


BMJ Open Deep learning-based carotid plaque vulnerability classification with multicentre contrast-enhanced ultrasound video: a comparative diagnostic study

Yang Guang,¹ Wen He ,¹ Bin Ning,¹ Hongxia Zhang,¹ Chen Yin,² Mingchang Zhao,² Fang Nie,³ Pintong Huang,⁴ Rui-Fang Zhang,⁵ Qiang Yong,⁶ Yanli Guo,⁷ Jianjun Yuan,⁸ Yicheng Wang,⁹ Lijun Yuan,¹⁰ Litao Ruan,¹¹ Tengfei Yu,¹ Haiman Song,¹ Yukang Zhang¹

To cite: Guang Y, He W, Ning B, *et al.* Deep learning-based carotid plaque vulnerability classification with multicentre contrast-enhanced ultrasound video: a comparative diagnostic study. *BMJ Open* 2021;**11**:e047528. doi:10.1136/bmjopen-2020-047528

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-047528>).

Received 02 December 2020
Accepted 03 August 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Wen He; hewen@bjtth.org

ABSTRACT

Objectives The aim of this study was to evaluate the performance of deep learning-based detection and classification of carotid plaque (DL-DCCP) in carotid plaque contrast-enhanced ultrasound (CEUS).

Methods and analysis A prospective multicentre study was conducted to assess vulnerability in patients with carotid plaque. Data from 547 potentially eligible patients were prospectively enrolled from 10 hospitals, and 205 patients with CEUS video were finally enrolled for analysis. The area under the receiver operating characteristic curve (AUC) was used to evaluate the effectiveness of DL-DCCP and two experienced radiologists who manually examined the CEUS video (RA-CEUS) in diagnosing and classifying carotid plaque vulnerability. To evaluate the influence of dynamic video input on the performance of the algorithm, a state-of-the-art deep convolutional neural network (CNN) model for static images (Xception) was compared with DL-DCCP for both training and holdout validation cohorts.

Results The AUCs of DL-DCCP were significantly better than those of the experienced radiologists for both the training and holdout validation cohorts (training, DL-DCCP vs RA-CEUS, AUC: 0.85 vs 0.69, $p < 0.01$; holdout validation, DL-DCCP vs RA-CEUS, AUC: 0.87 vs 0.66, $p < 0.01$), that is, also better than the best deep CNN model Xception we had performed, for both the training and holdout validation cohorts (training, DL-DCCP vs Xception, AUC: 0.85 vs 0.82, $p < 0.01$; holdout validation, DL-DCCP vs Xception, AUC: 0.87 vs 0.77, $p < 0.01$).

Conclusion DL-DCCP shows better overall performance in assessing the vulnerability of carotid atherosclerotic plaques than RA-CEUS. Moreover, with a more powerful network structure and better utilisation of video information, DL-DCCP provided greater diagnostic accuracy than a state-of-the-art static CNN model.

Trial registration number ChiCTR1900021846,

INTRODUCTION

Carotid atherosclerotic disease is currently a major health problem, accounting for

Strengths and limitations of this study

- This study used multicentre carotid plaque contrast-enhanced ultrasound (CEUS) to ensure its generalisation.
- Deep learning-based detection and classification of carotid plaque (DL-DCCP) shows significantly better overall performance than radiologists who manually examined the CEUS (CEUS).
- DL-DCCP used CEUS video to increase diagnostic accuracy and better generalisability compared with state-of-the-art deep convolutional neural network models for static image analysis.
- DL-DCCP can better use the interframe information provided by dynamic CEUS video to generalise from training cohort to unseen validation cohort.
- The study was currently limited to the population size, a larger sample size with more patients is required to better train the deep-learning model.

approximately 20% of all cases of cerebral ischaemia.¹ Such cases of cerebral ischaemia are likely due to carotid artery stenosis, resulting in poor central nervous system perfusion or plaque surface thrombosis and artery-artery embolisation. In the past few years, there has been increasing evidence that plaque characteristics are associated with neurological symptoms, leading to the concept of 'vulnerable plaques'.² Previous evidence has demonstrated that the vulnerability of atherosclerotic plaques may be related to their neovascularisation.³ Neovascularisation in plaques is an independent predictor of vulnerable plaque rupture. In addition, the number of neovessels in plaques was found to be closely related to clinical manifestations. A higher neovascularisation density was

associated with an increased risk of typical clinical presentation.^{4–6} Therefore, the detection of neovascularisation is of great significance in predicting plaque vulnerability, assessing stroke risk and guiding clinical treatment. With the rapid development of contrast-enhanced ultrasound (CEUS), we now have the ability to assess plaque vulnerability by observing the neovascularisation within plaques.^{3 7–9} CEUS has many advantages over conventional ultrasound, but the conventional strategy for using CEUS is insufficient for accurately assessing plaque vulnerability.

In recent years, an emerging technology named radiomics has provided the automated quantification of large amounts of imaging features from medical images, which has the potential to aid in uncovering disease characteristics that fail to be appreciated by the naked eye.¹⁰ Most medical images are static images, such as two-dimensional (2D) pictures and three-dimensional (3D) volumes, which lack dynamic information. Unlike conventional medical images, CEUS video data contain important dynamic information. However, traditional neural network methods only take separated frames as input for classification and ignore the continuity between frames.

We speculate that a distinctive radiomics technique might be capable of using more valuable information from CEUS. Our study will summarise the contribution of established imaging modalities to plaque vulnerability assessment and investigate the diagnostic performance of a new network structure in carotid plaque evaluation with CEUS video, thus potentially improving the diagnostic performance in evaluating the vulnerability of carotid plaques. The features extracted by the network may include not only neovascularisation features but also sonographic features of vulnerable plaques (surface, thrombus, heterogeneity, etc) and other dynamic features that are difficult to capture by eye. By combining all these features, the deep learning-based detection and classification of carotid plaque (DL-DCCP) may discover information that sonographers cannot recognise on sight and obtain a more accurate diagnosis than sonographers. In this study, we aimed to evaluate the performance of DL-DCCP in assessing neovascularisation in carotid plaque CEUS.

METHODS

Design and overview

This was a multicentre, prospective study. A new diagnostic approach named DL-DCCP was used to assess the vulnerability of carotid plaques. Histopathological or high-resolution MRI findings were used as the reference standard, and DL-DCCP was compared with the conventional evaluation of CEUS findings. MR scans were performed for carotid arteries using a 3.0 T MR scanner (Siemens Medical Systems, Erlangen, Germany; GE Healthcare, Waukesha, WI; Philips Medical System, Best, Netherlands) with an eight-channel carotid coil. The imaging protocol included three multicontrast 3D

blackblood sequences with T1, T2 and heavy T1 weightings. MR images were acquired in the coronal plane and covered the whole carotid artery, including the common carotid artery, bifurcation, internal carotid artery and external carotid artery. All MR images were evaluated on axially reformatted images.

