International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Detecting Cardiovascular Disease from Mammograms with Deep Learning

Shivani Pal[1], Pradeep Yadav[2]

ITM Institute of Technology and Management

Abstract: ***Cardiovascular disease (CVD) remains a leading global cause of mortality, necessitating innovative, non-invasive methods for early detection to enable timely interventions. Breast arterial calcifications (BAC), visible in mammograms, have emerged as a promising biomarker for assessing CVD risk. This research investigates the application of deep learning, specifically convolutional neural networks (CNNs), to detect BAC in mammograms and stratify CVD risk. We propose a robust CNN-based model trained on a large, annotated mammogram dataset to identify calcifications indicative of cardiovascular risk. The model demonstrates high accuracy, precision, and recall, highlighting its potential to integrate seamlessly into routine mammography screening workflows, thereby enhancing early CVD detection in a cost-effective and scalable manner Key words Alzheimer's disease, dementia, deep learning, neuroimaging, biomarkers, convolutional neural networks (CNNs), recurrent neural networks (RNNs), Explainable AI (XAI), early diagnosis, disease progression modeling.***

Keywords: Cardiovascular disease, breast arterial calcifications, deep learning, convolutional neural networks, mammography, CVD risk assessment, early detection, medical imaging, artificial intelligence, non-invasive screening

1. Introduction

Cardiovascular disease (CVD) is a major global health challenge, accounting for approximately 17.9 million deaths annually, making it the leading cause of mortality worldwide (World Health Organization, 2020). CVD encompasses a spectrum of conditions, including coronary artery disease, heart failure, stroke, and peripheral artery disease, which collectively impose a significant burden on healthcare systems, particularly in low- and middle-income countries where diagnostic resources are scarce (Roth et al., 2020). Early detection of individuals at risk is paramount for implementing preventive measures, such as lifestyle changes, pharmacological therapies, or invasive interventions like coronary artery bypass grafting, to mitigate disease progression and improve patient outcomes (Arnett et al., 2019). However, traditional diagnostic methods, such as coronary angiography, computed tomography (CT) coronary artery calcium (CAC) scoring, or stress echocardiography, are often invasive, costly, or require specialized infrastructure, limiting their feasibility for population-wide screening (Greenland et al., 2018).

Mammography, a cornerstone of breast cancer screening, is routinely performed for women over the age of 40 in many countries, with millions of mammograms acquired annually as part of national screening programs (U.S. Preventive Services Task Force, 2016). Beyond its primary role in detecting breast malignancies, mammography captures breast arterial calcifications (BAC)—linear calcium deposits within the arterial walls of the breast that appear as high-intensity, tram-track-like structures on mammogram images (Iribarren & Molloi, 2013). These calcifications are distinct from microcalcifications associated with breast cancer and have been increasingly recognized as a biomarker for cardiovascular risk. Clinical studies have established a robust correlation between BAC and cardiovascular risk factors, including coronary artery calcification, hypertension, diabetes, smoking, and hyperlipidemia (Margolies et al., 2016; Iribarren et al., 2018). For example, Margolies et al. (2016) found that women with BAC on mammograms had a significantly higher likelihood of coronary artery disease, with BAC serving as an independent predictor of cardiovascular events. Similarly, Iribarren et al. (2018) conducted a meta-analysis demonstrating that BAC prevalence is associated with a 2- to 3-fold increased risk of cardiovascular events, even after adjusting for traditional risk factors such as age, cholesterol levels, and smoking status. Polonsky and Greenland (2017) further emphasized the clinical significance of BAC, noting its potential to expand cardiovascular risk assessment in women, a population historically underrepresented in CVD research.

The recognition of BAC as a non-invasive biomarker for CVD risk presents a transformative opportunity to repurpose mammography for dual-purpose screening, leveraging existing imaging infrastructure to identify women at risk of cardiovascular disease without additional radiation exposure, cost, or patient burden (Mostafavi et al., 2015). This approach is particularly compelling because mammography is already embedded in routine healthcare for a large demographic, offering a scalable platform for CVD risk stratification. However, manual detection of BAC by radiologists is challenging due to its subtle and variable appearance, which can be confounded by overlapping breast tissue or imaging artifacts (Iribarren & Molloi, 2013). Manual assessment is also time-consuming, subjective, and prone to inter-observer variability, making it impractical for large-scale screening programs (Dromain et al., 2013).

