

A Novel 3D Recurrent R-CNN for Medical Imaging Feature Detection: A Case Study for Coronary Calcium Detection

Vikas Sarvasya
University of Pennsylvania
sarva@seas.upenn.edu

Robert Gotwals
North Carolina School of Science and
Math
gotwals@ncssm.edu

Liam Butler
Wake Forest University School of
Medicine
liabutle@wakehealth.edu

ABSTRACT

Deep neural networks (DNNs), when trained on high-quality, noise-free images, often underperform when applied to images with distortions or noise [22]. A key challenge lies in the architecture of convolutional neural networks (CNNs), where the layers are structured to prioritize features detected at the final layers, treating earlier features as latent or redundant [5]. This design hinders the ability of CNNs to effectively detect small, dynamic, and complex structures in images, thereby limiting their adaptability to image variations [5].

The primary objective of this study was to develop a novel network architecture capable of accurately detecting the bounding boxes of coronary arteries and subsequently calculating a calcium score. The model's bounding boxes were validated against manually annotated arterial outlines, and the calcium score was compared to clinician-adjudicated values. The proposed network introduced an innovative propagation mechanism, coupled with various derived algorithms and modifiers, to mitigate the impact of motion distortions on coronary artery tracking and detection.

Results demonstrated exceptional performance, with the testing set achieving an average Mean Squared Error (MSE) of less than 2%, and a deviation of less than 1.5 pixels for each coordinate within a 128x128-pixel image. The calcium score derived from the network's bounding boxes exhibited a strong correlation of $R^2 = 0.921$, with a region of interest (ROI) accuracy of 88% across all calcium score ranges. In contrast, a standard CNN used as a control struggled, yielding an $R^2 = 0.0091$ and an ROI accuracy of just 10%.

This student research project presents a pioneering network that utilizes specialized algorithms and propagation techniques to accurately identify small, dynamic structures in non-gated chest CT scans. The model's ability to provide reliable calcium scores enhances the clinical utility of chest CT scans, offering a promising tool for improving the diagnosis of coronary artery disease and optimizing the management of cardiac disease risk [19].

KEYWORDS

Deep neural networks, Convolutional neural networks, Bounding box regression, Backpropagation algorithm, CT scan, Coronary Artery Calcium score, Artery detection, Distortion

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Copyright ©JOCSE, a supported publication of the Shodor Education Foundation Inc.

© 2025 Journal of Computational Science Education
<https://doi.org/10.22369/jocse.2153-4136/16/2/6>

1 INTRODUCTION

Coronary calcium is calcium that builds up in coronary arteries. A substantial buildup of calcium deposits (calcium plaques) is called atherosclerosis and can narrow arteries and can be brought about by high blood pressure, obesity, diabetes, and high blood cholesterol [19]. A Coronary Artery Calcium (CAC) score of 0 shows that the individual has no coronary calcium buildup while an increasing calcium score (with virtually no limit) increases the risk of cardiovascular disease (CVD) [6]. Atherosclerosis is "the underlying cause of about 50% of all deaths in westernized society due to heart attacks, stroke, and peripheral arterial disease" [19]. CAC scores help clinicians assess risk of, amongst other cardiac diseases, heart attacks and stroke [19]. When CAC scores are moderate-high, clinical attention for disease prevention and/or treatment would be required. Cardiac gated computed tomography (CT) scans are conventionally used to calculate CAC. Gated CT scans may require specific drugs to slow down a patient's heart rate and imaging of the heart is done only during the mid to end diastolic phase, when cardiac movement is minimized [6]. This ensures that the gated CT contains no motion distortions and clearly highlights important features, such as coronary arteries [6]. This means that gated CT scans correct coronary motion and provide accurate values for the CAC score, since it is extremely easy to detect the coronary arteries [6].

Gated CTs are much more expensive than non-gated CTs [17] due to the use of drugs to modulate the heart rate. Consequently, cardiac gated CT scans are done very infrequently and only on patients who have already been diagnosed with some measure of CVD [7]. Roughly a quarter of patients who have heart attacks suffer from sudden cardiac death [16]. Of these patients, 80% have atherosclerosis, making it the most reliable indicator of cardiac arrest [26]. Therefore, it is imperative that better predictive tools are available in a larger group of patients to detect coronary atherosclerosis. Unlike gated CTs, non-gated chest CTs are more readily available and are routinely performed for a much wider range of pathologies other than coronary artery disease [13]. For example, chest CTs are used for detecting pulmonary embolism in patients presenting to the emergency room for trouble breathing.

If these non-gated CTs could be accurately evaluated for coronary artery disease, there would be less undiagnosed CVD patients, who could then be prescribed risk-reducing treatments and decrease the rate of instantly fatal heart attacks [13]. Non-gated CTs do not use drugs to modulate the heart rate and simply take cross-sectional images sequentially. Because images are taken during different stages of the heartbeat, these images contain motion distortion of the heart. In non-gated CTs, the coronary arteries are still in the field of view of the image, but the distortion of the image precludes

accurately identifying the arteries, making calculations of a CAC difficult [14]. This reduces the overall usability of these types of CTs for such cases. When non-gated CTs are used, the CAC score for these imaging scans is manually assigned into three categories: mild, moderate, and severe, rather than providing a numerical value [9]. This creates some issues, mostly due to variation and bias between physicians [9]. This unstandardized system also doesn't account for important factors such as age, gender, or race, making it unfeasible to use for diagnosis and treatment [8]. A mild CAC score for an elderly woman can be severe for a man in his 20s. There is a need for automated accurate coronary artery detection and CAC quantification, which can be achieved by the new 3D CNN's [24]. This is helpful for risk prediction and identifying patients who are at high-risk of cardiac diseases and require risk-reducing treatments [14].

Artificial Intelligence (AI) has made substantial advancements in its application within the medical field due to an increase in data availability. Deep learning methods have gained prominence in the use of medical imaging methods, such as magnetic resonance imaging (MRI) and computed tomography (CT) to develop AI-based prediction and detection models of multiple cardiovascular diseases. However, there remain caveats in their accuracy and generalizability, mostly due to the quality of the images used as inputs.

The literature into neural networks and brain structure suggests that this can be improved by altering the network to allow stacking maps on top of each other [1, 15]. Through each layer of a CNN, the filters contained in its nodes are applied to images to produce feature maps. By creating multiple sections of CNNs on top of each other, each layer will produce a set of stacked feature maps that can be analyzed sequentially and whose differences can be used to identify small and specific features. By using this approach, the CNN has more data it can train on (such as maps applied with sharpened kernels and sequentially subtracted images) but can process and find more features and increase the detail of examination when detecting features in an image [10, 21]. The overarching aim of this project is to address common issues in image regression within CNNs, which are distortion and variance in quality, size, and shape [11].

Traditionally, 3D Convolutional Neural Networks (CNNs) have been designed to handle RGB images by splitting the channels and training three distinct networks to learn features from each channel independently. This approach allowed each network to focus on a different set of criteria. However, by extending this methodology to more complex structures—such as networks with multiple stacks—our model can process grayscale or single-channel images effectively. This strategy increases the model's versatility, allowing it to handle not only color images but also those from modalities like MRI scans, which typically produce grayscale images.

