Deep Learning for CAD Prediction: X-ray Angiography Insights

Sankalp Srivastava[1], Rishi Matura[2], Sudhanshu Sharma[3], Hitesh[4], Chanpreet Singh[5]

Computer Science Engineering, Chandigarh University, Mohali, India.

Email:[1]ssankalp053@gmail.com, [2]rishigenius27@gmail.com, [3]Sudhanshusharma.cse@gmail.com,[4]hiteshrao7704@gmail.com, [5]Chanpreets920@gmail.com

Abstract

This study presents a deep learning-based approach to improve the prediction of coronary artery disease (CAD) using X-ray angiography images. The primary objective is to achieve accurate and automated CAD identification by employing a convolutional neural network (CNN) model. The methodology involves preprocessing the dataset through normalization and augmentation techniques and utilizes a U-Net architecture for precise detection of coronary stenosis. To ensure robustness and generalizability, hyperparameter tuning and dropout regularisation are applied during model training. The proposed model achieves high performance, with an average Dice coefficient of 0.57 and a Jaccard Index of 0.47 on a held-out test set, indicating its effectiveness in segmenting coronary artery stenosis. These findings support the integration of deep learning methods into clinical workflows for enhanced CAD diagnosis and early intervention.

Keywords: Coronary Artery Disease (CAD), X-Ray Angiography, Deep Learning, Convolutional Neural Network (CNN), U-Net, Segmentation, Stenosis Detection

1. Introduction

Due to its high morbidity and mortality rates, coronary artery disease (CAD) remains a significant global health challenge. CAD is characterized by the buildup of plaque leading to the narrowing or blockage of coronary arteries, which can result in severe outcomes such as heart attacks and sudden cardiac death. Early detection and accurate diagnosis are critical for

DOI: https://doi.org/10.36548/jaicn.2024.4.001
Received: 10.09.2024, received in revised form: 08.10.2024, accepted: 22.10.2024, published: 01.11.2024

effective disease control and intervention. Clinicians rely on imaging techniques, particularly X-ray angiography, to assess the shape and function of the coronary arteries and detect stenosis. Recent advancements in machine learning, especially convolutional neural networks (CNNs), have transformed medical image analysis, opening up new possibilities for risk stratification and CAD prediction. Researchers utilize large databases of X-ray angiography images to develop reliable prediction models. Among the various architectures, U-Net has been selected for this study due to its proven success in medical image segmentation tasks, particularly in delineating structures with limited and irregular shapes like coronary arteries. U-Net's encoder-decoder structure allows for precise localization and boundary detection, which is essential for identifying coronary artery stenosis. This study proposes a deep learning approach for CAD prediction using a CNN-based model trained on a well-annotated dataset of X-ray angiography images. The methodology includes precise data preprocessing, model selection, and comprehensive evaluation to ensure the model's accuracy. Additionally, a real-time CAD prediction tool is developed for healthcare professionals, allowing them to input images and receive interactive visualisations. This approach aims to enhance patient outcomes through early intervention and improved diagnosis, reducing the burden of CAD on healthcare systems.

Through this comprehensive analysis, the study highlights how machine learning-driven CAD prediction can revolutionize cardiovascular care and support personalized treatment strategies.

2. Literature Review

This section presents a thorough literature analysis, examining the earlier work and studies on heart disease prediction using machine learning algorithms. This study provides insights into existing approaches, findings, and gaps in the subject, serving as a platform for our investigation. To enhance coronary artery segmentation in cardiac CT angiography, the work[1] introduces DR-LCT-UNet, a modified U-Net architecture that incorporates Dense Residual and Local Contextual Transformer modules. The authors report better performance on multiple criteria when compared to current approaches. They credit a number of techniques, including deep supervision, multi-level feature retention, improved contrast, and noise reduction, for this improvement. Nevertheless, the study does not analyse cases of CAD that are very difficult (such as severe stenosis or widespread calcifications).