From September 2017 to September 2018, patients with carotid plaques who provided informed consent to participate in this study were enrolled from 10 Chinese hospitals in different regions. The overall flow chart of this study is shown in [figure 1](#).

Patient enrolments

Operative indications included either symptomatic or asymptomatic carotid plaques. Symptomatic carotid plaque was defined by the presence of symptoms associated with previous ischaemic events on the ipsilateral side within the preceding 6 months. Patients were classified as asymptomatic if they had not experienced any stroke or transitory ischaemic attack in the previous 6 months. The inclusion criteria were the presence of carotid plaques and a maximum plaque thickness ≥ 2.0 mm on grey scale ultrasound. The carotid plaques of all patients were imaged by high-resolution MRI or removed by carotid endarterectomy within 2 weeks of CEUS. The exclusion criteria were non-atherosclerotic disease, shadowing of more than 50% due to plaque calcification on grey scale ultrasound, previous endarterectomy or intravascular stent at the site of the index carotid artery, contraindications to CEUS, such as unstable angina, acute cardiac failure, acute endocarditis, known right-to-left shunts and known allergy for microbubble contrast agents, and physical or mental inability to participate in the study.

CEUS video acquisition

CEUS was performed using LOGIQ E9 (GE Medical Systems, Wauwatosa, Wisconsin, USA), Philips iU22 (Philips US, Bothell, Washington, USA) and Canon Aplio 500 devices (Canon Medical System Corporation, Tokyo, Japan) with a linear probe. Using grey scale images, we observed the intraplaque echogenicity (hypoechoic, isoechoic or mixed echoic). CEUS was performed on plaques whose intima–media thickness was greater than 2.0 mm to observe the intraplaque neovessels (IPNVs). Standard machine settings were used with a mechanical index of 0.05–0.08, and the focus was placed deeper than the plaque plane to avoid destroying the microbubbles. SonoVue (BRI, Bracco Imaging, Milan, Italy) was used as the contrast agent, supplied as a sterile, lyophilised powder contained in a septum-sealed vial. A suspension of the contrast agent was obtained by adding 5 mL of physiological saline to the powder (25 mg). A contrast bolus of 2.4 mL was injected manually at a flow rate of approximately 1 mL/s, followed by a 5 mL saline solution flush through a 20-gauge intravenous cannula placed in an antecubital vein by a nurse. Using a time-triggered acquisition over a period of 180 s, the imaging data obtained

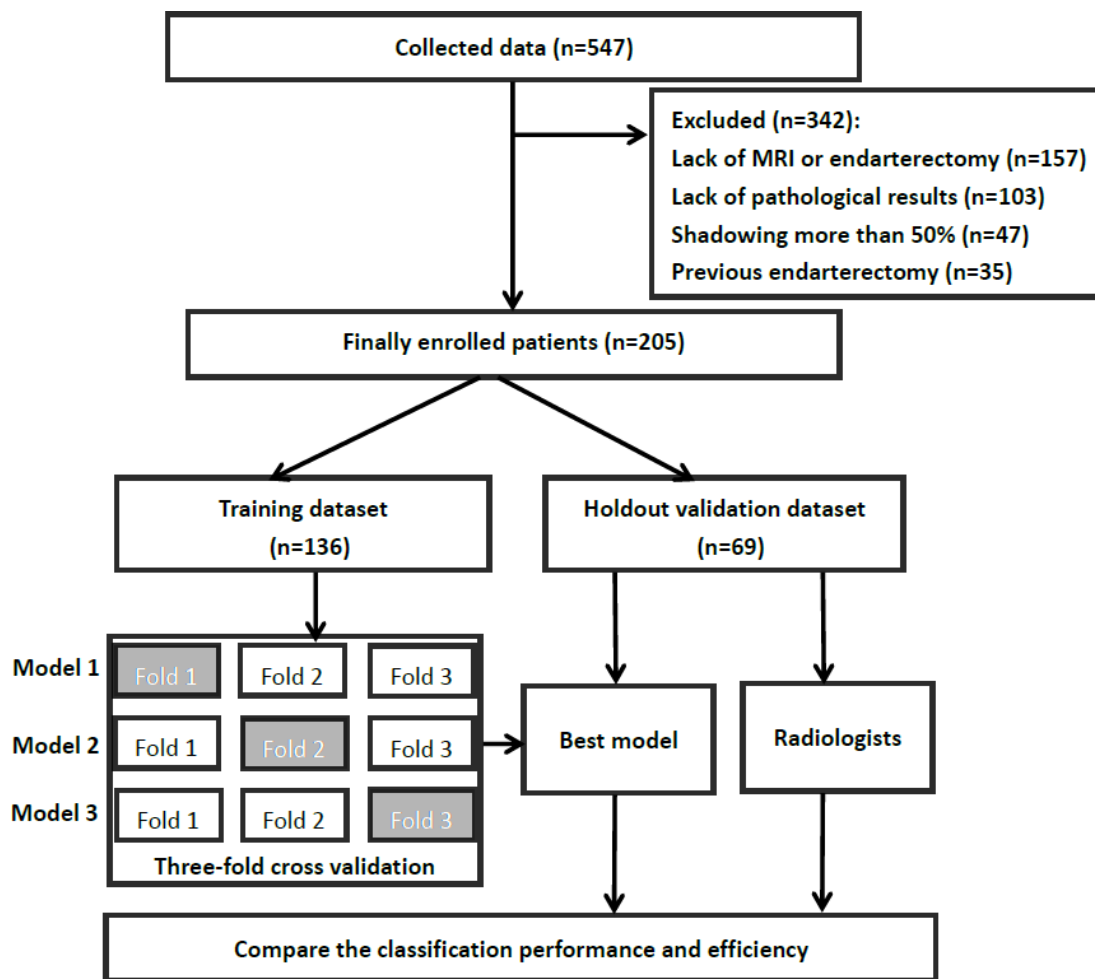


Figure 1 The overall flow chart of this study.

from CEUS were continuously stored on an integrated magnetic optical drive in the ultrasound machine.

