.757	0.522
L	0.841	0.942	0.893	0.892	0.181	0.944	0.563	0.563
M1	0.851	0.863	0.860	0.857	0.907	0.297	0.479	0.602
M2	0.826	0.887	0.856	0.856	0.554	0.529	0.542	0.541
M3	0.636	0.931	0.823	0.784	0.870	0.294	0.510	0.582
M4	0.803	0.860	0.846	0.831	0.057	0.940	0.726	0.499
M5	0.846	0.876	0.860	0.861	0.671	0.354	0.531	0.513
M6	0.600	0.910	0.796	0.755	1.000	0.011	0.385	0.506
All	0.772	0.910	0.864	0.841	0.559	0.586	0.600	0.572
Cortical	0.765	0.887	0.840	0.826	0.677	0.404	0.529	0.540
Subcortical	0.784	0.945	0.901	0.864	0.383	0.858	0.707	0.620
>6	**0.955**	**0.866**	**0.913**	**0.911**	0.272	0.816	0.538	0.544

Note: Ischemic regions are positive samples and normal regions are negative samples.
Bolded values indicate higher performance in our method compared to that of radiomics features in the binary results of ACPECTS>6.

CTP core volume, the ASPECTS obtained via our proposed method is more reflective of the core infarct area ($r = -0.6137$, $p < .001$). The higher ASPECTS score, the more probability of intravenous maintenance of antiplatelet drugs is used. However, the other indexes are not significantly related to ASPECTS score.

3.3.5 | Software interface with an example case

Our proposed method has been integrated into the software to provide ASPECTS scores for NCCT input images. It could provide radiologists with auxiliary quantification information, thus greatly improving the efficiency and performance in clinical workflow.

As shown in Figure 5, the first row shows an NCCT image of a subject, parameter maps of CTP, and maximum intensity projection (MIP) view of vessels. In the CBF image, the red circle region shows low-perfusion M5 and M6, compared with the opposite side. In the MIP, brain vessel of MCA shows that the distal end of M2 is not developed. The second row shows the interface of our software. Two key slices, that is, basal ganglia level and lateral ventricle (the body) level, are shown with masks of predicted ischemic regions. In the right column, structured results are provided.

TABLE 2 Demographic of subjects in the independent dataset, and characteristics of NCCT images

Variable	Cohort ($n = 207$)	p-value (ASPECTS >6 vs. ≤6)
Clinical variables		
Age, years; median (IQR)	72 (63–81)	0.7135
Female, no. (%)	81 (39.1)	0.2619
History of diabetes, no. (%)	35 (16.9)	1.0000
History of hyperlipidemia, no. (%)	49 (23.7)	0.6562
History of coronary heart disease, no. (%)	31 (15.0)	1.0000
History of atrial fibrillation, no. (%)	126 (60.9)	0.7135
History of hypertension, no. (%)	133 (64.3)	0.9973
History of tobacco use, no. (%)	56 (27.1)	0.9592
History of stroke, no. (%)	22 (10.6)	1.0000
Onset to CT time, minutes; median (IQR)	270 (210–360)	0.0033*
Occlusion site		0.8259
ICA	59 (28.5)	
M1	102 (49.3)	
M2	19 (9.2)	
ACA	3 (1.4)	
Tandem	24 (11.6)	
Prognosis related		
CTP core volume (CBF < 30%; ml), median (IQR);	22 (9.2–44) (subjects:144)	0.0000*
90DmRS; median (IQR)	3 (1–5)	0.0719
Preoperative NIHSS; median (IQR)	17 (13–20)	0.0081*
Postoperative 1D NIHSS changes; median (IQR)	−3 (−8–0)	0.1147
Symptomatic intracranial hemorrhage, no. (%)	22 (10.6)	0.7588
Intravenous infusion of Tirofiban, no. (%)	49 (23.7)	0.1228
NCCT ASPECTS score		
ASPECTS by radiologist, median (IQR)	6 (4–8)	Reference
ASPECTS by DA-net, median (IQR)	6 (4–9)	0.0000*
ASPECTS by Radiomics, median (IQR)	5 (4–6)	0.0395*

Note: The p-value was from Kolmogorov–Smirnov (K-S) test between ASPECTS score >6 and ASPECTS score ≤6 provided by radiologist scoring in each index.
*indicates p < .05.

