Automatic characterization of myocardial perfusion imaging polar maps employing deep learning and data augmentation

Ioannis D. Apostolopoulos[1] MSc,
Nikolaos D. Papathanasiou[2] MD, PhD,
Trifon Spyridonidis[2] MD,
Dimitris J. Apostolopoulos[2] MD, PhD

1. University of Patras, Medical School, Department of Medical Physics, Rio, Achaia, Greece
2. University of Patras, Medical School, Department of Nuclear Medicine, Rio, Achaia, Greece

Keywords: Coronary artery disease - Deep learning - Convolutional neural networks - Myocardial Perfusion Imaging

Corresponding authors:
Ioannis D. Apostolopoulos
University of Patras, Medical School, Department of Medical Physics, Rio, Achaia, PC 26504, Greece
Tel.: +30 6973866965
ece7216@upnet.gr

Received:
13 April 2020
Accepted revised:
22 June 2020

Abstract

Objective: To investigate a deep learning technique, more specifically state-of-the-art convolutional neural networks (CNN), for automatic characterization of polar maps derived from myocardial perfusion imaging (MPI) studies for the diagnosis of coronary artery disease. **Subjects and Methods:** Stress and rest polar maps corresponding to 216 patient cases from the database of the department of Nuclear Medicine of our institution were analyzed. Both attenuation-corrected (AC) and non-corrected (NAC) images were included. All patients were subjected to invasive coronary angiography within 60 days from MPI. As the initial dataset of this study was small to train a deep learning model from scratch, two strategies were followed. The first is called transfer learning. For this, we employed the state-of-the-art CNN called VGG16, which has been broadly exploited in medical imaging classification tasks. The second strategy involves data augmentation, which is achieved by the rotation of the polar maps, to expand the training set. We evaluated VGG16 with 10-fold cross-validation on the original set of images performing separate experiments for AC and NAC polar maps, as well as for their combination. The results were compared to the standard semi-quantitative polar map analysis based on summed stress and summed difference scores, as well as to the medical experts' diagnostic yield. **Results:** With reference to the findings of coronary angiography, VGG16 achieved an accuracy of 74.53%, sensitivity 75.00% and specificity 73.43% when the AC and NAC polar maps were incorporated into one single image set. Respective figures of MPI interpretation by experienced Nuclear Medicine physicians were 75.00%, 76.97% and 70.31%. The accuracy of semi-quantitative polar map analysis was lower, 66.20% and 64.81% for AC and NAC technique, respectively. **Conclusions:** The proposed deep learning model with data augmentation techniques performed better than the conventional semi-quantitative polar map analysis and competed with doctor's expertise in this particular patient cohort and ima-ge set. The model could potentially serve as an assisting tool to support interpretation of MPI studies or could be used for teaching purposes.

Hell J Nucl Med 2020; 23(2): 125-132 *Epub ahead of print: 27 July 2020* *Published online: 24 August 2020*

Introduction

Myocardial perfusion imaging (MPI) with single photon emission computed tomography (SPECT) is a well-established method for the identification of hemodynamically significant coronary artery disease (CAD), the assessment of major adverse cardiac event risk and the evaluation of myocardial viability [1, 2]. However, artifacts from tissue attenuation of radioactivity degrade the quality of MPI studies and pose diagnostic dilemmas. To address this matter, SPECT prone imaging, external radioactive sources providing a transmission map for attenuation correction (AC), as well as various software techniques, have been proposed [3]. The advent of hybrid SPECT/CT technology enabled computed tomography (CT) images to serve as transmission maps for the AC of SPECT data [4]. Images produced by CT hold many advantages: they are obtained in seconds or a few minutes (depending on the CT device) and generally provide high quality transmission maps, there is no radioactivity cross-talk with emission images and the life of the X-ray tube is very long. High-efficiency SPECT scanners, equipped with solid-state detectors and specialized collimators, have dramatically improved the count sensitivity and the image resolution. They enable shorter acquisition times facilitating imaging patients in multiple positions to assess image artifacts or the implementation of low radiation protocols, by performing standard time acquisitions or stress-only protocols [5].