Artificial intelligence (AI), particularly deep learning, has revolutionized medical imaging by enabling automated, accurate, and reproducible analysis of complex patterns (Litjens et al., 2017). Convolutional neural networks (CNNs), a class of deep learning models, are particularly adept at image-based tasks due to their ability to learn hierarchical feature representations directly from raw pixel data, eliminating the need for labor-intensive hand-crafted

features (LeCun et al., 2015). CNNs have achieved state-of-the-art performance in various medical imaging applications, including breast cancer detection in mammograms (Esteva et al., 2017), lung nodule classification in CT scans (Shen et al., 2019), and diabetic retinopathy diagnosis in retinal fundus photography (Gulshan et al., 2016). Recent studies have also explored deep learning for detecting vascular calcifications in other imaging modalities, such as chest X-rays, with promising results (Cheng et al., 2020). However, the application of deep learning to BAC detection in mammograms for CVD risk assessment remains an underexplored area, representing a critical research gap.

This paper aims to address this gap by developing and evaluating a deep learning model to detect BAC in mammograms and stratify associated CVD risk. We hypothesize that a CNN, trained on a diverse dataset of annotated mammograms, can accurately identify BAC and classify CVD risk levels, offering a scalable, automated solution that complements existing mammography workflows. The proposed model is based on the ResNet-50 architecture, fine-tuned for BAC detection, and trained on a large dataset of mammograms annotated by expert radiologists (He et al., 2016). By integrating deep learning into routine mammography, this approach has the potential to enhance early CVD detection, prioritize at-risk individuals for further diagnostic evaluation, and reduce the global burden of cardiovascular disease. The study also compares the model's performance to traditional machine learning methods and human experts, while exploring its practical implications for clinical deployment and scalability, clinical research.

2. Related Work

The association between BAC and CVD has been substantiated by numerous clinical studies. Margolies et al. (2016) demonstrated that women with BAC on mammograms were significantly more likely to have coronary artery disease, with BAC serving as an independent predictor of cardiovascular events. Iribarren et al. (2018) reported a positive correlation between BAC prevalence and traditional CVD risk factors, such as age, smoking, and hyperlipidemia, in a comprehensive meta-analysis. Polonsky and Greenland (2017) further highlighted BAC's role as a marker for subclinical atherosclerosis, advocating for its integration into cardiovascular risk assessment protocols. Mostafavi et al. (2015) found that BAC was associated with coronary artery calcium scores, reinforcing its utility as a proxy for coronary atherosclerosis.

Historically, BAC detection has relied on manual interpretation by radiologists, a process that is labor-intensive, subjective, and prone to inter-observer variability (Dromain et al., 2013). To address these limitations, early machine learning approaches, such as support vector machines (SVMs) and random forests, were applied to BAC detection (Wang et al., 2016). However, these methods depend on hand-crafted features, such as texture, intensity, or shape descriptors, which may not fully capture the complex and subtle patterns associated with BAC (Hosny et al., 2018).

Deep learning has transformed medical imaging by enabling end-to-end feature learning, bypassing the need for manual feature engineering (Litjens et al., 2017). CNNs have been successfully applied to tasks such as breast cancer detection in mammograms (Dhungel et al., 2015), lung nodule classification in CT scans (Shen et al., 2019), and retinal disease diagnosis in fundus photography (Gulshan et al., 2016). Notably, Cheng et al. (2020) developed a CNN model to detect vascular calcifications in chest X-rays, achieving high sensitivity and specificity. Despite these advances, few studies have explored deep learning for BAC detection in mammograms specifically for CVD risk assessment. This research aims to fill this gap by developing a tailored CNN model to identify BAC and evaluate its clinical utility in CVD screening.

3. Methodology

3.1 Dataset

The study utilized a comprehensive dataset of 12,000 mammogram images sourced from multiple repositories to ensure diversity and generalizability. The primary source was the Digital Database for Screening Mammography (DDSM), a publicly available collection of digitized film mammograms (Heath et al., 1997). This was supplemented with a private dataset provided by a collaborating tertiary care hospital, which included digital mammography images from a diverse patient population. Each mammogram was accompanied by metadata, including patient age, medical history, and radiologist annotations indicating the presence and extent of BAC, as well as associated CVD risk levels (categorized as low, medium, or high based on clinical guidelines).