This student research project proposed a novel backpropagation algorithm which would allow for the training of the 3D matrices/filters that build convolutional layers. This algorithm would advance current scientific knowledge on how to scale up stacked networks and how to isolate specific changes throughout numerous different variables used, allowing for more specific detected features and kernel changes. This novel method is superior to the current 2D CNNs for medical image analysis because of their inability to

adapt to different complex structures while 3D CNNs will be able to make accurate arterial detection possible despite motion distortion in non-gated CTs [5].

In this project, we seek to apply a novel algorithm to radiological scans to isolate image sections of increased coronary calcification and output a calcium score. More specifically this research addresses the following aims:

- (1) Develop a novel backpropagation neural network algorithm to allow 3D image processing
- (2) Mathematically explain how this recurrent convolutional neural network (RCNN) can be integrated with this algorithm
- (3) Maintain a higher accuracy level for the 3D CNN compared to 2D CNNs

2 METHODS

The data for this experiment consisted of 76 non-gated chest CT scans containing over 8,000 images in total. 53 of the 76 scans contained coronary calcium, while 23 did not. The 76 scans were divided into 80% training dataset and 20% testing dataset for 60 training scans and 16 testing scans. Of the training dataset, there were 41 with coronary calcium and 19 without. Of the testing dataset, there were 12 with coronary calcium and 4 without.

2.1 Backpropagation Algorithm

2.1.1 Forward Pass and Loss Function. The forward pass in a CNN involves propagating data through the layers, where each layer consists of weights (filters) and biases that are adjusted during training. The output of each layer is a feature map that represents certain patterns or features in the input data. The loss function, typically quantified using cross-entropy or mean squared error, measures the deviation between the predicted output and the true label of the input data [3]. In this case, the loss function is represented by Equation 1, shown below.

$$L = \frac{1}{XYZ} \sum_{x,y,z=1}^{X,Y,Z} (v - P_{xyz})^2 \quad (1)$$

where x , y , and z are the input variables (pixels or features), L is the loss value, v is the result of the forward pass, and P is the predicted value. The deviation from the true values is calculated by computing the partial derivatives of the loss function with respect to each of the input variables, shown in Equation 2.

$$\frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}, \frac{\partial L}{\partial z} \quad (2)$$

These derivatives indicate the sensitivity of the loss function to changes in the input values, providing crucial information for adjusting filters and weights during backpropagation. Next, we have the convolution operation. The central operation in CNNs is the convolution, where an input matrix (image) is convolved with a filter to generate an output matrix (feature map). The convolution is performed by multiplying corresponding elements of the input matrix and filter, summing these products, and sliding the filter over the image to generate the output matrix O , calculated by Equation 3 [4].

$$O_{11} = \sum_{i=1}^n \sum_{j=1}^n X_{ij} F_{ij} \quad (3)$$

where n is the size of the convolution kernel, F is the filter, and X is the input matrix/image. A similar convolution operation is repeated for the rest of the input matrix, generalizing the previous equation to make Equation 4.

$$O_{kl} = \sum_{i=1}^n \sum_{j=1}^n X_{i+k,j+l} F_{ij} \quad (4)$$

The convolution operation extracts local features from the input image, which are then passed through activation functions like ReLU to introduce non-linearity and allow the network to learn complex patterns.

2.1.2 Backward Pass. The backpropagation algorithm relies on the chain rule of calculus to compute gradients of the loss function with respect to the network parameters. Specifically, we need to calculate the derivative of the loss function with respect to the output of the convolution operation. This can be achieved by summing up the product of the partial derivatives of the loss with respect to the individual components of the output.

$$\frac{\delta L}{\delta O} = \sum \left(\frac{\delta L}{\delta v} \cdot \frac{\delta v}{\delta O} \right) \quad (5)$$

In Equation 5, v represents the loss function, and O is the result of the convolution applied to the image. These partial derivatives help in adjusting the weights and biases of the network during training, ensuring that the network learns from the error at each layer. With these variables, we derive the gradient with respect to the filter. Then, we compute the gradient of the loss function with respect to the filter. The derivative provides the information needed to adjust the filter weights. The gradient is calculated as the sum of the input values multiplied by the gradient with respect to the filter in Equation 6.

$$\frac{\delta O}{\delta F} = \sum X_{ij} \cdot \Delta f(x, y, z) \quad (6)$$

This step calculates how much the output of the convolution operation changes when the filter is modified, guiding the network to adjust the filter for better feature extraction in subsequent iterations.

The modification of the backpropagation algorithm to improve CNN performance is well-supported in the literature. Several studies have explored enhancing the convolution operation and backpropagation process to optimize feature extraction, especially in the context of medical imaging [6, 7]. A major challenge in medical image analysis is the presence of noise and distortions due to factors like patient movement and varying scan quality. One effective approach to overcoming these issues is with image augmentation techniques, which artificially expand the training dataset by applying transformations such as rotation, translation, and scaling to the input images [15]. Image augmentation helps prevent overfitting, improves generalization, and ensures that the network learns robust features even in the presence of distortions.

Moreover, research has shown that incorporating additional layers in CNNs, such as 3D convolutional networks, can significantly

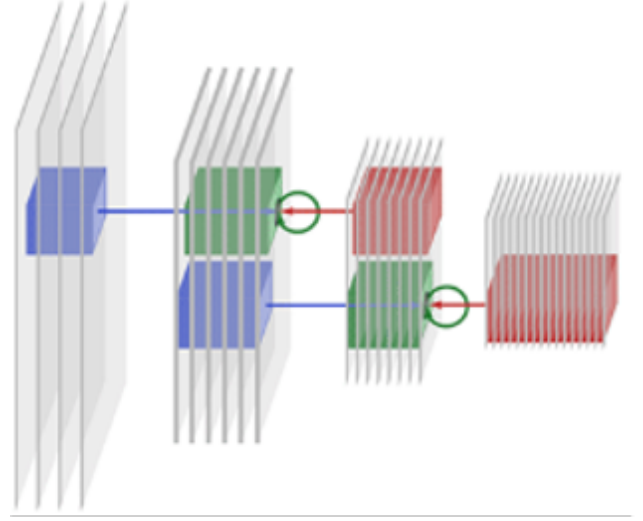


Figure 1: Types of Propagation. The blue cubes indicate forward propagation, which generates predictions. The red cubes indicate backward propagation, which is the method used by neural networks to learn and improve. The green cubes represent lateral propagation, or exchange of information between slices in the same layer, leading to augmentation and feature addition in consecutive image slices.

improve the model's ability to handle complex structures in medical images, like the detection of coronary artery calcium (CAC) in CT scans [9]. By stacking filters and applying them sequentially across multiple layers, the network can better capture hierarchical patterns in the data and identify smaller, more specific features. The proposed modification of the backpropagation algorithm and convolution operation is consistent with these findings, as it aims to enhance the network's ability to process noisy, distorted images with high sensitivity and specificity.

2.2 Network Structure

The second objective of this research centers around the development and functional integration of a novel Recurrent RCNN architecture. This new structure is designed to fully exploit all three spatial dimensions (x, y, z), which significantly enhances its ability to process and analyze complex 3D medical images. This three-dimensional approach is detailed in Figures 1 and 2, where each individual line represents a separate network processing one slice of the image set, and each "dot" in the broken line corresponds to a complete layer of the network, encompassing the entire x, y , and z dimensions.

