Differentiating between stable and progressive carotid atherosclerotic plaques from in-vivo ultrasound images using texture descriptors

Martin Kostelanský[a], Ana Manzano Rodríguez[b], Jan Kybic[a], Miroslav Hekrdla[a], Ondřej Dvorský[c], Jiří Kozel[d], Patricie Baurová[d], David Pakizer[d] and David Školoudík[e]

[a]Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic;
[b]Barcelona School of Telecommunications Engineering, Universitat Politècnica de Catalunya, Spain;
[c]Faculty of Biomedical Engineering, Czech Technical University in Prague, Czech Republic;
[d] Faculty of Health Sciences, Palacký University Olomouc, Czech Republic;
[e]Faculty of Medicine, University of Ostrava, Czech Republic

ABSTRACT

We describe an automatic pipeline for processing ultrasound images of the carotid artery, consisting of image type classification, carotid artery localization, segmentation, feature descriptor extraction, and plaque stability classification. The aim is to distinguish between stable (safe) and progressive (dangerous) atherosclerotic plaques from a single standard ultrasound transversal or longitudinal B-mode examination. The processing pipeline uses modern deep CNN techniques, while the descriptors are based on geometry and wavelets to characterize texture. When testing on a large dataset of 28718 images from 413 patients, we found that our automatically calculated descriptors are statistically significantly different between the two classes with a very high significance level, $p < 10^{-3}$. We have also created a random forest-based classifier to distinguish between progressive and stable plaques, although its accuracy remains low ($61 \sim 62\%$).

Keywords: Atherosclerosis, plaque, ultrasound, progression, automatic, analysis, texture, classification

1. INTRODUCTION

Atherosclerosis is among the most frequent cause of cardiac death worldwide and the most common cause of stroke,[1] a medical emergency condition occurring when the blood supply to the brain is seriously reduced. Stroke is one of the leading causes of morbidity, mortality and long-term disability, it can also cause dementia and epilepsy. Atherosclerotic plaques are detected by ultrasonographic screening[2] in about 50% of the people at the age of 50.[3] Most ($\sim 60\%$) of these people do not experience any adverse effects and their plaques remain *stable*. To detect the minority with potentially dangerous *unstable* plaques, regular follow-up examination need to be performed. It would therefore be useful to determine the risk of a specific plaque being unstable from a single examination. It would not only allow reducing the number of follow-up examinations but it might also improve the indication criteria for pharmacological treatment, carotid endarterectomy, or stenting. Unfortunately, despite intensive research, the mechanism of the atherosclerotic plaque growth and the associated risk factors are not yet fully understood.

In this work, we attempt to predict the plaque stability from image-based features calculated from standard transversal and longitudinal B-mode ultrasound images of the carotid artery. While it has been established previously[1,4,5] that there is a relationship between the carotid artery appearance in the ultrasound image and the plaque stability, the studies so far do not agree on what the relevant quantitative features are.

There are two key contributions with respect to previous work. First, our proposed pipeline is fully automatic and does not require any manual steps. Second, we not only establish a statistical significance but also create a classifier to determine whether a particular image corresponds to a stable or progressive plaque.