Commonly, Nuclear Medicine experts analyze a considerable number of diagnostic tests daily. Besides MPI findings, the electrocardiogram (ECG) and ECG changes during stress, as well as the results of previous tests are also reviewed [6]. Despite its advances,

MPI still performs sub-optimally for the prediction of CAD [2], obliging the doctors to take into consideration other relevant information also [7]. This includes symptoms, age, gender, previous CAD history, CAD predisposing factors, concomitant diseases, etc [8]. Consequently, MPI interpretation in everyday practice is often somehow subjective. Moreover, the whole diagnostic procedure is time-consuming. Utilizing an MPI automated image analysis and characterization system, could potentially liberate doctors from lengthy procedures or act as an auxiliary tool for educational purposes.

The recent advances in artificial intelligence, the availability of extensive medical data, and the continuous improvements in computational capabilities, enable engineers to design complex and more accurate computer-aided diagnostic systems [9].

Over the last five years, the field of deep learning is thoroughly explored, and new techniques have emerged [10]. Convolutional neural networks (CNN) [11], which are the essence of deep learning, are considered to be a promising method for medical image analysis [12]. Deep learning is an automatic feature extractor method, which is aided by machine learning algorithms to distinguish the significant features from images.

Before the emergence of deep learning, computer-aided knowledge extraction from medical images was performed with handmade feature extraction, which was painful and time-consuming. Despite the reliability of the results, defining the most significant features from an image usually required statistical analysis, as well as extensive data preprocessing.

Our aim in this work was to investigate the effectiveness of CNN to characterize MPI images and to compare with experts' diagnostic yield.

Patients and Methods

Patients

Over a period of one year (from 16/2/2018 to 27/2/19) 1078 consecutive patients underwent gated-SPECT myocardial perfusion imaging (MPI) with ^{99m}Tc-tetrofosmin in the department of Nuclear Medicine of our hospital. Two-hundred and twenty-nine of these (21.2%) were subsequently subjected to invasive coronary angiography (ICA) within 60 days from MPI. After excluding 13 patients with inconclusive MPI results or missing ICA reports, the studies of 216 patient cases were finally included in the present work. The study was approved by the Ethical Committee of our Institution. The nature of the study waives the requirement to obtain informed consent. Patients' characteristics are listed in Table 1. A hybrid SPECT/CT gamma-camera (Infinia, Hawkey-4, GE Healthcare) was used for MPI imaging. Computed tomography-based AC in both stress and rest images was applied in all cases. Vasodilator stress with adenosine was implemented in 155 patients and treadmill exercise or combined vasodilator-exercise test in the remaining.

Methods

MPI interpretation

Three experienced Nuclear Medicine physicians (NDP, TS and DJA of the authors) had been employed prospectively in the interpretation and reporting of MPI studies. Their diagnosis relied on the inspection of the full set of tomographic slices and polar maps obtained from both AC and non-attenuation corrected (NAC) studies. Pretest probability of CAD, patient's history, previous tests results, baseline ECG and ECG changes during stress were incorporated in the final MPI interpretation. Myocardial perfusion imaging reports were considered positive for CAD if they described a reversible tracer defect of any extent or a fixed defect not normalized by AC implementation, accompanied by normal/near normal wall motion or thickening. Depending on the clinical context, small defects or defects attributed to other conditions (e.g. recent myocardial infarction with residual peri-infarct ischemia or specific ECG conduction abnormalities producing mild non-reversible defects) were considered suggestive of no flow-limiting CAD.

Table 1. *Patient characteristics.*

Clinical characteristics	Frequency
No	216
Age (mean±sd)	66.2±10.4 years
Sex (male/female)	88%/12%
History of CAD	49%
Previous myocardial infarction	20%
Previous revascularization (PCI/CABG)*	36%
CAD predisposing factors	
Previous stroke	2%
Hypertension	82%
Dyslipidemia	68%
Smoking	39%
Diabetes	34%
Peripheral arteriopathy	6%
End stage renal failure	3%
Family history of premature CAD	16%

**PCI: percutaneous coronary intervention; CABG: coronary artery bypass grafting*

Semi-quantitative polar map analysis

The standard semi-quantification method of polar map analysis, depending on summed stress, rest and difference scores (SSS, SRS, SDS) was also implemented. The positivity criterion for NAC studies was SSS>3 and SDS>1, while for AC SSS>0 and SDS>0.

