CNN-CASS: CNN for Classification of Coronary Artery Stenosis Score in MPR Images

Mariia Dobko, Bohdan Petryshak, Oles Dobosevych
The Machine Learning Lab, Ukrainian Catholic University, Lviv, Ukraine
{dobko_m, petryshak, dobosevych}@ucu.edu.ua

Abstract. *To decrease patient waiting time for diagnosis of the Coronary Artery Disease, automatic methods are applied to identify its severity using Coronary Computed Tomography Angiography scans or extracted Multiplanar Reconstruction (MPR) images, giving doctors a second-opinion on the priority of each case. The main disadvantage of previous studies is the lack of large set of data that could guarantee their reliability. Another limitation is the usage of handcrafted features requiring manual preprocessing, such as centerline extraction. We overcome both limitations by applying a different automated approach based on ShuffleNet V2 network architecture and testing it on the proposed collected dataset of MPR images, which is bigger than any other used in this field before. We also omit centerline extraction step and train and test our model using whole curved MPR images of 708 and 105 patients, respectively. The model predicts one of three classes: 'no stenosis' for normal, 'non-significant' — 1-50% of stenosis detected, 'significant' — more than 50% of stenosis. We demonstrate model's interpretability through visualization of the most important features selected by the network. For stenosis score classification, the method shows improved performance comparing to previous works, achieving 80% accuracy on the patient level. Our code[1] is publicly available.*

1. Introduction

According to the American Heart Association report [2], approximately 17.6 million deaths were attributed to Cardiovascular Diseases (CVD) globally in 2016, making it the leading cause of death in

---
[1] https://github.com/ucuapps/
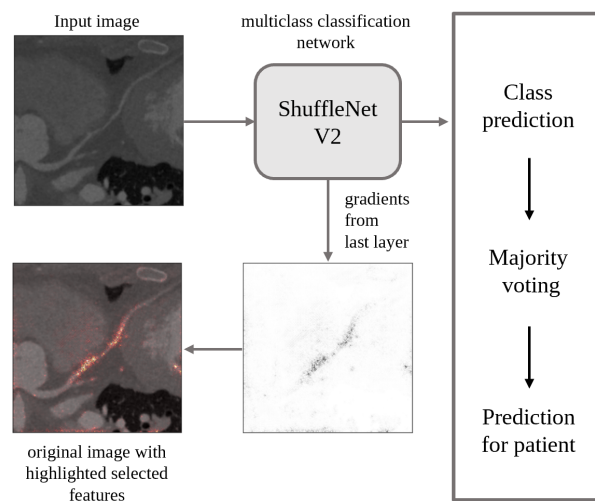CoronaryArteryStenosisScoreClassification/

Figure 1: The pipeline of stenosis classification on an MPR image. For each 2D image, ShuffleNet V2 predicts probabilities of stenosis score. Activation regions of the last layer in the model are shown and overimposed on the input image.

the world. By 2030 it is estimated that CVD will be responsible for over 23.6 million deaths [2], and thus the ability to get early diagnosis becomes crucial. Cardiovascular diseases affect the heart or blood vessels. They include Coronary Heart Disease, or Coronary Artery Disease (CAD), which occurs when plaque (a combination of cholesterol, calcium, fat and other substances) builds up in the arteries and clogs them. This narrowing of the arteries, called stenosis, interferes in the healhy blood flow by hindering oxygen-rich blood cells transportation to the heart.

Coronary Computed Tomography Angiography (CCTA) is one of the components used by radiologists in diagnostics of the coronary artery disease. In the recent studies, CCTA was proven to be clinically effective in combination with functional test-

ing (SCOT-HEART [1]), or even as an alternative to it (PROMISE [7]). Coronary CTA helps doctors in evaluation of the degree of artery stenosis. For patients in the risk group, it is vital to get the diagnosis in a short period of time, however, nowadays it takes up to two weeks to receive analysis results after the scanning procedure. In order to decrease the waiting time, the previous works [12], [22] applied various semi-automated algorithms to identify the severity of disease in medical images, giving doctors a second-opinion on the priority of each case. Such automated computer-aided systems are capable to increase the access to diagnostics and eventually reduce mortality by faster recognition of critical cases.

Due to recent advances in applications of machine learning to medical domain, it is now possible to use neural networks for assessment of the severity of coronary artery stenosis. Among the main disadvantages of previous approaches are the lack of large set of data that could guarantee their reliability and usage of handcrafted features during the preprocessing steps. Our contribution to this field lies in application of a different approach and testing it on the created dataset, which is bigger than any other used in this field before. We also propose a fully automated method to classify stenosis score, that utilizes whole curved Multiplanar Reconstruction (MPR) images without manual preprocessing or centerline extraction, see Figure 1.