In this paper [2], a useful deep learning-based tool called CoMoFACT—which can mimic coronary artery motion artefacts in CT images—is shown. This model offers a new method for producing realistic training data, which is essential for creating deep learning models that are reliable for CAD applications where motion artefacts can seriously deteriorate image quality. The study shows that deep learning models trained on CoMoFACT data can precisely measure motion artefacts. This feature is clinically significant since it may improve imaging procedures and may activate real-time motion correction algorithms to improve diagnostic efficacy.

This research [3] targets motion artefact correction, a major problem in cardiac CT angiography, by presenting CoMPACT, a deep learning-driven method. The technique measures motion vectors by utilising convolutional neural networks (CNNs) and incorporates them into a specific motion compensation algorithm. Although the study's results on clinical photos are intriguing, it admits its limits when it comes to handling extremely complicated motion patterns. This emphasises how these techniques must be further improved to handle the entire spectrum of possible motion artefacts that arise in clinical CAD diagnosis.

3. Convolutional Neural Networks (CNN)

CNNs are a powerful deep-learning architecture specifically designed for image analysis [4]. Their ability to learn increasingly complex hierarchical features makes them exceptionally well-suited for medical imaging tasks, including those in coronary artery disease analysis. The core principle of convolution, inspired by the visual system, enables CNNs to extract patterns at various scales while preserving spatial relationships within the image. This is crucial in medical images[5] where subtle variations in texture, shape, and intensity often encode important diagnostic information. The convolutional layer is a basic component of CNNs. These layers use convolution to process the input image through a range of flexible filters. Every filter is made to find certain elements in isolated areas of the input images such as edges, textures, or forms. These filters create feature maps that encode more abstract representations of the input data as they go across the input image. By using this hierarchical feature extraction method, CNNs learn to recognize complex variations and patterns seen in the pictures. CNNs frequently use pooling layers in addition to convolutional layers to decrease the spatial dimensions of feature maps, which lowers computational costs and enhances translation invariance. CNNs can capture complex associations among characteristics because of the introduction of nonlinearity through activation functions such as ReLU. Fully connected layers integrate the learned features from the convolutional and pooling layers to yield the final output, which in image classification applications, represents the class probabilities.

4. CNN Classification Algorithms

LeNet, AlexNet, VGGNet, R-CNN, U-Net, and SegNet are just a few of the groundbreaking CNN designs that have greatly advanced computer vision. Although the depth, intricacy, and performance of these designs differ, they all aim to maximize CNNs' ability to learn hierarchical representations of visual data. Table 1 provides a comparison of these CNN models, detailing their architectures, along with their respective advantages and disadvantages.

Table 1. CNN Classification Algorithm Comparison

Algorithm	Model Architecture	Advantages	Disadvantages	Applications
LeNet (LeCun et al.,1998)[6]	Convolutional Neural Network (CNN) with convolutional layers, pooling layers, and fully connected layers for digit recognition.	Simple architecture, easy to implement. Effective feature extraction via convolutional layers.	Shallow architecture lacks depth for complex tasks. Limited generalisation to larger and more varied datasets.	Handwritten digit recognition (MNIST), early image classification tasks.
AlexNet (Krizhevsky et al., 2012)[7]	Deep CNN with stacked convolutional layers, ReLU activations, and dropout regularisation.	Introduced deeper network layers and ReLU for faster training. Improved feature learning via larger convolutional kernels.	High computational cost due to deeper architecture. Requires a large amount of labelled data.	Image classification (e.g., ILSVRC), object detection, CAD image analysis.