TABLE 3 Correlation coefficients between ASPECTS and clinical indexes

Metrics	Mean (std)	Expert score (r, p-value)	Proposed method (r, p-value)	Radiomics score (r, p-value)
CTP core volume	28.42 (±30.54)	−0.6086***	−0.6137***	−0.2201**
90mRS	3.14 (±2.04)	−0.2352***	−0.2131**	−0.2160**
Preoperative NIHSS	16.96 (±5.73)	−0.2041**	−0.1613*	−0.1044
Postoperative NIHSS	13.54 (±8.54)	−0.1916**	−0.1470*	−0.1304
IIT	–	0.1899**	0.1915**	0.1960**
sICH	–	−0.1172	−0.1079	−0.019

Note: The symbol r is the correlation coefficient between metrics and scores. The p-value denotes significance of correlation between metrics and scores.
Abbreviations: IIT, Intravenous infusion of Tirofiban; NIHSS, National Institutes of Health Stroke Scale; sICH, symptomatic intracranial hemorrhage.
*p < .05; **p < .01; ***p < .001.

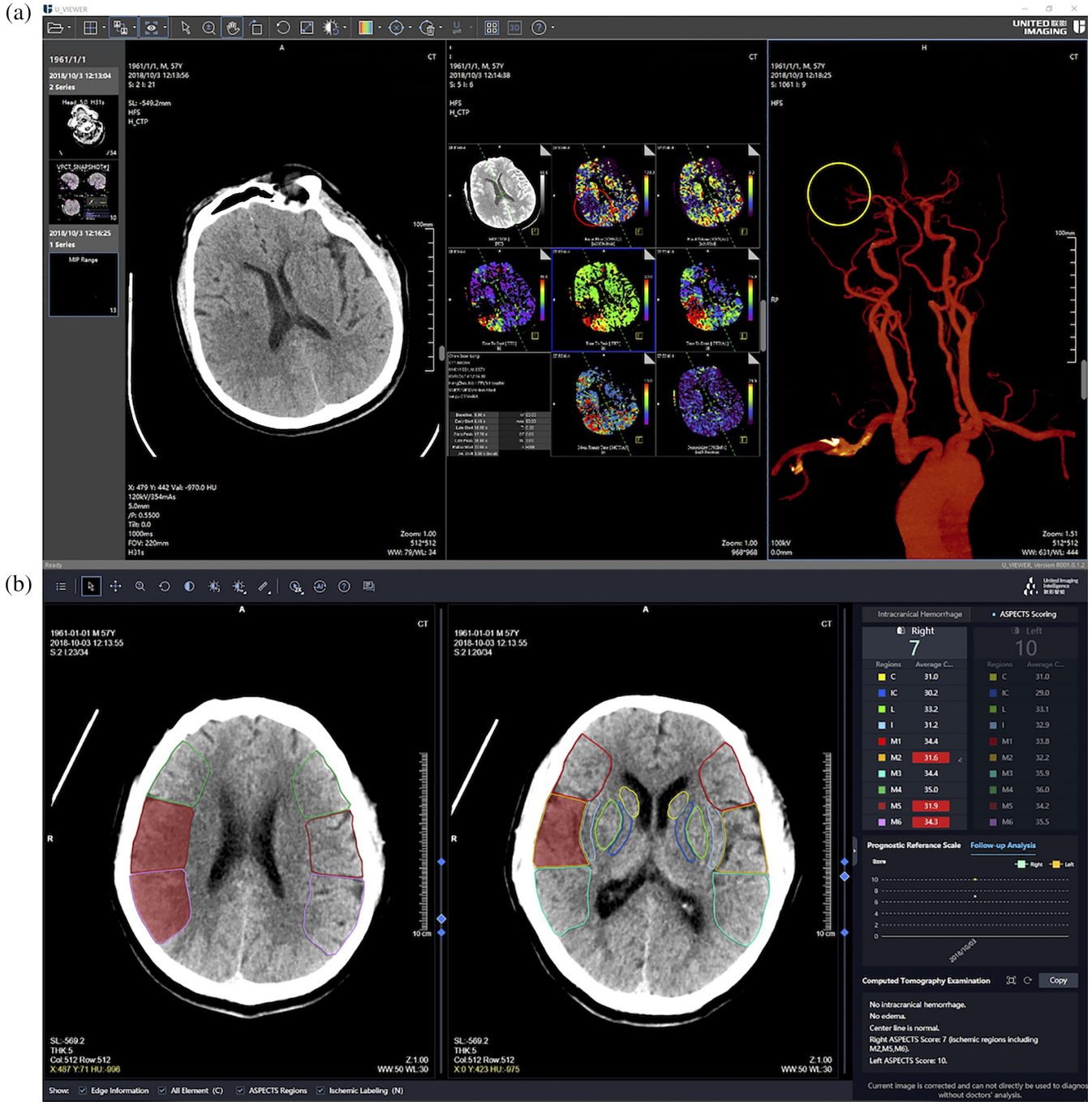

FIGURE 5 Illustration of (a) the results of NCCT, CTP parameter maps, and MIP of vessel for a given subject, and (b) the interface of our proposed stroke software, where the ASPECTS results of NCCT were provided

4 | DISCUSSION

Subjective scoring of CT ASPECTS from clinicians may suffer from large inconsistency, and so machine learning approaches are promising to provide a standardized and rapid evaluation. In this article, we proposed an automatic ASPECTS scoring framework for processing CT images and utilizing the asymmetry nature of ischemic regions into the network design, by simulating the process of radiologists' reading. The experimental results show that, by using CTP lesion as a gold standard, our proposed method achieved high agreement than the conventional Radiomics-based method. Also, our results are promising in the independent dataset, by using manual labels by experts as ground truth. On the other hand, although the specificity and accuracy of caudate region are high, the sensitivity is relatively low, mostly due to a small number of positive samples in caudate region. Including more caudate ischemia samples may improve the performance. Also, we observe that the performance of our model in subcortical region is superior to that in cortical region, which is consistent with Kuang et al.'s findings (Kuang et al., 2020).