Coronary angiography interpretation

Invasive coronary angiography was considered positive for flow-limiting CAD if there was >50% stenosis of the left main artery or if lesions causing >70% lumen stenosis of the main coronary arteries or their major branches were detected. Fractional flow reserve (FFR) measurements were undertaken in case of intermediate (50%-70%) stenoses. An FFR value <0.8 was considered positive for flow-limiting CAD.

Convolutional neural networks

Convolutional neural networks is a deep neural network, which utilizes convolution layers and pooling layers, to extract useful information from the input data, usually connecting the extracted features to a fully connected layer (i.e. a neural network) [13]. They exhibit impressive performance on a variety of machine learning tasks [14].

A convolution operation is performed as a filter, which is actually a table of weights, is sliding throughout the image, which is converted to an array format, containing pixel values. An output pixel produced at every position is a weighted sum of the input pixels (the pixels that the filter has passed from). The weights of the filter, as well as the size of the table (usually called filter size or receptive field), are constant for the duration of the scan. Therefore, convolutional layers can seize the shift-invariance of visible patterns and depict robust features [15]. Usually, after a set of convolutional layers, pooling layers follow.

The pooling operation computes a specific norm over small areas, thereby achieving translation invariance. This operation aggregates small pitches of pixels and down-samples the photo aiming to reduce the parameters of the data. This operation reduces the computational cost dramatically, especially when the input images are large, and the network's trainable parameters are exponentially increased, due to the network's depth. The most common pooling operation used in developing CNN is max-pooling, which outputs the most pixel value of the area.

Non-linearities are desirable for CNN to detect complex and non-linear features. Therefore, commonly, after a convolutional operation, an activation function follows. An activation function is a non-linear transformation that has been traditionally used in neural networks. Typical activation functions for CNN are rectified linear unit (ReLU) [16], leaky ReLU [17], and ELU [18].

After several convolutional and pooling layers, there may be one or more fully connected layers aiming to perform high-level reasoning [19]. Fully connected layers connect all previous layer's neurons with every neuron of the fully connected layer. In this way, every extracted feature, which comes with a specific weight, is processed independently, or in relation to the other features. The last layer of CNN is an output layer. Softmax [20] operator is a conventional classifier for CNN. The aim of the last layer is to classify the extracted features according to their relation with the desired outputs, called classes.

Training CNN is a problem of global optimization. A loss function is used to evaluate the difference between the output of a CNN and the desired label (i.e., the loss). Training is intended to achieve the minimization of the value of the loss function. Among many types of loss functions, the cross-entropy loss, along with softmax output activations, are the most typically used for classification tasks.

The CNN developed by virtual geometry group (VGG) for transfer learning

Instead of developing a new CNN architecture to train from scratch, which is the traditional pipeline in deep learning and its results depend highly on the size of the data provided, an alternative process called transfer learning was preferred. With transfer learning, it is possible to exploit the knowledge of a specific CNN, which was trained on a specific task, to make predictions for another task. The CNN developed by the visual geometry group (VGG) [21] is widely utilized for medical image classification tasks. Thus, employing the VGG was preferred for the experiments of this work.

During the process of pre-training, VGG has learned to extract specific characteristics from images assigning them constant weights. For our experiments, we retained some of those weights constant. The extracted features are connected to a classic neural network, which learns their significance, based on the supplement of images to train the model. The evaluation of the classification accuracy was performed with 10-fold cross-validation.

For the classification task, we use transfer learning to exploit the capabilities of pre-trained networks. From the derivatives of the main architecture of VGG, two CNN architectures called VGG16 and VGG19 were selected. Both are named after their number of total trainable weight layers, which are 16 and 19 accordingly. The number of the weight layers affects the depth of the network, and the feature extraction steps. Since little difference between those two architectures exists, it was preferable to utilize both.

VGG16

VGG16 was utilized by retaining some of the learned weights across its convolutional layers. At the top of the CNN, a global max pooling layer is added to reinforce the generalization capability of the network. The global max pooling aids to dimensionality reduction. The extracted features are connected to a fully connected layer of 500 nodes, followed by a dropout layer, which randomly disconnects 50% of the learned features. The dropout layers are commonly used to prevent the CNN from learning too specific characteristics of the images, by forcing the CNN to ignore randomly a proportion of the learned features.