2. Related Works

The problem of stenosis score classification on CCTA images of coronary arteries is insufficiently studied. The major difficulty is the absence of publicly available structured and professionally labeled sets of data. Another one is domain specificity, which requires certain expertise in analysing medical data.

Datasets of coronary arteries are usually formed by CCTA scans, from which MPR images can be extracted and transformed to either straight or curved representation of arteries. In the related paper [3], CCTA scans of only 163 patients were collected, and the proposed network was trained and tested using images of 98 and 65 patients, respectively. The authors first straightened MPR images by applying the centerline extraction technique [20] and then used the transformed data to simplify classification of stenosis level. The centerline extraction step used in this approach requires manual placement of a single seed point in the artery of interest, so that the

method is not fully automated. 3D convolutional neural network was utilized to extract features which are used by recurrent neural network for classification. While the achieved accuracy shows good performance and feasibility of deep learning methods for stenosis score classification, the reliability of obtained results can not be justified on such small data sample. Another drawback is its poor stability: as the authors admit, even small errors in centerline extraction may essentially increase the overall error.

Centerline extraction is a common preprocessing method [15] although it often requires manual assistance. It is used in the previous study [22], where user interaction is needed to localize the artery by annotating the start and end points of the vessel. As the authors described, the start point was placed in the coronary ostium of the corresponding arterial tree and the end point was placed at the most distal point inside the vessel. Some parameters were chosen manually, for example, contrast filled (foreground) regions are defined by empirically determining a lower and upper bound values of intensity. Other handcrafted features include the mean and dispersion values of the artery radius, as well as the mean and dispersion of a rotation angle corresponding to a typical location of the artery.

In order to avoid errors caused by centerline extraction, in our approach we use curved MPR images instead. These MPRs are generated from CCTA with the help of radiologist assistant during a general pipeline of the coronary artery diagnosis. Thus, our method does not require any handcrafted features, but utilizes the whole MPR image, where artery is curved.

3. Data

For training our algorithms, we used curved MPR images of the coronary artery with stenosis levels annotated by professional radiologists from a well-renowned Future Medical Imaging Group (FMIG) in Australia. Our current dataset consists of 160,000 MPR images which were extracted from CCTA scans of 828 unique patients; see data statistics in Table 1.

3.1. MPR generation process

The CCTA stack of images representing the patient's heart, is produced by the CT scanner. The raw CT scans require thorough and long analysis as the coronary artery is represented in each slice as a small circular region, similar to a dot. To increase the in-

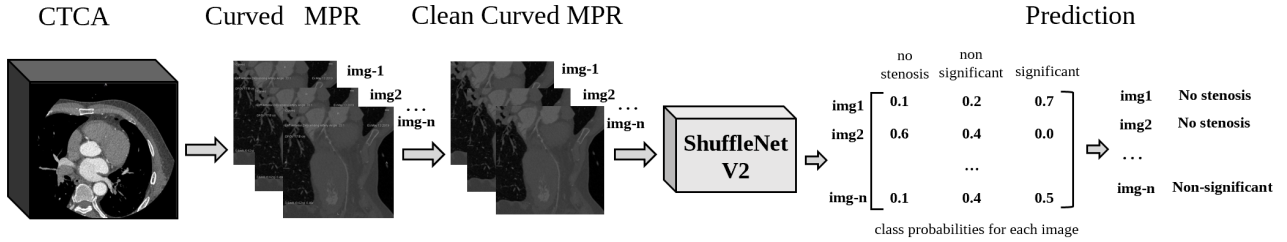CTCA   Curved MPR   Clean Curved MPR   Prediction

Figure 2: **Prediction process.** First, the whole CCTA is converted into sets of MPR images for each artery branch (completed by a radiologist assistant). Then, our method automatically cleans the image from the text and meta information, and feeds the obtained preprocessed images to the ShuffleNet V2. As an output, the probability predictions of each class are produced.

Arteries	Arteries	Sections	Sections
LAD	824	**LAD**	822
		D-1	729
		D-2	356
		D-3	68
LCX	722	**LCX**	639
		PLV-LCX	15
		PDA-LCX	17
RCA	721	**RCA**	91
		OM	6
		OM-1	81
		OM-2	281
		OM-3	75
		PLV-RCA	609
		PDA-RCA	71

Table 1: **Collected dataset statistics.** Number of cases containing certain arteries and branches for all 828 patients.

terpretability of the data, the clinicians use the MPR technique [10, 5]. MPR is the process of using the data from axial CT images to create a more anatomical representation of the coronary artery by tracking the whole specific artery branch along the CT volume and generating its two-dimensional image.

Each branch is represented by 50 MPR images, where one image corresponds to the specific view angle (180 degrees in total). The reason for that is that plaque might be located anywhere along the vessel, and be invisible only from some view angles.