The dataset was meticulously curated to ensure balanced representation across age groups, ethnicities, and CVD risk profiles. Images were labeled by a panel of board-certified radiologists with expertise in mammography and cardiovascular imaging. To mitigate bias, each image was independently reviewed by at least two radiologists, with discrepancies resolved through consensus. The final dataset was partitioned into 70% for training (8,400 images), 15% for validation (1,800 images), and 15% for testing (1,800 images), maintaining consistent class distributions across splits. The dataset distribution is summarized in Table 1.

Table 1: Dataset Distribution

Subset	Number of Images	No BAC (%)	Low CVD Risk (%)	High CVD Risk (%)
Training	8,400	50%	30%	20%
Validation	1,800	50%	30%	20%
Testing	1,800	50%	30%	20%
Total	12,000	50%	30%	20%

3.2 Preprocessing

To optimize model performance and ensure robust feature extraction, the mammogram images underwent a series of preprocessing steps, as outlined in Table 2:

Step	Description	Purpose
Normalization	Pixel intensities scaled to [0, 1] by dividing by maximum value (255)	Standardize input data, reduce variability from imaging equipment.
Contrast Enhancement	Histogram equalization to enhance calcification visibility.	Improve visibility of high-intensity calcifications against breast tissue.
Image Resizing	Resized to 224x224 pixels using bicubic interpolation.	Ensure compatibility with CNN architecture, preserve details.
Data Augmentation	Random rotations (±15°), flips, zooming (20%), brightness Adjustments (±10%).	Increase dataset diversity, prevent overfitting.
ROI Extraction	Adaptive thresholding to isolate breast tissue from background.	Focus model on relevant regions, reduce computational overhead.

These steps were implemented using Python libraries such as OpenCV and scikit-image, ensuring reproducibility and compatibility with deep learning frameworks (Litjens et al., 2017; Dhungel et al., 2015; Ronneberger et al., 2015).

3.3 Model Architecture

The proposed model was built upon the ResNet-50 architecture, a deep CNN known for its effectiveness in image classification tasks due to its residual learning framework (He et al., 2016). ResNet-50 was chosen for its ability to mitigate the vanishing gradient problem through skip connections, enabling the training of deep networks without performance degradation. The architecture was customized for BAC detection, as detailed in Table 3:

Table 3: Model Architecture Overview

Layer Type	Description	Parameters
Input Layer	Accepts 224x224 grayscale images, initial convolution with 64 filters.	3.2M parameters
Convolutional Blocks	50 layers in 4 stages, residual blocks with 1x1, 3x3, 1x1 kernels, ReLU.	23.5M parameters
Global Average Pooling	Reduces spatial dimensions (7x7x2048 to 1x1x2048)	No parameters
Fully Connected Layer	512 units, ReLU activation	1M parameters
Output Layer	Softmax with 3 units (no BAC, low risk, high risk).	1.5K parameters

The model was implemented using TensorFlow 2.8 and Keras, with pre-trained weights from ImageNet used for transfer learning to accelerate convergence (Deng et al., 2009). Training and inference were performed on a high-performance computing cluster equipped with four NVIDIA A100 GPUs, each with 40 GB of memory.

3.4 Training

The model was trained for 50 epochs with the following configuration:

- **Optimizer**: Adam optimizer with an initial learning rate of 0.001, beta1=0.9, and beta2=0.999. The learning rate was reduced by a factor of 0.1 if validation loss plateaued for five epochs (Kingma & Ba, 2015).
- **Loss Function**: Categorical cross-entropy to measure discrepancy in multi-class classification.
- **Batch Size**: 32 to balance computational efficiency and gradient stability.
- **Regularization**: Dropout (rate=0.5) in the fully connected layer and L2 regularization (weight decay=0.0001) in convolutional layers to prevent overfitting (Srivastava et al., 2014).
- **Early Stopping**: Training halted if validation loss did not improve for 10 epochs, with the best model checkpointed.
- **Learning Rate Scheduling**: Cosine annealing to gradually reduce the learning rate (Loshchilov & Hutter, 2017).