Deep learning-based detection and classification of carotid plaques

Because of the dynamic characteristics of CEUS video, it is difficult to achieve good results by using traditional convolutional neural networks (CNNs) for static image analysis. The method used in this study (DL-DCCP) is a specially designed CNN combined with video information aggregation that can extract rich dynamic features from CEUS video. As shown in [figure 2](#), the first step (A) ([figure 2A](#)) is to preprocess the video to remove measurements or other annotated markers from the image. In step (B) ([figure 2B](#)), we blend the CEUS image and B-mode image into a dual channel image and then determine the size and location of the region of interest (ROI) containing the plaque through a CNN in step (C) ([figure 2C](#)). In the next step (D) ([figure 2D](#)), the ROI information is used to crop every frame in the CEUS video, and then we employ the Xception¹¹ network to extract the features of each cropped frame. To select better features and reduce the number of subsequent calculations, in step E ([figure 2E](#)), we compress the feature dimensions of each frame from 2048 to 1024. Thus far, we have only dealt

with each frame independently. To make full use of the rich dynamic information contained in the CEUS video, in step F ([figure 2F](#)), we further aggregate all the frame features of the video. Finally, these fused features are fed into a classifier in step G ([figure 2G](#)) to obtain the type of plaque (stable or vulnerable).

There are often some measurement marks or other annotation information left by ultrasound physicians in ultrasound images, which can make subsequent automatic image processing and analysis difficult. To solve this problem, we designed a CNN as a preprocessing (step A) to remove these markers. As shown in [figure 3](#), from the raw image (with unwanted markers as indicated by yellow arrow) and the mask image (generated manually), an initial rough predicted image (coarse result) is generated first by using a CNN called Coarse Network. Then, we use a second CNN called Refinement Network to obtain a refined and cleaned image (inpainting result). Dilated convolutions are used to use the contextual information around the masked area and reconstruct the missing pixels. With this method, we can obtain clean and usable images for subsequent processing.

ROI detection and blending of CEUS image and B-mode image are important steps (step B and step C)

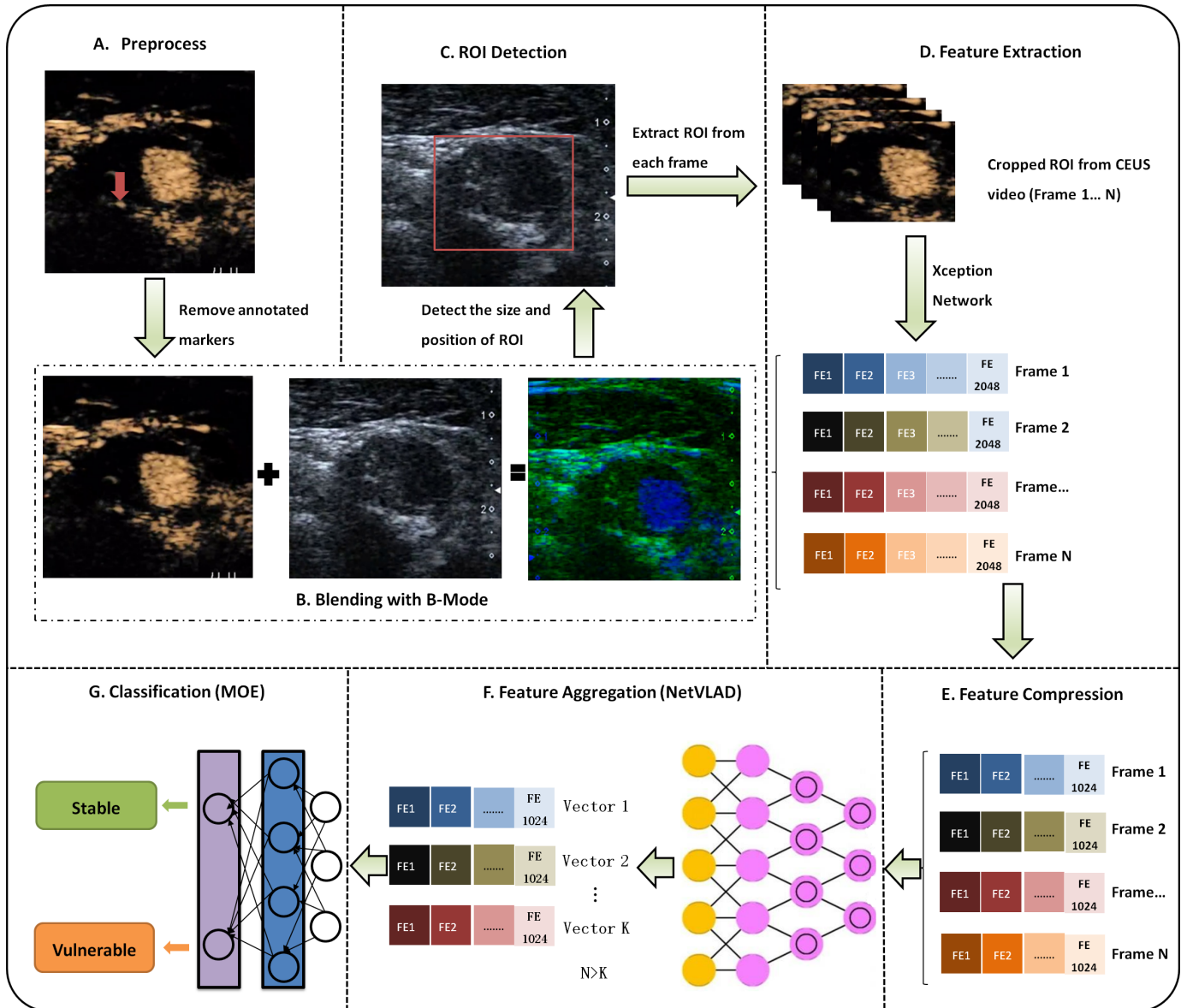


Figure 2 Overall algorithm diagram for deep learning-based detection and classification of carotid plaque (DL-DCCP). CEUS, contrast-enhanced ultrasound; MOE, mixture of experts; ROI, region of interest;

in the carotid plaque classification system. The ROI containing the plaque is only a small part of the entire image, which makes the precise detection and localisation of the plaque very difficult. Instead of detecting

the plaque, we detect the adventitia of the carotids by combining two state-of-the-art methods: ResNet50¹² is used to extract features, and RetinaNet¹³ is used to detect a rectangular region around the adventitia. The workflow

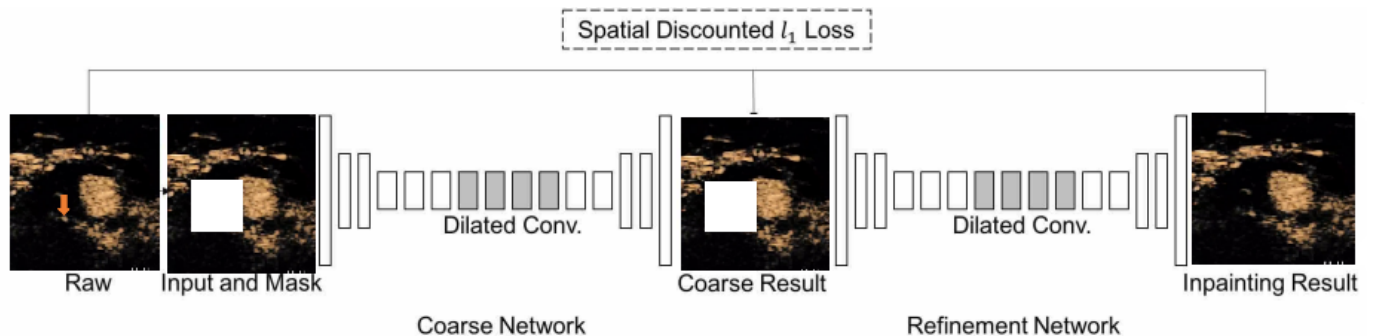


Figure 3 Network architecture for preprocessing ultrasound images to remove markers.

