Stenosis detection in coronary angiography images using deep learning models

1st Rafael M. Luque-Baena
Department of Computer Languages and Computer Science
University of Malaga
Malaga, Spain
rmluque@lcc.uma.es

2nd Irene Romero Granados
Department of Computer Languages and Computer Science
University of Malaga
Malaga, Spain
ireero99@uma.es

3rd Ariadna Jiménez-Partinen
Department of Computer Languages and Computer Science
University of Malaga
Malaga, Spain
ariadnapartinen@uma.es

4th Esteban J. Palomo
Department of Computer Languages and Computer Science
University of Malaga
Malaga, Spain
ejpalomo@lcc.uma.es

5th Manuel Jiménez-Navarro
Hospital Universitario Virgen de la Victoria
C/ Doctor Miguel Díaz Recio, 28, 29010
Malaga, Spain
jimeneznavarro@secardiologia.es

***Abstract*—The emergence of deep learning has caused its massive application to different fields in industry and research, among which is the clinical field, especially in those where the data is structured in the form of images or video. The present proposal intends to develop a coronary angiography image analysis system based on artificial intelligence. These images are radiocontrast X-ray images of the coronary arteries. The proposed system will be able to analyze these coronary angiography images of patients with no obstructive coronary lesions to detect and characterize smooth and irregular coronary arteries and predict the presence of cardiovascular events during follow-up. Deep learning convolutional artificial neural networks will support the algorithmic basis of the proposed system.**

***Index Terms*—stenosis detection, coronary angiography images, convolutional neural networks**

I. Introduction

Nowadays, cardiovascular diseases remain to be the leading cause of death in several countries, especially in developed countries [1], making it necessary to continue and maintain prevention protocols and correct diagnosis of patients with cardiovascular heart disease (CHD). One of the many possible causes of a heart attack is suffering a cardiovascular artery disease (CAD), such as the presence of stenosis in the coronary arteries. Stenosis is understood as the narrowing or reduction of the lumen of a conduit or orifice in the human body, in this case, the coronary arteries, which prevent or hinder blood flow to the heart [2]. Invasive Coronary Angiography (ICA) is considered the gold standard when there is suspected a CAD [3].

ICA is based on introducing a specialized catheter of about 2 millimeters in diameter through a percutaneous incision in the femoral or brachial artery, situated in the groin and the arm, respectively. Once the catheter reaches the coronary arteries, it injects a radiopaque contrast media, staining them. In this way, it allows the angiographer, an X-ray medical team, to define the coronary anatomy of the patient, showing the state of the coronary arteries and if there is a luminal obstruction, its degree. The physicians analyzed the image acquired and the information obtained includes the identification of the area in which is located the lesion, its length and diameter, and the nature of the obstruction, such as a thrombus or a dissection [4]. This precise information allows for making a complete diagnosis.

This analysis requires a highly experienced physician with extensive knowledge of coronary artery anatomy because commonly the assessment of the percentage of lesions is done visually. Therefore, the interpretation of the ICAs has an important subjective component and interobserver variability, that is accompanied by other facts, such as the presence

This work is partially supported by the Ministry of Science, Innovation and Universities of Spain under grant RTI2018-094645-B-I00, project name Automated detection with low-cost hardware of unusual activities in video sequences. It is also partially supported by the Autonomous Government of Andalusia (Spain) under project UMA18-FEDERJA-084, project name Detection of anomalous behavior agents by deep learning in low-cost video surveillance intelligent systems. It is also partially supported by the Autonomous Government of Andalusia (Spain) under project UMA20-FEDERJA-108, project name Detection, characterization and prognosis value of the non-obstructive coronary disease with deep learning. All of them include funds from the European Regional Development Fund (ERDF). It is also partially supported by the University of Malaga (Spain) under grants B1-2019_01, project name Anomaly detection on roads by moving cameras, and B1-2019_02, project name Self-Organizing Neural Systems for Non-Stationary Environments.

