

Mitigating Diagnostic Errors in Coronary Angiography Using CNN-Based Deep Learning Models

NALLAPERUMALREDDI JYOTHI¹, POTHUREDDY GOWTHAMI², REDDAPPA PUDI³, NALUKURTHI SUMALATHA⁴

^{1,2,3,4}Academic Consultant, Department of ECE, Sri Venkateswara University College of Engineering, SVU, Tirupati, India.

Email: nalukurthisumalatha@gmail.com, gowthami0055@gmail.com, redappa401@gmail.com, Jyothi.jkl@gmail.com

ABSTRACT- Cardiology is a critical domain within medical sciences that increasingly relies on advanced technology for accurate diagnosis and treatment planning. One of the primary diagnostic procedures in cardiology is coronary angiography, employed to detect blockages or narrowing of the coronary arteries. However, despite its widespread use, traditional angiography combined with Fractional Flow Reserve (FFR) and Digital Subtraction Angiography (DSA) often faces limitations in accurately detecting the number and severity of blockages, leading to potential misdiagnoses and treatment delays. This paper highlights the existing drawbacks faced by cardiologists and coronary artery disease (CAD) patients associated with current angiographic practices. To address these challenges, we propose the integration of Deep Learning techniques, particularly Convolutional Neural Networks (CNNs), into the angiographic analysis workflow. By leveraging CNN-based models for automated feature extraction and classification of angiographic images, it is possible to significantly enhance diagnostic accuracy, potentially achieving near 100% reliability in detecting coronary blockages. This approach aims to assist clinicians in making faster and more accurate decisions, ultimately improving patient outcomes in cardiac care.

KEY WORDS:- Fractional flow reserve (FFR), Deep Learning ,Convolutional Neural Network (CNN),Angiography,Digital Subtraction Angiography (DSA) ,Artificial Neural Networks (ANN) , Natural Language Processing (NLP) ,Medical Image Computing , Magnetic Resonance Imaging.

1. INTRODUCTION

Angiography or arteriography is a medical imaging technique used to visualize the inner content of blood vessels and organs of the body, in particular the arteries, veins, and the heart chambers. This is actually done by injecting a radio-opaque contrast agent into the blood vessel and imaging using X-ray based techniques such as fluoroscopy and FFR. For instance a person suffering from coronary artery disease first goes through this angiography it gives us the image results the results might be like 80%blockage in a particular artery and 70% blockage is there in left artery and 35%blockage in another artery so total of three blocks are there with some percentage of blockage so this angiography gives us the results to surgeon like three blocks should be operated and after this during the operation if they are doing in real the surgeon has found one more block so what happens here is the waste of time occurs and also mental stress might occur. Depending on the type of angiogram, access to the blood vessels is gained most commonly through the femoral artery, to look at the left side of the heart and at the arterial system; or the jugular or femoral vein, to look at the right side of the heart and at the venous system. Using a system of guide wires and catheters, a type of contrast agent (which shows up by absorbing the X-rays), is added to the blood to make it visible on the X-ray images. The X-ray images taken may either be still images or motion images. The images are usually taken using a technique called digital subtraction angiography or DSA. In this images are generally taken at 2–3 frames per second, which allows the interventional radiologist to evaluate the flow of the blood through a vessel or vessels. This technique "subtracts" the bones and other organs such that only the vessels filled with contrast agent can be seen. The heart images are taken at 15–30 frames per second. Both the techniques Fluoroscopy and FFR enable the interventional radiologist or cardiologist to see blockages or narrowings. The heart images can be taken with the Fractional flow reserve (FFR) technique which is a technique used in coronary catheterization to measure pressure differences across a coronary artery blockages or narrowings to determine the likelihood that the blockage impedes oxygen delivery to the heart muscle.

Fractional flow **reserve** (FFR) is a technique used in coronary catheterization to measure pressure differences across a coronary artery stenosis (narrowing, usually due to atherosclerosis) to determine the likelihood that the stenosis impedes oxygen delivery to the heart muscle (myocardial ischemia).

Fractional flow reserve is defined as the pressure after (distal to) a stenosis relative to the pressure before the stenosis. The result is an absolute number; an FFR of 0.80 means that a given stenosis causes a 20% drop in blood pressure. In

other words, FFR expresses the maximal flow down a vessel in the presence of a stenosis compared to the maximal flow in the hypothetical absence of the stenosis.