VGGNet (Simonyan & Zisserman, 2014)[8]	Very deep CNN with small (3x3) convolutional filters, max-pooling layers, and fully connected layers.	Consistent use of small filters enhances network depth. Strong feature extraction power.	Computationally expensive due to depth. High memory usage for training.	Image classification, feature extraction for medical images, CAD analysis.
R-CNN (Girshick et al., 2014)[9]	Region-based CNN that generates region proposals, applies CNN feature extraction, and classifies regions.	High accuracy for object detection in complex scenes. Adaptable to various detection tasks.	Multi-stage pipeline is computationally intensive. Slow processing compared to newer methods.	Object detection, medical image lesion detection, CAD region analysis.
U-Net (Ronneberger et al., 2015)[10]	Encoder-decoder architecture with skip connections, specifically for image segmentation tasks.	Excellent for pixel-level classification and segmentation. Combines high-level features with spatial details via skip connections.	Can be excessive for simpler tasks. Requires substantial computational resources.	Biomedical image segmentation, artery lumen segmentation in CAD.
SegNet (Badrinarayanan et al., 2017)[11]	Encoder-decoder architecture with efficient memory usage, optimised for segmentation tasks.	Highly efficient with memory-saving techniques in the decoder. Suitable for real-time tasks.	Sacrifices some segmentation accuracy for efficiency.	Real-time image segmentation, large-scale CAD dataset analysis.

5. Methodology

5.1Dataset and Preprocessing

This study utilised a dataset of 1,200 angiography images from the ARCADE[12](Atlas of Coronary Artery Disease and its Effects) dataset, focusing on cases with coronary artery stenosis (Figure 1). The dataset was divided into 1,000 images for training, with 200 images held out for validation. Image masks were generated using the provided information on closed polygons, which marked the areas of interest in the angiography images. The dataset was organized based on critical selection criteria, including stenosis severity distribution and image quality standards. To address common variations in brightness and contrast found in angiography images, min-max scaling was applied for intensity normalisation. Dataset-specific parameters were optimised to ensure consistency, which is crucial for accurate CAD analysis. Furthermore, data augmentation techniques like random flips and rotations were applied to simulate realistic variations in coronary artery orientation. These augmentations aimed to enhance model generalisation and reduce the risk of overfitting, particularly due to the limited dataset size.

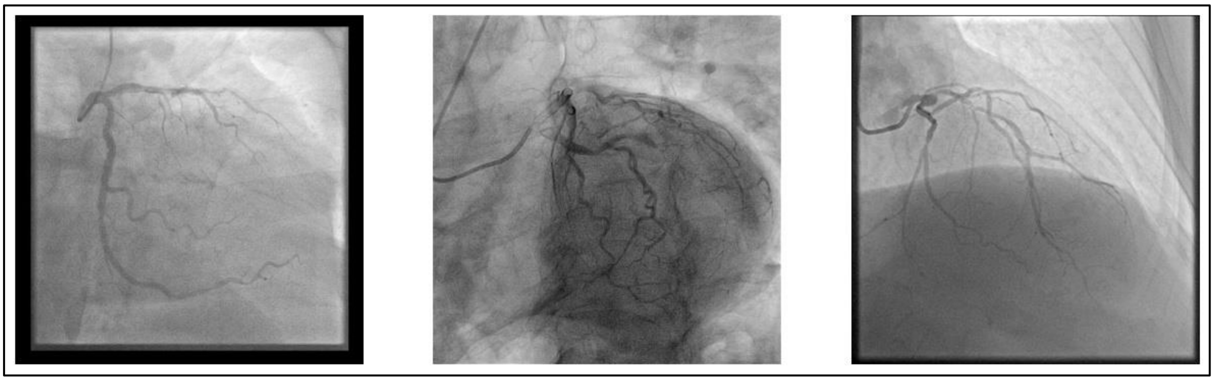

Figure 1. Sample Training Images from the ARCADE[12] Dataset.

5.2Model Architecture and Summary: U-Net for Coronary Artery Segmentation

The proposed U-Net model leverages a multi-level convolutional structure for precise coronary artery stenosis segmentation. The architecture is structured with specific components that include the following layers:

- DoubleConv Module: Each DoubleConv module consists of two consecutive 3x3 convolutional layers, each followed by batch normalisation and a ReLU activation. This module captures complex features at every stage of downsampling and upsampling, enhancing the U-Net’s ability to discern stenosis-related details.
- Encoder Path: Comprising the DoubleConv modules and max-pooling layers, the encoder path has four levels that progressively downsample the input, preserving essential feature information at various scales.
- Bottleneck Layer: Positioned between the encoder and decoder paths, the bottleneck layer compresses and encodes global context across multiple scales, essential for accurately segmenting multi-scale stenotic regions.
- Decoder Path: The decoder mirrors the encoder path with four levels, each upsampling the input via transposed convolutions followed by a DoubleConv

module. This path restores the spatial resolution while capturing fine-grained details through concatenation with encoder feature maps using skip connections.

