ConvNeXt Model for Breast Cancer Image Classification

1st Devin Setiawan
Department of Electrical Engineering and Computer Science
University of Kansas
Lawrence, United States
devinryandi.s@ku.edu

2nd Andrea Stevens Karnyoto
Bioinformatics and Data Science Research Center
Bina Nusantara University
Jakarta, Indonesia
andrea.stevens@binus.edu

3rd Indo Intan
Department of Informatics Engineering
Universitas Dipa Makassar
Makassar, Indonesia
indo.intan@undipa.ac.id

4th Bens Pardamean
Computer Science Department, BINUS Graduate Program – Master of Computer Science Program,
Bina Nusantara University
Jakarta, Indonesia
bpardamean@binus.edu

***Abstract*— Breast cancer is an occurrence of cancer that attacks breast tissue and is the most common cancer among women worldwide, affecting one in eight women. In this modern world, breast cancer image classification simplifies the process of analyzing, providing objective and accurate results. By leveraging machine learning algorithms and computer vision techniques, we developed breast cancer detection. The dataset is histopathology dataset from BreakHis and UNHAS Hospital. We chose the ConvNeXt-Tiny model then modified the classifier head as the proposed method. Before the dataset is processed by the model, we augment the images by applying random horizontal and vertical flips, adjustments to brightness, contrast, saturation, and hue using color jitter. The augmentation process simulates real-world variance and enhances the model's ability to generalize to unseen data. Our proposed model gained better performance (accuracy, F1-Score) results compared two other techniques to VGG16 and SVM. According to our experiments, the F1-Score for the ConvNeXt-Tiny model with classifier head modification is higher at 0.9516, than the gain for VGG16 at 0.9292, and the gain for the SVM at 0.83.**

Keywords—Breast Cancer, Image Classification, ConvNeXT, CNN, Tranfer Learning

I. Introduction

Breast cancer is the kind of cancer that occurs in the breast tissue and it occurs in both men and women[1], [2]. However, women are more likely to struggle with this cancer[3]. This disease manifests when abnormal cells divide uncontrollably, resulting in a tumor that spreads to other tissues and organs[4]. Breast cancer can spread to other parts of the body, such as the bones, liver, and lungs if not treated immediately. It also leads to metastasis and organ failure, which can cause death[5], [6], [7]. Early detection and prompt treatment are crucial to improving survival chances and reducing complications risk.

For successful treatment and improved patient outcomes, early detection is essential[8], [9], [10]. Mammograms and physical assessments may not identify all cases of breast cancer[11], [12]. In this case, Artificial Intelligence (AI) is extremely beneficial[13]. AI has the potential to revolutionize breast cancer detection by improving accuracy, efficiency, and accessibility[14], [15]. It can analyze histopathology images in very large volumes and also faster than humans[16], allowing for faster screening and reduced waiting times[17], [18]. The other benefit is that we can reduce false positives in diagnoses. Also, the systems can be developed for use with mobile devices, enabling remote screening and monitoring.

In this paper, we develop breast cancer detection using the ConvNeXt model. ConvNeXt is a state-of-the-art convolutional neural network architecture that has shown promising results in transfer learning. ConvNeXt is the modernization of a standard ConvNet (ResNet[19], [20]) in the direction of designing a Hierarchical Vision Transformer using Shifted Windows (Swin Transformer[21]), which does not include modules that are based on attention[22]. By simplifying the traditional ResNet architecture, ConvNeXt scales better with the number of parameters compared to traditional CNNs and adopts some of the design principles from a Vision Transformer (ViTs) [23], [24]. It allows ConvNeXt to improve performances in various benchmarks. We chose ConvNeXt model because it is more computationally efficient than some transformer-based models and has simple structure compared to traditional CNNs like ResNets, incorporating principles from ViTs.

The contributions of this research are as follows:

1. We applied the combination of traditional convolutional networks (standard ResNet) and modern technique vision transformer which name is ConvNeXt for breast cancer image classification.
2. We present a data processing pipeline integrating histopathology images from two diverse sources: BreakHis and UNHAS Hospital. This combined dataset represents real-world breast cancer histopathology, improving model robustness. To our knowledge, this is the first study utilizing these datasets together for breast cancer classification.
3. We compared the results between traditional machine learning (SVM, VGG16 [25], [26]) and transfer learning (ConvNeXt-Tiny and its finetuning).