Prior work in 3D convolutional neural networks (3D CNNs) has demonstrated the potential of extending traditional 2D CNNs to three dimensions for tasks such as medical image segmentation, object detection, and classification. Algorithms like 3D U-Net and 3D ResNet have significantly advanced volumetric data analysis [25].

These models utilize 3D convolutions to capture spatial relationships across adjacent image slices, which is essential for understanding structures in volumetric medical data. However, these earlier models typically focus on applying convolution layers independently across slices, without fully exploiting inter-slice information or the potential benefits of recurrent connections [3].

The proposed 3D Recurrent CNN structure, in contrast, integrates not only the spatial dimensions (x, y, z) but also introduces a novel recurrent mechanism to enable lateral propagation, a feature not commonly seen in traditional 3D CNNs. As shown in Figure 1, the red arrows represent the outward flow of data from the central image, while the blue arrows indicate the inward movement of data during backpropagation. This highlights the conventional forward and backward propagation steps within the training process, which are fundamental to most CNN architectures [1]. However, unlike standard models, the recurrent structure incorporated into our design allows for information to be shared laterally across neighboring nodes. This lateral propagation facilitates more complex interactions between adjacent layers, enabling the network to leverage contextual information from both the current slice and neighboring slices in the 3D volume.

The green cubes in Figure 1 illustrate the development of the recurrent state, where the core innovation of our approach is implemented. This recurrent mechanism allows the network to continuously refine its understanding of spatial relationships across the 3D volume, which is particularly valuable in medical imaging where subtle spatial dependencies between structures are crucial for accurate diagnosis. Previous work has explored recurrent networks in the context of medical image analysis, showing that recurrent connections can improve the network's ability to maintain long-term dependencies and capture contextual information across time-series data [25].

Beyond forward propagation, the new RCNN architecture employs lateral propagation between nodes during the prediction process, enabling a more robust exchange of information across layers. This allows the model to leverage techniques traditionally not possible in single-image or 2D CNN-based models, such as Intersection over Union (IoU) analysis, parity coloring, and feature overlap detection. For example, IoU, commonly used in object detection, can be more accurately applied in 3D volumes by taking advantage of the network's enhanced ability to assess spatial overlaps across slices. Similarly, parity coloring assigns higher values to pixels that are closer to corresponding pixels in the original image, further enhancing the model's sensitivity to fine-grained spatial patterns. Feature overlap, identified through Region Proposal Networks (RPNs), enables the network to focus on potential object areas, improving its detection performance in complex, highly dimensional datasets [20].

While previous algorithms for 3D CNNs have established a strong foundation for volumetric image analysis, the introduction of recurrent connections in our proposed RCNN structure offers a new approach to capturing and refining spatial dependencies across the entire 3D volume [1]. By combining forward and lateral propagation, as well as advanced image set techniques like IoU, parity coloring, and feature overlap, our model provides a more robust framework for handling the complexities of medical image data.

2.3 Modifiers

While the new backpropagation algorithm ensures that the network structure operates correctly while training on the stacked maps, the network training has not yet been optimized. One of the main issues with using 2D CNNs to detect these kinds of small, complex objects is that CNNs can only detect a singular feature or match a singular criterion for what it detects. Detecting a coronary artery in a chest CT scan requires the identification of several criteria to determine it as an artery, as shown in Figure 3 [24].

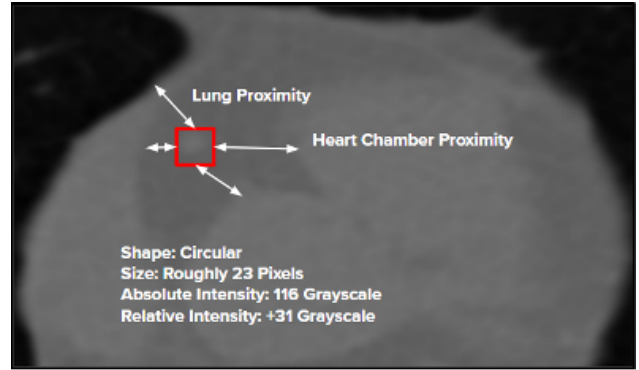


Figure 2: Arterial Conditions. Due to the complexity and variance of arteries, all the conditions labeled in the figure must be considered to detect an artery for feature detection.

2.3.1 Events with States & Plane Recurrence. Events with States is a mathematical framework designed to explore how sets of numbers and probabilities influence each other. In the context of this research, the technique treats each image layer as a separate set and works across five layers, represented as Images = (A, B, C, D, E), where each component corresponds to a distinct section of the network. As the algorithm progresses through these sets, a similarity coefficient is computed (Eq. 8). Here, A and B represent two sample images, and i and j are the pixel coordinates of these images. The calculation follows these key steps:

$$\text{Logarithmic Normalization: } \frac{1}{2} \left(\frac{A_{ij} + B_{ij}}{255} \right)$$

The term computes the average pixel intensity of corresponding pixels from both images, normalized by the maximum intensity value (255 for standard 8-bit images). Applying the natural logarithm ($\ln$) ensures that brighter regions of the image receive higher precedence. This helps highlight important features, especially in medical imaging, where lighter structures (like bones or lesions) are often crucial for analysis.

$$\text{Absolute Pixel Difference: } \frac{A_{ij} - B_{ij}}{A_{ij} + B_{ij}}$$

The absolute difference measures the pixel-level variance between the two images, revealing how much the pixel intensities differ. This term is then divided by their sum to normalize the difference, so that variations are on a similar scale for all comparisons. Squaring this difference amplifies larger discrepancies more, ensuring that significant differences between images (such as the

presence of abnormal features) are emphasized. Once the pixel-level similarity has been calculated for each pair of images, the total similarity coefficient for the image set is obtained by summing the individual pixel-wise similarities and normalizing the total number of pixels. Where N is the total number of pixels across all layers (images), and s is the similarity measure between corresponding pixels in images AA and B , the coefficient becomes:

$$\sum_{i,j=0}^N s(A, B, C, D, E) \quad (7)$$

This coefficient quantifies how similar the images are and is used as a feedback mechanism during training. During the forward pass of the network, the calculated similarity coefficients are aggregated for each image. This approach helps the network learn more complex interrelationships between different layers and variations of images. In addition to forward propagation, plane recurrence facilitates the exchange of information between layers during back-propagation, refining the model's understanding of spatial features across the 3D image. This recurrence ensures that the network doesn't treat each slice in isolation but instead considers the relationships between slices, promoting a more holistic understanding of the image data.

The determinant of the filter matrix is key to establishing connections between nodes in different layers. During backpropagation, the filter matrices are adjusted by calculating their determinants to ensure that high-importance features are transferred effectively between layers. The relationship between nodes can be mathematically modeled.

$$F_{l,adj.} = \frac{F}{\|F\|} \left(\sum_{k=1}^5 (-1)^{l-k} \frac{\|F_k\|}{2^{l-k}} \right) \quad (8)$$

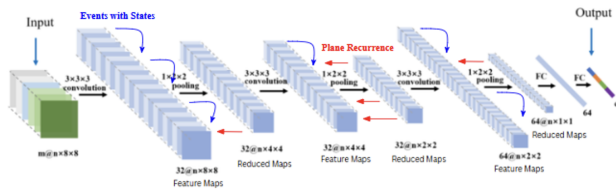


Figure 3: Network Architecture. The base layers are similar to an extended CNN architecture. The convolutional filters and feature maps are 3D, and there are modifiers applied to the network structure.