- Final Convolution and Output Layer: A 1x1 convolution maps the output to segmentation classes, followed by a sigmoid activation to yield probabilistic outputs for mask segmentation.
- Dice Loss for Segmentation: A customised Dice loss function is used to calculate the overlap between predicted and target masks, supporting single mask segmentation, and ensuring effective learning even with imbalanced classes.

Model Summary: The model consists of 7.8 million trainable parameters, with no non-trainable parameters, bringing the total parameters to 7.8 million. The trained model size is 31.062 MB. The model has 82 modules in train mode, indicating the complex data tasks of image segmentation.

This implementation optimises the U-Net for complex stenosis segmentation, supported by advanced feature extraction and context-aware design. The Dice coefficient loss with Sigmoid-adjusted prediction enables accurate segmentation masks, leveraging this custom architecture for improved clinical applicability in CAD analysis.

5.3Training Strategy and Hyperparameters

The model was trained using the Adam optimizer with an initial learning rate of 10-3 , which was dynamically adjusted using a learning rate scheduler (reducing the rate by 0.1 when the validation loss plateaued). The Dice coefficient loss function was chosen, given its suitability for imbalanced datasets, particularly in medical imaging like CAD analysis.

Key hyperparameters were set as follows:

- Epochs: 50
- Batch Size: 16
- Dropout Rate: 0.5 (applied in both the encoder and decoder to prevent overfitting)

Additionally, data augmentation techniques were employed to simulate real-world variations, including random rotations, flips, and scaling, enhancing generalisation and reducing overfitting risks.

5.4Preprocessing Methodology

The preprocessing pipeline in this study is designed to optimise the X-ray angiography images, improving model performance and ensuring consistent input for effective CAD segmentation. The preprocessing consists of two main steps:

- Data Normalisation: Images are scaled to a uniform range of [0, 1] by dividing each pixel value by the maximum intensity (255 for grayscale images). This helps stabilise and expedite training by providing standardised input data.
- Data Augmentation:
    - Custom Augmentation Transformer: This transformer applies random rotations (90°, 180°, and 270°) and horizontal flips with a 50% probability. The randomised transformations introduce variability to the dataset, making the model more robust to orientation and position changes.
    - Preprocessing Image Pipeline: The function applies a binary threshold to the image, identifying non-black pixels and facilitating the detection of contours. Using these contours, the area of interest is defined for cropping. The image and its corresponding mask are then cropped based on these coordinates. A median blur is applied to the cropped image to reduce noise, and CLAHE (Contrast Limited Adaptive Histogram Equalization) is utilised to enhance contrast. The processed image and mask are resized to a target dimension of (512, 512) to ensure uniform input for the deep learning model.

These preprocessing techniques significantly enhance the U-Net's ability to detect stenosis in varying angiography images by standardising intensity and improving data diversity. Figure 2, which shows sample images processed using normalisation, noise reduction, and CLAHE can visually demonstrate the effectiveness of this pipeline.