3.2. Labels extraction from medical reports

The condition of coronary artery of each patient is described by the report. It contains the meta information about the person (age, gender, heart rate, etc.), characterization of stenosis score, type of the plaque

and calcium score to all artery sections and branches. The raw reports are not suitable for training classification machine learning algorithms as they do not have any specific category attached to the particular image or at least to a stack of images.

We created the parsing pipeline, which takes the report of the patient as an input and extracts all information relevant to our task. The parsed data include the description of all important artery sections and branches with corresponding stenosis categories. The latter are grouped according to the standard defined by the Society of Cardiovascular Computed Tomography (SCCT) and Coronary Artery Disease - Reporting and Data System (CAD-RADS) [4]: 0% - Normal, 1-24% - Minimal stenosis or plaque with no stenosis, 25-49% Mild stenosis, 50-69% - Moderate stenosis, 70-99% - Severe stenosis, 100% - Total Occlusion. In the reports, three main artery sections are presented: LAD (Left Anterior Descending Artery) with D-1, D-2, D-3 branches; RCA (Right Coronary Artery) with PDA-RCA, PLV-RCA branches; LCx (Left Circumflex Artery) with OM-1, OM-2, OM-3, LCx-PDA, LCx-PLV branches.

3.3. Data labeling process

Specific recommendation for further patient treatment depends largely on the identified level of stenosis [4]. No further cardiac investigation is required unless moderate (50-69%) or higher stage was reported. Preventive therapy and risk factor modification is suggested for minimal or mild stenosis. Due to these specific regulations, we assign one of the three classes for each MPR image: 'no stenosis' for normal cases, 'non-significant' — 1-50% of stenosis detected, 'significant' — critical cases where more than 50% of stenosis is present and instant doctor's atten-

tion is required.

3.4. Challenges in the dataset

After all described preprocessing steps, we obtained a structured labeled dataset, but it is still incomplete and contains noise. One of the main issues is the appearance of several branches on the same image, see example in Figure 3 (b). This increases the risk of mislabeled data as each set of images representing one branch has just one label, which is then assigned to every single image in the set. For example, for the healthy LAD branch all the 50 images retrieved from different viewpoint angles are labeled as being normal. However, if on some of those images a neighbouring vessel (e.g., LCx) with over 50% score of stenosis is partly present, then the model may detect stenosis and misclassify a healthy LAD.

(a) Small portion of useful signal

(b) Several branches on one MPR
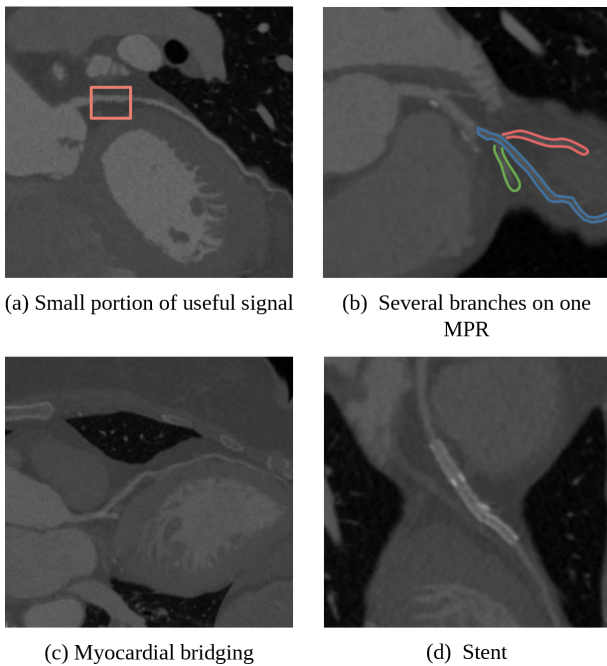
(c) Myocardial bridging

(d) Stent

Figure 3: **Examples of difficult images.** Examples of the hard cases which are present in collected dataset. **(a)** The amount of pixels responsible for region of stenosis is many times smaller than the entire sample. **(b)** While the label for one MPR image corresponds to only one artery segment, several of them might be present on the image. **(c)** Physically natural narrowing can be visually similar to stenosis called myocardial bridging. **(d)** Example of artery with inserted stent, which can be mistakenly classified as stenosis with plaques.

Another issue is an inconsistency between medical reports written by doctors and the labeling sys-

tem prescribed by CAD-RADS. The problem arises when the annotation provided by a radiologist is on the borderline between two classes from the CAD-RADS system. For example, the doctor might mark a specific branch by a "50%" of the stenosis. While it is satisfactory in medical terms, it becomes a challenge for us to choose which group this annotation belongs to - whether it should be considered as a significant or non-significant case.

Also we pre-process MPR images before inference. Each MPR image contains text information (meta-information about the picture), which has the highest intensity on the image. We remove it by assigning the average intensity of the neighbor pixels. Other examples of data challenges and difficulties are represented in Figure 3.