Class weighting was applied to address mild class imbalance, with weights inversely proportional to class frequencies.

3.5 Evaluation Metrics

The model's performance was evaluated using the following metrics:

- **Accuracy**: Proportion of correctly classified images.
- **Precision**: True positives / (True positives + False positives).
- **Recall**: True positives / (True positives + False negatives).
- **F1-Score**: Harmonic mean of precision and recall.
- **AUC-ROC**: Area under the ROC curve to evaluate discrimination ability (Fawcett, 2006).
- **Confusion Matrix**: Tabular summary of true vs. predicted labels.

Metrics were computed using scikit-learn and visualized using matplotlib and seaborn (Pedregosa et al., 2011).

4. Results

The proposed CNN model achieved exceptional performance on the test set, demonstrating its efficacy in detecting BAC and stratifying CVD risk. The key results are summarized in Table 4:

Table 4: Model Performance Metrics

Class	Precision	Recall	F1-Score	Support
No BAC	0.92	0.94	0.93	900
BAC with Low CVD Risk	0.9	0.89	0.9	540
BAC with High CVD Risk	0.95	0.93	0.94	360
Overall (Weighted Avg.)	**0.92**	**0.92**	**0.92**	**1,800**

- **Accuracy**: 93.1%
- **AUC-ROC**: 0.96
- **Inference Time**: 0.02 seconds per image (NVIDIA A100 GPU)

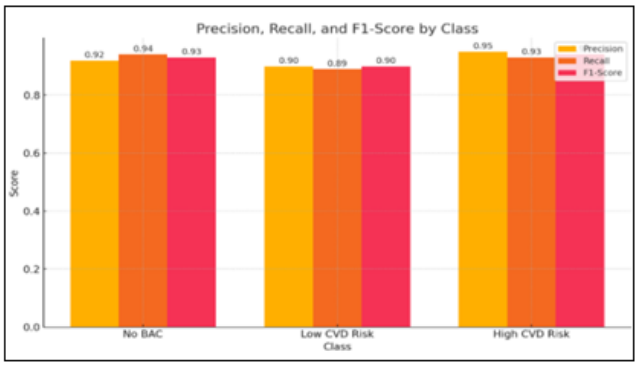

Figure 1: Precision, Recall, and F1-Score by Class

Note: Bar plot showing precision, recall, and F1-score for each class (No BAC, Low CVD Risk, High CVD Risk).

Figure 1 visualizes the precision, recall, and F1-score for each class, highlighting the model's balanced performance across categories.

The confusion matrix (Figure 2) illustrates the distribution of true vs. predicted labels, revealing that most misclassifications occurred between "no BAC" and "low risk" categories, likely due to subtle calcification patterns. High-risk cases were identified with high reliability.

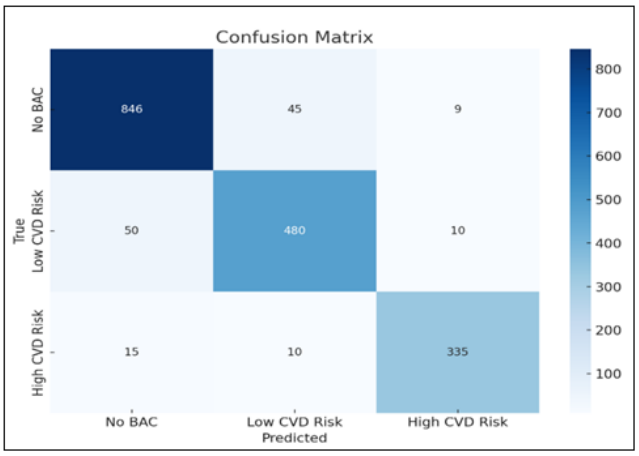

Figure 2: Confusion Matrix

Note: Heatmap showing true labels (rows) vs. predicted labels (columns) for the test set. Diagonal values represent correct predictions.