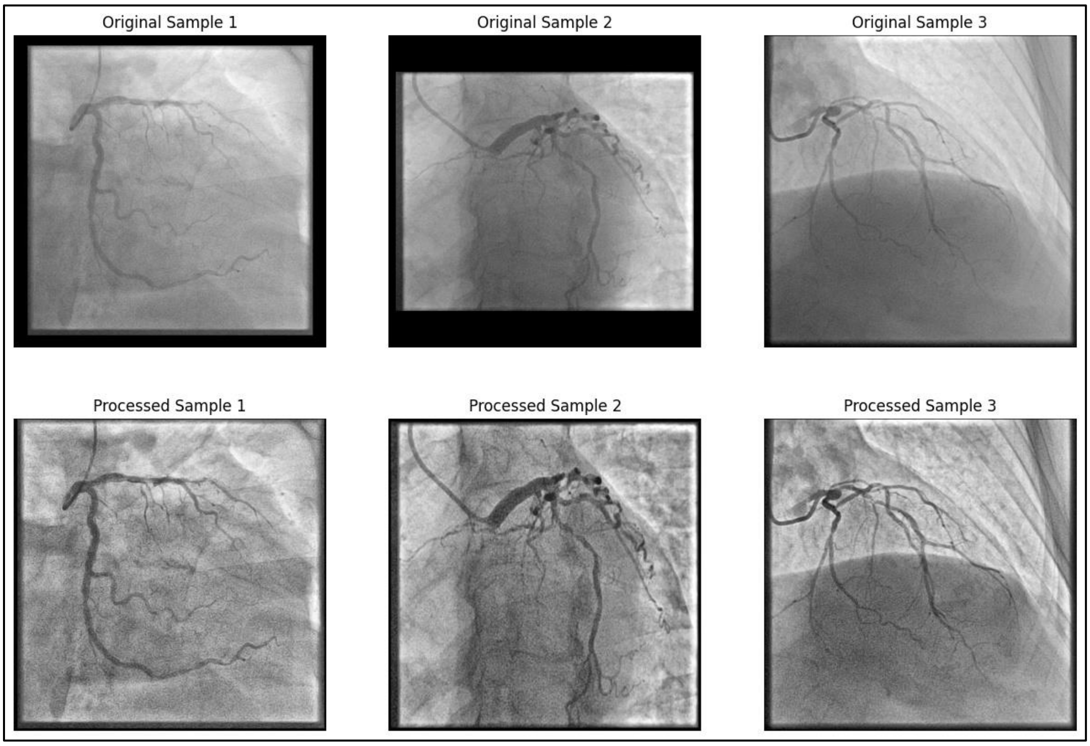

Figure 2. Pre-processed Sample Images Displaying Cropping, Contrast Enhancement and Noise Reduction

5.5Tools Used for Implementation

- TensorFlow/Keras: For model design, training, and evaluation.
- PyTorch Lightning: For simplifying PyTorch code structure and managing model training workflows.
- Python: The primary programming language for implementation and data preprocessing.
- NumPy: For numerical computations and matrix operations.
- OpenCV: Used for image processing and augmentation tasks.
- Matplotlib/Seaborn: For visualising training metrics and segmentation results.

6. Evaluation Metrics

Metrics for assessment are statistical instruments that compare the predicted masks with the actual masks in the dataset to determine how successfully a model segments masks. The Sørensen–Dice index, or Dice Coefficient, measures the overlap between the ground truth and predicted masks to quantify segmentation accuracy. A perfect segmentation score is one. It is essential for CAD analysis since it is extremely sensitive to segmentation boundaries[13]. One further way to quantify segmentation overlap is through the Jaccard

Index (Intersection over Union). Regarding coronary artery segmentation models, this is pertinent and conceptually similar to the Dice Coefficient. The sensitivity (recall) of the model measures its ability to identify actual artery stenosis pixels[14]. In order to diagnose CAD, even minor plaque or mild stenosis must be detected using a very sensitive model.

7. Result

Careful normalisation using the normalisation method ensured consistent intensity values across the ARCADE[12] angiography dataset, improving the U-Net model's feature learning process. Data augmentation, specifically random flips and rotations, successfully diversified the training set. This enhanced the model's ability to recognize stenoses in images acquired with varying orientations, ultimately boosting its robustness. The U-Net model was trained using the Dice coefficient loss function to optimise its segmentation capabilities, and the Adam optimiser proved effective in guiding the learning process. Analysis of the loss curve over training epochs provided insights into model convergence behaviour and helped identify any potential overfitting issues. Evaluation on a held-out test set demonstrated the model's strong performance in segmenting coronary artery stenosis. The projected segmentations and expert-annotated ground truth masks demonstrated strong alignment, with an average Dice coefficient of 0.57 and a Jaccard Index of 0.47, indicating effective distinction between stenotic and non-stenotic regions. With sensitivity of 0.86, the model is detecting about 86% of actual stenosis regions. Compared to the validation results in Kjerland's 2017 [15] study, which achieved a Dice coefficient of 0.59 on CT scans, our slightly lower performance is attributed to image challenges such as background noise and distortions, which affect segmentation accuracy, especially in smaller artery structures and stenotic regions.