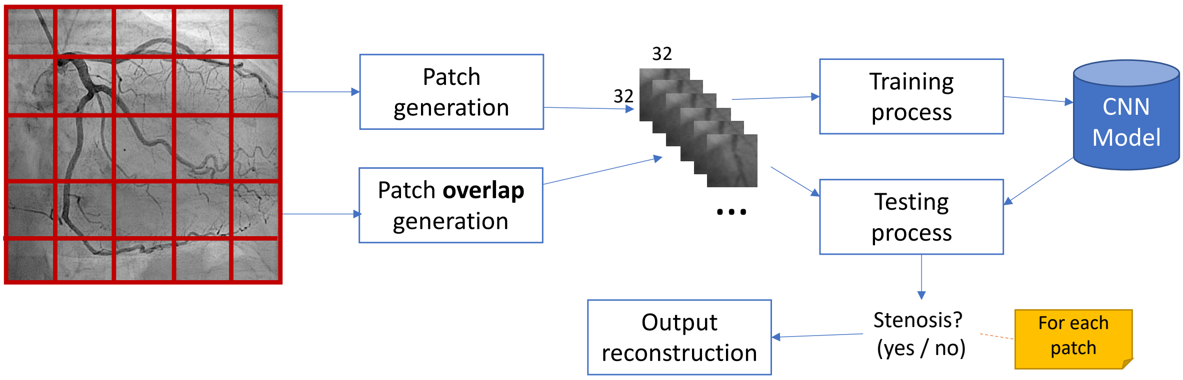

Fig. 1. The framework of the proposed methodology.

of alterations by drugs administered [4]. In this context, a misrepresentation could have severe clinical consequences, because during the catheterization procedure the cardiologist has to evaluate if the patient present stenosis and its diagnosis, to decide if an angioplasty, placing a stent, would be necessary or not, which depends on the degree of the lesion.

To help improve the clinical diagnosis of these cardiovascular heart diseases, deep learning techniques are used, which are being widely used for the analysis of medical images. Thanks to this computational paradigm, it is possible to know if a patient has stenosis and, more importantly, where it is present [5], [6]. Information that is very valuable to place the stent in the right place [7].

The main objective of this proposal is to automatically detect and localize the CAD from ICA images through deep learning techniques. This way, we could help cardiologists detect those lesions that have an uncertain prognosis and is currently visually carried out. Additionally, a comparative study will be carried out between different classification neural networks based on convolutional models.

II. Methodology

Figure 1 shows a schematic of the framework developed for this work. To take advantage of some public data sets to detect stenosis [6], [7], whose images are divided into patches or regions, the proposal will be framed within a classification system and not a detection and localization one. Thus, the first task will be to divide the image into 32x32 patches or regions. This division will not entail overlap for the training of the convolutional network models. We will compare known architectures, such as Resnet, EfficientNet or MobileNet, adapting the input layer to take reduced images (32x32 grayscale patches) and the output layer to work with a binary classification problem. Subsequently, these already defined models will be trained to classify new regions as stenosis or non-stenosis. These models will be stored to be used in the inference phase of new coronary angiography images.

A. Data augmentation

Having unbalanced sets would make the model results unacceptable or unreliable, as the model will accurately predict the majority class, while being relatively inefficient at classifying the minority class. This situation arises in our proposal. There is a much higher number of regions labeled as non-stenosis than regions labeled as stenosis in any of the available public data sets. Different data augmentation techniques have been applied to the training set to solve this issue. In addition, an stratified cross-validation will be used to improve the results further.

Therefore, using this methodology the initial set will be divided into 5 boxes with samples from both classes. For each K the initial set will be divided into training (80%) and test (20%) and different data augmentation techniques will be applied only to the training set of the minority class (stenosis). In this way, for each K, the training set, as well as the test set, will be different and the model will have more different data with which to learn and with which to predict.

1) Sampling and augmentation: The sets of stenosis and non stenosis have been balanced by equalizing their samples to a midpoint, so that the minority class is increased and the number of samples of the majority class is reduced. Multiple images with different transformations from the original ones are generated for this data augmentation technique. The modifications applied are 15 degrees of rotation, 0.20 displacement in width, and horizontal and vertical flip. This process provides five hundred new images from the training set with stenosis. An under-sampling is performed in the majority class, preserving six hundred images with no stenosis randomly selected. The final training set will consist of a balanced number of samples for each class, 500 augmented + 90 actual stenosis, and 600 non-stenosis regions.

2) Minority class augmentation: For this data augmentation technique, a more aggressive balancing has been performed than the one used in the previous section since the minority class will be equalized to the majority class. The same types of transformations are used as in the last technique. We start from about 90 images that present stenosis of the training set and generate 1000 new images. The result is a balanced training set with 1000 images for each class.