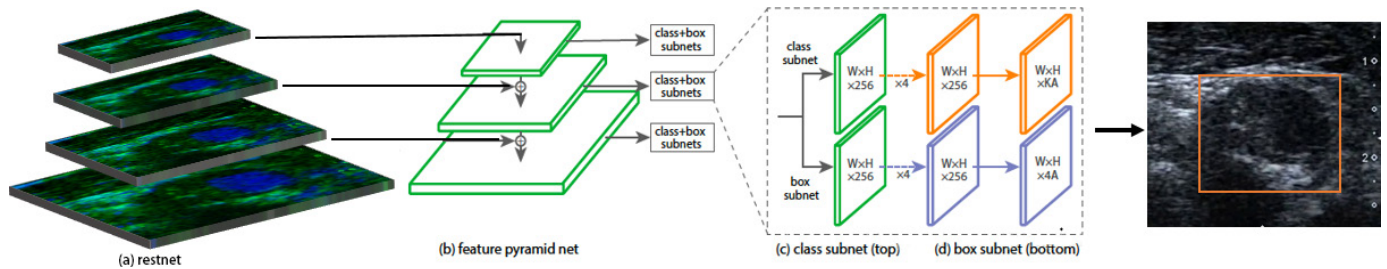


Figure 4 Network architecture for ROI detection.

of RetinaNet is shown in figure 4. To improve the robustness of the detection algorithm, the B-mode (grey scale) and CEUS images (figure 2B) are used as the input simultaneously.

In this study, the method of assessing the vulnerability of carotid plaque was dependent on CEUS video. To reduce the storage space required and accelerate the calculation, the videos are resampled to a fixed frame rate (frames per second, FPS) while maintaining the original spatial resolution. According to the length of the video, the number of frames (N , figure 2D) of the resampled video ranges from 70 to 120. After video resampling, each frame is cropped using the detected ROI information and then input to the Xception network to calculate the frame features (step D). The shallow layers of the Xception network extract detailed features such as edges and textures, and the deep layers extract abstract features such as shapes and contours. Finally, the Xception network outputs a 1D feature vector with a fixed length of 2048 that integrates the shallow-layer and deep-layer features. To select better features and reduce the number of subsequent calculations, we applied the commonly used principal component analysis algorithm to compress the feature dimensions from 2048 to 1024 (step E). Since the number of frames (N) can be different for different videos, we aggregate the video features ($N \times 1024$) to the fixed-length features ($K \times 1024$, $K < N$) by a learnable pooling layer (step F). As shown in figure 5, a Generalized Vector of Locally Aggregated Descriptors layer (NetVLAD)¹⁴ is one of the implementations of the learnable pooling layer. It can learn to extract the motion dynamic features of multiple frames. After these steps, one feature vector (VLAD) with a fixed length ($K \times D$, $D = 1024$) for each video is obtained. In the last step (step G), a classifier named mixture of experts, as shown in figure 6, is applied to make the prediction for each video.

Given the input feature vector VLAD (represented by x), k expert networks generate k outputs O_1, O_2, \dots, O_k . At the same time, the gating network with learnable parameters v also generates k weighting vectors g_1, g_2, \dots, g_k using the following formula:

$$g_i = \frac{\exp(\text{dot}(x, v_i))}{\sum_{j=1}^k \exp(\text{dot}(x, v_j))}, \quad i \text{ from } 1 \text{ to } k, \quad \text{dot} \text{ is the dot product of two vectors and } v_s \text{ is the } s^{\text{th}} \text{ row of } v.$$

The output vector Y is calculated by mixing the output of the k experts: $Y = \sum g_i O_i$, $i = 1, \dots, k$. Y is a 1-D vector with length 2. The first number of Y (Y_0) represents the probability that the plaque is stable, and the second number of Y (Y_1) represents the probability that the plaque is vulnerable. The final classified result (stable or vulnerable plaque) is obtained from the following equation:

$$\text{Classified result} = \begin{cases} \text{Stable plaque, } Y_0 > Y_1 \\ \text{Vulnerable plaque, otherwise} \end{cases}$$

Training and validation of DL-DCCP

Two-thirds of the enrolled patients were randomly selected, and their corresponding CEUS videos were used as the training cohort for DL-DCCP. The CEUS videos of the remaining patients were used as the holdout validation cohort to evaluate the diagnostic accuracy of DL-DCCP. In the training stage, threefold cross-validation was used to iteratively adjust the network architecture (hyperparameters, number of iterations, regularisation method and class weights). For each fold, one model was trained with a subset of 2/3 of the patients, and the remaining 1/3 of the patients were used for validation. After cross-validation, the best performing model was selected for the final evaluation with the holdout validation cohort. We directly used the original architecture for all the networks. Transfer learning was used to finely

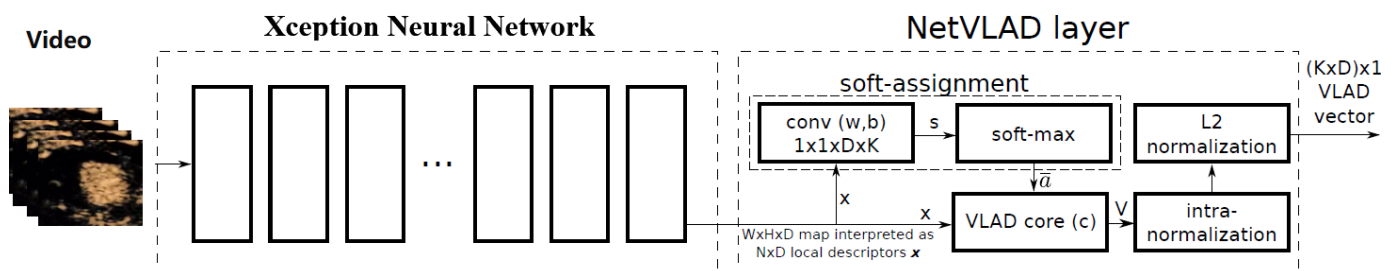


Figure 5 Network architecture for feature extraction and NetVLAD layer. VLAD, Vector of Locally Aggregated Descriptors layer.