Finally, another fully connected layer of 250 nodes is utilized, also supported by a dropout layer. Adam optimizer and categorical cross-entropy loss are proposed to handle the learning rate and weight updates, as well as the calculation of

the loss. Adam optimizer and categorical cross-entropy loss are the default hyper parameters of the initial VGG, and it was preferable to retain them. An overview of the proposed VGG16 is presented in the Figure 1.

VGG19

To compare both VGG16, and VGG19, it was preferable to add the exact same neural network at the top of both CNN. VGG19 contains a sixth set of convolutional layers to further process the input image, while the remainder of its structure is the same with VGG16, as it is presented in Figure 1.

The dataset of the study and data preprocessing

We utilized a collection of polar maps from 216 patient cases, from the database of the department of Nuclear Medicine of our hospital. Attenuation correction and NAC polar maps in stress and in rest condition were extracted from the SPECT scan in jpeg format. The AC and NAC images of stress and rest were rescaled to 150x150, to reduce the computational cost. The four polar maps corresponding to each case were concatenated into one image. For training deep neural networks, large - scale datasets are necessary. Since the dataset for this experiment was not large enough to effectively train a deep CNN, data augmentation to generate new instances, was performed.

We created three different datasets, named as AC, NAC, and ACNAC. The AC and NAC datasets contained two polar maps for each patient case (stress and rest condition), while the ACNAC dataset contained 4polar maps concatenated into one image for each patient case.

Data augmentation

For data augmentation, we simultaneously and slightly rotated the polar maps (by a maximum of 45 degrees). The rotated polar maps were then concatenated, thus generating new images. In this way, the CNN ignored slight spatial diversity between the polar maps, and focused on the color variations, regardless of their relative position. During the augmentation step, 1000 new images of both classes (diseased/healthy) were generated and utilized for training. To balance the dataset, we generated more instances belonging to the no-CAD class. Thus, the training was performed with balanced data, despite the initial slight imbalance problem. Figure 2 presents the data preprocessing and data augmentation steps related to the ACNAC dataset. The same procedure was held for the AC and the NAC datasets.

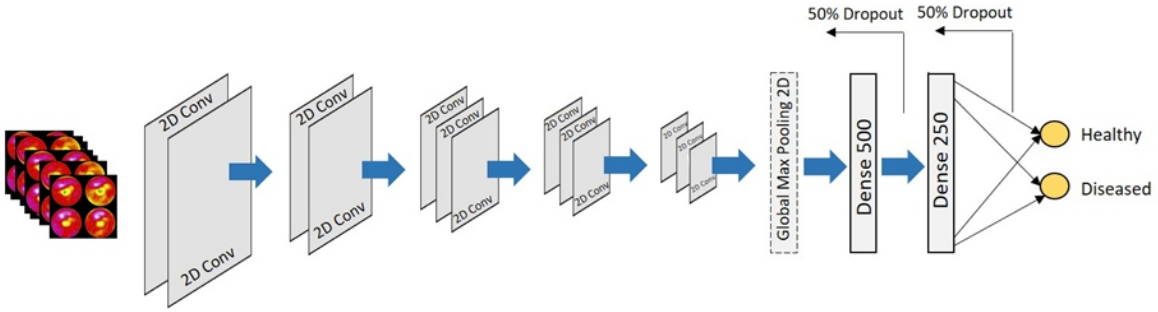

Figure 1. The architecture of VGG16 for transfer learning.