In this paper, we demonstrate that our approach leads to better results in every performance indicator, especially F1-Score, the ConvNeXt-Tiny model with classifier head modification gain higher of 0.9516, then the VGG16 gain of 0.9292, and the SVM gain of 0.83.

The remaining section of this article provides an overview of the methodology, which includes a flowchart, ConvneXt, and its modification. Our experiment findings are described in the result section. And the last section is the conclusion, which contains a summary of our article.

II. METHODOLOGY

We addressed an image classification problem to distinguish between malignant and benign breast cancer using CNN. To tackle this, we employed ConvNeXt-Tiny, a modern CNN architecture known for its robust performance and computational efficiency. We leveraged transfer learning with pretrained weights from the ImageNet1K dataset, followed by fine-tuning the model on our specific dataset. We then evaluate the performance of ConvNeXt-Tiny with VGG16 and SVM models across various metrics to evaluate their effectiveness in this classification task. Both the ConvNeXt-Tiny model and VGG16 follow a similar training procedure highlighted in Fig. 1. First, we transformed the data to fit the expected input of both CNN models. Data augmentation was then performed on the training set to simulate real-world variance and enhance generalization. Transformations included random flips, color adjustments, Gaussian blur, resizing, and normalization consistent with ImageNet1K preprocessing. Validation and test sets underwent the same resize, crop, and normalization without random components. The classifier head of both models was modified for binary classification. Finally, the training was conducted in two phases: first, only the classifier head was trained highlighted in Fig. 2A, followed by fine-tuning all layers highlighted in Fig. 2B. The same parameters were applied to both ConvNeXt-Tiny and VGG-16 models. The performance of both models was then compared on various performance metrics. In addition, we trained a SVM model to compare the CNNs to a more classical machine learning approach to image classification using support vectors.

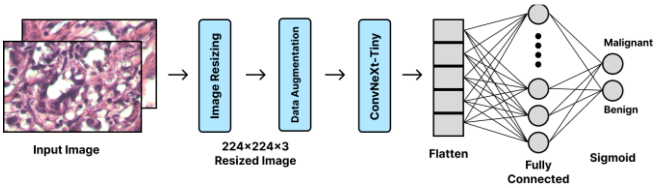

Fig. 1. **Machine learning pipeline for breast cancer classification using ConvNeXt-Tiny.** Schematic depicting the machine learning pipeline for breast cancer classification using ConvNeXt-Tiny model. Input images undergo resizing and data augmentation before passing through ConvNeXt-Tiny, followed by a flattening step and fully connected layers. The final sigmoid activation outputs probabilities for malignant or benign classification.

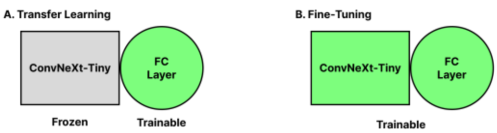

Fig. 2. **Transfer learning and fine-tuning in ConvNeXt-Tiny for breast cancer classification.** (A) Transfer learning involves freezing the ConvNeXt-Tiny convolutional layers (gray rectangle) while only training the fully connected layer (green circle). (B) Fine-tuning proceeds by training both the ConvNeXt-Tiny convolutional layers (green rectangle) and the fully connected layer (green circle), allowing the model to adapt its learned features to the specific characteristics of the breast cancer dataset.

A. Dataset

Breast cancer datasets originated from BreakHis[27] and UNHAS Hospital, with data compositions of 80% and 20%, respectively. We sorted the selected datasets based on our research needs. This data set comprises 1693 images, divided into three categories: train data, validation data, and test data, arranged in a row with 918, 230, and 545 images respectively. The data train has a data distribution of 621 malignant images and 297 benign images. We process and analyze these data, and then transform them into model input. Changes in data composition, from such data becoming numerical values to finding the best hyperparameter tuning models, affect model performance results. However, this is albeit with classifications of significant or insignificant.

B. Data Augmentation on Training Set

Data augmentation is performed on the training set to simulate real-world variance and enhance the model's ability to generalize to unseen data. The transformations applied include random horizontal and vertical flips, adjustments to brightness, contrast, saturation, and hue using color jitter, and the application of Gaussian blur with a kernel size of 5 and sigma ranging from 0.1 to 2, which simulates out-of-focus areas and noise often present in real-world images. The images are then resized to 236 pixels using bilinear interpolation, followed by a center crop to 224 pixels. These resize and crop settings are consistent with the transformations applied to the ImageNet1K dataset before feeding it into the ConvNeXt model, ensuring compatibility and optimal performance. Finally, the images are converted to tensors and normalized with a mean of [0.485, 0.456, 0.406] and a standard deviation of [0.229, 0.224, 0.225], following the same normalization parameters used for the ImageNet1K dataset.