3) Proportional data augmentation: In this technique, the training and validation set has been augmented without considering the labels, i.e., they have been augmented regardless of whether they present stenosis or not. A rotation of 10 degrees, an offset of 0.05 in width, and a horizontal flip have been applied to both sets. Five hundred images have been generated for training and one hundred for validation. Note that the unbalanced ratio will still exist when generating new images of both classes.

4) Synthetic data: Another technique that has been employed is an algorithm that generates synthetic regions from scratch [7]. This algorithm first draws random background gradients, uses Bézier curves to represent the veins, adds white noise, and applies Gaussian blur to the whole image. This technique makes it possible to generate any number of images for both classes. These synthetic images are easily distinguishable from the real ones. Especially the natural set of images without stenosis has many images with borders, while in the synthetic set, there is always at least one vein. Therefore, it is necessary to train the model first with the synthetic data and a second time with the actual training set.

B. Convolutional Neural Networks

Convolutional Neural Networks (CNN) is a type of model of deep learning which incorporate at least one convolutional layer, whose proposal is to extract the features of the input image. The former convolutional layers extract the basic features and the later layers extract detailed features from the basis. The convolutional layers are based on applying convolutional kernels, which work as a kernel filter being activated with a specific condition [8]. CNNs are mainly composed of a sequence of convolutional layers and other basic operations, such as pooling layer, based on simplifying the given input; activation, which returns a featured map with a selection of a threshold; fully connected layer, which is a classification layer, whose output is formed of as many nodes as there are classes; and softmax layer that gives a probability classification [9].

Residual Networks (*ResNets*) are characterized by the introduction of shortcut connections, which allow skipping some blocks of convolutional layers of the original network and then normalized by batches and not linearly. Specifically, in this study *ResNet50* is used, which is characterized by being composed of 50 layers deep [10].

EfficientNet is considered as a group of convolutional models, that are between B0 and B7. These models stood out because they are based on the implementation of a different activation function called Swish, instead of the commonly used Rectifier Linear Unit (ReLU), improving the performance of classification problems [11]. Also, *EfficentNet* is characterized by achieving more efficient results by scaling the three dimensions of a CNN: depth, which is the number of layers; width, that is the number of filters; and resolution, which is the size of the feature map; instead of focusing only on one of them, as other models do [12].

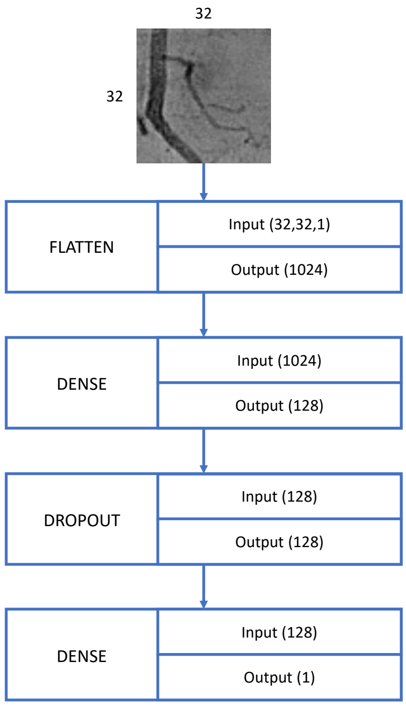

Fig. 2. The structure of the custom CNN proposed.

MobileNet is characterized by less number of parameters in comparison with other known architectures, which decreases the computational load. The MobileNet architecture is based on depthwise separable convolutions, which are composed of two layers: depthwise convolution, which applies a single filter to the input without extracting features; and then pointwise convolution, which creates a linear combination output with new features, while standard convolutions do it in one step [13].

A simpler neural network is also considered in comparison with the previous approaches (Figure 2). In it, we will not take into account the neighborhood of the pixels, therefore using only dense layers. Thus, the input image will initially be vectorized, subsequently applying a dense layer, followed by a ReLU activation. Next is a dropout layer, which helps avoid overfitting by randomly setting 50% of the input to 0. Finally, the last hidden layer is dense with one output node, which applies a sigmoid activation as a function output.