4. Experiments

4.1. General pipeline

We took 708 unique patients for training, 15 for validation and 105 for testing. We include the MPR images of every coronary artery and its branches described in Subsection 3.2 in training and testing phases.

We fed the MPR images into an optimized network architecture ShuffleNet V2. Each branch of the artery in most of the cases is represented by 50 MPR images (see Subsection 3.1). Thus for one branch, we get 50 predictions describing its stenosis score. Then using the majority rule, we assign the final stenosis score for the branch. The prediction pipeline is illustrated in Figure 2, and it is followed by evaluation technique, shown in Figure 4.

4.2. Methodology

The technique which is widely used for optimizing the neural network architectures is 1x1 convolution. The authors of ShuffleNet [21] approached it and managed to reduce the time for this operation. The main idea behind ShuffleNet was to use separable depthwise convolutions [3], grouped convolutions on 1x1 convolution layers - pointwise group, followed by channel shuffle operation. In this paper, we use an improved version of this architecture - ShuffleNet V2 [14], which is not only faster, but also more accurate.

We use ShuffleNet V2, pretrained on ImageNet [6], to extract features from curved MPR images and classify them into three classes: 'no stenosis', 'non-significant stenosis', 'significant stenosis'.
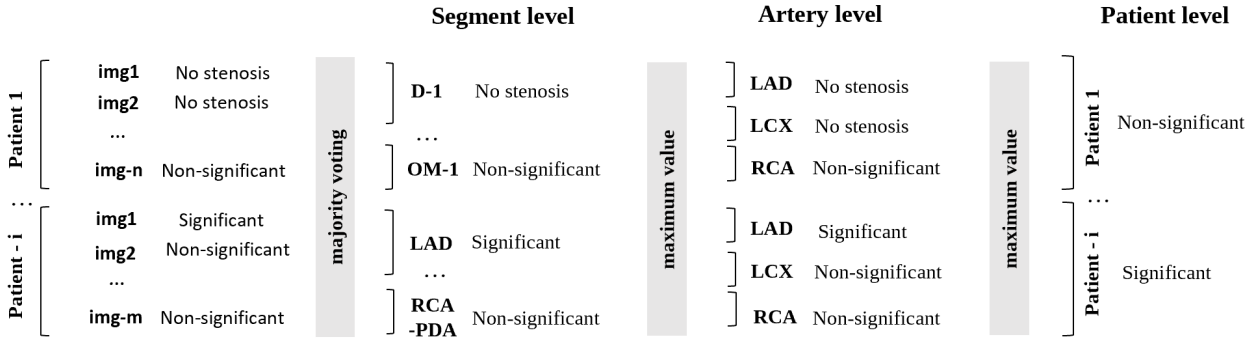
Figure 4: **Evaluation process.** The performance of proposed method is evaluated as suggested in [19] on segment-, artery- and patient-levels using F_1 score and accuracy. For each image, we predict one of the three classes, and then assign the class for the segment (branch) by applying the majority rule to all 50 images corresponding to it. The maximal (most critical) class of all of the branches from one artery is assigned to this artery. The final prediction is calculated for each patient by choosing the maximal class out of all patient's arteries.

The structure of the basic building block of ShuffleNet V2 [14] with residual is displayed in Figure 5. There are several building blocks which are stacked to construct the network. The input of feature channels is split into two branches at the beginning. In each block one branch directly goes through the block and joins the next one. The other branch has three convolution layers with the same input and output channels. Only one from the three 1x1 convolutions is group-wise. The two branches are concatenated after convolutions. For spatial downsampling the block is modified by the removal of split operator. The Channel Shuffle improves accuracy by enabling information communication between different groups of channels.

We trained the model on our dataset using Adam optimizer [11] with 10^{-4} learning rate. We chose the best value of learning rate for our model using LR Range Test [17]. It is a method that implies running the model for a few iterations with initially very small learning rate and then increasing it linearly between low and high learning rate values after each epoch. This allows to estimate the minimum and maximum boundary learning rates. The gradient accumulation and batch normalization [9, 8] were used to increase the batch size and provide a better direction towards a minimum of the loss function.

To introduce robustness properties and desired invariance in our model, we employed standard data augmentation techniques. In the case of MPR images, we primarily need a scale, rotate, blur, bright-

ness, and transpose invariance.