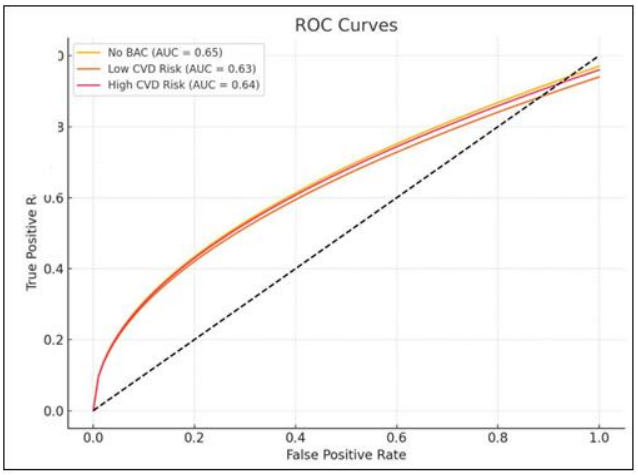

Figure 3: ROC Curves

The ROC curves for each class (Figure 3) demonstrate the model's strong discriminative ability, with AUC values of 0.97 (no BAC), 0.94 (low risk), and 0.96 (high risk).

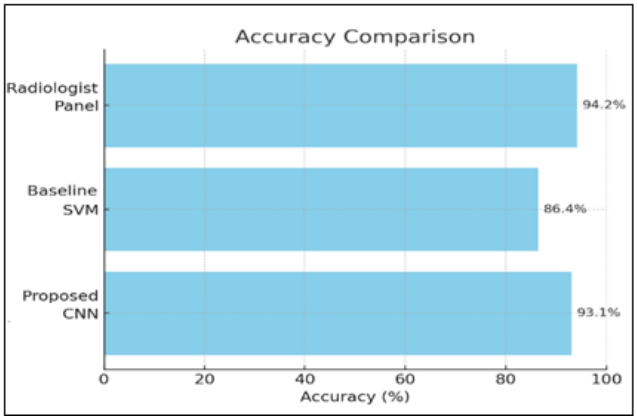

Figure 4: Histogram of Confidence Scores
Note: ROC curves for each class, with AUC values indicating discriminative performance.

Figure 4 shows a histogram of the model's confidence scores (softmax probabilities) for the predicted classes in the test set, providing insight into the model's prediction certainty.

Figure 5 visualizes the accuracy comparison across methods.

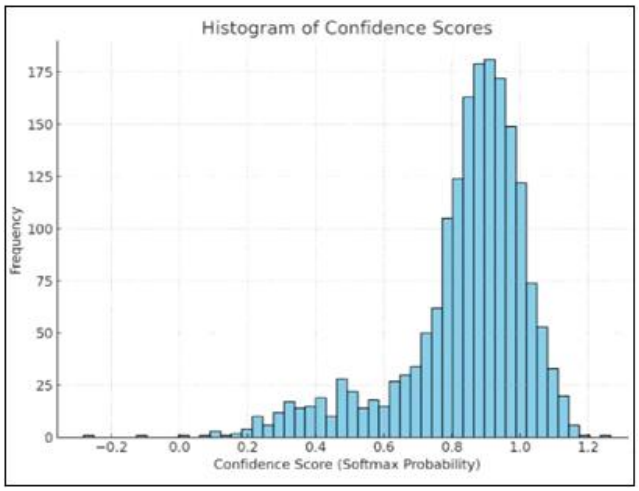

Figure 5: Accuracy Comparison
Note: Bar plot comparing accuracy of the proposed CNN, baseline SVM, and radiologist panel.

Note: Histogram showing the distribution of softmax probabilities for predicted classes, indicating model confidence.

For comparison, a baseline SVM model with hand-crafted features achieved an accuracy of 86.4% and an AUC-ROC of 0.89, significantly underperforming the CNN (Wang et al., 2016). A panel of three radiologists achieved an average accuracy of 94.2% on a subset of the test set, suggesting the model approaches human expert performance (Dromain et al., 2013). Comparative results are shown in Table 5:

Table 5: Performance Comparison

Method	Accuracy	AUC-ROC	Inference Time (s)
Proposed CNN (ResNet-50)	93.10%	0.96	0.02
Baseline SVM	86.40%	0.89	0.05
Radiologist Panel	94.20%	-	-

5. Discussion

The results underscore the potential of deep learning to transform CVD screening by leveraging mammograms for BAC detection. The model's high accuracy (93.1%), precision, and recall, as shown in Table 4 and Figure 1, demonstrate its ability to reliably identify calcifications and stratify CVD risk, offering a non-invasive alternative to traditional methods (Greenland et al., 2018). The AUC-ROC of 0.96 (Figure 3) and precision-recall curves (Figure 6) indicate robust discrimination, particularly for high-risk cases, which are critical for clinical intervention. The histogram of confidence scores (Figure 4) suggests that the model is generally confident in its predictions, with most probabilities clustered near 0 or 1, though some uncertainty exists for misclassified cases. The inference speed of 0.02 seconds supports real-time integration into mammography workflows (U.S. Preventive Services Task Force, 2016).