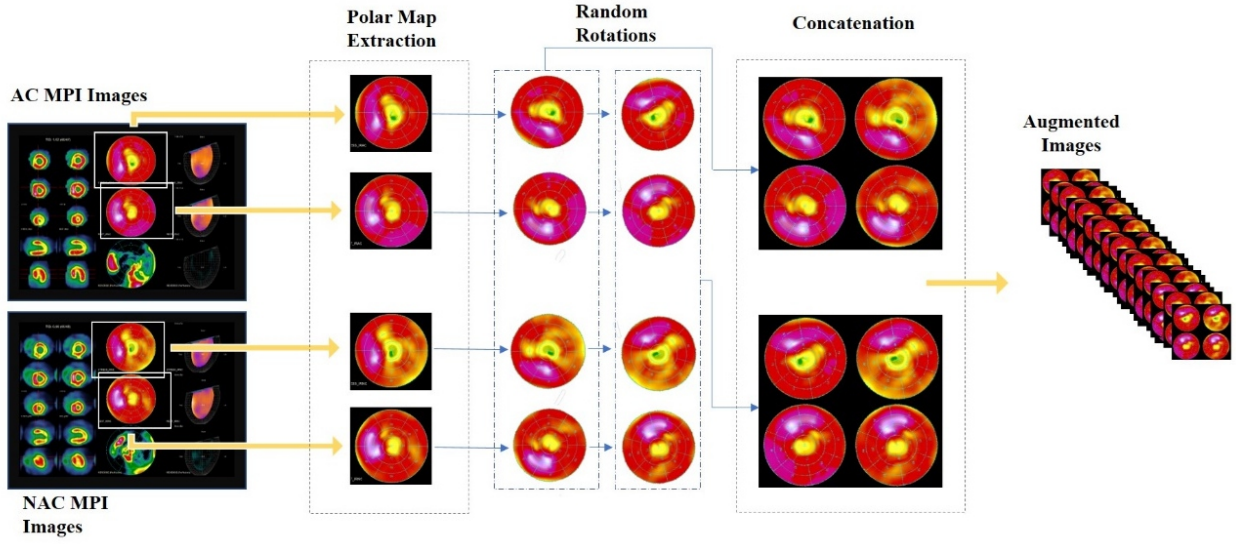

Figure 2. The process of data augmentation. The extracted polar maps from each case are slightly and simultaneously rotated, to achieve diversity and to slightly expand the training dataset at each training step.

Evaluation procedure

Due to dataset scale limitations, 10-fold cross-validation was performed in order to achieve proper training and trustworthy evaluation of the CNN. During this procedure, the initial dataset was split into ten parts, where the first nine parts were the training set, while the tenth part remained hidden from the CNN for testing, which was held after the completion of the training. Before the training, the nine training folds were augmented, as explained. This procedure was repeated for the remaining folds, until every fold was selected for test set. Hence, the CNN underwent ten separate training processes and ten evaluations. The overall experiment setup is represented in Figure 3.

Results

Based on the findings of coronary angiography, 152 patients (70.4%) were classified in the flow-limiting CAD, while the remaining 64 (29.6%) in the no-CAD group. The results of the CNN compared with the human expertise and the semi-quantitative polar map analysis, in terms of accuracy, sensitivity, and specificity for each dataset are presented in Table 2.

The semi-quantitative polar map analysis exhibited low diagnostic accuracy (64.2%-64.81%). More specifically, NAC showed low specificity (46.88%), while AC low sensitivity (65.13%). Interpretation of MPI images by the medical staff incorporated both NAC and AC findings in the final diagnosis, as well as all other relevant information from patients' history, resulting in higher accuracy (75%).

The CNN model focused on the extraction of possibly significant biomarkers from the polar maps only. The datasets containing only AC, or only NAC polar maps impeded the CNN from learning significant information, which was reflected on their evaluation metrics. Specifically, by separate NAC and AC data manipulation we observed a significant drop in terms of accuracy (-16%). However, by combining NAC and AC polar maps in one image set (ACNAC), the CNN model managed to increase its diagnostic efficiency, reaching values of sensitivity (75%), specificity (73.44%) and accuracy (74.53%) which were not substantially different from the experts' decision. In Table 3, the confusion matrices and some common machine learning metrics are provided.