4.3. Evaluation metrics

The CAD-RADS classification is applied on a per-patient basis and represents the highest grade of stenosis from the coronary tree [19]. Taking this into account, we evaluate our method perfor-
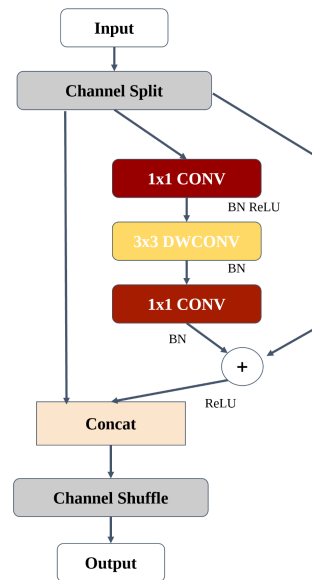

Figure 5: Structure of the basic building block of ShuffleNet V2 with residual [14]. CONV: convolution layer. DWCONV: depthwise convolution. BN: batch normalization. Channel Shuffle: crucial operation for ShuffleNet architectures.

mance on segment-, artery- and patient-levels, and define stenosis score according to the maximum value found at the current level. On the artery level, the highest grade of stenosis is selected out of all the grades on segments. On the patient level, the maximal stenosis stage is chosen among all the arteries. The accuracy and F_1 score for multiclass classification are computed on each level.

We apply weighted averaging for F_1 score measure because we deal with multiclass labels. For each label metrics are calculated and then average weighted by the number of true instances for each class:

$$\sum_{i=1}^{3} F_1(class_i) * W_i$$

where i - iterator over the 3 classes, F_1 is $F_1 - score$ and W - weight for the current class .

Results	Accuracy		F_1 score	
	Our	RNN-based[22]	Our	RNN-based[22]
Segment-level	**0.81**	0.80	**0.81**	0.75
Artery-level	**0.81**	0.76	**0.82**	0.77
Patient-level	**0.80**	0.75	**0.80**	0.75

Table 2: Accuracy and F_1 score on test sets: 105 patients (approximately 25,000 MPR images in total) in our test, 65 patients in Zreik et al.[22]

The final results and comparison to previous study are reported in Table 2. We also display confusion matrices in Table 3 for every level separately in order to show results across the classes. This allows to observe the model's sensitivity (True Positive rate) - cases where the model correctly predicts the positive class, and specificity (True Negative rate) - cases where the model correctly predicts the negative class. For medical problems the False Positive error is always less dangerous than False Negative, these metrics are shown on confusion matrices. None of the patients with significant stenosis were classified as having no stenosis, and none of the healthy patients were put in the significant stenosis category. The model makes mistakes between 'no stenosis' and 'non-significant' classes, as well as between 'non-significant' and 'significant', which could be caused by the noise in data (see Figure 3) and weak labels.

4.4. Results interpretability

In order to achieve model interpretability we use Captum [13] library containing implemented meth-

Segment level	No stenosis	Predicted Non-Significant	Significant
No stenosis	0.92	0.07	0.01
Non-Significant	0.32	0.6	0.08
Significant	0.18	0.26	0.55

Artery level	No stenosis	Predicted Non-Significant	Significant
No stenosis	0.91	0.08	0.01
Non-Significant	0.24	0.67	0.09
Significant	0.13	0.26	0.61

Patient level	No stenosis	Predicted Non-Significant	Significant
No stenosis	0.81	0.19	0.00
Non-Significant	0.07	0.80	0.13
Significant	0.00	0.21	0.79

Table 3: **Confusion matrices.** For each level: segment, artery and patient we calculate confusion matrix to see the number of False Positives, False Negatives and compare them.

ods that identify which training features are important for the model. We visualized the features from the last layer of our model to understand which image regions have the largest impact on the model. The attribution of the network prediction to its input features was performed by applying axiomatic attribution method – Integrated Gradients [18] implemented in Captum [13]. In this approach, the integral of the gradients of the output prediction for the specified class is computed with respect to the input image pixels. We observe that for all types of arteries and all levels of stenosis the model pays attention mostly to the artery zone, while the background with noise does not play a role in classification. Some examples of the features visualization using heatmaps are shown in Figure 8. It is noticeable that for images which contain stenosis, model is more confident in the regions where plaques are located. This demonstrates model reliability, since plaques presence directly correlates with stenosis.