In this study, we selected ASPECTS = 6 as a threshold, which is commonly used to assist clinical decisions regarding whether a patient needs treatment of thrombolysis. Although studies have suggested that patients with low ASPECT scores (ASPECTS ≤6) may benefit from thrombectomy (Cagnazzo et al., 2020), the HERMES data show that ASPECTS >6 is associated with good outcome, and there is no significant benefit for patients with the 0-to-5 ASPECTS (Goyal et al., 2016). It can be seen that, when using ASPECTS >6 or all regions as a criterion, our model outperforms the radiomics-based method in all metrics, including sensitivity, specificity and accuracy, and AUC. On the other hand, the threshold of 4 was also picked as a criterion in other studies (Kuang et al., 2019).

We identified a moderate correlation between the baseline automatic ASPECTS scores and the CTP core volumes within 24 hr. Our results are consistent with other studies (Olive-Gadea et al., 2019; Siegler et al., 2020; Voleti et al., 2021). The results of the DAWN and DEFUSE 3 trials extend the treatment window to 24 hr (Albers et al., 2018; Nogueira et al., 2018). The automatic ASPECTS may help select decisions in primary hospitals, that is, whether transferring patients to comprehensive stroke centers. Since CTP and automatic post-processing are not universally available, our study shows that automatic ASPECTS helps to establish treatment decisions when CTP is not available.

This study has several limitations. First, there exist previous studies that employed DWI as ground truth to define the ischemic core, where DWI was acquired within 1 hr after NCCT scan. That short period of image acquisition is quite challenging in many clinics, and thus we alternatively used CTP to define ischemic core and did not compare the results with those obtained by DWI. Besides, in patients with acute stroke or even hyperacute stroke, the image signs in CT are not obvious. Thus, the CT detection results may be false positive. Also, further studies are needed to provide interpretable descriptions for better understanding of the results.