Send correspondence to Jan Kybic, e-mail: kybic@fel.cvut.cz

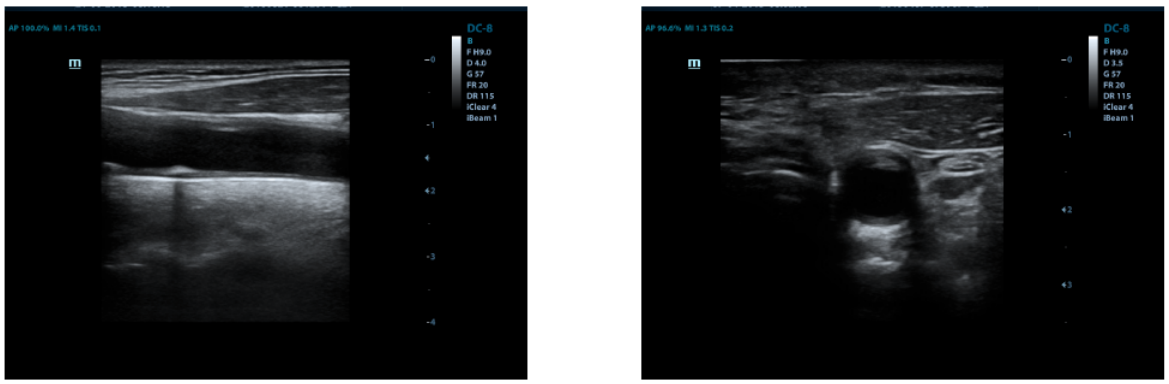

Figure 1. Examples of the longitudinal (left) and transversal (right) ultrasound images of the carotid artery.

2. DATA

The data comes from an approved clinical study ANTIQUE.[6] There were 413 patients with the degree of stenosis higher than 30%. Several transversal, longitudinal, conical, and Doppler ultrasound images of the carotid artery were acquired for each patient using Mindray DC8 ultrasound scanner with a linear 3–12 MHz probe. The patients were followed every six month for three years, corresponding to 1322 individual examinations and 28178 ultrasound scans. The plaques were labeled as *progressive*, if their maximum width increased by more than 2σ in 3 years, and as *stable* if its maximum width increased by less than σ, where $\sigma = 0.2$ mm was the measurement error, estimated as the 99% percentile of the difference between two measurements two weeks apart. The task attempted here is to automatically predict this label (stable or progressive) from the transversal and longitudinal ultrasound images from the first examination (Fig. 1)

The images were anonymized by cropping out the patient-related textual fields and changing the file names.

3. METHODS

The processing pipeline[7,8] consists of automatic identification of transversal and longitudinal images (Section 3.1), localization of the carotid artery (Section 3.2), segmentation (Section 3.3), calculation of the shape and texture descriptors (Section 3.4), and binary classification of each image as progressive or stable (Section 3.5). The pipelines for the transversal and longitudinal images have the same structure but are trained independently. The code was developed in Python using the PyTorch library.

3.1 Image type classification

Since in our dataset the images are not tagged by type (longitudinal, transverse, conical, or Doppler), the first step is to classify them into these four classes, in order to avoid examining all the 28178 images by hand. We have manually tagged 1679 images from the ANTIQUE dataset and we have also used 1282 already tagged images (84 longitudinal and 1214 transversal) from the SPLab datasets*.[9] All images were intensity normalized and scaled to the same size. We used the ResNet50[10] convolutional neural network with the cross-entropy loss function, pretrained on the ImageNet dataset and fine-tuned on our data using stochastic gradient descent with learning rate 10^{-4} and Nesterov momentum 0.95 for 30 epochs. On a test set of 514 images from the ANTIQUE dataset (not used for training), only 4 were misclassified, corresponding to accuracy of 99.22%. Applying the classifier to the whole of the ANTIQUE dataset yielded 7532 longitudinal (3986 progressive and 3546 stable) and 9709 transversal (4982 progressive and 4727 stable) images for subsequent processing. The other two image types were not used.

*`http://splab.cz/en/download/databaze/ultrasound`, `http://splab.cz/en/research/zpracovani-medicinskych-signalu/databaze/artery`