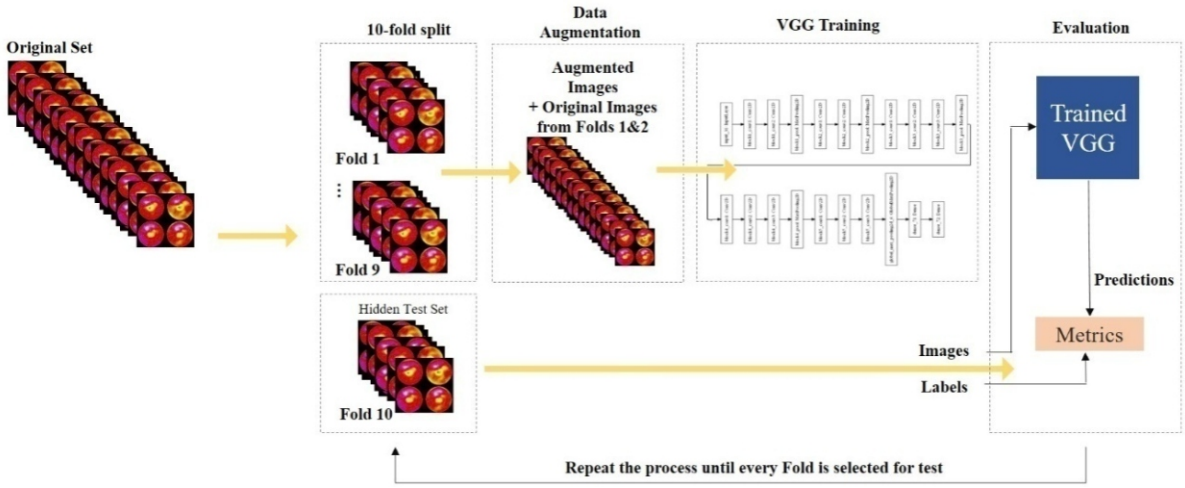

Figure 3. Training and evaluation of VGG for the classification task.

Table 2. *Accuracy, Sensitivity, and Specificity of the Networks, medical experts and semi-quantitative polar map analysis.*

Datasets	ACNAC			AC			NAC		
Model	**ACC (%)**	**SEN (%)**	**SPE (%)**	**ACC (%)**	**SEN (%)**	**SPE (%)**	**ACC (%)**	**SEN (%)**	**SPE (%)**
Experts	75.00	76.97	70.31	-	-	-	-	-	-
Semi-Q (NAC)	-	-	-	-	-	-	64.81	72.37	46.88
Semi-Q (AC)	-	-	-	66.20	65.13	68.75	-	-	-
VGG16	74.53	75.00	73.43	57.40	61.18	48.43	41.66	47.36	28.12
VGG19	74.53	75.00	73.43	51.85	57.89	28.12	58.33	62.50	48.43

Semi-Q: classical semi-quantitative polar map scoring; AC: attenuation correction; NAC: non-AC; ACNAC: AC-NAC combination; ACC: accuracy; SEN: sensitivity; SPE: specificity; Medical experts did not evaluate NAC and AC studies separately.

Table 3. *Confusion matrices, precision and recall metrics.*

Model	Actual values		Precision TP/(TP+FP)	Recall TP/(TP+FN)
Experts	**CAD**	**No-CAD**		
Predicted CAD	117 (TP)	19 (FP)	86.03%	76.97%
Predicted no CAD	35 (FN)	45 (TN)		
Semi-Q analysis	**CAD**	**No-CAD**		
Predicted CAD (NAC ; AC)	110 ; 99	34 ; 20	76.39%	72.37%
Predicted no CAD (NAC ; AC)	42 ; 53	30 ; 44	83.19%	65.13%
VGG16 and VGG19	**CAD**	**No CAD**		
Predicted CAD	114	17	87.02%	75.00%
Predicted no CAD	38	47		

To evaluate the transfer learning strategy, and the specific fine-tuning strategy for the specific VGG architecture, we conducted another experiment. Firstly, we trained the VGG16 from scratch, i.e. without retaining its learned weights, which is considered the conventional training procedure. This model is referred to as VGG16-16. Secondly, we experiment with altering the number and layout of the convolutional layers we keep trainable and untrainable each time. This process can be considered as part of a process called fine-tuning, and involves experimenting with the parameters of the model to achieve the best possible result. Starting from the top of the network, we gradually adjust the convolution blocks to be trainable. The different modifications of the base VGG19 are named after the number of layers made trainable, that is VGG19-1, VGG19-4, VGG19-8, VGG19-12. Finally, we employ VGG for the classic feature extraction via transfer learning, where every layer is untrainable (VGG16-0). Table 4 presents the accuracy, sensitivity, and specificity on the ACNAC dataset for each case.