5 | CONCLUSION

In this article, we presented an automated framework to estimate ASPECTS score via NCCT image. We designed a deep asymmetry network (DA-Net) on asymmetrical structure to detect the difference of left–right hemispheres, simulating the process of radiologists reading films of stroke patients. We used different models for cortical and subcortical regions to better capture the image difference. Besides, we evaluated the performance of estimating whether the ASPECTS score is higher than 6, and the results demonstrated that the deep-learning derived results are comparable to the experienced radiologist. Furthermore, we found that the ASPECTS score correlates with the CTP core volume and 90DmRS, which could be useful in assisting future clinic treatment and prognosis.

ACKNOWLEDGMENTS

This work was supported in part by the National Key Research and Development Program of China under grant 2018YFC0116400, the National Science Foundation of China (62131015, 81871337), and Science and Technology Commission of Shanghai Municipality (STCSM) (grant number 21010502600).

CONFLICT OF INTEREST

Zehong Cao, Tianyang Sun, Yichu He, Ying Wei, Feng Shi, and Dinggang Shen are employees of Shanghai United Imaging Intelligence Co., Ltd. The company has no role in designing and performing the surveillances and analyzing and interpreting the data. All other authors report no conflicts of interest relevant to this article.

AUTHOR CONTRIBUTIONS

Dinggang Shen, Feng Shi, and Zhongxiang Ding designed the study and supervised the project. Zehong Cao, Yichu He, and Tianyang Sun developed the methodology and ran the experiments. Zehong Cao and Jiaona Xu wrote the draft. Bin Song, Zhongxiang Ding, and Lizhou Chen collected the data and performed the image evaluations. Guozhong Niu, Yu Zhang, and Qianjin Feng contributed to the interpretation of the data and the revision of the manuscript. All authors have approved the final version of the submission.

DATA AVAILABILITY STATEMENT

The main and independent datasets are part of on-going studies at the Huaxi Hospital of Sichuan University, China, and Hangzhou First People's Hospital of Zhejiang University, China, respectively, and we are bound by the policies of the data providers. The code was released at https://github.com/simonsf/ct-aspects

REFERENCES

Albers, G. W., Marks, M. P., Kemp, S., Christensen, S., Tsai, J. P., Ortega-Gutierrez, S., … Investigators, D. (2018). Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *The New England Journal of Medicine*, *378*(8), 708–718. https://doi.org/10.1056/NEJMoa1713973

Barber, P. A., Demchuk, A. M., Zhang, J., & Buchan, A. M. (2000). Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. *The Lancet*, *355*(9216), 1670–1674. https://doi.org/10.1016/s0140-6736(00)02237-6

Brinjikji, W., Abbasi, M., Arnold, C., Benson, J. C., Braksick, S. A., Campeau, N., … Kallmes, D. F. (2021). E-ASPECTS software improves interobserver agreement and accuracy of interpretation of aspects score. *Interventional Neuroradiology*, *27*, 781–787. https://doi.org/10.1177/15910199211011861

Cagnazzo, F., Derraz, I., Dargazanli, C., Lefevre, P.-H., Gascou, G., Riquelme, C., … Costalat, V. (2020). Mechanical thrombectomy in patients with acute ischemic stroke and ASPECTS ≤6: A meta-analysis. *Journal of Neurointerventional Surgery*, *12*(4), 350–355. https://doi.org/10.1136/neurintsurg-2019-015237

Chen, X., Zhang, H., Zhang, L., Shen, C., Lee, S. W., & Shen, D. (2017). Extraction of dynamic functional connectivity from brain grey matter and white matter for MCI classification. *Human Brain Mapping*, *38*(10), 5019–5034.