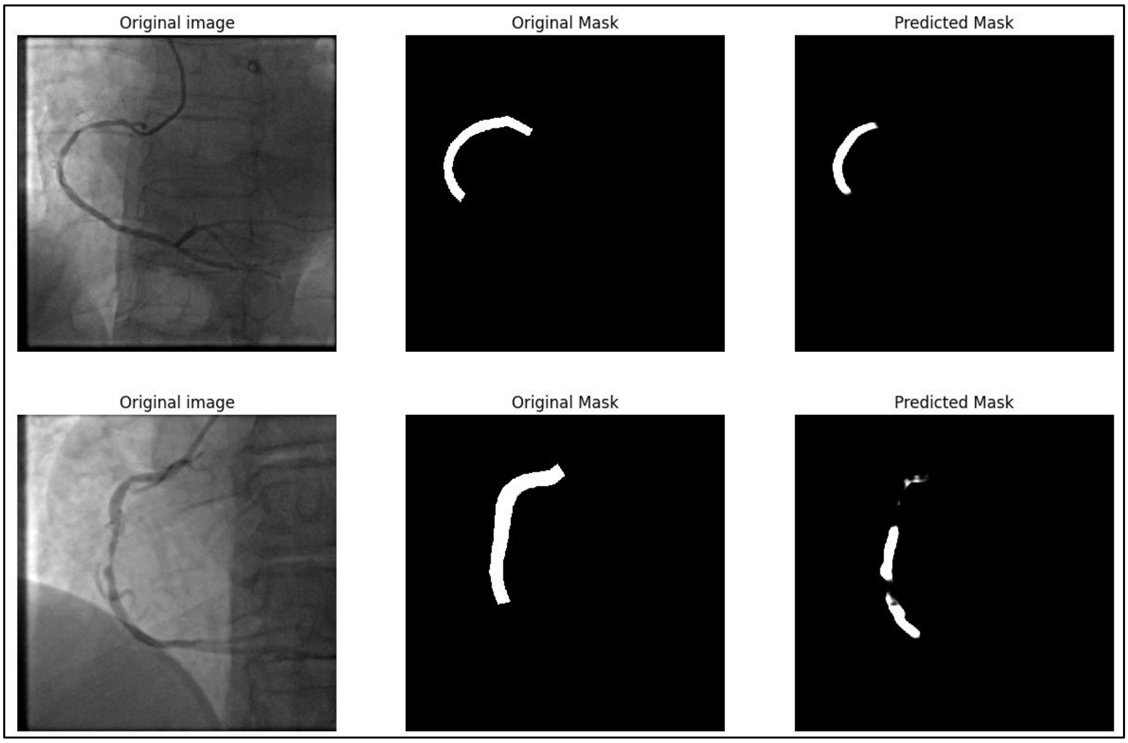

Figure 3. Coronary Artery stenosis in Angiography Image, Stenosis Region Predicted by the Model.

Visual inspection of segmentation results confirmed the quantitative findings. The model reliably delineated stenotic areas in a variety of images but also offered a nuanced view of specific challenging cases where refinement may be needed. Figure 3, which shows a coronary artery stenosis in an angiography image with the stenosis region predicted by the model, illustrates this performance. These insights will be crucial in guiding further model development