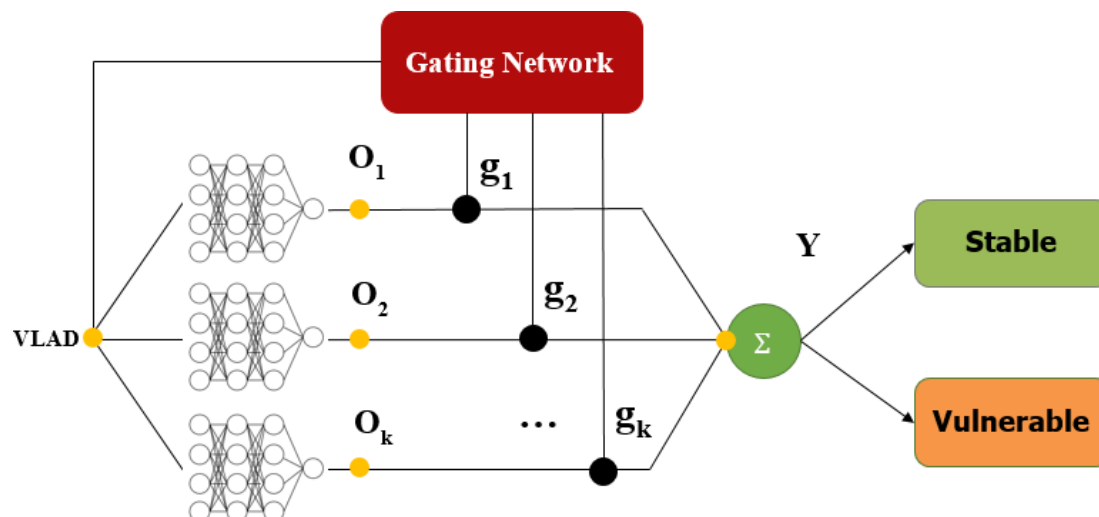


Figure 6 The mixture-of-experts classifier.

tune the parameters of Xception,¹¹ ResNet50¹² and RetinaNet,¹³ while the other network parameters were learnt from scratch. We used the Adam optimizer to fine tune the entire network end-to-end. The initial learning rate was set to 0.001, and the learning rate was reduced 50% when the validation loss on 10 consecutive epochs did not decrease. The minimum learning rate was set to 0.00001. A total of 500 periods were trained and the batch size was set to 16.

We also separately trained an Xception¹¹ model to classify each frame of CEUS video and obtained the video level result by majority voting. The training and validation protocol was the same as with DL-DCCP.

Assessing the diagnostic accuracy of radiologists-CEUS

Using CEUS images, we observed the location, number and enhanced intensity of IPNVs. We classified the enhanced intensity of the IPNVs into four grades according to the report of Shah *et al.*³ Grade 0 indicated no appearance of neovessels within the plaque. Grade 1 indicated a limited appearance of neovessels within the plaque, and grade 2 was considered to represent the moderate appearance of neovessels within the plaque. Grade 3 was used to indicate the presence of a pulsating arterial vessel within the plaque. Grade 2 and Grade 3 plaques were classified as vulnerability plaques. Radiologists who had performed more than 100 carotid plaque CEUS scans or more than 50 supervised carotid plaque CEUS examinations were enrolled in this multicentre

study, and they were all strictly trained on carotid plaque CEUS using a uniform procedure. Two carotid plaque CEUS operators with more than 10 years of ultrasound operating experience were employed as quality controllers for reviewing all carotid plaque CEUS images and excluding unqualified acquisitions.

Carotid plaque CEUS images were independently reviewed offline by two radiologists with more than 10 years of experience with carotid plaque CEUS. Plaque echogenicity, thickness and IPNV grading were performed by two independent investigators blinded to the patients' characteristics and a clinical research fellow after the CEUS examination. Intraobserver and interobserver agreements were evaluated (table 1).

Statistical analysis

Patient age is summarised using the mean±SD. Student's t-test was used to compare patient age between the training and holdout validation groups. The χ^2 test was applied to compare the differences in other characteristics between the two groups. The area under the receiver operating characteristic (ROC) curve (AUC), sensitivity, specificity, positive/negative predictive values and positive/negative diagnostic likelihood ratios were all calculated to assess the accuracy of the classification of carotid plaque vulnerability. The DeLong test was used to compare the differences between AUCs. The probability threshold of DL-DCCP was determined based on the Youden index, which indicates the point at which the sensitivity and

Table 1 Intraobserver and interobserver agreement assessment of the carotid plaques

	ICC (95% CI)		
	Plaque echogenicity	Plaque thickness	IPNV grading
Intraobserver 1	0.912 (0.846 to 0.952)	0.897 (0.804 to 0.921)	0.895 (0.783 to 0.925)
Intraobserver 2	0.892 (0.796 to 0.962)	0.901 (0.811 to 0.943)	0.906 (0.835 to 0.947)
Interobserver	0.903 (0.804 to 0.934)	0.894 (0.801 to 0.937)	0.863 (0.793 to 0.921)

ICC, intraclass correlation coefficient; IPNV, intraplaque neovessels.

Table 2 Baseline characteristics of patients (n=205)

Variables	All patients	Training cohort	Validation cohort	P value
Clinical characteristics				
No of patients, n (%)	205	136 (66.3)	69 (33.7)	—
Age (years)	61.6±8.4	61.2±8.7	62.3±7.6	0.340
Gender (male)	166 (80.9)	113 (83.0)	53 (76.8)	0.279
Symptomatic patients, n (%)*	67 (32.7)	48 (35.3)	19 (27.5)	0.263
Cardiovascular risk factors				
Family history, n (%)	20 (9.8)	11 (8.1)	9 (13.0)	0.259
Overweight/obesity, n (%)	84 (41.0)	55 (40.4)	29 (42.0)	0.827
Dyslipidaemia, n (%)	88 (42.9)	61 (44.9)	27 (39.1)	0.434
Hypertension, n (%)	115 (56.1)	71 (52.2)	44 (63.8)	0.115
Diabetes mellitus, n (%)	47 (22.9)	34 (25.0)	13 (18.8)	0.322
Smoking, current, n (%)	120 (58.5)	76 (55.9)	44 (63.8)	0.279
Plaque echogenicity, n (%)				0.329
Hypoechoic	59 (28.8)	42 (30.9)	17 (24.6)	
Isoechoic	51 (24.9)	36 (26.5)	15 (21.7)	
Mixed echoic	95 (46.3)	58 (42.6)	37 (53.6)	
Plaque thickness, n (%)				0.263
<3 mm	55 (26.8)	36 (26.5)	19 (27.5)	
3–5 mm	85 (41.5)	52 (38.2)	33 (47.8)	

Note: Values are given as mean ± SD and the percentage of patients.