Farzin, B., Fahed, R., Guilbert, F., Poppe, A. Y., Daneault, N., Durocher, A. P., … Raymond, J. (2016). Early CT changes in patients admitted for thrombectomy: Intrarater and interrater agreement. *Neurology*, *87*(3), 249–256. https://doi.org/10.1212/WNL.0000000000002860

Goyal, M., Menon, B. K., van Zwam, W. H., Dippel, D. W. J., Mitchell, P. J., Demchuk, A. M., … Jovin, T. G. (2016). Endovascular thrombectomy after large-vessel ischaemic stroke: A meta-analysis of individual patient data from five randomised trials. *Lancet (London, England)*, *387*(10029), 1723–1731. https://doi.org/10.1016/S0140-6736(16)00163-X

Han, M., Yao, G., Zhang, W., Mu, G., Zhan, Y., Zhou, X., & Gao, Y. (2019). Segmentation of CT thoracic organs by multi-resolution VB-nets. *Proceedings of the 2019 Challenge on Segmentation of THoracic Organs at Risk in CT Images, CEUR Workshop Proceedings*, CEUR Workshop Proceedings, 2349. http://ceur-ws.org/Vol-2349/SegTHOR2019_paper_1.pdf

Herzog, N. J., & Magoulas, G. D. (2021). Brain asymmetry detection and machine learning classification for diagnosis of early dementia. *Sensors (Basel)*, *21*(3), 778. https://doi.org/10.3390/s21030778

Hoelter, P., Muehlen, I., Goelitz, P., Beuscher, V., Schwab, S., & Doerfler, A. (2020). Automated ASPECT scoring in acute ischemic stroke: Comparison of three software tools. *Neuroradiology*, *62*(10), 1231–1238.

Hua, R., Huo, Q., Gao, Y., Sui, H., Zhang, B., Sun, Y., … Shi, F. (2020). Segmenting brain tumor using cascaded V-nets in multimodal MR images. *Frontiers in Computational Neuroscience*, *14*, 9. https://doi.org/10.3389/fncom.2020.00009

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). *Densely connected convolutional networks*. Paper presented at the proceedings of the IEEE conference on computer vision and pattern recognition.

Jia, H., Wu, G., Wang, Q., & Shen, D. (2010). ABSORB: Atlas building by self-organized registration and bundling. *NeuroImage*, *51*(3), 1057–1070.

Jones, A., O'Connell, N., David, A. S., & Chalder, T. (2020). Functional stroke symptoms: A narrative review and conceptual model. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *32*(1), 14–23.

Kaesmacher, J., Chaloulos-Iakovidis, P., Panos, L., Mordasini, P., Michel, P., Hajdu, S. D., … Friedrich, B. (2019). Mechanical thrombectomy in ischemic stroke patients with Alberta stroke program early computed tomography score 0–5. *Stroke*, *50*(4), 880–888.

Kalavathi, P., Senthamilselvi, M., & Prasath, V. (2017). Review of computational methods on brain symmetric and asymmetric analysis from neuroimaging techniques. *Technologies*, *5*(2), 16. https://doi.org/10.3390/technologies5020016

Kim, J., Thayabaranathan, T., Donnan, G. A., Howard, G., Howard, V. J., Rothwell, P. M., … Thrift, A. G. (2020). Global stroke statistics 2019. *International Journal of Stroke*, *15*(8), 819–838. https://doi.org/10.1177/1747493020909545

Kuang, H., Najm, M., Chakraborty, D., Maraj, N., Sohn, S. I., Goyal, M., … Qiu, W. (2019). Automated ASPECTS on noncontrast CT scans in patients with acute ischemic stroke using machine learning. *AJNR. American Journal of Neuroradiology*, *40*(1), 33–38. https://doi.org/10.3174/ajnr.A5889

Kuang, H., Qiu, W., Najm, M., Dowlatshahi, D., Mikulik, R., Poppe, A. Y., … Collaborators, I. N. (2020). Validation of an automated ASPECTS method on non-contrast computed tomography scans of acute ischemic stroke patients. *International Journal of Stroke*, *15*(5), 528–534. https://doi.org/10.1177/1747493019895702

Li, A., Li, F., Elahifasaee, F., Liu, M., Zhang, L., & Alzheimer's Disease Neuroimaging, I. (2021). Hippocampal shape and asymmetry analysis by cascaded convolutional neural networks for Alzheimer's disease diagnosis. *Brain Imaging and Behavior*, *15*, 2330–2339. https://doi.org/10.1007/s11682-020-00427-y

Liu, M., Zhang, J., Adeli, E., & Shen, D. (2019). Joint classification and regression via deep multi-task multi-channel learning for Alzheimer's disease diagnosis. *IEEE Transactions on Biomedical Engineering*, *66*(5), 1195–1206. https://doi.org/10.1109/TBME.2018.2869989

Milletari, F., Navab, N., & Ahmadi, S.-A. (2016). *V-net: Fully convolutional neural networks for volumetric medical image segmentation*. Paper presented at the 2016 fourth international conference on 3D vision (3DV).