The actual procedure of it is during coronary catheterization, a catheter is inserted into the femoral (groin) or radial arteries (wrist) using a sheath and guidewire. FFR uses a small sensor on the tip of the wire (commonly a transducer) to measure pressure, temperature and flow to determine the exact severity of the lesion. This is done during maximal blood flow (hyperemia), which can be induced by injecting products such as adenosine or papaverine. A pullback of the pressure wire is performed, and pressures are recorded across the vessel. An example of real-time FFR assessment in clinical use is shown here. There is no absolute cut-off point at which FFR becomes abnormal; rather, there is a smooth transition, with a large grey zone of insecurity. In clinical trials however, a cut-off point of 0.75 to 0.80 has been used; higher values indicate a non-significant stenosis, whereas lower values indicate a significant lesion.

Digital Subtraction Angiography (DSA) Provides an image of the blood vessels in the brain to detect a problem with blood flow. The procedure involves inserting a catheter (a small, thin tube) into an artery in the leg and passing it up to the blood vessels in the brain. A contrast dye is injected through the catheter and X-ray images are taken of the blood vessels.

DSA is primarily used to image blood vessels. It is useful in the diagnosis and treatment of arterial and venous occlusions, including carotid artery stenosis, pulmonary embolisms, and acute limb ischaemia; arterial stenosis, which is particularly useful for potential kidney donors in detecting renal artery stenosis (DSA is the gold standard investigation for renal artery stenosis); cerebral aneurysms and arteriovenous malformations (AVM).

Medical image computing (MIC) is an interdisciplinary field at the intersection of computer science, information engineering, electrical engineering, physics, mathematics and medicine. This field develops computational and mathematical methods for solving problems pertaining to medical images and their use for biomedical research and clinical care.

The main goal of MIC is to extract clinically relevant information or knowledge from medical images. While closely related to the field of medical imaging, MIC focuses on the computational analysis of the images, not their acquisition. The methods can be grouped into several broad categories: image segmentation, image registration, image-based physiological modeling, and others. Medical image computing typically operates on uniformly sampled data with regular x-y-z spatial spacing (images in 2D and volumes in 3D, generically referred to as images). At each sample point, data is commonly represented in integral form such as signed

and unsigned short (16-bit), although forms from unsigned char (8-bit) to 32-bit float are not uncommon. The particular meaning of the data at the sample point depends on modality: for example a CT acquisition collects radiodensity values, while a MRI acquisition may collect T1 or T2-weighted images. Longitudinal, time-varying acquisitions may or may not acquire images with regular time steps. Fan-like images due to modalities such as curved-array ultrasound are also common and require different representational and algorithmic techniques to process. Other data forms include sheared images due to gantry tilt during acquisition; and unstructured meshes, such as hexahedral and tetrahedral forms, which are used in advanced biomechanical analysis example tissue deformation, vascular transport, bone implants.

Deep learning is a next version of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised. Deep learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance. Basically we have three types of agents they are like smart agent, strong agent, weak agent. The smart agent means it can predict whatever the information given by the user like the Virtual assistants like Siri, Ok Google etc., If you give any query to Siri like who is your father it gives you answer like Apple inc., so in this way strong agents work. The strong agent means it cannot predict everything but it can answer few queries given by user we can consider and ATM machine as an example for this. The weak agent means it doesn't have the smart and strong agent features. Deep learning can be applied in many applications like industries, in industries what it can do is if we implement that for the machinery then without the need of workmen it can perform the whole tasks inside it. It can also be used in medical for operating during surgeries and also post surgeries and many more facilities in a hospital. In educational application it can be used in a way like automatic attendance of the students entering a class room or an examination hall etc., In agricultural application it can be used in way that automatic updates regarding weather and other climatic conditions. In commercial applications like in shops it can predict the true buyer with his gestures these are few applications in this deep learning.

A neural network is a network or circuit of neurons, or in a modern sense, an artificial neural network, composed of artificial neurons or nodes. Thus a neural

network is either a biological neural network, made up of real biological neurons, or an artificial neural network, for solving artificial intelligence (AI) problems. The connections of the biological neuron are modeled as weights. A positive weight reflects an excitatory connection, while negative values mean inhibitory connections. All inputs are modified by a weight and summed. This activity is referred to as a linear combination. Finally, an activation function controls the amplitude of the output. For example, an acceptable range of output is usually between 0 and 1, or it could be -1 and 1 .

These artificial networks may be used for predictive modeling, adaptive control and applications where they can be trained via a dataset. Self-learning resulting from experience can occur within networks, which can derive conclusions from a complex and seemingly unrelated set of information.