*Neurological symptoms within 6 months.

specificity are maximised. All statistical tests were two sided, and p values less than 0.05 were considered statistically significant. All statistical analyses procedures were carried out using STATA V.15 software (StataCorp).

RESULTS

Baseline characters

From September 2017 to September 2018, up to 547 potentially eligible patients from 10 Chinese hospitals were prospectively enrolled in this study. According to the inclusion and exclusion criteria, 205 patients with CEUS videos were finally enrolled for analysis. No significant adverse events occurred as a result of CEUS. After patient randomisation, 136 patients were assigned to the training cohort, and the remaining 69 patients were assigned to the holdout validation cohort. Their characteristics are summarised in [table 2](#). Between the training and validation cohorts, there were no significant differences in any baseline characteristics (p>0.05).

Comparison of the diagnostic accuracy of DL-DCCP, Xception and radiologists-CEUS

We compared the diagnostic performance of the proposed model (DL-DCCP) with that of experienced radiologists (RA-CEUS) independently for the training and holdout validation cohorts. The AUCs of DL-DCCP

were significantly better than those of RA-CEUS for both the training and holdout validation cohorts (training, DL-DCCP vs RA-CEUS, AUC: 0.85 vs 0.69, p<0.01; holdout validation, DL-DCCP vs RA-CEUS, AUC: 0.87 vs 0.66, p<0.01, [table 3](#)). Xception, one of the best deep CNN models for static image analysis, also performed better than RA-CEUS for both the training and holdout validation cohorts (training, Xception vs RA-CEUS, AUC: 0.82 vs 0.69, p<0.01; holdout validation, Xception vs RA-CEUS, AUC: 0.77 vs 0.66, p<0.01, [table 3](#)).

For DL-DCCP and Xception, threefold cross-validation was performed to evaluate their robustness. The ROC curves for the training and holdout validation cohorts are shown in [figure 7](#). The DeLong tests showed that there were no significant differences among the threefold models (all p>0.05) for both DL-DCCP and Xception. The ROC curves of all threefold models were within a reasonable deviation of the respective mean ROC curves for the two networks ([figure 7A–D](#)). However, DL-DCCP performed significantly better than Xception for both the training and holdout validation cohorts (training, DL-DCCP vs Xception, AUC: 0.85 vs 0.82, p<0.01; holdout validation, DL-DCCP vs Xception, AUC: 0.87 vs 0.77, p<0.01, [table 3](#), [figure 7E–F](#)). Interestingly, the difference between the AUCs for the validation cohort was much higher than that for the training cohort. This fully

Table 3 Diagnostic performances of DL-DCCP, Xception and RA-CEUS for the vulnerability evaluation of carotid plaque

	N	AUC (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	LR+ (95% CI)	LR- (95% CI)	
DL-DCCP	Train	136	0.85 (0.80 to 0.90)	80.5 (69.9 to 88.7)	76.3 (63.4 to 86.4)	81.6 (71.0 to 89.6)	75.0 (62.1 to 85.3)	3.4 (2.1 to 5.4)	0.3 (0.2 to 0.4)
	Holdout validation	69	0.87 (0.82 to 0.92)	79.2 (57.9 to 92.9)	84.4 (70.5 to 93.5)	73.1 (52.2 to 88.4)	88.4 (74.9 to 96.1)	5.1 (2.5 to 10.4)	0.2 (0.1 to 0.5)
Xception	Train	136	0.82 (0.76 to 0.88)	74.0 (62.8 to 83.4)	78.0 (65.3 to 87.7)	81.4 (70.3 to 89.7)	69.7 (57.2 to 80.4)	3.4 (2.0 to 5.5)	0.3 (0.2 to 0.5)
	Holdout validation	69	0.77 (0.71 to 0.83)	75.0 (53.3 to 90.2)	75.6 (60.5 to 87.1)	62.1 (42.0 to 79.3)	85.0 (67.0 to 94.4)	3.1 (1.7 to 5.4)	0.3 (0.2 to 0.7)
RA-CEUS	Train	136	0.69 (0.61 to 0.75)	89.6 (82.8 to 96.4)	47.5 (34.7 to 60.2)	69.0 (59.9 to 78.1)	77.8 (64.2 to 91.4)	1.7 (1.3 to 2.2)	0.2 (0.1 to 0.4)
	Holdout validation	69	0.66 (0.56 to 0.76)	87.5 (74.3 to 100)	44.4 (29.9 to 59.0)	45.7 (31.3 to 60.0)	87.0 (73.2 to 100)	1.5 (1.1 to 2.1)	0.2 (0.1 to 0.8)

AUC, area under the curve; DL-DCCP, deep learning-based detection and classification of carotid plaque; LR-, negative diagnostic likelihood ratio; LR+, positive diagnostic likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; RA-CEUS, classification of carotid plaque by radiologists manually examination contrast-enhanced ultrasound video.

demonstrates that although a deep CNN that considers only static image information could achieve good results with the training cohort, it had great difficulty in generalising well to the holdout validation cohort that had not been seen before.

DISCUSSION

Previous studies have applied computer-assisted ultrasound images to evaluate the vulnerability of carotid plaques.^{15 16} Traditional neural networks are usually formed from the superposition of multiple convolution layer networks and directly predict their outputs.^{17 18} Due to the complexity, dynamics and low signal-to-noise ratio of carotid plaque images, traditional methods might make incorrect predictions when they depend only on separated frames, while doctors often use contextual and dynamic information from images for diagnosis. Therefore, the neural networks described in this study take dynamic video as input and finally fuse the features from all the frames to assess the vulnerability of carotid plaques. To our knowledge, this study was the first to present a novel neural network approach for automatically characterising the vulnerability of carotid plaques in dynamic carotid CEUS video.

In this multicentre prospective study, we presented a deep learning approach for automatically characterising the vulnerability of carotid plaque in CEUS video. The diagnostic accuracy of DL-DCCP and the conventional manual evaluation of CEUS (RA-CEUS) in assessing the vulnerability of carotid plaque were compared against histopathological or high-resolution MRI findings. In assessing the vulnerability of carotid plaque, DL-DCCP demonstrated significant improvements over RA-CEUS. With the training cohort, the AUC of DL-DCCP reached 0.85, and with the holdout validation cohort, it was 0.87, which indicated that DL-DCCP provided good diagnostic efficacy and generalisability simultaneously. RA-CEUS showed lower diagnostic accuracy, with an AUC of 0.69 with the training cohort and 0.66 with the holdout validation cohort. The diagnostic accuracy of DL-DCCP was significantly better than that of RA-CEUS. This means that DL-DCCP can be a potential breakthrough in making clinical predictions on the vulnerability of carotid plaques.