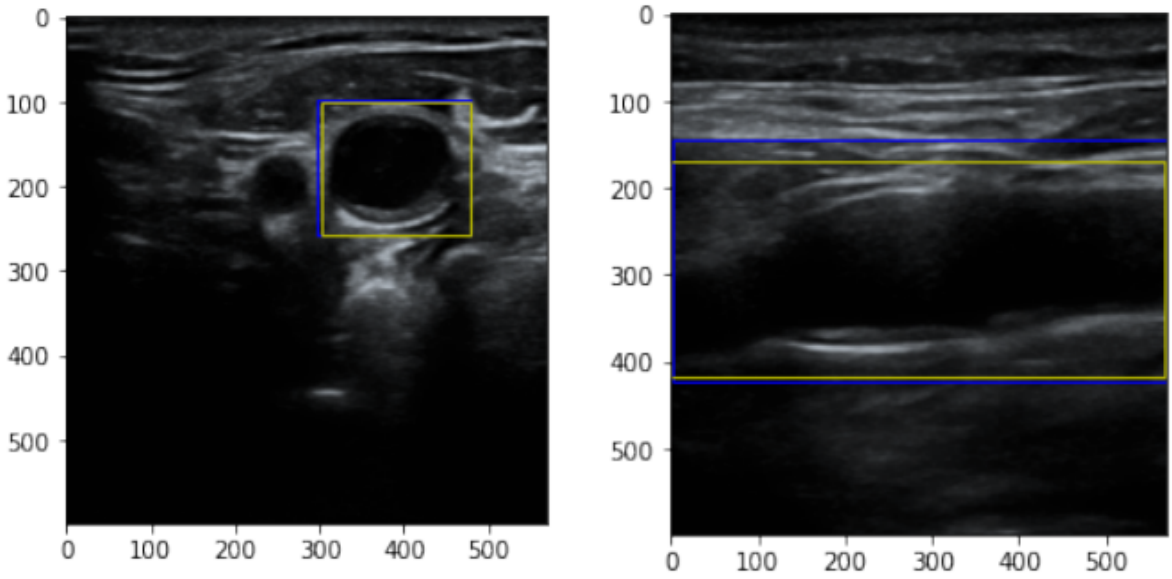

Figure 2. Example manual (yellow) and automatic (blue) ROI in the transversal (left) and longitudinal (right) images.

3.2 Carotid artery localization

The carotid artery is usually just a part of the input image. To simplify the task of the subsequent steps, we automatically identify a small rectangular region of interest (ROI) containing the carotid artery in both longitudinal and transversal images. If both internal (ICA) and external (ECA) branches of the carotid artery are present in the image, ICA is chosen, since it supplies the brain and the consequences of the stenosis may therefore be more serious.

We created a dataset by selecting one longitudinal and one transversal image from 75 patients from the *progressive* and 75 patients from the *stable* group and manually created ROIs (see Fig. 2), i.e. 300 images in total. We have also used transversal images with ROIs from the SPLab dataset (1214 images). The test subset consisted of 50 transversal and 50 longitudinal ANTIQUE images (different from the training images), the rest was used for training. Annotation was performed by the CVAT web-based tool[†]. Images were intensity normalized and random crops and scale changes ($\pm 20\%$) were used for augmentation.

We used the Faster R-CNN architecture with the initial convolutional blocks pretrained on ImageNet and optimizing the sum of the classification and localization losses.[11] A single highest-probability ROI was selected. Training was performed during 40 epochs with the Adam optimizer and an initial learning rate 10^{-4} with gradual decay. Best results for the transversal images were obtained by first training the network on the SPLab dataset and then fine-tuning on the ANTIQUE dataset—on the test data the IoU (intersection over union) was larger than 0.6 in 98% of cases, larger than 0.75 in 90% of cases, and larger than 0.85 in 68% of cases. Examining the results shows that differences from the reference ROI were often caused by annotation variability and the ROIs were almost always usable for further processing even if the IoU was low. Only in one image (out of 50) the network failed to find the carotid artery completely.

For longitudinal images, where only 150 annotated training images were available, we achieved IoU better than 0.6 in 100% test images, IoU larger than 0.75 in 90% and IoU larger than 0.85 in 62% of test images. All identified ROIs for the 50 test images were suitable for further processing, even if sometimes less tight than the reference ROIs.

The trained network was applied to all transverse and longitudinal images of the ANTIQUE dataset to extract the ROIs.

[†]`https://github.com/openvinotoolkit/cvat`

3.3 Segmentation

In the segmentation step, the transversal ROIs are segmented to the following classes — background, vessel wall, plaque and lumen (Fig. 3).

For longitudinal images (Fig. 5), it turned out that the experts could not reliably separate the plaque and vessel wall regions. Furthermore, they could reliably determine the wall thickness only in some of the regions. We have therefore extrapolated the vessel wall thickness to other regions, creating a new class *maybe vessel* (yellow in Fig. 4), which the network is allowed to classify without penalty as either *vessel wall* or *background*.