Although our model is capable of handling the background noise information and mainly takes into account only relevant areas, there are some corner cases. One of the most common types of plaque is calcified plaque. In computed tomography scans, it is represented as the pixels with a high level of intensity. There are structures and tissues like calcified ribs, sternum, costal cartilages, ventricle walls, etc., which are in the same range of radiodensity. The visualization showed that our models pay attention to these regions and associate them with the stenosis presence, which causes the lower specificity of our
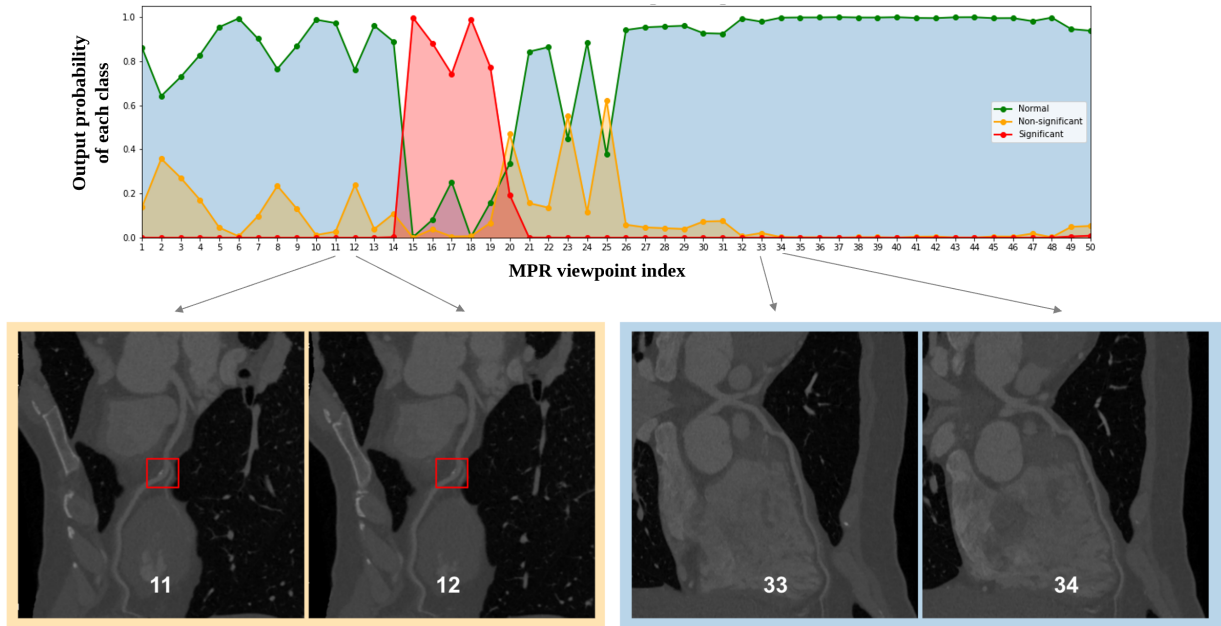
Figure 6: **Prediction for one segment.** The particular example has 25% of stenosis in LAD artery. The labels for all 50 images representing this segment are the same - 'non-significant'. While it is true for some of the MPR view points (see two images at the bottom left), from most of the angles stenosis is not seen either by a human eye, nor by a model (see two images at the bottom right). Thus, it is a direct illustration of weak labeling. It is important to treat every segment as a set of 50 images which are related.
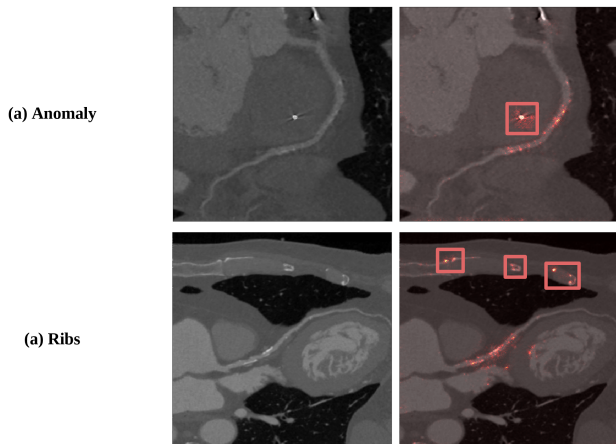
algorithms (see Figure 7).

Figure 7: **Visualization of the model's confusion.** Scans with the structures, which are similar to the calcified plaque. **(a)** The anomaly, with the high level intensity pixels. **(b)** Calcified ribs, which caused the model confusion.

5. Conclusion

We propose a simple automated framework, which is capable of detecting the stenosis score in curved MPR images. Our method shows improvements over the previous results [22] reporting 80% accuracy for the multiclass classification of stenosis level.

Our main contribution lies in creating new dataset of Cardiac CT scans of more than 800 patients, which is larger than any previous dataset, and suggesting a new approach for stenosis level classification. The proposed method omits centerline extraction and does not require any handcrafted features. Furthermore, we obtain explainable results and display features which impacted network's decisions.

The model interpretability through visualization of feature importance is very helpful in medical imaging as radiology specialists may use it to build trust in the model's predictions, refine classification of borderline cases, as well as gather observations for future testing.