The processing of new images not used for model training will be similar to that performed in the training phase. Initially, and from the new input image to locate and detect stenosis, a set of overlapping patches will be generated, storing the image pixel to which each one is associated. The previously trained model will evaluate each of these regions to determine if it corresponds to a region where stenosis is observed or not. After analyzing all the patches, a reconstruction process of the resulting image is carried out, where the pixels of each region with stenosis will be marked. Since a pixel in the image can be associated with several regions or patches due to overlapping

TABLE I

RESULTS OBTAINED ON THE TEST SET USING THREE MEASURES: BALANCED ACCURACY, SENSITIVITY AND SPECIFICITY. THE HIGHER THE BETTER THE PERFORMANCE OF THE MODEL. FOUR CONVOLUTIONAL NETWORK ARCHITECTURES HAVE BEEN TRAINED APPLYING FIVE DIFFERENT DATA AUGMENTATION STRATEGIES.

Data Augmentation Strategy	Measures	Convolutional network model			
		Custom	*ResNet50*	*EfficientNetB4*	*MobileNet*
None	Balanced Acc	0.500 ± 0.00	0.499 ± 0.00	0.500 ± 0.00	0.500 ± 0.00
	Specificity	1.000 ± 0.00	0.999 ± 0.00	1.000 ± 0.00	1.000 ± 0.00
	Sensibility	0.000 ± 0.00	0.000 ± 0.00	0.000 ± 0.00	0.000 ± 0.00
Sampling and data augmentation	Balanced Acc	0.773 ± 0.02	0.933 ± 0.01	0.915 ± 0.02	0.872 ± 0.19
	Specificity	0.686 ± 0.04	0.956 ± 0.02	0.963 ± 0.03	0.976 ± 0.02
	Sensibility	0.860 ± 0.05	$\mathbf{0.909 \pm 0.03}$	0.868 ± 0.05	0.767 ± 0.39
Minority class augmentation	Balanced Acc	0.808 ± 0.03	$\mathbf{0.940 \pm 0.03}$	0.931 ± 0.03	0.902 ± 0.05
	Specificity	0.715 ± 0.04	0.980 ± 0.00	0.986 ± 0.01	$\mathbf{0.987 \pm 0.00}$
	Sensibility	0.901 ± 0.06	0.900 ± 0.07	0.875 ± 0.05	0.817 ± 0.09
Proportional data augmentation	Balanced Acc	0.500 ± 0.00	0.842 ± 0.08	0.835 ± 0.14	0.784 ± 0.17
	Specificity	1.000 ± 0.00	0.984 ± 0.02	0.994 ± 0.00	0.991 ± 0.01
	Sensibility	0.000 ± 0.00	0.701 ± 0.17	0.675 ± 0.29	0.576 ± 0.35
Synthetic data [7]	Balanced Acc	0.500 ± 0.00	0.500 ± 0.00	0.792 ± 0.20	0.764 ± 0.14
	Specificity	1.000 ± 0.00	1.000 ± 0.00	0.991 ± 0.01	0.989 ± 0.01
	Sensibility	0.000 ± 0.00	0.000 ± 0.00	0.592 ± 0.40	0.540 ± 0.27

in creating the set of regions, there will be sets of pixels that are more likely to have stenosis than others, given that they have been marked more frequently. It should note that this final image will be the same size as the input coronary angiography image.

III. EXPERIMENTAL RESULTS

The data set used [7] has 1394 patches without stenosis and 125 with stenosis, with a size of 32x32. It is pretty evident that it suffers from a class imbalance problem and requires some process of balancing or augmenting the minority class, which has previously described in section II-A. Experiments have been carried out using a k-fold methodology, studying different convolutional architectures and applying different data augmentation strategies.

A. Metrics

In order to obtain good results in the model evaluation stage, it is necessary to define metrics that are suitable for our dataset. The first thing to consider when choosing such metrics is to know which characteristics of the dataset to be evaluated. In this study, the data is highly unbalanced, with a majority class with 1394 samples corresponding to negative in stenosis and a minority class with only 125 regions corresponding to positive in stenosis.