By adjusting filters based on the determinant, the network can more accurately identify and pass significant features across layers, enhancing the detection and segmentation of complex structures in medical images.

2.3.2 Calcium Score Calculation. The original images were in DICOM format, they were extracted and then converted into PNG to allow for better manipulation. This was done by setting the window level (center pixel value) and width (range of pixel intensities). Then, only the desired intensities on the DICOM image are saved. By

narrowing the range, finer details are shown to give a more detailed image [17].

The Hounsfield Units (HU) of DICOM images are often used for CAC scoring. Since such values cannot be obtained from converted PNG images, a new calcium scoring algorithm must be developed based upon two values: threshold (t) and increment (i). The usual calcium score algorithm is as follows: each lesion of calcium detected is assigned a certain scaling factor, based on its maximum intensity. If the maximum intensity was between 130 HU and 200 HU, it was assigned a factor of 1 [12]. Between 200 HU and 300 HU was given a factor of 2, between 300 HU and 400 HU was given a factor of 3, and anything above 400 HU was assigned a factor of 4 [12]. For HU values, it has an equivalent threshold $t = 130$ and varying increment $i = 70$ or 100. To find the t and i values for PNG, we had to test out all possible combinations to find the optimal one.

2.3.3 Network Validation. The network was validated with 76 non-gated chest CTs, of which the first 60 were designated as the training dataset and the last 16 as the testing dataset. These images were non-contrast and originally 512 by 512 pixels. They were later shrunk down through max pooling to 128 by 128 pixels, which is the size they were input into the network. The images were all taken from the same source (Wake Forest Baptist Medical Center) and had standardized values for the kilovoltage peak (KVP) of 120. The field of view was also noted for each scan for accurate calcium scoring.

For validation of the bounding boxes, each image from each patient was taken and annotated for the location of the Left Anterior Descending (LAD) artery, the artery that was detected in this project. Four-pixel values were assigned per coordinate: the Left X, Right X, Top Y, and Bottom Y, with the location of the artery being determined by these four measurements and the tuple (0, 0, 0, 0) being given if the artery was not located. After validation with bounding boxes, the network used the calcium scoring algorithm on its generated bounding boxes to create a calcium score that was validated against a manually calculated one.

This project was undertaken using the Python programming language, and the files were held in two places: an Ubuntu 20 virtual machine stored on a local disk and a personal account to the Bridges2 Pittsburgh Supercomputer. The virtual machine was used for most of the debugging, data analysis, data curation, and initial writing of the code, while these files were transferred to Bridges2 to utilize its parallel processing capabilities. The library mpi4py and its submodule Message Passing Interface (MPI) were used to implement parallel processing with 80 processors on the Python network scripts, allowing the network to run at six seconds per image sample.

2.4 Standard Network

The standard CNN used as a control was built using the Keras/TensorFlow libraries. It used the optimizer Adam, trained on the training set for 20 epochs and had the following architecture:

Input Layer
Convolutional Layer (64 Filters of 5x5) with ReLU Activation
Max Pooling (2x2)
Convolutional Layer (16 Filters of 5x5) with ReLU
Max Pooling (2x2)

Convolutional Layer (4 Filters of 5x5) with ReLU
 Max Pooling (2x2)
 Flattening Layer
 Dense (Fully Connected) Output Layer

3 RESULTS

3.1 Results of the Modified Network

3.1.1 Bounding Box Accuracy. To evaluate the spatial accuracy of the model, we assessed predicted bounding box coordinates using two primary metrics: the Mean Squared Error (MSE) and absolute coordinate-wise deviations from the ground truth.

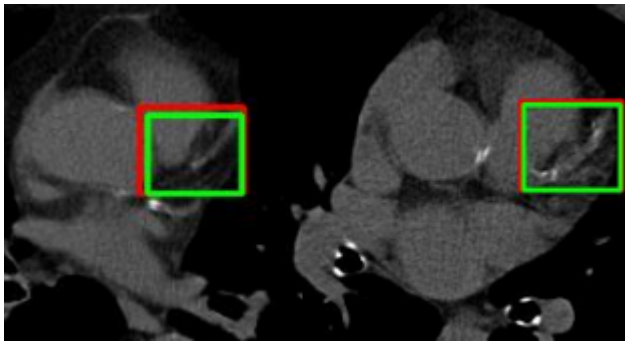


Figure 4: Annotated & Predicted Arterial Bounding Boxes in the Absence of Calcification. The green boxes are the annotated boxes, and the red are the network-generated predictions.

Coordinate deviations were defined as DLT_X (Left X), DLT_Y (Top Y), DRB_X (Right X), and DRB_Y (Bottom Y). As expected with neural network optimization, the accuracy of bounding box predictions improved progressively over the course of training. Final correlation coefficients reached 0.87 for DLT_X, 0.80 for DLT_Y, 0.97 for DRB_X, and 0.96 for DRB_Y, indicating high spatial alignment between predicted and annotated regions. Linear regression analyses further demonstrated the reliability of the model's predictions. The regression slopes were calculated as follows:

$$\text{Left X: } y = 1.0083x$$

$$\text{Top Y: } y = 1.0136x$$

$$\text{Right X: } y = 0.9955x$$

$$\text{Bottom Y: } y = 0.9931x$$

These results indicate that the predicted coordinates consistently deviate from the ground truth by less than 1.5%, with the largest deviation observed in the Top Y coordinate. The mean deviation across all four coordinates was within 1% of the ideal $y = x$ line, underscoring the model's high precision in spatial localization. On average, this is a two-pixel difference in each measurement. This is a variance of eight pixels in the perimeter and 74 pixels in the area, or an 11% decrease in the predicted area.

On a per-image basis, the average positional error between the predicted and true bounding boxes was approximately three pixels. Considering the input image resolution of 512×512 pixels and that the region of interest (i.e., the LAD artery) typically occupies a 100×100 pixel subregion, this error represents a highly accurate

localization. Moreover, the model demonstrated robustness across variations in cardiac orientation, field of view, and anatomical differences, consistently producing bounding boxes within two pixels of ground truth in most cases. Visual examples of this performance are shown in Figure 4, where the model's predicted boxes closely overlap with manually annotated references, even under varying image conditions.

3.1.2 Calcium Scoring Accuracy. Beyond anatomical localization, the model's clinical relevance was assessed through automated calcium scoring, using predicted bounding boxes as the basis for score calculation. Based on the method described earlier, threshold and increment parameters were set to $t=141$ and $i=25$, respectively. Predicted calcium scores were compared to ground truth values obtained through a validated algorithm, with results summarized in Figures 7 and 8. The null hypothesis for the data is that they are correlated significantly, and not just because of random chance.

Clinically, calcium scores are stratified into risk categories: 0, 1–10, 11–100, 101–400, and >400. The model accurately placed over 90% of test cases into their correct risk category. Among misclassified cases, the average percent error remained below 20%, suggesting that even when a score was placed outside its true bin, it remained clinically proximate. Correlation analysis between predicted and true calcium scores yielded an R^2 value of 0.921 with a p -value of 0.86, indicating a strong and statistically robust relationship. Since the p -value threshold for significance is 0.05, we fail to reject the null hypothesis.