Our experts created manual segmentations of the same 300 training images as in Section 3.2, evenly split between the transversal and longitudinal types. Manual ROIs were used. Segmentations were done using the CVAT software and converted to bitmaps. Images were intensity normalized and flips, random shifts, rotation, contrast and brightness changes used for augmentation.

Several CNN architectures were tested. The best performing CNN architecture for transversal images turned out to be a U-net[12] with 4 levels, the log-cosh Dice loss[13] with increased weight for minority classes (vessel wall 1.5, lumen 1.0, plaque 1.75, found using the validation data), and dropout with probability 0.25. Input ROIs were resized to the 256×256 input image size and the output segmentation rescaled back. We used padding for same size input and output, PReLU activation functions[14] and batch normalization[15] after every second convolutional block. The training took 500 epochs. This architecture reached a pixel-wise accuracy of 85% on the test transversal images. Background and lumen were distinguished well (IoU 84% and 86%, respectively), while wall and plaque segmentation were less reliable (IoU 66% and 59%, respectively).

For longitudinal images, we used the same U-net architecture but for 5×5 kernels. Training took 800 epochs. This architecture achieved a pixel-wise test accuracy of 88.0%. Lumen is segmented very well (average IoU 91.9%), background is acceptable (IoU 81.4%) but the plaque segmentation could be better (IoU 57.5%), while wall segmentation often failed (IoU 10.4%). This was understandable given that even the experts could not delineate the wall reliably.

3.4 Descriptors

We calculate a descriptor vector[8] of length 35 and 15 for transverse and longitudinal ROI images, respectively, as described in Tables 1 and 2. The first 5 parameters are related to the geometry and calculated from the automatic segmentation. In particular, we define

$$\text{relative stenosis} = \frac{\text{plaque area}}{\text{total vessel area}} \tag{1}$$

The *plaque boundary irregularity* is calculated by first approximating the plaque-lumen boundary with equidistant points $\mathbf{x}_i$ and then calculating the mean orientation change per point:

$$\text{plaque boundary irregularity} = \text{mean}_i \left| \angle\left(\mathbf{x}_i\mathbf{x}_{i+1}, \mathbf{x}_{i-1}\mathbf{x}_i\right) \right| \tag{2}$$

The rest of the geometric descriptors are given in Tables 1 and 2 and should be self-explanatory.

The remaining descriptors characterize the texture of the plaque (for transversal images also of the vessel wall and lumen) using Haar wavelet frames.[16,17] At each level $i = 1, 2, 3$, the ROI image $f^{(i)}$ is filtered by four separable filter combinations H_xH_y, H_xG_y, G_xH_y, and G_xG_y, where

$$H(z) = \frac{1 + z^l}{2} \quad \text{and} \quad G(z) = \frac{z^l - 1}{2} \tag{3}$$

are the low- and high-pass *à trous* Haar filters, respectively, of size $l = 2^i$. The low-pass band is used as input to the next level, $f^{(i+1)} = H_xH_yf^{(i)}$, with $f^{(1)}$ being the input image. No subsampling is performed. Each level provides three descriptor values — the mean energies in the high-pass subbands

$$d_{3i-2} = \frac{1}{|\Omega|} \left\| H_xG_yf^{(i)} \right\|^2_{\ell_2(\Omega)}, \quad d_{3i-1} = \frac{1}{|\Omega|} \left\| H_yG_xf^{(i)} \right\|^2_{\ell_2(\Omega)}, \quad d_{3i} = \frac{1}{|\Omega|} \left\| G_xG_yf^{(i)} \right\|^2_{\ell_2(\Omega)} \tag{4}$$