Convolutional neural network (CNN, or ConvNet) is a concept of deep neural networks, most commonly applied for analyzing the images. They are also known as shift invariant or space invariant artificial neural networks. They have applications in image and video recognition, recommender systems, image classification, medical image analysis, natural language processing, and financial time series. For instance when we consider some query and answer based online systems they familiarly use this convolutional neural networks because it is an expert in predicting which is correct and which is wrong. This CNN is completely about verifying and giving the true information regarding the images so the input for this CNN would be images. This can predict 100% correct information with the images it self so its very advantageous to use this technique in image assessment analysis.

2. EXISTING SYSTEM

With the FFR technology we can assess the blockages or narrowing in the heart upto 90%. Sometimes it is 100% but not all the time. So the results with the FFR and Fluoroscopy weren't 100% accurate that's why we can say that the drawback is inaccuracy in finding exact number of blockages or narrowings in the heart as we have taken the example of heart organ in the body.

3. PROPOSED SYSTEM

With the implementation of CNN image recognition technique to this FFR and Fluoroscopy results, we can easily find the exact number of blockages in the heart which results in time saving during surgery and also the doctors can get complete assessment of the heart before the surgery. In this we have considered the organ that is heart but not only heart we can also apply this technique to any organ of the body because this can be done and analysed with the still images or

motion images. The brain scans will also have the still images and motion images etc.,. So with this we can easily assess the correct result as it checks the hidden layers of the all types of images for all types of organs in a human body.

4. METHODOLOGY

In the assessment of number of blockages in the heart with the help of FFR, during coronary catheterization, a catheter that is a tube like structure is inserted into the radial arteries of wrist using a sheath and guidewire. FFR uses a small sensor on the tip of the wire to measure pressure, temperature and flow to determine the exact severity of the affected area. A pullback of the pressure wire is performed, and pressures are recorded across the vessel in the form of images or motion images. The results are images so if we use the CNN technique that is image recognition to assess the images. In this way, we can easily find out the exact number of blockages in the heart.

The convolutional neural network (CNN) indicates that the network takes a mathematical operation called convolution. Convolution is a special kind of linear operation. Convolutional networks are simply neural networks that use convolution in place of general matrix multiplication in at least one of their layers. This consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN consist of a series of convolutional layers that convolve with a multiplication or other dot product. The activation function is commonly a RELU layer, and is subsequently followed by additional convolutions such as pooling layers, fully connected layers and normalization layers, referred to as hidden layers because their inputs and outputs are masked by the activation function and final convolution. In general the heart images will have 15 to 20 frames so sometimes the present process may fail to define number of blockages as it has to check the hidden layers of all these 20 frames but with the involvement of this technique it can define many number of hidden layers of the images of more than 20 frames too.

In this, each level learns to transform its input data into a slightly more abstract and composite representation. In an image recognition application, the input will be a matrix of pixels. The first representational layer abstracts the pixels and encode edges, the second layer composes and encode arrangements of edges, the third layer encodes a the contrast part of the blood vessel, and the fourth layer recognizes a number of blocks in heart as of our considered example. Importantly, a deep learning process can make learn which features work to optimally place and in which level on its own.

5. STEPS OF IMAGE RECOGNITION

1. **IMAGE SIZE** : The image size is the first thing which we need to consider while doing the process. The images with higher quality give the model more information but requires more computing process time and also more number of network layers. It first generalizes the pixels and counts number of pixels present in it so if it generalizes in a way that it divides the image into length, breadth, width so with this it can easily define the size of the image.
2. **DEFINING THE TOTAL NUMBER OF IMAGES** : The training set represents the total population of the image taken. It actually analyses the data taken and gives us the exact how much amount of data present in it.
3. **THE NUMBER OF CHANNELS** : Grayscale images have two channels and color images have three channels.
4. **IMAGE TRANSFORMATIONS** : After considering the images they will be transformed respectively.
5. **DATA SCANNING** : The images whatever you have considered they are scanned to wellness. Which means it checks the all hidden layers by passing through their channels.

To program a CNN, the input is a tensor with shape

(number of images) x (image width) x (image height) x (image depth).

Then after passing through a convolutional layer, the image becomes abstracted to a feature map, with shape (number of images) x (feature map width) x (feature map height) x (feature map channels). The CNN attributes are Convolutional kernels defined by a width and height. The number of input channels and output channels. The depth of the Convolution filter must be equal to the number channels of the input feature map. Convolutional layers convolve the input and pass its result to the next layer. This is similar to the response of a neuron in the visual cortex to a specific stimulus. Each convolutional neuron processes data only for its receptive field. The fully connected feedforward neural networks can be used to learn features as well as classify data though. A very high number of neurons would be necessary, even in a shallow architecture, due to the very large input sizes associated with images, where each pixel is a relevant variable. For example, a fully connected layer for a small image of size 100 x 100 has 10,000 weights for each neuron in the second layer. The convolution operation brings a solution to this problem as it reduces the number of free parameters, allowing the network to be deeper with fewer parameters. For example, regardless of image size, tiling regions of size 5 x 5, each with the same shared weights, requires only 25 learnable parameters. In this way, it resolves the vanishing or exploding gradients problem in training traditional multi-layer neural networks with many layers by using backpropagation. The convolutional layer is the core building