Together, these results highlight the model's high spatial precision in localizing coronary artery calcification and its ability to generate clinically meaningful calcium scores. The minimal pixel-level error in bounding box regression, combined with strong agreement in score stratification, supports the utility of this approach for automated cardiovascular risk assessment.

3.2 Comparison to Standard Network

3.2.1 Bounding Box Comparison. The standard network had an average bounding box deviance of 25 pixels per measurement, and 100 pixels in the perimeter. This led to an average difference of 1200 pixels in the area, an increase of 171%. While the modified network had a net decrease in area, indicating that the network was too precise with its detection, the standard network was unable to properly identify the complex structures surrounding the artery and often defaulted to simply encompassing all of the nearby structures, including valves, coronary veins, and the aorta. These structures can also contain calcium, resulting in diseases such as aortic stenosis, and muddle the measure of the arterial calcium.

Since the coronary arteries have a general area where they move throughout during a CT scan, the standard network was able to create bounding boxes in a reasonable region. However, it struggled to follow and detect longer segments of the artery. It would recognize the artery well at the beginning but was unable to track it throughout the slices and instead identified alternative structures that moved into the place where coronary arteries were before. For example, the Left Anterior Descending artery (LAD) splits in the lower slices, with the main segment moving to the dorsal side of the heart, while the circumflex (CX) branch takes its place. The standard network would continually identify the CX branch in the

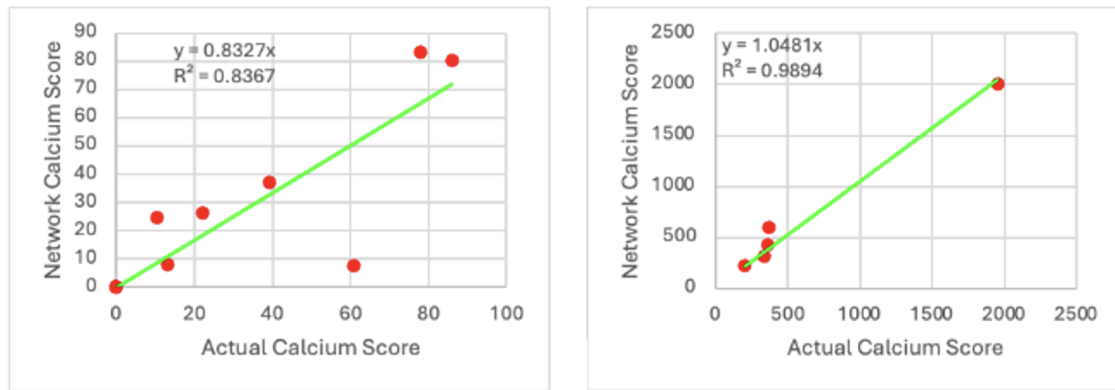


Figure 5: Actual v. Network Calcium Scores. The left graph shows calcium scores under 100, generally defined as mild or moderate. The right graph shows calcium scores above 100, defined as severe. The scores under 100 are 84% accurate, while those above 100 are 99% accurate. The scores are in the correct bin more than 90% of the time.

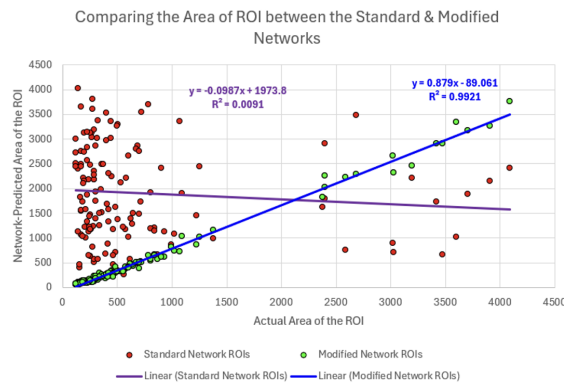


Figure 6: Analyzing the area of ROIs generated by the Standard and Modified Network. The Standard Network had the tendency to inflate bounding boxes, especially for small, precise arteries and therefore has a correlation close to 0, indicating that it was highly inaccurate. The Modified Network has a high correlation of 0.992 but does deflate the bounding boxes to 87.9% of their original size.

LAD section, and was unable to correctly identify the CX section during the CX detection process.

3.2.2 Calcium Scoring Comparison. Unlike the modified network, the standard network also fails in classifying scores into bins. The network frequently outputted a 100+ score for below 100 scores, a major issue since 100 is often used as the threshold for determining severity, especially in older patients. The network also fails for the lower scores: it is incredibly important to correctly determine 0 scores, since the presence of any arterial calcium is highly detrimental for younger patients and give false diagnoses for congenital heart disease or require unnecessary further imaging or procedures.

Scan	Actual Score	Predicted Modified Score	Predicted Standard Score
83	369	566.932	1898.117
84	361	425.294	610.465
85	0	0	23.204
86	338	318.859	480.779
87	0	0	4.397
88	86	80.124	142.551
89	1954	2004.065	2783.224
90	0	0	2.395
91	205	226.524	451.141
92	13	7.987	7.368
93	78	83.304	166.157
94	61	7.471	174.446
95	22.1	26.448	56.637
96	0	0	0
97	10.3	24.782	40.183
99	39.3	37.079	102.235

Figure 7: The arterial calcium scores of the scans, accompanied by the generated values from each network. Of the 14 scans in the testing set, the modified network fails to classify the score into the correct bin in 2 scans and the standard fails in 10 scans.

3.2.3 Reasoning for Improvements. The modified network demonstrated significantly improved performance in both artery detection and coronary calcium scoring compared to the standard architecture. These improvements were particularly evident in cases involving motion-distorted data, where the standard network consistently underperformed due to its inability to robustly account for anatomical variability and motion artifacts.

Specifically, the standard network exhibited a tendency to accurately localize arterial bounding boxes within the initial few slices; however, its performance degraded rapidly in subsequent slices. This inconsistency suggests a failure to model the temporal and spatial continuity of arterial structures across the imaging volume.

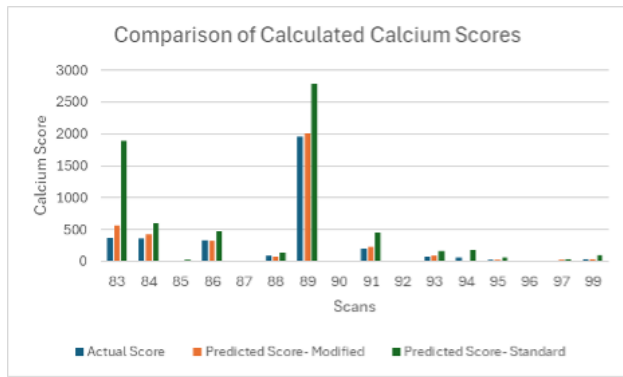


Figure 8: Comparison of Calculated Calcium Scores. The standard network had an average calcium error of 90.7%, compared to the modified network error of 24.7%. Correlation analysis between the standard score and actual score gave an R^2 value of 0.814 and a p -value of 0.408. Since the p -value threshold for significance is 0.05, we reject the null hypothesis.

In contrast, the modified network leveraged plane recurrence to maintain coherence across slices, enabling it to more reliably track arterial trajectories even in the presence of significant motion.

A key limitation of the standard convolutional neural network (CNN) approach lies in its reliance on structural uniformity. CNN filters are typically optimized to detect patterns that are consistent in size, shape, and intensity—an assumption that fails in the context of cardiac imaging where arterial geometries are often distorted due to cardiac motion. This inherent nonuniformity contributed to the standard network’s inability to detect and delineate arteries accurately.