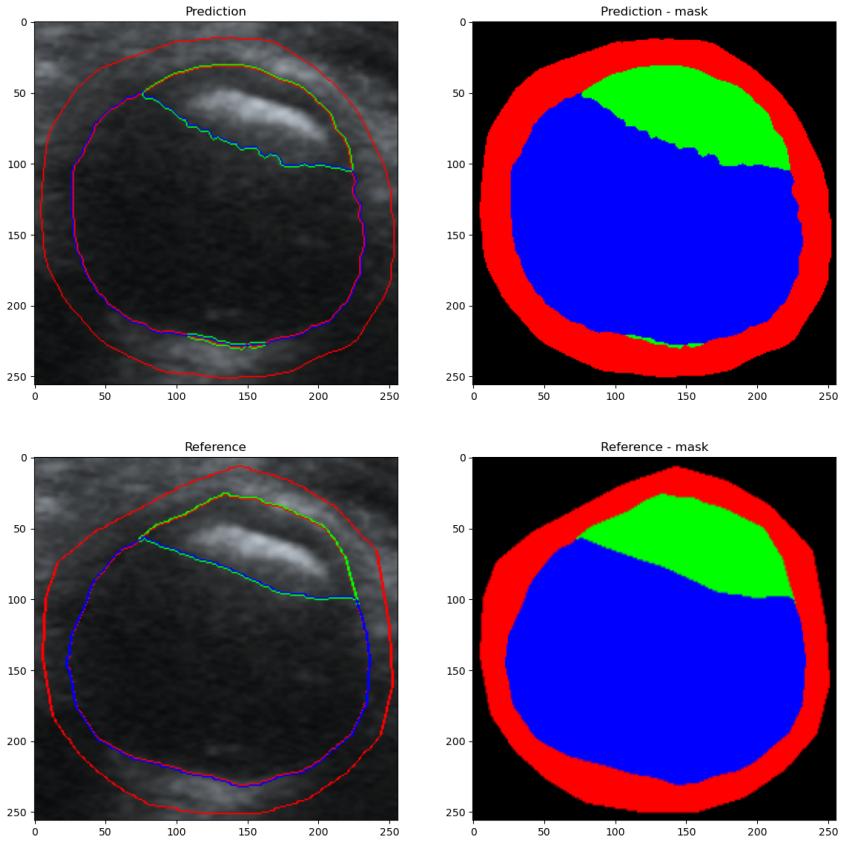

Figure 3. Example automatic (top) and manual (bottom) segmentation of a transversal ROI from the test set to background (black), vessel wall (red), plaque (green) and lumen (blue).

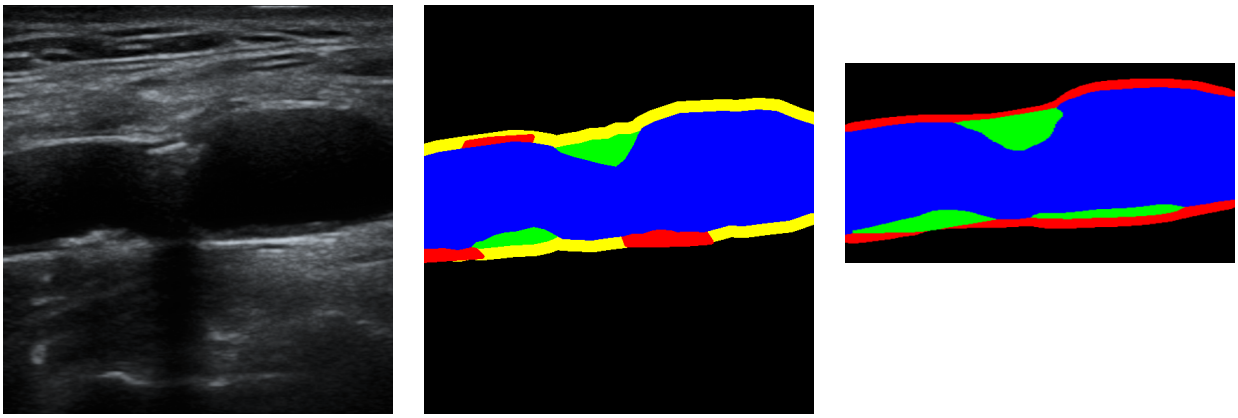

Figure 4. Example longitudinal test image: original image (left), manual segmentation with extrapolated classes (middle) and automatic segmentation (right). (The automatic segmentation was done on the cropped ROI, hence the difference in size.) Color code: background (black), vessel wall (red), extrapolated/maybe vessel wall (yellow), plaque (green) and lumen (blue).