Since the accuracy measure is unsuitable for unbalanced data, the balanced accuracy metric is considered for evaluation. This measure is especially useful in classification problems where the classes are significantly unbalanced. It reflects the model's accuracy in correctly classifying both classes without the number of samples of each class being relevant:

$$BalancedAccuracy = \frac{Sensitivity + Specificity}{2} \quad (1)$$

where:

- *Sensitivity* (TPR): measures the ability of the model to detect as positive those samples that are actually positive.

$$Sensitivity = \frac{True_{positive}}{True_{positive} + False_{negative}} \quad (2)$$

- *Specificity* (TNR): measures the ability of the model to detect as negative those samples that are actually negative.

$$Specificity = \frac{True_{negative}}{True_{negative} + False_{positive}} \quad (3)$$

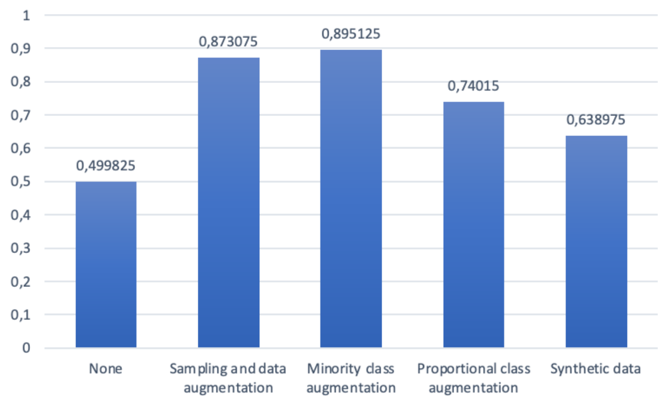

Fig. 3. Balanced accuracy with different data augmentation strategies.

B. Results

The experimental results using different neural models and diverse data augmentation strategies are observed in Table I. The starting point is the *Custom* network that, due to the class imbalance problem, is not able to predict any case of stenosis. Nevertheless, an using data augmentation techniques, the combination that provided the best results was the one that performed minority class augmentation, obtaining a balanced accuracy of 0.808, and the one that performed sampling and data augmentation, obtaining a balanced accuracy of 0.773.

Attempts have been made to further improve the results by making use of pre-trained networks. The three combinations that provided the best results were the augmentation of the minority class with *ResNet50*, reaching the maximum balanced accuracy obtained in the experiment with 0.940. With sampling and data augmentation and *ResNet50* we obtained a balanced accuracy of 0.933 and finally, augmenting the minority class with *EfficientNetB4*, we reached 0.931 of balanced accuracy.

Figures 3 and 4 show the result of the balanced accuracy measure analyzing the data augmentation strategies and the convolutional network proposals independently. It can be stated then that the best data augmentation technique for this problem is the minority class augmentation combined with the pre-trained network *ResNet50*. The loss and balanced accuracy during the epochs in the training phase can be observed in Figure 6. By analyzing the sensitivity measure, a 0.90 is reached. That implies the model is very good at predicting the stenosis class, which is the most relevant for our study. Figure 5 shows the confusion matrix of this experiment. The values are the classification results over the test set, on average, from the cross-validation strategy. It can be seen that out of the total average samples in the test set (303 regions), there are only two false negatives (classified as non-stenosis when it is stenosis) and six false positives (classified as stenosis when it is non-stenosis).

The final step of the proposed methodology involves reconstructing the input image after splitting it into tiles to infer each region over the trained neural model. Thus, the process consists in resizing the input image to 128x128, dividing it into non-overlapping tiles/regions of 32x32, and generating 16 regions to analyse. The best pre-trained network is used to infer each region and determine if it presents stenosis or not. Finally, a reconstruction of the output image is performed considering the class (stenosis or not) associated with each tile. Figure 7 shows an example of the output image from a coronary angiography image.

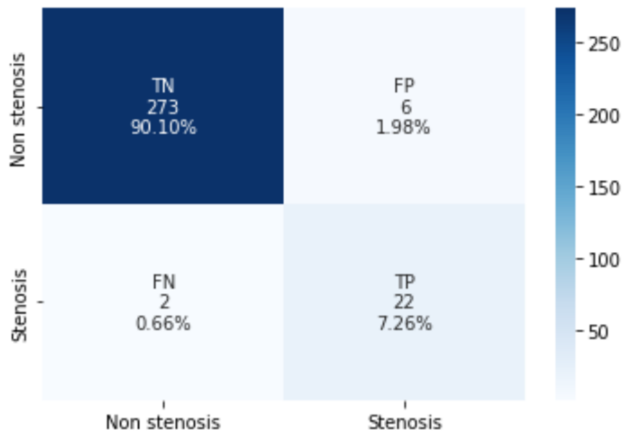

Fig. 5. Average confusion matrix using a K-Fold methodology (K=5) with ResNet50 model and minority class augmentation strategy.