To investigate whether directly using CEUS video as input affected the performance of DL-DCCP, a separate deep CNN (Xception) for static image analysis was trained. To ensure a fair comparison, both DL-DCCP and Xception used the same backbone network and the same training and validation protocol. Each frame of the CEUS video was processed separately by Xception, and then the video-level classification result was obtained by majority voting. The results revealed that DL-DCCP significantly outperformed Xception, especially with the holdout validation cohort (training, DL-DCCP vs Xception, AUC: 0.85 vs 0.82; holdout validation, DL-DCCP vs Xception, AUC: 0.87 vs 0.77). These findings suggest that DL-DCCP

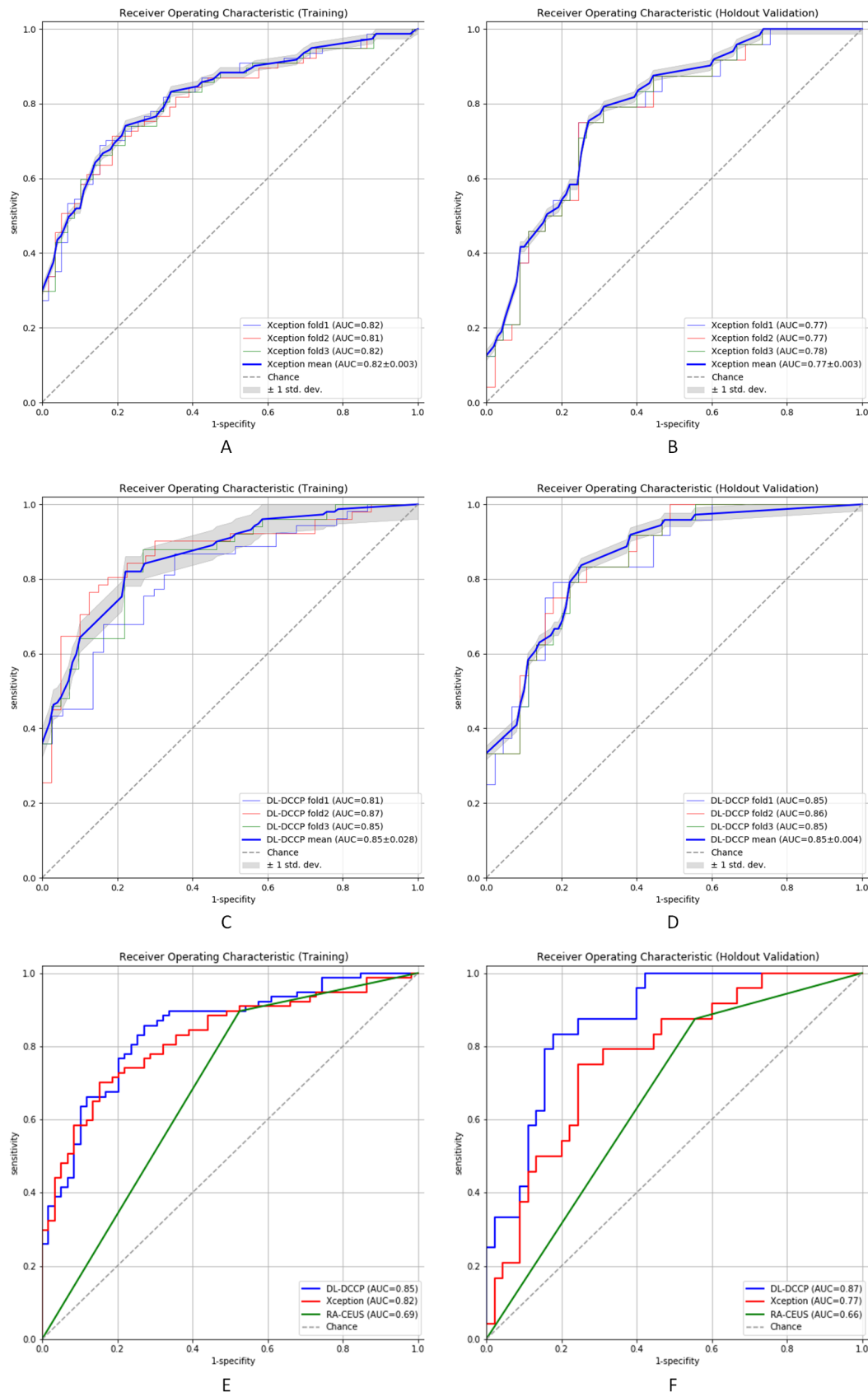


Figure 7 Comparison of diagnostic performance with ROC analysis. AUC, area under the curve; DL-DCCP, deep learning-based detection and classification of carotid plaque; RA-CEUS, radiologists who manually examined the contrast-enhanced ultrasound; ROC, receiver operating characteristic curve.



can better use the interframe information provided by dynamic CEUS video to generalise the diagnostic ability learnt from the training cohort to the unseen validation cohort, while the traditional static CNN, which analyses only intraframe information, has a risk of overfitting to the training cohort.

This present study has some limitations that are important to mention. First, the major limitation of our study was the limited population size. With the holdout validation group, the sensitivity and specificity of DL-DCCP of CEUS video were 79.2% and 84.4%, respectively. The diagnostic accuracy of DL-DCCP still needs to be improved for clinical application. Future studies should involve more patients with carotid plaque in a larger sample size to better train the deep-learning model. Second, in the present study, all ultrasound videos were manually trimmed by a radiologist. The video of the whole carotid angiography process consisted of contrast agent fill-in, diffusion and wash-out. However, only the portion of the videos that contained contrast agent diffusion was used in this study. In future studies, the fill-in and wash-out processes for the contrast agents should also be considered and distinguished automatically. Furthermore, in addition to grey scale ultrasound and CEUS images, elastography and superb microvascular imaging (SMI) are also applied in carotid plaque assessment.^{19–22} However, elastography and SMI were not taken into consideration in this study. The area of interest may capture different sections in different imaging modes, which can lead to temporal and spatial differences in these areas; thus, it would be very difficult to fuse elastography and SMI images into CEUS videos. Therefore, we applied only CEUS videos of carotid plaque in this study. To improve the diagnostic accuracy of our model, multimode ultrasound diagnosis that includes elastography and SMI would be worthy of further study in the future. In addition, clinically relevant analyses comparing the application value of DL-DCCP versus RA-CEUS for patients and their related risk of recurrent stroke are needed for long-term follow-up study.