6. Discussion

There are several ways to improve our approach. The curved MPR images, which were used as input in this work, contain not only artery, but also the background, where there is no useful information for determining the level of stenosis. We believe that the results of the proposed approach can be improved by using segmentation as an additional step during pre-
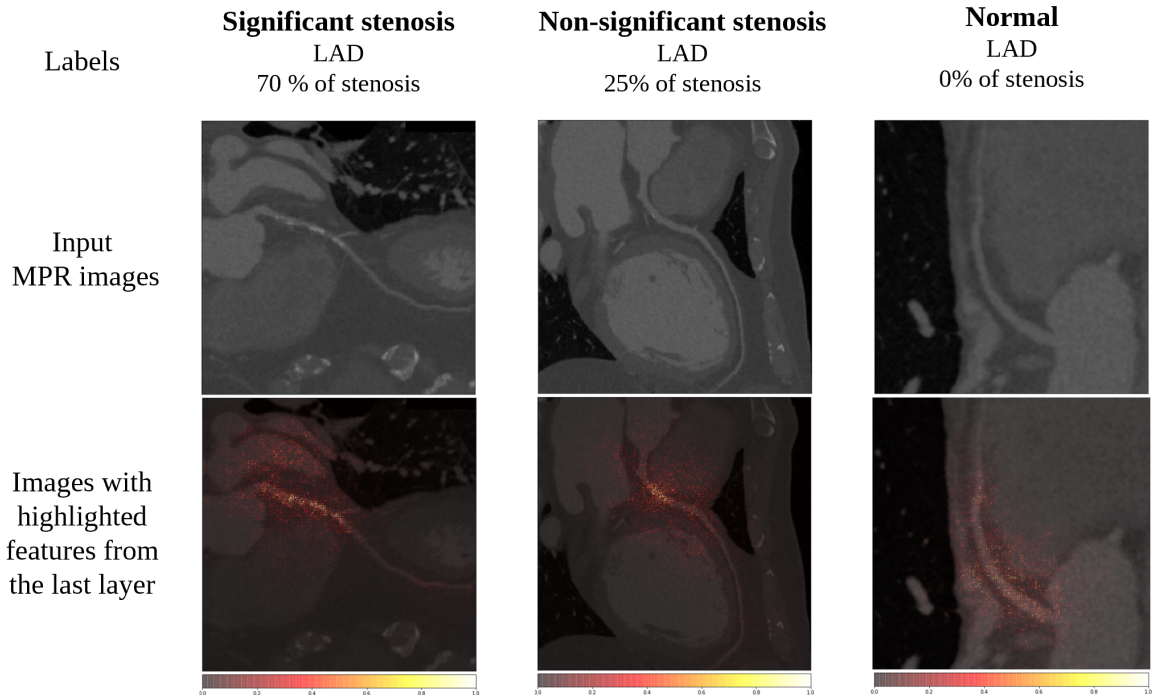
	Significant stenosis	**Non-significant stenosis**	**Normal**
Labels	LAD	LAD	LAD
	70 % of stenosis	25% of stenosis	0% of stenosis

Figure 8: **Visualizations of the last layer's most important features for correctly predicted cases**. These were created by using Integrated Gradients [18]. Each class is represented for three examples of Left Anterior Descending artery: significant, non-significant, no stenosis. **Top**: The input two-dimensional MPR images of three different patients generated from CCTA scans. **Bottom**: The gradient visualization of most impactful features for the model. The brighter the pixels are, the more importance they have in prediction.

processing and feeding the network with segmented images, where only the artery region is present.

Taking into account the problems shown on Figure 7, we might improve the performance of our models by adding the attention gate [16] to the current network architecture. It will automatically learn the relevant areas for our task and suppress the unrelated target structures.

To obtain the final stenosis score for one branch, we take 50 predictions of our network for each corresponding MPR image representing the artery and assign the prevailing class. With this approach, we do not take into account the spatial relationship between MPR images. One possible improvement might be to apply the 3D CNNs to catch patterns across three spatial dimensions. One of the options to skip the step of MPR extraction is to create a new method, which would directly use 3D images of CCTA for stenosis score classification.

Due to the difficulty in collecting reliable labels for medical data, unsupervised or weakly-supervised approaches should be considered. We believe that one of the possible ways of implementing such a so-lution is to train autoencoder exclusively on normal images with noise, then, to decide whether the particular MPR image represents the normal case based on its distance to the corresponding image generated by the model.

The other area for research is the extraction of MPR images from CCTA. This task is handled semi-manually by a radiologist at the clinics, therefore, it is costly and takes long time. Our dataset already contains extracted MPR data. We think this process can be simplified by application of deep learning in building multiplanar reconstruction images based only on data from CCTA scans.