Table 4. *Metrics of transfer learning and training from scratch strategies on the ACNAC dataset.*

Model	ACC (%)	SEN (%)	SPE (%)
VGG16-0	54.16	59.21	42.18
VGG16-1	55.55	58.55	48.43
VGG16-4	74.53	75.00	73.43
VGG16-8	62.03	61.84	62.50
VGG16-12	47.22	50.65	39.06
VGG16-16	43.05	49.34	28.12

ACC, SEN, SPE abbreviations as denoted in Table 2.

The results demonstrate that the transfer learning strategy is preferable over the conventional training of CNN. This is proved by the low accuracy of VGG16-16 (43.05%). Besides, the performance of VGG16-0 (54.16%), which is the CNN used for classic feature extraction via transfer learning is confirming that the fine-tuning procedure is vital to tune the parameters of the CNN to obtain more accurate results. In the particular experiment, the optimal adjustment of the trainable layers of the VGG16 network involved allowing 4 convolutional layers at the top of the network to be trainable.

Discussion

Due to the shortage of large-scale publicly available datasets containing SPECT images for the detection of CAD, the application of deep learning has not been thoroughly explored. Spier et al. (2019) [22] proposed a method for automatic classification of polar maps based on graph CNN. The proposed process was evaluated using 946 labeled datasets and compared quantitatively to three other neural-network-based methods. The proposed model achieved an agreement with the human observer on 89.3% of rest test polar maps and on 91.1% of stress test polar maps. Localization performed on a fine 17-segment division of the polar maps achieved an agreement of 83.1% with the human observer, while localization on a coarse 3-segment division based on the vessel beds of the left ventricle had an agreement of 78.8% with the human observer. Bentacur et al. (2019), [23] applied deep learning (DL) to combine upright and supine MPI polar maps automatically and to predict obstructive CAD. The performance of DL was compared to the clinically established combined perfusion quantification by upright and supine total perfusion deficit (TPD) using the rule of the concomitant location of the defect, and information, all unknown by DL and stan-

dard clinical combined TPD (cTPD). The DL procedure learned image statistics from supine and upright maps and integrated them with sex information to compute a score for obstructive CAD outperforming cTPD prediction. The DL procedure was able to capture complex relationships that were not easily captured with the rule for upright and supine concomitancy. Seven hundred and eighteen (62%) patients and 1,272 of 3,480 (37%) arteries had obstructive disease. The area under the receiver operating characteristics curve for prediction of disease on a per-patient and per-vessel basis by DL was higher than for combined TPD (per-patient, 0.81 vs. 0.78; per-vessel, 0.77 vs. 0.73; P<0.001). With the DL cut-off set to exhibit the same specificity as the standard cut-off for combined TPD, per-patient sensitivity improved from 61.8% (TPD) to 65.6% (DL) (P<0.05), and per-vessel sensitivity improved from 54.6% (TPD) to 59.1% (DL) (P<0.01). With the threshold matched to the specificity of a normal clinical read (56.3%), DL had a sensitivity of 84.8%, versus 82.6% for an on-site clinical read (P=0.3).

The results of the current work highlight the capabilities of deep learning for medical image classification tasks. On the specific dataset, blind VGG performed better than the conventional semi-quantitative polar map scoring and competed with doctors' abilities, obtaining high accuracy (74.53%), compared to expertise efficiency (75.00%). Moreover, it is highlighted that transfer learning, and more specifically, transfer learning via fine-tuning, is an effective strategy for classification of the specific images, capable of dealing with small datasets.

Limitations

Some limitations of the experiment have to be mentioned. First, the initial dataset is small impeding the generalization capabilities of the trained CNN. It is fair to assume that the features learned by the CNN, are features extracted from the specific polar maps, and do not necessarily depict significant factors that could be considered as biomarkers. This is a common issue of machine learning and deep learning face, especially when the datasets are small contain insignificant, or irrelevant relations between their attributes. The better accuracy obtained by the VGG does not suggest better knowledge from the medical staff, due to its inability to generalize, and due to the high experience the medical staff possess during the years. Secondly, the data augmentation method performed to circumvent the shortage of data is not robust enough to generate realistic images. Moreover, by rotating polar maps the site of abnormal findings was attributed to another territory of the left ventricular wall. However, in this work we examined the global diagnostic efficiency of the method and not its ability to localize results to a specific coronary vessel distribution.