block of a CNN. The layer's parameters consist of a set of filters, which have a small receptive field, but extend through the full depth of the input volume. During the forward pass, each filter is convolved across the width and height of the input volume, computing the dot product between the entries of the filter and the input and producing a 2-dimensional activation map of that filter. As a result, the network learns filters that activate when it detects some specific type of feature at some spatial position in the input. Stacking the activation maps for all filters along the depth dimension forms the full output volume of the convolution layer. Every entry in the output volume can thus also be interpreted as an output of a neuron that looks at a small region in the input and shares parameters with neurons in the same activation map.

The pooling is an important concept in this CNN, which is a form of non-linear down-sampling. This divides the input image into a set of non-overlapping rectangles and, for each such sub-region, outputs the maximum. The exact location of a feature is less important than its rough location relative to other features. This is the main important thing behind the use of pooling in CNN. The pooling layer reduces the spatial size of the representation, to reduce the number of parameters, memory footprint and amount of computation in the network, it also controls overfitting. It is quite common to periodically insert a pooling layer between successive convolutional layers in a CNN architecture. The pooling layer operates every depth slice of the input and resizes it spatially.

The rectified linear unit, which applies the non-saturating activation function. It removes negative values by setting them to zero. It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields of the convolution layer.

After several convolutional and max pooling layers, the high-level reasoning in the neural network is done by fully connected layers. Neurons in a fully connected layer have connections to all activations in the previous layer, as seen in regular artificial neural networks. Their activations can thus be computed as an affine transformation, with matrix multiplication followed by a bias offset.

This layer specifies how training penalizes the deviation between the predicted and true labels and is normally the final layer of a neural network. Various loss functions appropriate for different tasks may be used. So if we implement this deep learning concept of CNN layers to FFR technology we can achieve good results of it and also can save the time.

6. CONCLUSION

With the help of this Deep Learning technique that is CNN we can easily assess the good results and along with this we can also apply this technique for getting the good results of Interventional Radiology area which includes X-ray fluoroscopy, ultrasound, computed tomography or magnetic resonance imaging.

REFERENCES

- [1]. Kang, S. J., Ahn, J. M., & Park, S. J. (2017). Fractional flow reserve: Current status and future directions. *Heart*, 103(14), 1109–1114. <https://doi.org/10.1136/heartjnl-2015-308827>
- [2]. Koo, B. K., Erglis, A., Doh, J. H., Daniels, D. V., Jegere, S., Kim, H. S., ... & Fearon, W. F. (2011). Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary CT angiograms. *JACC: Cardiovascular Imaging*, 4(4), 406–413. <https://doi.org/10.1016/j.jcmg.2010.11.013>
- [3]. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- [4]. Zreik, M., Lessmann, N., van Hamersvelt, R. W., Wolterink, J. M., Voskuil, M., Viergever, M. A., ... & Isgum, I. (2018). Deep learning analysis of coronary arteries in cardiac CT angiography for detection of patients with functionally significant coronary artery stenosis. *Medical Image Analysis*, 44, 72–85. <https://doi.org/10.1016/j.media.2017.12.005>
- [5]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- [6]. Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., ... & Ng, A. Y. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Medicine*, 15(11), e1002686. <https://doi.org/10.1371/journal.pmed.1002686>
- [7]. Guo, Y., Dong, H., & Liu, G. (2021). Deep learning for cardiovascular disease diagnosis using medical imaging: A review. *Physics in Medicine & Biology*, 66(4), 04TR01. <https://doi.org/10.1088/1361-6560/abce65>
- [8]. Zhou, S. K., Greenspan, H., & Shen, D. (Eds.). (2019). *Deep learning for medical image analysis*. Academic Press. <https://doi.org/10.1016/C2017-0-03529-4>
- [9]. Ahn, J. M., Park, D. W., Shin, E. S., Koo, B. K., Nam, C. W., Doh, J. H., ... & Park, S. J. (2017). Fractional flow reserve or optical coherence tomography to guide management of angiographically intermediate



- coronary stenosis: A prospective randomized trial. *Circulation*, 136(13), 1280-1290. <https://doi.org/10.1161/CIRCULATIONAHA.117.029095>
- [10]. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29. <https://doi.org/10.1038/s41591-018-0316-z>