To address these challenges, the modified network incorporated the Events with States mechanism, which facilitates dynamic modeling of structural similarity across a five-slice window. This approach allows the network to infer and model patterns of motion-related distortion, thereby improving its ability to localize arteries despite significant inter-slice displacement. While the network does not generate a full parametric mask of the distortion, it effectively contextualizes the motion and compensates for it during inference.

Furthermore, the Events with States module enhances spatial awareness by prompting the network to evaluate the anatomical context surrounding candidate arterial regions. By incorporating constraints such as proximity to cardiac and pulmonary structures, relative intensity, and geometric orientation (as illustrated in Figure 3), the network applies a more rigorous spatial validation process. This not only increases detection precision but also reduces false positives in regions with overlapping intensities or ambiguous morphology.

4 DISCUSSION

This study attempted to identify coronary arteries using novel 3D networks and propagation algorithms. In addition to accurately

identifying the coronary arteries, it automatically calculated a calcium score. Additionally, it correctly identified patients with calcium scores less than 100 and greater than 100. The results from this research demonstrate the potential of such 3D networks. These networks are capable of filtering through noisy and cluttered data and detecting objects, such as arteries, despite changes in orientation, position, and field of view, and can even account for distortion [20]. The innovative techniques described in this research as well as the novel structure that account for this accuracy are byproducts of the expansion to three dimensions modeling. We also show that the network can improve its results by using several images at once. This network is capable of accurately detecting bounding boxes of small, mobile, complex image components and can have substantial and important use in the medical field for quick and reliable analyses as part of routine care, helping in clinical decision making and health management.

While the bounding boxes around arteries (in each image) provide an extremely high level of accuracy, there can be improvement in the calcium score algorithm. Most of the observed limitations are due to the use of a new calcium score calculation algorithm since the network requires PNG images instead of full-resolution DICOM files. In a clinical setting, the network would run on the DICOM images taken straight from the scan, so the calcium score calculation could be more accurate due to its overall higher resolution. Consequently, this research can be further enhanced in a large-scale study with higher Graphical Processing Unit capacities, which can allow the network to continue learning using DICOM files. However, the accuracy achieved in this research when using PNG files has some merits. Predominantly, using the current model with PNG images allows for many more scans to be passed through the algorithm and generate calcium scores simultaneously. This is because the chest CT scans used are somewhat cheaper and much more commonly performed compared to cardiac CT scans, which are currently used for calcium score estimation. Not only does this research showcase a potential tool for clinicians to use while diagnosing a patient, but it also expands to a more general patient population that could receive a calcium score test, reducing potential underdiagnosis. Since chest CTs are taken on patients often reporting cardiac, pulmonary, or abdomen-related issues or symptoms [23], the calcium score could reveal an underlying disease or unnoticed atherosclerosis, helping clinicians find problems fast and more accurately, enabling them to provide more effective care [16].

The precision of the calcium scores maintaining the same numerical bin and the accuracy reducing the average percent error means that this network not only succeeds in detecting arteries within the image but can also identify hotspots of dense calcification [8]. These capabilities not only enable it to provide a useful calcium score but use many of the factors used in the calcium score calculation to identify parts of the artery that require the most focus and decide on which coronary bioimplants can be used for maximum effect. For example, two cases may have similar calcium scores, but different distributions of fatty plaque. A more spread-out plaque (possibly reducing its visibility) distribution along a long segment of the artery would most likely warrant the use of a stent to minimize the partial blockages and maintain the long-distance blood flow [18]. However, if the plaque was highly concentrated

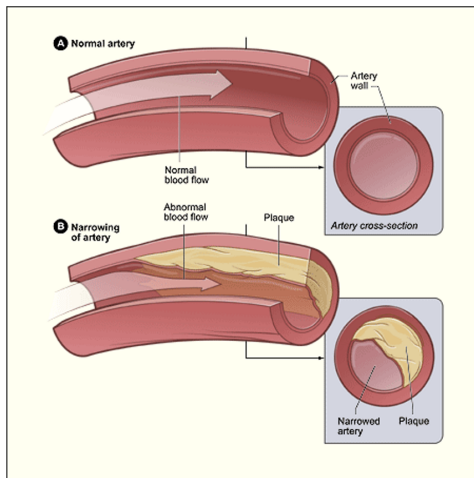


Figure 9: The Effect of Atherosclerotic Plaque on Arterial Blood Flow. Plaque buildup causes the artery to narrow, forcing blood flow to be constricted, increasing blood pressure. Even small deposits of plaque can cause disruptions in blood flow.

in a few completely blocked portions, bypass surgery would be much more effective [18]. The detection of the artery bounding boxes can be used beyond calcium scores to create 3D models of the artery through the chest and find narrow points or measure rate of blood flow, which can be used to prescribe a variety of life-saving medications such as Statins, blood thinners or beta blockers.

The implications of this research extend far beyond just CT scans and could be instrumental in revolutionizing computer vision applications in general practice. While the work focused on CT images, the approach can be generalized to a wide variety of imaging modalities, such as MRI, X-ray, and even ultrasound. In medical imaging, several common challenges, including distortion, obstruction, variance in shape, and size, can significantly degrade the quality of automated diagnoses. The algorithm developed in this research can help address these issues, particularly through transfer learning. By training on a diverse set of images, the model improves its robustness, making it more capable of handling variations inherent in medical imaging. A particularly promising extension of this research is the application to MRI scans, where images are taken from multiple angles (360 degrees) around the organ of interest.

In such cases, enabling the neural network to process images from any angle greatly expands the dataset and enhances the depth of learning during the training phase. This adaptability leads to more accurate models capable of interpreting complex anatomical structures from various perspectives, resulting in improved diagnostic reliability. In cases involving rare diseases or conditions with limited patient samples, a small dataset can hinder the network's ability to make informed decisions [22]. However, by increasing the dataset's variance through angle diversity and data augmentation, the model's robustness improves, allowing it to make more reliable and precise predictions [22]. This approach also facilitates the development of multi-modal models, capable of integrating different

imaging techniques, such as combining CT with MRI or ultrasound, further enhancing diagnostic capabilities [24].

An essential feature of the proposed model is its ability to handle multiple images at once without slowing down the training process. The system can incorporate various perspectives simultaneously, making it more efficient while expanding its knowledge base. This multi-image processing approach allows different sections of the network to specialize in different aspects, such as varying angles or distinct image features, contributing to a more refined and accurate understanding of the data. The speed and scalability of training are maintained even as the dataset grows, which is particularly crucial in real-world applications.

This research also demonstrates the potential of combining multiple network structures to address challenges like artery detection and calcium scoring in chest CT scans. These tasks have traditionally been difficult to automate due to the complexity and variability of the anatomical features. The novel network architecture proposed here—by integrating different image perspectives and applying advanced backpropagation techniques—has shown promising results, particularly in bounding box regression accuracy. This suggests that the network, although still in its early stages, holds significant potential for real-world clinical applications, where precision in detecting subtle features like arterial calcifications can directly impact patient care.