Despite the issues mentioned, it is undeniable that the results suggest that the automatic characterization of polar maps can be achieved in future works, assuming that larger datasets emerge. Also, prospective application of deep learning methods in everyday practice is a challenge and a future goal.

In conclusion, it is possible to classify a small dataset of MPI images at a level competing with experts' diagnostic ability with deep learning methods, employing transfer learning and data augmentation techniques. The proposed model can be easily and immediately deployed clinically for the evaluation of NAC and AC images. The value of the model could be two-fold: a) to support clinical diagnosis of CAD as an assisting tool and b) to serve teaching purposes. In the future, more extensive MPI image databases are needed to confirm our results.

Bibliography

1. Willerson JT, Wellens HJJ, Cohn JN, Holmes DR, editors. Cardiovascular Medicine. London: Springer London; 2007. Available from: http://link.springer.com/10.1007/978-1-84628-715-2.
2. Apostolopoulos DJ, Savvopoulos C. What is the benefit of CT-based attenuation correction in myocardial perfusion SPET. *Hell J Nucl Med* 2016; 19: 89-92.
3. Kalantari F, Rajabi H, Saghar M. Quantification and reduction of attenuation related artifacts in SPET by applying attenuation model during iterative image reconstruction: a Monte Carlo study. *Hell J Nucl Med* 2011; 14: 278-83.
4. Xin W, Shao X, Wang Y et al. Is there an incremental value to use myocardial perfusion imaging with or without CT attenuation for the diagnosis of coronary artery disease? A study in Chinese patients. *Hell J Nucl Med* 2018; 21: 48-54.
5. Betancur J, Commandeur F, Motlagh M et al. Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study. *JACC Cardiovasc Imaging* 2018; 11: 1654-63.
6. Montalescot G, Sechtem U, Achenbach S et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. *Eur Heart J* 2013; 34: 2949-3003.
7. Einstein A. Effects of radiation exposure from cardiac imaging: how good are the data? *J Am Coll Cardiol* 2012; 59: 553-65.
8. Scanlon P, Faxon D, AudetA et al. ACC/AHA guidelines for coronary angiography. *J Am Coll Cardiol* 1999; 33: 1756-824.
9. Litjens G, Kooi T, Bejnordi B et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017; 42: 60-88.
10. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Process Syst* 2012: 1097-105.
11. LeCun Y, Boser B, Denker J et al. Handwritten digit recognition with a back-propagation network. *Adv Neural Inform Process Syst* 1990: 396-404.
12. Ravi D, Wong C, Deligianni F et al. Deep Learning for Health Informatics. *IEEE J Biomed Health Inform* 2017; 21: 4-21.
13. LeCun Y, Kavukcuoglu K, Farabet C. Convolutional networks and applications in vision. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. Paris, France: IEEE; 2010, cited 2019 Nov 18: 253-6. Available from:http://ieeexplore.ieee.org/document/5537907/.
14. Koushik J. Understanding Convolutional Neural Networks. arXiv: 160509081, 2016, cited 2019 Nov 18, Available from: http://arxiv.org/abs/1605.09081.
15. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521: 436-44.
16. Nair V, Hinton G. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th international conference on machine learning (ICML-10). 2010; 807-14.
17. Maas A, Hannun A, Ng A. Rectifier nonlinearities improve neural network acoustic models. Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA, 2013.
18. Clevert D, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (elus). arXiv:151107289, 2015.
19. Liu M, Shi J, Li Zet al. Towards Better Analysis of Deep Convolutional Neural Networks. arXiv:160407043, 2016, cited 2019 Dec 12, Available from: http://arxiv.org/abs/1604.07043.

20. Bishop CM. Pattern recognition and machine learning. Springer Publishing, 2006.
21. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:14091556, 2015, cited 2020 Jan 10; Available from:http://arxiv.org/abs/1409.1556.
22. Spier N, Nekolla S, Rupprecht Cet al. Classification of Polar Maps from Cardiac Perfusion Imaging with Graph-Convolutional Neural Networks. *Sci Rep* 2019; 9: 756-9.
23. Betancur J, Hu L, Commandeur F et al. Deep Learning Analysis of Upright-Supine High-Efficiency SPECT Myocardial Perfusion Imaging for Prediction of Obstructive Coronary Artery Disease: A Multicenter Study. *J Nucl Me* 2019; 60: 664-70.