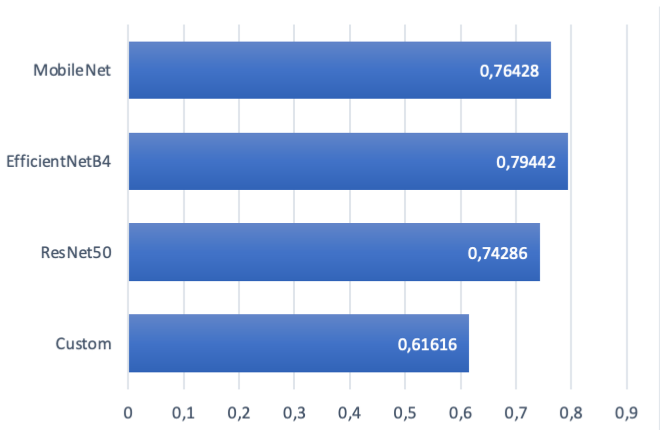

Fig. 4. Balanced accuracy over four convolutional neural networks.

IV. Conclusions

This paper presents a study for the classification of images with stenosis. The image is divided into 32x32 regions to address the problem, and each region is classified using a deep convolutional network. For the study to be complete, different models and strategies for data augmentation have been tested since it is a non-balanced problem. It is observed that the best combination consists of the *ResNet50* pre-trained network together with the minority class augmentation strategy, with 0.94 of balanced accuracy in the test set.

References

[1] F. Sanchis-Gomar, C. Perez-Quilis, R. Leischik, and A. Lucia, "Epidemiology of coronary heart disease and acute coronary syndrome," *Annals of translational medicine*, vol. 4, no. 13, 2016.

[2] W. C. Little, M. Constantinescu, R. Applegate, M. Kutcher, M. Burrows, F. Kahl, and W. Santamore, "Can coronary angiography predict the site of a subsequent myocardial infarction in patients with mild-to-moderate coronary artery disease?" *Circulation*, vol. 78, no. 5, pp. 1157–1166, 1988.

[3] J.-P. Collet, H. Thiele, E. Barbato, O. Barthélémy, J. Bauersachs, D. L. Bhatt, P. Dendale, M. Dorobantu, T. Edvardsen, T. Folliguet *et al.*, "2020 esc guidelines for the management of acute coronary syndromes in patients presenting without persistent st-segment elevation: the task force for the management of acute coronary syndromes in patients presenting without persistent st-segment elevation of the european society of cardiology (esc)," *European heart journal*, vol. 42, no. 14, pp. 1289–1367, 2021.

[4] P. J. Scanlon, D. P. Faxon, A.-M. Audet, B. Carabello, G. J. Dehmer, K. A. Eagle, R. D. Legako, D. F. Leon, J. A. Murray, S. E. Nissen *et al.*, "Acc/aha guidelines for coronary angiography: a report of the american college of cardiology/american heart association task force on practice

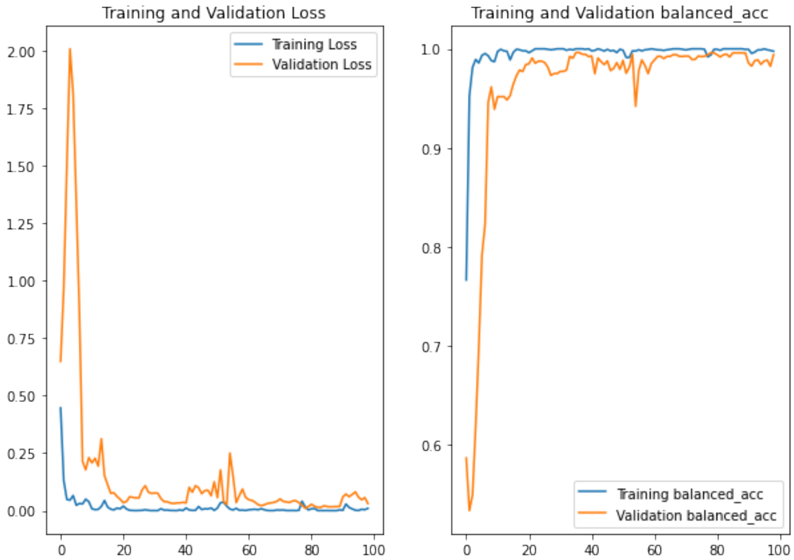

Fig. 6. Loss and balanced accuracy for *ResNet50* architecture with minority class augmentation strategy.