CONCLUSION

This study demonstrated that DL-DCCP was more accurate than RA-CEUS in assessing the vulnerability of carotid plaques. Compared with a state-of-the-art deep CNN model for static image analysis, DL-DCCP provided increased diagnostic accuracy and better generalisability. All of these results suggest the promising clinical potential of DL-DCCP for the accurate non-invasive assessment of carotid plaque vulnerability.

Author affiliations

¹Department of Ultrasound, Beijing Tiantan Hospital, Beijing, China

²Department of R&D, CHISON Medical Technologies Co Ltd, Wuxi, China

³Department of Ultrasound, Lanzhou University Second Hospital, Lanzhou, Gansu, China

⁴Department of Ultrasound, Zhejiang University School of Medicine Second Affiliated Hospital, Hangzhou, Zhejiang, China

⁵Department of Ultrasound, Zhengzhou University First Affiliated Hospital, Zhengzhou, Henan, China

⁶Department of Ultrasound, Beijing An Zhen Hospital, Chaoyang-qu, Beijing, China

⁷Department of Ultrasound, Third Military Medical University Southwest Hospital, Chongqing, China

⁸Department of Ultrasound, Henan Provincial People's Hospital, Zhengzhou, Henan, China

⁹Department of Ultrasound, Hebei North University Basic Medical College, Zhangjiakou, Hebei, China

¹⁰Department of Ultrasound, Tangdu Hospital Fourth Military Medical University, Xi'an, Shaanxi, China

¹¹Department of Ultrasound, Xi'an Jiaotong University Medical College First Affiliated Hospital, Xi'an, Shaanxi, China

Contributors All authors planned the study. Study conception: WH and YG. Data collection: BN, FN, PH, R-FZ, QY, YG, JY, YW, LY and LR. DL-DCCP system management and the data extraction: CY and MZ. Data analysis: CY and MZ. Analysis and prepared the figures and tables: YG and CY. Administrative support: HZ, HS and TY. Manuscript drafting: YG and YZ. All authors read and approved the final version of the manuscript. All persons named in the Acknowledgments section have provided the corresponding author with permission to be named in the manuscript.

Funding The National Natural Science Foundation of China (grant no.81730050); The National Natural Science Foundation of China (grant no. 81901744); Beijing Natural Science Foundation (grant no. 7204255).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval Ethics committee of the principal investigator's hospital. Ethics committee of Beijing Tiantan hospital (No. KY2017-050-01).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information. See.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Wen He <http://orcid.org/0000-0002-8427-971X>

REFERENCES

- 1 Grau AJ, Weimar C, Buggle F, *et al*. Risk factors, outcome, and treatment in subtypes of ischemic stroke: the German stroke data bank. *Stroke* 2001;32:2559–66.
- 2 Saba L, Anzidei M, Marincola BC, *et al*. Imaging of the carotid artery vulnerable plaque. *Cardiovasc Intervent Radiol* 2014;37:572–85.
- 3 Shah F, Balan P, Weinberg M, *et al*. Contrast-enhanced ultrasound imaging of atherosclerotic carotid plaque neovascularization: a new surrogate marker of atherosclerosis? *Vasc Med* 2007;12:291–7.
- 4 Naghavi M, Libby P, Falk E, *et al*. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: Part I. *Circulation* 2003;108:1664–72.
- 5 Hellings WE, Peeters W, Moll FL, *et al*. Composition of carotid atherosclerotic plaque is associated with cardiovascular outcome: a prognostic study. *Circulation* 2010;121:1941–50.
- 6 Hingwala D, Kesavadas C, Sylaja PN, Thomas B, *et al*. Multimodality imaging of carotid atherosclerotic plaque: going beyond stenosis. *Indian J Radiol Imaging* 2013;23:26–34.
- 7 Feinstein SB. Contrast ultrasound imaging of the carotid artery vasa vasorum and atherosclerotic plaque neovascularization. *J Am Coll Cardiol* 2006;48:236–43.
- 8 Granada JF, Feinstein SB. Imaging of the vasa vasorum. *Nat Clin Pract Cardiovasc Med* 2008;5:S18–25.
- 9 Staub D, Partovi S, Imfeld S, *et al*. Novel applications of contrast-enhanced ultrasound imaging in vascular medicine. *Vasa* 2013;42:17–31.

- 10 Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;278:563–77.
- 11 Chollet F. Xception: deep learning with depthwise separable convolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition* 2017:1800–7.
- 12 He K, Zhang X, Ren S. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition* 2016:770–8.
- 13 Lin TY, Goyal P, Girshick R. Focal loss for dense object detection. *2017 IEEE International Conference on Computer Vision* 2017:2980–8.
- 14 Arandjelovic R, Gronat P, Torii A, et al. NetVLAD: CNN architecture for weakly supervised place recognition. *IEEE Trans Pattern Anal Mach Intell* 2018;40:1437–51.
- 15 Grønholdt ML, Nordestgaard BG, Schroeder TV, et al. Ultrasonic echolucent carotid plaques predict future strokes. *Circulation* 2001;104:68–73.
- 16 Hoogi A, Adam D, Hoffman A, et al. Carotid plaque vulnerability: quantification of neovascularization on contrast-enhanced ultrasound with histopathologic correlation. *AJR Am J Roentgenol* 2011;196:431–6.
- 17 Gao L, Liu R, Jiang Y, et al. Computer-aided system for diagnosing thyroid nodules on ultrasound: a comparison with radiologist-based clinical assessments. *Head Neck* 2018;40:778–83.
- 18 Wang K, Lu X, Zhou H, et al. Deep learning radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut* 2019;68:729–41.
- 19 Cloutier G, Cardinal M-HR, Ju Y, et al. Carotid plaque vulnerability assessment using ultrasound elastography and echogenicity analysis. *AJR Am J Roentgenol* 2018;211:847–55.
- 20 Di Leo N, Venturini L, de Soccio V, et al. Multiparametric ultrasound evaluation with CEUS and shear wave elastography for carotid plaque risk stratification. *J Ultrasound* 2018;21:293–300.
- 21 Forsberg F, Machado P, Stanczak M, et al. Assessing carotid plaque neovascularity and calcifications in patients prior to endarterectomy. *J Vasc Surg* 2019;70:1137–44.
- 22 Chen X, Wang H, Jiang Y, et al. Neovascularization in carotid atherosclerotic plaques can be effectively evaluated by superb microvascular imaging (SMI): initial experience. *Vasc Med* 2020;25:328–33.