number	descriptor(s)	Welch	Mann-Whitney
1	relative stenosis	yes	yes
2	absolute lumen area	yes	yes
3	lumen minimum enclosing circle radius	no	no
4	plaque boundary irregularity	no	no
5	plaque area	yes	yes
6–15	plaque texture wavelet descriptors	4 of 10	4 of 10
16-25	vessel wall texture wavelet descriptors	0 of 10	2 of 10
26–35	lumen texture wavelet descriptors	3 of 10	4 of 10

Table 1. Descriptors for transverse ROIs and two statistical tests: whether the class means are significantly different according to the Welch's test and whether the class distributions are different according to the Mann-Whitney U rank test, both at significance level $p = 0.01$.

where Ω are the pixels of the class being evaluated. The final texture descriptor d_{10} is the mean energy of the last low-pass subband

$$d_{10} = \frac{1}{|\Omega|} \left\| H_x H_y f^{(3)} \right\|^2_{\ell_2(\Omega)} \tag{5}$$

All descriptors can be calculated very quickly, with just a few operations per pixel.

3.5 Plaque stability classification

The precalculated descriptor vectors from the 7532 longitudinal and 9709 transversal ROI images are fed into a random forest[18] classifier with 100 trees using Python `scikit-learn` library[‡]. Only nodes with at least 30 samples are split. The dataset was divided to 80% training and 20% testing by stratified sampling. Training is fast, it takes ≈ 1 s, with an even faster prediction taking ≈ 20 ms.

4. RESULTS

The main goal of the experiments is to find out whether the descriptors provided by the automatic pipeline described above can be used to distinguish between progressive and stable plaques.

[‡] `scikit-learn.org`

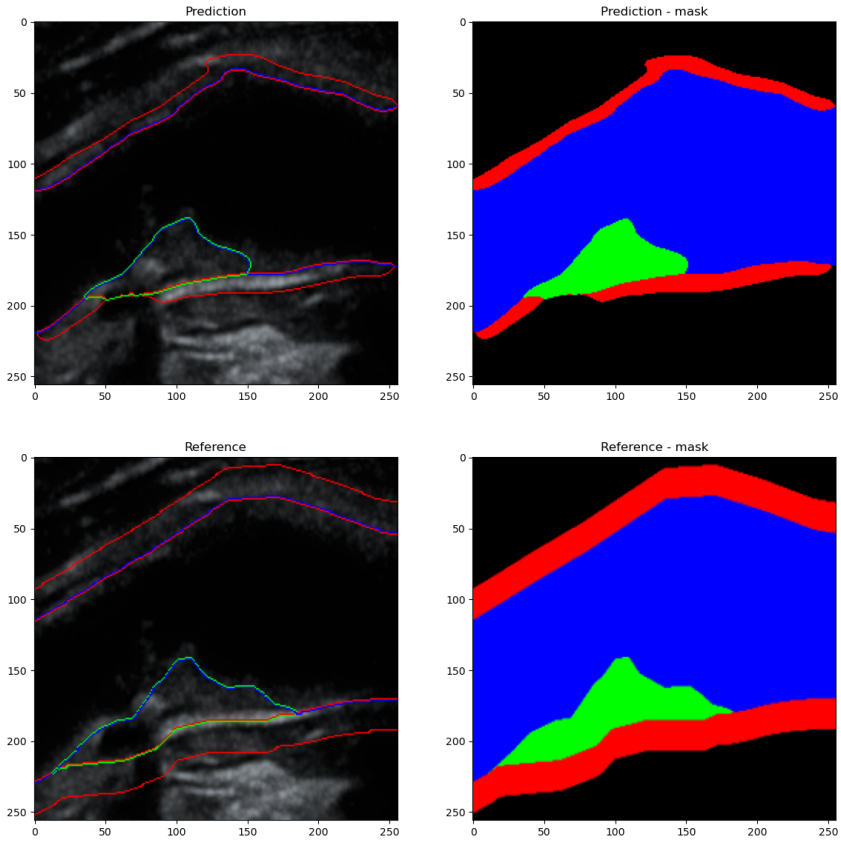

Figure 5. Example automatic (top) and manual (bottom) segmentation of a longitudinal ROI from the test set to background (black), vessel wall (red), plaque (green) and lumen (blue).