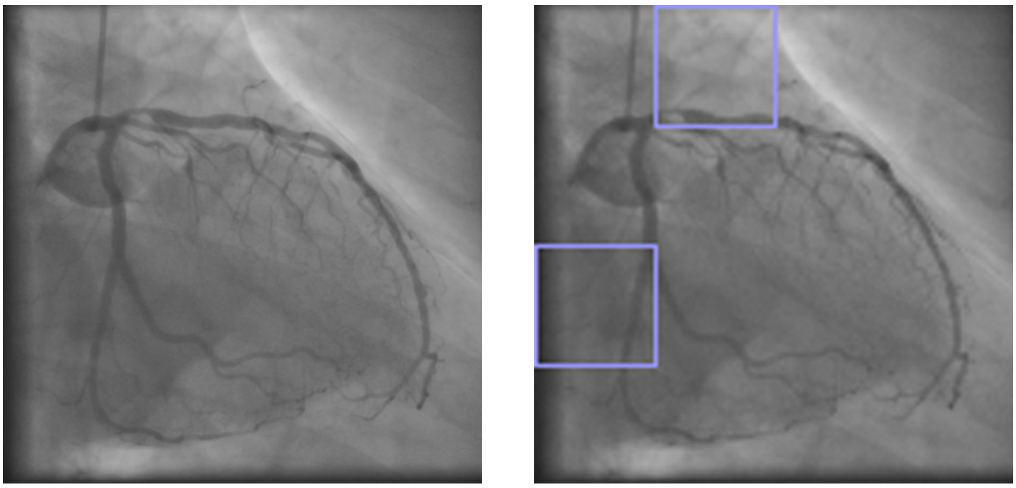

Fig. 7. Input coronary angiography image at left and output reconstruction after applying our proposed methodology at right. Two stenosis regions are detected and highlighted with blue squares.

guidelines (committee on coronary angiography) developed in collaboration with the society for cardiac angiography and interventions," *Journal of the American College of Cardiology*, vol. 33, no. 6, pp. 1756–1824, 1999.

[5] D. Liang, J. Qiu, L. Wang, X. Yin, J. Xing, Z. Yang, J. Dong, J. Dong, J. Dong, and Z. Ma, "Coronary angiography video segmentation method for assisting cardiovascular disease interventional treatment," *BMC Medical Imaging*, vol. 20, no. 1, 2020.

[6] E. Ovalle-Magallanes, J. G. Avina-Cervantes, I. Cruz-Aceves, and J. Ruiz-Pinales, "Transfer learning for stenosis detection in x-ray coronary angiography," *Mathematics*, vol. 8, no. 9, 2020.

[7] Antczak, Karol and Liberadzki, Lukasz, "Stenosis detection with deep convolutional neural networks," *MATEC Web Conf.*, vol. 210, p. 04001, 2018. [Online]. Available: https://doi.org/10.1051/matecconf/201821004001

[8] J. Wang, H. Zhu, S.-H. Wang, and Y.-D. Zhang, "A review of deep learning on medical image analysis," *Mobile Networks and Applications*, vol. 26, no. 1, pp. 351–380, 2021.

[9] Y. Fujisawa, S. Inoue, and Y. Nakamura, "The possibility of deep learning-based, computer-aided skin tumor classifiers," *Frontiers in medicine*, vol. 6, p. 191, 2019.

[10] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. De Geus, "Malicious software classification using transfer learning of resnet-50 deep neural network," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2017, pp. 1011–1014.

[11] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using efficientnet deep learning model," *Ecological Informatics*, vol. 61, p. 101182, 2021.

[12] L. T. Duong, P. T. Nguyen, C. Di Sipio, and D. Di Ruscio, "Automated fruit recognition using efficientnet and mixnet," *Computers and Electronics in Agriculture*, vol. 171, p. 105326, 2020.

[13] W. Sae-Lim, W. Wettayaprasit, and P. Aiyarak, "Convolutional neural networks using mobilenet for skin lesion classification," in *2019 16th international joint conference on computer science and software engineering (JCSSE)*. IEEE, 2019, pp. 242–247.