Mokin, M., Levy, E. I., Saver, J. L., Siddiqui, A. H., Goyal, M., Bonafé, A., … Albers, G. W. (2017). Predictive value of RAPID assessed perfusion thresholds on final infarct volume in SWIFT PRIME (solitaire with the intention for thrombectomy as primary endovascular treatment). *Stroke*, *48*(4), 932–938.

Munsell, B. C., Wee, C.-Y., Keller, S. S., Weber, B., Elger, C., da Silva, L. A. T., … Bonilha, L. (2015). Evaluation of machine learning algorithms for treatment outcome prediction in patients with epilepsy based on structural connectome data. *NeuroImage*, *118*, 219–230. https://doi.org/10.1016/j.neuroimage.2015.06.008

Murphy, S. J., & Werring, D. J. (2020). Stroke: Causes and clinical features. *Medicine (Abingdon)*, *48*(9), 561–566. https://doi.org/10.1016/j.mpmed.2020.06.002

Naganuma, M., Tachibana, A., Fuchigami, T., Akahori, S., Okumura, S., Yi, K., … Yonehara, T. (2021). Alberta stroke program early CT score calculation using the deep learning-based brain hemisphere comparison algorithm. *Journal of Stroke and Cerebrovascular Diseases*, *30*(7), 105791.

Neuberger, U., Nagel, S., Pfaff, J., Ringleb, P. A., Herweh, C., Bendszus, M., … Kickingereder, P. (2020). Impact of slice thickness on clinical utility of automated Alberta Stroke Program Early Computed Tomography Scores. *European Radiology*, *30*(6), 3137–3145.

Neuhaus, A., Seyedsaadat, S. M., Mihal, D., Benson, J., Mark, I., Kallmes, D. F., & Brinjikji, W. (2020). Region-specific agreement in ASPECTS estimation between neuroradiologists and e-ASPECTS software. *Journal of Neurointerventional Surgery*, *12*(7), 720–723. https://doi.org/10.1136/neurintsurg-2019-015442

Nogueira, R. G., Jadhav, A. P., Haussen, D. C., Bonafe, A., Budzik, R. F., Bhuva, P., … Hanel, R. A. (2018). Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *New England Journal of Medicine*, *378*(1), 11–21.

Ojaghihaghighi, S., Vahdati, S. S., Mikaeilpour, A., & Ramouz, A. (2017). Comparison of neurological clinical manifestation in patients with hemorrhagic and ischemic stroke. *World Journal of Emergency Medicine*, *8*(1), 34–38. https://doi.org/10.5847/wjem.j.1920-8642.2017.01.006

Olive-Gadea, M., Martins, N., Boned, S., Carvajal, J., Moreno, M. J., Muchada, M., … Rubiera, M. (2019). Baseline ASPECTS and e-ASPECTS correlation with infarct volume and functional outcome in patients undergoing mechanical thrombectomy. *Journal of Neuroimaging*, *29*(2), 198–202.

Powers, W. J., Rabinstein, A. A., Ackerson, T., Adeoye, O. M., Bambakidis, N. C., Becker, K., … Tirschwell, D. L. (2019). Guidelines for the early Management of Patients with Acute Ischemic Stroke: 2019 update to the 2018 guidelines for the early Management of Acute Ischemic Stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, *50*(12), e344–e418. https://doi.org/10.1161/STR.0000000000000211

Ruby, A. U., Theerthagiri, P., Jacob, I. J., & Vamsidhar, Y. (2020). Binary cross entropy with deep learning technique for image classification. *International Journal of Advanced Trends in Computer Science and Engineering*, *9*(4), 5393–5397. https://doi.org/10.30534/ijatcse/2020/175942020

Rudkin, S., Cerejo, R., Tayal, A., & Goldberg, M. F. (2018). Imaging of acute ischemic stroke. *Emergency Radiology*, *25*(6), 659–672.