number	descriptor(s)	Welch	Mann-Whitney
1	lumen maximum absolute height	no	yes
2	lumen mean absolute height	no	no
3	lumen minimum absolute height	no	no
4	plaque maximum height	yes	yes
5	plaque mean height	yes	yes
6–15	plaque texture wavelet descriptors	7 of 10	7 of 10

Table 2. Descriptors for longitudinal ROIs and two statistical tests: whether the class means are significantly different according to the Welch's test and whether the class distributions are different according to the Mann-Whitney U rank test, both at significance level $p = 0.01$.

type	class	precision	recall	F_1	support	accuracy	AUC
transversal	stable	60%	62%	61%	1418	61%	60.4%
	progressive	63%	60%	61%	1495		
longitudinal	stable	60%	55%	57%	1064	62%	61.3%
	progressive	63%	68%	65%	1196		

Table 3. Random forest classification performance using all descriptors on both transversal and longitudinal test data. 'support' is the number of images in a class.

4.1 Significance testing

We have first used Welch's unequal variance t-test to test the hypothesis that the probability distribution of individual descriptors has the same mean for both progressive and stable classes. As a non-parametric alternative not requiring normality, we have also performed the Mann-Whittney U test to test whether one of the two distributions is stochastically greater than the other.

As shown in Tables 1 and 2, both tests give similar results and the null hypothesis was in fact rejected at the $p = 0.01$ significance level for many of the calculated descriptors, namely related to the plaque dimensions and plaque texture. In fact $p < 10^{-3}$ for these descriptors. Interestingly, some of the vessel wall and lumen texture descriptors in transversal images also showed significant differences.

4.2 Plaque stability descriptor classification evaluation

Let us first clarify that while the differences between descriptor distributions and means for the two classes are significant, it does do not imply that a single descriptor would be a particularly good predictor. The best single predictor is the relative stenosis calculated from transversal images, which has a mean and standard deviation 34.74 ± 4.02 for the *stable* class and 36.25 ± 5.14 for the *progressive* class. The optimal thresholding classifier based on this descriptor has accuracy of only 54%.

Classification performance using all calculated descriptors on all data is summarized in Table 3. We see that the accuracy for both transversal and longitudinal images is $61\% \sim 62\%$.

To evaluate the effect of the automatic segmentation, we have performed the classification training and testing using only the 150 manually segmented longitudinal and 150 transversal images, split $120 : 30$ between training and testing. The accuracy has dropped to 58% and 47% (for transversal and longitudinal images, respectively). This shows that the loss due to a small dataset is bigger than the loss from automatic segmentation errors.

5. CONCLUSIONS

We have confirmed the finding from Školoudík et al.[1] and other studies[4,5] that the appearance of ultrasound images of stable and progressive carotid atherosclerotic plaques differ. Unlike,[1,4] our descriptors are calculated fully automatically using a newly developed image processing pipeline, which makes it possible to analyze much larger datasets with minimum effort. The differences have very high statistical significance.

On the other hand, distinguishing between stable and progressive plaques based on individual images seems to be an unsolved problem and turned out to be rather difficult. The best classification accuracy we have achieved was only $61\% \sim 62\%$, which is far from the level needed for clinical applications.

In future work, we plan to apply deep learning techniques, in the hope of capturing additional image characteristics beyond our hand-designed descriptors. We shall also integrate other know risc (not image-based) factors such as high blood pressure, smoking, or obesity.

ACKNOWLEDGMENTS

The authors acknowledge the support of the OP VVV funded project "CZ.02.1.01/0.0/0.0/16_019/0000765 Research Center for Informatics" and the Czech Health Research Council project NV19-08-00362.