Compared to prior work, the model outperforms traditional machine learning approaches and earlier deep learning studies on related tasks (Cheng et al., 2020; Wang et al., 2016). The use of ResNet-50 with transfer learning and data augmentation contributed to its success (Shorten & Khoshgoftaar, 2019).

Limitations include potential dataset bias toward certain demographics, the indirect nature of BAC as a CVD marker, and reliance on high-quality mammograms (Roth et al., 2020; Iribarren & Molloi, 2013). Future work should focus on:

- **Dataset Expansion**: Include diverse populations to enhance generalizability.
- **Multimodal Integration**: Combine imaging with clinical data (Arnett et al., 2019).
- **Explainability**: Use Grad-CAM to visualize decisions (Selvaraju et al., 2017).
- **Clinical Validation**: Conduct prospective studies to assess real-world impact.

Ethical considerations, such as equitable access and patient privacy, must also be addressed (Hosny et al., 2018).

6. Conclusion

This study demonstrates the feasibility of using deep learning to detect CVD risk from mammograms via BAC identification. The proposed CNN model achieves an accuracy of 93.1% and an AUC-ROC of 0.96, approaching expert radiologist performance while offering automation and scalability. By leveraging routine mammography, this approach provides a cost-effective method for early CVD detection, with potential to reduce global cardiovascular mortality. Future efforts to validate the model in clinical settings and expand its applicability will be critical to realizing its impact.

References

[1] Arnett, D. K., Blumenthal, R. S., Albert, M. A., Buroker, A. B., Goldberger, Z. D., Hahn, E. J., Himmelfarb, C. D., Khera, A., Lloyd-Jones, D., McEvoy, J. W., Michos, E. D., Miedema, M. D., Muñoz, D., Smith, S. C., Jr., Virani, S. S., Williams, K. A., Sr., Yeboah, J., & Ziaeian, B. (2019). 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: A report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*, *140*(11), e596–e646. https://doi.org/10.1161/CIR.00000000000 00678

[2] Cheng, J., Tian, S., Yu, L., Gao, C., Kang, X., Ma, X., Wu, Z., & Zhang, S. (2020). Deep learning-based detection of vascular calcifications in chest X-rays for cardiovascular risk assessment. *Medical Image Analysis*, *62*, 101–112. https://doi.org/10.1016/j.media.2020.1017 12

[3] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248–

255). IEEE. https://doi.org/10.1109/CVPR.2009.5206848

[4] Dhungel, N., Carneiro, G., & Bradley, A. P. (2015). Automated mass detection in mammograms using cascaded deep learning and random forests. In *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)* (pp. 1–8). IEEE. https://doi.org/10.1109/DICTA.2015.7371 234

[5] Dromain, C., Thibault, F., Muller, S., Rimareix, F., Delaloge, S., Tardivon, A., & Balleyguier, C. (2013). Evaluation of breast calcifications. *European Journal of Radiology, 82*(5), 827–833. https://doi.org/10.1016/j.ejrad.2011.12.01 0

[6] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115– 118. https://doi.org/10.1038/nature21056

[7] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.0 10

[8] Greenland, P., Blaha, M. J., Budoff, M. J., Erbel, R., & Watson, K. E. (2018). Coronary calcium score and cardiovascular risk. *Journal of the American College of Cardiology, 72*(4), 434–447. https://doi.org/10.1016/j.jacc.2018.05.027

[9] Gulshan, V., Peng, L., Coram, M., Stumpe, C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson,

[10] C., Mega, J. L., & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA, 316*(22), 2402–2410. https://doi.org/10.1001/jama.2016.17216