One of the core innovations of this research lies in its attempt to guide neural networks in navigating complex environments by incorporating advanced image segmentation techniques [5]. The primary goal was to assist the neural network in recognizing specific patterns rather than relying solely on trial-and-error learning [5]. Through the implementation of advanced training algorithms, we provided the network with concrete guidance, helping it better understand the intricacies of arterial structures in CT scans.

While the methods employed here are promising, they represent only a fraction of the possible algorithms that could further optimize artery detection. Future work could involve exploring the potential of hybrid models that combine CNNs with other machine learning techniques, such as reinforcement learning or attention-based models, which are already showing promise in other areas of medical imaging. Additionally, cross-validation with other algorithms could provide a clearer comparison of performance, ensuring that the most efficient and accurate methods are chosen for clinical deployment.

This research also introduced a new optimizer and a modified backpropagation algorithm to improve the training process. The flexibility of the network structure allowed for adjustments between different stack sections and creates opportunities for fine-tuning performance for specific imaging tasks. Furthermore, the incorporation of parallel processing methods enhances the computational efficiency of the network, enabling it to scale across larger datasets and handle more complex tasks in real time.

The strong Pearson correlations across all four bounding box coordinates and sub-1% average deviation from ground truth trends support the spatial robustness of the network. Moreover, the regression slopes approaching unity confirm that the model is not systematically under- or over-estimating any specific boundary. These findings compare favorably to prior literature in anatomical detection using CNNs or object detection networks, which often

exhibit reduced localization precision when applied to small or irregular targets like coronary arteries [13]. Our network's superior accuracy may be attributed to task-specific training, optimized loss functions, and the constrained spatial variability of the LAD within the dataset.

Importantly, the model's performance remained consistent despite physiological and positional variations, including changes in heart orientation, image noise, and field of view shifts across test images. This robustness suggests strong generalization capacity, an essential feature for clinical deployment where real-world data often present with high variability.

From a clinical perspective, automated coronary calcium scoring remains a crucial yet underutilized component of cardiovascular risk assessment. Traditional methods rely on manual segmentation and measurement, which are time-intensive and subject to interobserver variability. Our model demonstrated high concordance with conventional calcium scoring methods, with an R^2 value of 0.921 and over 90% bin-level accuracy in clinical calcium score categories. These results align with or exceed those reported in other automated calcium scoring systems, particularly in terms of stratifying risk within accepted clinical thresholds (e.g., 0, 1-10, 11-100, 101-400, >400) [2].

Notably, the model maintained an average score deviation of less than 20% even when misclassified, highlighting its potential utility not only in screening but also in longitudinal risk monitoring. These findings are significant, given the growing recognition of calcium scoring as an independent predictor of cardiac events, particularly in asymptomatic patients or those with equivocal risk profiles based on traditional metrics [3].

In summary, this work lays the foundation for a new generation of neural networks capable of handling a diverse range of medical imaging challenges. The adaptability of the network, combined with the ability to process multi-perspective images and apply advanced training methods, promises to improve the speed, accuracy, and scalability of medical diagnoses, particularly in environments where data is sparse or complex.

5 LIMITATIONS AND FUTURE WORK

The limitations of this study include training on only one artery: the Left Anterior Descending (LAD) instead of all four available arteries: the RCA, LAD, CX, and LM. Training on these arteries would make the model more robust and versatile. The second limitation is that the proof-of-concept model architecture was developed to run on PNG, not DICOM images. While the process of window leveling tried to ensure that the PNG image kept most of the important cardiac features in the DICOM images, testing on the DICOM images is still necessary to either increase the models' accuracy and/or assess whether there are any differences when using PNGs vs DICOMs.

Additionally, the dataset used, while diverse in anatomy and image orientation, was limited to a single imaging modality and may not reflect the full range of acquisition protocols seen in broader clinical practice. Another limitation was inaccuracy in the ground truth annotations: they were manually curated and, while reviewed

by experts, remain subject to human bias. Incorporating multi-institutional datasets with broader demographic and imaging diversity could further strengthen model generalizability.

Future directions for this project include expansion to all four coronary arteries, a larger dataset of chest CT scans, the addition of more convolutional layers, and the implementation of new techniques to further optimize the network. These adjustments would improve and optimize the network; however, it is computationally expensive. To further improve the network, images can be augmented through the application of convolutional kernels such as sharpening or modulating the relative intensity of surrounding structures, the network is able to explore in a wider range of images and environments to become better optimized and robust [10].

On CT scans, there is also the possibility of calculating the risk of valve disease. This type of network would specifically be helpful in datasets with continuous data, such as a functional MRI in the medical field. Beyond medicine, this could have wide-ranging applications in real-time detection and imaging by utilizing its third dimension to analyze images over time [25]. Another focus is integrating this detection module into a larger end-to-end diagnostic framework capable of processing raw DICOM images, performing artery segmentation, calcium quantification, and risk prediction without the need for manual intervention.

6 CONCLUSION

This project created a three-dimensional R-CNN to address common issues with 2D CNNs with respect to distortion and variance in size and shape. Through techniques such as new optimizers or events with states as well as alterations in the network structure and propagation, a five-section 3D CNN was created to train on non-gated chest CT scans, which contain distorted coronary artery motion, making it an example of the improvement of 3D CNNs over 2D CNNs. The network was used to identify bounding boxes of arteries in the chest CT scan and use those boxes to calculate a coronary artery calcium score. Both the network-generated bounding boxes and calculated calcium scores showed high accuracy, making this network capable of detecting small complex structures and providing meaningful clinical info. Applications of this network include datasets with multiple continuous images, such as a functional MRI or real-time camera data, especially in a setting with high variance and movement. The expansion of an AI network to 3D analysis using stacked images would further advance medical imaging analysis knowledge by providing information on how to analyze hyper-specific differences between consecutive images to quantify both shared and isolated features. The stacked feature maps may lead to new experiments with deriving different feature maps in the same convolutional layer. This leads to better feature isolation and a wider range of proposed features.

7 STUDENT REFLECTION

This project provided an opportunity to develop both technical and organizational skills. One of the main challenges was managing all aspects of the project independently, from planning and research to execution and problem-solving. Time management became crucial, as balancing the workload with other commitments required careful prioritization. A key solution was setting clear milestones

and breaking the project into smaller, manageable tasks, which helped maintain focus and ensure steady progress. The experience enhanced my ability to work autonomously, making me more confident in my problem-solving and decision-making abilities. From an educational perspective, it was a valuable exercise in applying theoretical concepts to practical situations. Specifically, I was inspired by concepts I learned in my classes for Multivariable Calculus and Linear Algebra, both taught by Dr. Michael Lavigne. This project demonstrated my initiative and capacity to handle complex tasks independently.

ACKNOWLEDGEMENTS

This work used Bridges-2 at Pittsburgh Supercomputing Center through allocation CHE160071 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296 [2].