Runchey, S., & McGee, S. (2010). Does this patient have a hemorrhagic stroke?: Clinical findings distinguishing hemorrhagic stroke from ischemic stroke. *JAMA*, *303*(22), 2280–2286.

Schroder, J., & Thomalla, G. (2016). A critical review of Alberta stroke program early CT score for evaluation of acute stroke imaging. *Frontiers in Neurology*, *7*, 245. https://doi.org/10.3389/fneur.2016.00245

Shi, F., Xia, L., Shan, F., Song, B., Wu, D., Wei, Y., … Gao, Y. (2021). Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification. *Physics in Medicine & Biology*, *66*(6), 065031.

Shi, F., Yap, P.-T., Gao, W., Lin, W., Gilmore, J. H., & Shen, D. (2012). Altered structural connectivity in neonates at genetic risk for schizophrenia: A combined study using morphological and white matter networks. *NeuroImage*, *62*(3), 1622–1633. https://doi.org/10.1016/j.neuroimage.2012.05.026

Siegler, J. E., Messe, S. R., Sucharew, H., Kasner, S. E., Mehta, T., Arora, N., … Mistry, E. A. (2020). Noncontrast CT versus perfusion-based core estimation in large vessel occlusion: The blood pressure after endovascular stroke therapy study. *Journal of Neuroimaging*, *30*(2), 219–226. https://doi.org/10.1111/jon.12682

van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., … Aerts, H. J. W. L. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, *77*(21), e104–e107. https://doi.org/10.1158/0008-5472.Can-17-0339

Van Seeters, T., Biessels, G. J., Niesten, J. M., van der Schaaf, I. C., Dankbaar, J. W., Horsch, A. D., … Velthuis, B. K. (2013). Reliability of visual assessment of non-contrast CT, CT angiography source images and CT perfusion in patients with suspected ischemic stroke. *PLoS One*, *8*(10), e75615.

Voleti, S., Vidovich, J., Corcoran, B., Zhang, B., Khandwala, V., Mistry, E. A., … Vagal, A. (2021). Correlation of Alberta stroke program early computed tomography score with computed tomography perfusion core in large vessel occlusion in delayed time windows. *Stroke*, *52*(2), 498–504.

Yew, K. S., & Cheng, E. (2009). Acute stroke diagnosis. *American Family Physician*, *80*(1), 33.

Yoo, A. J., Berkhemer, O. A., Fransen, P. S., van den Berg, L. A., Beumer, D., Lingsma, H. F., … van Walderveen, M. A. (2016). Effect of baseline Alberta stroke program early CT score on safety and efficacy of intra-arterial treatment: A subgroup analysis of a randomised phase 3 trial (MR CLEAN). *The Lancet Neurology*, *15*(7), 685–694.

Yoshimoto, T., Inoue, M., Yamagami, H., Fujita, K., Tanaka, K., Ando, D., … Toyoda, K. (2019). Use of diffusion-weighted imaging-Alberta stroke program early computed tomography score (DWI-ASPECTS) and ischemic Core volume to determine the malignant profile in acute stroke. *Journal of the American Heart Association*, *8*(22), e012558. https://doi.org/10.1161/JAHA.119.012558

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. & Torralba, A. (2016). Learning deep features for discriminative localization. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2921–2929) Las Vegas: IEEE Computer Society. http://doi.ieeecomputersociety.org/10.1109/CVPR.2016.319

How to cite this article: Cao, Z., Xu, J., Song, B., Chen, L., Sun, T., He, Y., Wei, Y., Niu, G., Zhang, Y., Feng, Q., Ding, Z., Shi, F., & Shen, D. (2022). Deep learning derived automated ASPECTS on non-contrast CT scans of acute ischemic stroke patients. *Human Brain Mapping*, *43*(10), 3023–3036. https://doi.org/10.1002/hbm.25845