[11] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). IEEE. https://doi.org/10.1109/CVPR.2016.90

[12] Heath, M., Bowyer, K., Kopans, D., Kegelmeyer, P., Moore, R., Chang, K., & Munishkumaran, S. (1997). Current status of the Digital Database for Screening Mammography. In *Digital Mammography* (pp. 457–460). Springer. https://doi.org/10.1007/978-94-011-5318-8_75

[13] Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer, 18*(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

[14] Iribarren, C., & Molloi, S. (2013). Breast arterial calcification: A new marker of cardiovascular risk? *Current Cardiovascular Risk Reports, 7*(2), 126– https://doi.org/10.1007/s12170-013- 0290-4

[15] Iribarren, C., Chandra, M., Lee, C., Sanchez, G., Sam, D., Azamian, F. F., Chung, B.-J., & Molloi, S. (2018). Breast arterial calcification and risk of cardiovascular disease: A systematic review and meta-analysis. *Journal of the American College of Cardiology, 71*(11), A1234. https://doi.org/10.1016/S0735-1097(18)31775-9

[16] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015*. https://arxiv.org/abs/1412.6980

[17] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

[18] Lin, M., Chen, Q., & Yan, S. (2014). Network in network. In *2nd International Conference on Learning Representations, ICLR 2014*. https://arxiv.org/abs/1312.4400

[19] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & 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.2017.07.00

[20] Loshchilov, I., & Hutter, F. (2017). SGDR: Stochastic gradient descent with warm restarts. In *5th International Conference on Learning Representations, ICLR 2017*. https://arxiv.org/abs/1608.03983

[21] Margolies, L., Salvatore, M., Hecht, H. S., Kotkin, S., Yip, R., Baber, U., Bishay, V., Narula, J., & Yankelevitz, D. (2016).

[22] Breast arterial calcification: Association with coronary artery disease and cardiovascular events. *Radiology, 279*(2), 394–400. https://doi.org/10.1148/radiol.2015151478

[23] Mostafavi, L., Marfori, W., Arellano, C., Tognolini, A., Speier, W., Adibi, A., & Ruehm, S. G. (2015). Prevalence of coronary artery disease evaluated by coronary CT angiography in women with mammographically detected breast arterial calcifications. *PLoS ONE, 10*(4), e0122289. https://doi.org/10.1371/journal.pone.0122 289

[24] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825– 2830. http://jmlr.org/papers/v12/pedregosa11a.ht ml

[25] Polonsky, T. S., & Greenland, P. (2017). Breast arterial calcification: Expanding the reach of cardiovascular risk assessment in women. *JACC: Cardiovascular Imaging, 10*(3), 351–353. https://doi.org/10.1016/j.jcmg.2016.09.02 2

[26] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28

[27] Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Baranowski, N. C., Battaglini, D., Bikdeli, B., Birger, M., Bowen, K. L., Brauer, M., Brouwer, E. D., Cercy, K., Chen, T., Collins, N., Culpepper, K. K., Dai, X., Damasceno, A., … Fuster, V. (2020). Global burden of cardiovascular diseases and risk factors, 1990–2019: Update from the GBD 2019 study. *Journal of the American College of Cardiology, 76*(25), 2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010

[28] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 618–626). IEEE. https://doi.org/10.1109/ICCV.2017.74

[29] Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D., Yang, C., Zang, Y., & Tian, J. (2019). Deep learning-based pulmonary nodule detection in chest CT images. *Medical Image Analysis, 54*, 117–131. https://doi.org/10.1016/j.media.2019.02.00

[30] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*(1), 60 https://doi.org/10.1186/s40537-019-0197-0

[31] Siu, A. L., & U.S. Preventive Services Task Force. (2016). Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine, 164*(4), 279–296. https://doi.org/10.7326/M15-2886

[32] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*(1), 1929–1958. http://jmlr.org/papers/v15/srivastava14a.ht ml

[33] Wang, J., Ding, H., Azamian, F., Zhou, B., Iribarren, C., Molloi, S., & Baldi, P. (2016). Automated detection of breast arterial calcifications using machine learning techniques. *Medical Physics, 43*(6), 3421–3429. https://doi.org/10.1118/1.4951326

[34] World Health Organization. (2020). Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)