Acknowledgements

References

[1] CT coronary angiography in patients with suspected angina due to coronary heart disease (SCOT-HEART): an open-label, parallel-group, multicentre trial. *The Lancet*, 385(9985):2383–2391, June 2015. 2

[2] E. J. Benjamin, P. Muntner, A. Alonso, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, A. R. Chang, S. Cheng, S. R. Das, F. N. Delling, L. Djousse, M. S. Elkind, J. F. Ferguson, M. Fornage, L. C. Jordan, S. S. Khan, B. M. Kissela, K. L. Knutson, T. W. Kwan, D. T. Lackland, T. T. Lewis, J. H. Lichtman, C. T. Longenecker, M. S. Loop, P. L. Lutsey, S. S. Martin, K. Matsushita, A. E. Moran, M. E. Mussolino, M. O'Flaherty, A. Pandey, A. M. Perak, W. D. Rosamond, G. A. Roth, U. K. Sampson, G. M. Satou, E. B. Schroeder, S. H. Shah, N. L. Spartano, A. Stokes, D. L. Tirschwell, C. W. Tsao, M. P. Turakhia, L. B. VanWagner, J. T. Wilkins, S. S. Wong, and S. S. V. and. Heart disease and stroke statistics — a report from the american heart association. *Circulation*, 139(10), Mar. 2019. 1

[3] F. Chollet. Xception: Deep learning with depthwise separable convolutions, 2016. 4

[4] R. C. Cury, S. Abbara, S. Achenbach, A. Agatston, D. S. Berman, M. J. Budoff, K. E. Dill, J. E. Jacobs, C. D. Maroules, G. D. Rubin, F. J. Rybicki, U. J. Schoepf, L. J. Shaw, A. E. Stillman, C. S. White, P. K. Woodard, and J. A. Leipsic. CAD-RADSTM coronary artery disease – reporting and data system. an expert consensus document of the society of cardiovascular computed tomography (SCCT), the american college of radiology (ACR) and the north american society for cardiovascular imaging (NASCI). endorsed by the american college of cardiology. *Journal of Cardiovascular Computed Tomography*, 10(4):269–281, July 2016. 3

[5] N. C. Dalrymple, S. R. Prasad, M. W. Freckleton, and K. N. Chintapalli. Introduction to the language of three-dimensional imaging with multidetector CT. *RadioGraphics*, 25(5):1409–1428, Sept. 2005. 3

[6] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009. 4

[7] P. S. Douglas, U. Hoffmann, M. R. Patel, D. B. Mark, H. R. Al-Khalidi, B. Cavanaugh, J. Cole, R. J. Dolor, C. B. Fordyce, M. Huang, M. A. Khan, A. S. Kosinski, M. W. Krucoff, V. Malhotra, M. H. Picard, J. E. Udelson, E. J. Velazquez, E. Yow, L. S. Cooper, and K. L. Lee. Outcomes of anatomical versus functional testing for coronary artery disease. *New England Journal of Medicine*, 372(14):1291–1300, Apr. 2015. 2

[8] J. Hermans, G. Spanakis, and R. Möckel. Accumulated gradient normalization, 2017. 5

[9] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift, 2015. 5

[10] A. Kanitsar, D. Fleischmann, R. Wegenkittl, P. Felkel, and E. Groller. CPR - curved planar reformation. In *IEEE Visualization, 2002. VIS 2002*. IEEE. 3

[11] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization, 2014. 5

[12] Kirişli et al. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. *Medical image analysis*, 17(8):859–876, 2013. 2

[13] N. Kokhlikyan, V. Miglani, M. Martin, E. Wang, J. Reynolds, A. Melnikov, N. Lunova, and O. Reblitz-Richardson. Pytorch captum. `https://github.com/pytorch/captum`, 2019. 6

[14] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun. ShuffleNet v2: Practical guidelines for efficient CNN architecture design. In *Computer Vision – ECCV 2018*, pages 122–138. Springer International Publishing, 2018. 4, 5

[15] C. T. Metz, M. Schaap, A. C. Weustink, N. R. Mollet, T. van Walsum, and W. J. Niessen. Coronary centerline extraction from CT coronary angiography images using a minimum cost path approach. *Medical Physics*, 36(12):5568–5579, Nov. 2009. 2

[16] J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker, and D. Rueckert. Attention gated networks: Learning to leverage salient regions in medical images, 2018. 8

[17] L. N. Smith. Cyclical learning rates for training neural networks. In *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, Mar. 2017. 5

[18] M. Sundararajan, A. Taly, and Q. Yan. Axiomatic attribution for deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 3319–3328, 2017. 6, 8

[19] A. R. van Rosendael, I. J. van den Hoogen, A. M. Bax, S. J. Al'Aref, O. A. H. Alawamlh, D. Larine, and J. K. Min. CT and calcification. In *Coronary Calcium*, pages 83–123. Elsevier, 2019. 5

[20] J. M. Wolterink, R. W. van Hamersvelt, M. A. Viergever, T. Leiner, and I. Išgum. Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier. *Medical Image Analysis*, 51:46–60, Jan. 2019. 2

[21] X. Zhang, X. Zhou, M. Lin, and J. Sun. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In *2018 IEEE/CVF Conference*

on Computer Vision and Pattern Recognition. IEEE, June 2018. 4

[22] Zreik et al. A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography. *IEEE Transactions on Medical Imaging*, 38(7):1588–1598, 2019. 2, 6, 7