8. Conclusion

This study demonstrates how well a properly modified U-Net design performs when segmenting coronary artery stenosis in angiography images. Important contributions include showing how customised data augmentation and dataset-specific intensity normalisation affect model resilience, especially when managing the unpredictability found in clinical CAD analysis. Thorough hyperparameter adjustment, directed by the Dice coefficient loss, produced precise segmentation, as demonstrated by robust assessment metrics. These results highlight how deep learning may improve and expedite CAD workflows. The created model provides a good foundation for automated stenotic area delineation techniques. This approach could assist medical professionals in diagnosing and quantifying stenosis severity more efficiently, which is crucial for timely treatment decisions. Future research will explore several new directions. One key area will be validating the clinical utility of the model by examining additional CAD-specific parameters. The next step involves investigating the model’s capabilities for stenosis quantification, enabling automated evaluation of disease burden. Ultimately, these advancements could enhance patient care by integrating the findings into decision support systems. Additionally, the reviewed literature emphasizes the growing interest in such models, further highlighting their potential for clinical implementation.

References

[1] Sarker, Iqbal H. "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions." SN computer science 2, no. 6 (2021): 420.

[2] Bhatt, Dulari, Chirag Patel, Hardik Talsania, Jigar Patel, Rasmika Vaghela, Sharnil Pandya, Kirit Modi, and Hemant Ghayvat. "CNN variants for computer vision: History, architecture, application, challenges and future scope." Electronics 10, no. 20 (2021): 2470.Z.

[3] Lipton, Zachary Chase. "A Critical Review of Recurrent Neural Networks for Sequence Learning." arXiv Preprint, CoRR, abs/1506.00019 (2015).

[4] Pan, Li-Syuan, Chia-Wei Li, Shun-Feng Su, Shee-Yen Tay, Quoc-Viet Tran, and Wing P. Chan. "Coronary artery segmentation under class imbalance using a U-Net based architecture on computed tomography angiography images." Scientific Reports 11, no. 1 (2021): 14493.

[5] Wang, Qianjin, Lisheng Xu, Lu Wang, Xiaofan Yang, Yu Sun, Benqiang Yang, and Stephen E. Greenwald. "Automatic coronary artery segmentation of CCTA images using UNet with a local contextual transformer." Frontiers in Physiology 14 (2023): 1138257.

[6] LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86, no. 11 (1998): 2278-2324.

[7] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." Communications of the ACM 60, no. 6 (2017): 84-90.

[8] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014)..

[9] Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587. 2014.

[10] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, pp. 234-241. Springer International Publishing, 2015.

[11] Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation." IEEE transactions on pattern analysis and machine intelligence 39, no. 12 (2017): 2481-2495.

[12] Maxim Popov, A. A. "ARCADE: Automatic Region-based Coronary Artery Disease diagnostics using x-ray angiography imagEs Dataset Phase." (2023).https://zenodo.org/records/8386059

[13] Gharleghi, Ramtin, Nanway Chen, Arcot Sowmya, and Susann Beier. "Towards automated coronary artery segmentation: A systematic review." Computer Methods and Programs in Biomedicine 225 (2022): 107015.

[14] Arroyo-Espliguero, Ramón, Pablo Avanzas, Juan Quiles, and Juan Carlos Kaski. "Predictive value of coronary artery stenoses and C-reactive protein levels in patients with stable coronary artery disease." Atherosclerosis 204, no. 1 (2009): 239-243.

[15] Kjerland, Øyvind. Segmentation of Coronary Arteries from CT-Scans of the Heart Using Deep Learning. Master's thesis, Norwegian University of Science and Technology (NTNU), 2017.

Author's biography

Sankalp Srivastava - An undergraduate student at Chandigarh University pursuing a B.E. in computer science and engineering (CSE). Currently in the program's final year.

Rishi Matura - An Undergraduate student currently studying at Chandigarh University, pursuing a B.E. in Computer Science and Engineering (CSE), and currently in the program's final year.

Sudhanshu Sharma - A dedicated faculty member at Chandigarh University, specialising in Computer Science and Engineering (CSE), committed to fostering student learning and innovation.

Hitesh - An undergraduate student at Chandigarh University pursuing a B.E. in computer science and engineering (CSE). Currently in the program's final year.

Chanpreet Singh - An undergraduate student in the final year of the B.E. program in Computer Science and Engineering (CSE) at Chandigarh University.