REFERENCES

[1] Školoudík, D., Kešnerová, B., Hrbáč, T., Netuka, D., Vomáčka, Langová, K., Roubec, M., Herzig, R., and Belšan, T., "Visual and digital analysis of the ultrasound image in a stable and progressive carotid atherosclerotic plaque," *Česká a slovenská neurologie a neurochirurgie* (1), 38–44 (2021). in Czech with English abstract.

[2] Nicolaides, A., Beach, K. W., Kyriacou, E., and Pattichis, C. S., [*Ultrasound and carotid bifurcation atherosclerosis*], Springer Science & Business Media (2012).

[3] Salonen, R., Seppänen, K., Rauramaa, R., et al., "Prevalence of carotid atherosclerosis and serum cholesterol levels in eastern finland," *Arteriosclerosis* **6**(8), 788–792 (1988).

[4] Doonan, R. J., Gorgui, J., Veinot, J. P., Lai, C., Kyriacou, E., Corriveau, M. M., Steinmetz, O. K., and Daskalopoulou, S. S., "Plaque echodensity and textural features are associated with histologic carotid plaque instability," *Journal of Vascular Surgery* , 671–677 (2016).

[5] Kakkos, S., Nicolaides, A., Kyriacou, E., Daskalopoulou, S., Sabetai, M., Pattichis, C., Geroulakos, G., Griffin, M., and Thomas, D., "Computerized texture analysis of carotid plaque ultrasonic images can identify unstable plaques associated with ipsilateral neurological symptoms," *Angiology* **62**(4), 317–328 (2011).

[6] Školoudík, D., "Atherosclerotic plaque characteristics associated with a progression rate of the plaque in carotids and a risk of stroke.," (Feb. 2015). Clinical trial NCT02360137.

[7] Kostelanský, M., *Localization and segmentation of in-vivo ultrasound carotid artery images*, master thesis, Faculty of Electrical Engineering, Czech Technical University in Prague (2021).

[8] Manzano Rodríguez, A., *Processing in vivo ultrasound images of the carotid artery*, master thesis, Escola Tècnica d'Enginyeria de Telecomunicacio de Barcelona, Universitat Politècnica de Catalunya, Spain (June 2021).

[9] Říha, K., Mašek, J., Burget, R., Beneš, R., and Závodná, E., "Novel method for localization of common carotid artery transverse section in ultrasound images using modified Viola-Jones detector," *Ultrasound in Medicine and Biology* **39**, 1887–1902 (2013).

[10] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition," in [*CVPR: IEEE Conference on Computer Vision and Patter Recognition*], (2106).

[11] Ren, S., He, K., Girshick, R. B., and Sun, J., "Faster R-CNN: Towards real-time object detection with region proposal networks.," in [*NIPS: Neural information processing systems conference*], 91–99 (2015).

[12] Ronneberger, O., Fischer, P., and Brox, T., "U-Net: Convolutional networks for biomedical image segmentation," in [*MICCAI: Medical Image Computing and Computer-Assisted Intervention*], (May 2015).

[13] Jadon, S., "A survey of loss functions for semantic segmentation," in [*IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology*], 1–7, IEEE (2020).

[14] He, K., Zhang, X., Ren, S., and Sun, J., "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in [*ICCV: IEEE International Conference on Computer Vision*], 1026–1034, IEEE Computer Society (2015).

[15] Ioffe, S. and Szegedy, C., "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in [*ICML: International Conference on Machine Learning*], Bach, F. R. and Blei, D. M., eds., *JMLR Workshop and Conference Proceedings* **37**, 448–456, JMLR.org (2015).

[16] Unser, M., "Local linear transforms for texture measurements," *Signal Processing* **11**, 61–79 (1986).
[17] Svoboda, T., Kybic, J., and Hlaváč, V., [*Image Processing, Analysis and Machine Vision — A MATLAB Companion*], Thomson, Toronto, Canada, 1. ed. (September 2007).
[18] Breiman, L., "Random forests," *Machine learning* (45), 5–32 (2001).