REFERENCES

- [1] Luis A. Alexandre. 2014. 3D Object Recognition Using Convolutional Neural Networks with Transfer Learning Between Input Channels. In *Intelligent Autonomous Systems 13 - Proceedings of the 13th International Conference IAS-13, Padova, Italy, July 15-18, 2014 (Advances in Intelligent Systems and Computing)*, Emanuele Menegatti, Nathan Michael, Karsten Berns, and Hiroaki Yamaguchi (Eds.), Vol. 302. Springer, 889–898. https://doi.org/10.1007/978-3-319-08338-4_64
- [2] Shawn T. Brown, Paola Buitrago, Edward Hanna, Sergiu Sanielevici, Robin Scibek, and Nicholas A. Nystrom. 2021. Bridges-2: A Platform for Rapidly-Evolving and Data Intensive Research. In *Practice and Experience in Advanced Research Computing 2021: Evolution Across All Dimensions (PEARC '21)*. Association for Computing Machinery. <https://doi.org/10.1145/3437359.3465593>
- [3] Giuseppe Pio Cannata. 2025. Backpropagation in fully convolutional networks (fcns). <https://towardsdatascience.com/backpropagation-in-fully-convolutional-networks>
- [4] Kevin Clark. [n. d.]. Natural language processing with deep learning [Lecture Notes]. <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/lectures/lecture5.pdf>
- [5] Laya Das, Abhishek Sivaram, and Venkat Venkatasubramanian. 2020. Hidden representations in deep neural networks: Part 2. Regression problems. *Computers & Chemical Engineering* 139 (2020). <https://doi.org/10.1016/j.compchemeng.2020.106895>
- [6] Benoit Desjardins and Ella A. Kazerooni. 2004. ECG-Gated Cardiac CT. *American Journal of Roentgenology* 182, 4 (2004), 993–1010. <https://doi.org/10.2214/ajr.182.4.1820993>
- [7] Alexey Dosovitskiy, Jost Tobias Springenberg, Maxim Tatarchenko, and Thomas Brox. 2017. Learning to Generate Chairs, Tables and Cars with Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 4 (2017), 692–705. <https://doi.org/10.1109/TPAMI.2016.2567384>
- [8] Philip Greenland, Michael J Blaha, Matthew J Budoff, Raimund Erbel, and Karol E Watson. 2018. Coronary calcium score and cardiovascular risk [review]. *Journal of the American College of Cardiology* 72.4 (2018): 434–447. 72, 4 (2018), 434–447. <https://doi.org/10.1016/j.jacc.2018.05.027>
- [9] Harvey S. Hecht, Paul Cronin, Michael J. Blaha, Matthew J. Budoff, Ella A. Kazerooni, Jagat Narula, David Yankelevitz, and Suhny Abbata. 2017. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. *Journal of Cardiovascular Computed Tomography* 11, 1 (2017), 74–84. <https://doi.org/10.1016/j.jcct.2016.11.003>
- [10] Alex Hernández-García and Peter König. 2018. Further Advantages of Data Augmentation on Convolutional Neural Networks. In *Artificial Neural Networks and Machine Learning – ICANN 2018*, Věra Kůrková, Yannis Manolopoulos, Barbara Hammer, Lazaros Iliadis, and Ilias Maglogiannis (Eds.). Springer International Publishing, 95–103.
- [11] Hossein Hosseini, Baicen Xiao, Mayoore Jaiswal, and Radha Poovendran. 2017. On the Limitation of Convolutional Neural Networks in Recognizing Negative Images. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 352–358. <https://doi.org/10.1109/ICMLA.2017.0-136>
- [12] Abdul Rahman Idayhid, Nick S. R. Lan, Michelle Williams, David Newby, Julien Flack, Simon Kwok, et al. 2023. Evaluation of an artificial intelligence coronary artery calcium scoring model from computed tomography. *Cardiac* 33 (2023), 321–329. <https://doi.org/10.1007/s00330-022-09028-3>
- [13] Ivana Išgum, Annemarieke Rutten, Mathias Prokop, Marius Staring, Stefan Klein, Josien P. W. Pluijm, Max A. Viergever, and Bram van Ginneken. 2010. Automated aortic calcium scoring on low-dose chest computed tomography. *Medical Physics* 37, 2 (2010), 714–723. <https://doi.org/10.1118/1.3284211>
- [14] Ivana Išgum, Annemarieke Rutten, Mathias Prokop, and Bram van Ginneken. 2007. Detection of coronary calcifications from computed tomography scans for automated risk assessment of coronary artery disease. *Medical Physics* 34, 4 (2007), 1450–1461. <https://doi.org/10.1118/1.2710548>
- [15] Scott P. Johnson, J. Gavin Bremner, Alan Slater, Uschi Mason, Kirsty Foster, and Andrea Cheshire. 2003. Infants' perception of object trajectories. *Child Development* 74, 1 (2003), 94–108. <https://doi.org/10.1111/1467-8624.00523>
- [16] Nicole Karam, Sophie Bataille, Eloi Marijon, Muriel Tafflet, Hakim Benamer, Christophe Caussin, et al. 2019. Incidence, mortality, and outcome-predictors of sudden cardiac arrest complicating myocardial infarction prior to hospital admission. *Circulation: Cardiovascular Interventions* 12, 1 (2019), e007081. <https://doi.org/10.1161/CIRCINTERVENTIONS.118.007081>
- [17] Andrew Murphy. 2023. Cardiac gating (CT). *Radiopaedia* (2023). <https://radiopaedia.org/articles/cardiac-gating-ct?lang=us>
- [18] Khurram Nasir and Miguel Cainzos-Achirica. 2021. Role of coronary artery calcium score in the primary prevention of cardiovascular disease. *BMJ* (2021). <https://doi.org/10.1136/bmj.n776>
- [19] Roma Pahwa and Ishwarlal Jialal. 2023. *Atherosclerosis*. Statpearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK507799/>
- [20] K. Raju, B. Chinna Rao, K. Saikumar, and Nalajala Lakshman Pratap. 2022. *An optimal hybrid solution to local and global facial recognition through machine learning*. Springer International Publishing, 203–226. https://doi.org/10.1007/978-3-030-76653-5_11
- [21] Jia Shijie, Wang Ping, Jia Peiyi, and Hu Siping. 2017. Research on data augmentation for image classification based on convolution neural networks. In *2017 Chinese Automation Congress (CAC)*. 4165–4170. <https://doi.org/10.1109/CAC.2017.8243510>
- [22] Daniel Soukup. 2020. The necessity and pitfall of augmentation in deep learning: Observations during a case study in triplet learning for coin images. In *Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods - Volume 1 (ICPRAM 2020)*. SciTePress, 387–394. <https://doi.org/10.5220/0008910303870394>
- [23] Conrad Wittram, Michael M. Maher, Albert J. Yoo, Mannudeep K. Kalra, JoAnne O. Shepard, and Theresa C. McLoud. 2004. CT angiography of pulmonary embolism: Diagnostic criteria and causes of misdiagnosis. *RadioGraphics* 24 (2004). <https://doi.org/10.1148/rg.245045008>
- [24] Guanyu Yang, Yang Chen, Xiufang Ning, Qiaoyu Sun, Huazhong Shu, and Jean-Louis Coatrieux. 2016. Automatic coronary calcium scoring using noncontrast and contrast CT images. *Medical Physics* 43, 5 (2016), 2174–2186. <https://doi.org/10.1118/1.4945045>
- [25] Bendong Zhao, Huanzhang Lu, Shangfeng Chen, Junliang Liu, and Dongya Wu. 2017. Convolutional neural networks for time series classification. *Journal of Systems Engineering and Electronics* 28, 1 (2017), 162–169. <https://doi.org/10.21629/JSEE.2017.01.18>
- [26] Douglas P Zipes and Hein JJ Williams. 1998. Sudden cardiac death. *Circulation* 98, 21 (1998). <https://doi.org/10.1161/01.CIR.98.21.2334>