C. Transformation on Validation and Test Set

The transformation applied to the validation and test sets involves the same resize, crop, and normalization settings used for the training set, but without the random components. Specifically, the images are resized to 236 pixels using bilinear interpolation, followed by a center crop to 224 pixels. These settings align with the preprocessing steps used for the ImageNet1K dataset before inputting it into the ConvNeXt model. Finally, the images are converted to tensors and normalized with a mean of [0.485, 0.456, 0.406] and a standard deviation of [0.229, 0.224, 0.225], maintaining the same normalization parameters as used for the ImageNet1K dataset.

D. ConvNeXt-Tiny model training

The ConvNeXt-Tiny model was chosen for its balance of performance and computational efficiency, making it suitable for a variety of tasks in image classification. The motivation behind this choice is to find a small enough model that can perform the classification task well enough. ConvNeXt is a convolutional neural network (CNN) architecture that combines the strengths of traditional CNNs with improvements inspired by vision transformers (ViTs). ConvNeXt-Tiny is a smaller variant in the ConvNeXt family, specifically designed to offer high accuracy while maintaining lower computational and memory requirements compared to larger models. This makes ConvNeXt-Tiny an ideal choice for environments with limited resources or for applications requiring faster inference times.

We chose the model with the 'DEFAULT' pretrained weights, which are trained on the ImageNet1K dataset, to speed up the training process through transfer learning. The

ConvNeXt-Tiny model achieves 82.52% top-1 accuracy and 96.146% top-5 accuracy on the ImageNet-1K dataset. The model has 28,589,128 parameters and requires 4.46 GFLOPS. The file size of the model is 109.1 MB, and the minimum input size for the model is 32x32 pixels. By utilizing the ConvNeXt-Tiny model, this study leverages a state-of-the-art architecture that provides robust performance and efficiency, ensuring the ability to handle complex image classification tasks with limited computational resources.

a) Model Modification: To adapt the pretrained model for binary classification, the classifier head is modified accordingly (as shown in Table 1). The original classifier head, designed for multiclass classification with 1,000 output classes, is replaced with a new classifier head suitable for binary classification. The 'LayerNorm2d' and 'LayerNorm' are essentially performing the same normalization operation. The slight difference in the syntax (LayerNorm2d vs. LayerNorm) does not change the underlying functionality, as both are configured to normalize across the channel dimension (768). The 'Flatten' layer remains unchanged in both configurations. It converts the multi-dimensional tensor into a 1D tensor. The most significant change is in the 'Linear' layer. In the original classifier head, the linear layer outputs 1,000 classes, suitable for the multiclass classification problem the network was originally trained on. For binary classification, this layer is modified to output only 2 classes (out_features=2), corresponding to the two possible classes in the binary classification task.

TABLE I. MODEL MODIFICATION

	Original Classifier Head	**Modified Classifier Head**
Layer 0	LayerNorm2d((768,), eps=1e-06, elementwise_affine=True)	LayerNorm((768, 1, 1), eps=1e-06, elementwise_affine=True)
Layer 1	Flatten(start_dim=1, end_dim=-1)	Flatten(start_dim=1, end_dim=-1)
Layer 2	Linear(in_features=768, out_features=1000, bias=True)	Linear(in_features=768, out_features=2, bias=True)

b) Training Parameter: The training process is conducted in two phases: the first training pass focuses on the classifier head, and the second training pass fine-tunes all layers. Additionally, the SVM model's settings are specified for clarity (as shown in Table 2).

TABLE II. TRAINING PARAMETERS

	1st Training Pass (Transfer Learning)	**2nd Training Pass (Fine-Tuning)**
Layers Trained	Classifier Head	All layers
Epoch	20	20
Batch Size	32	32
Criterion	Class Weighted Cross Entropy	Class Weighted Cross Entropy
Optimizer	Adam (lr = 0.001)	Adam (lr = 0.0001)
Scheduler	StepLR (step_size = 7, gamma = 0.1)	StepLR (step_size = 7, gamma = 0.1)

During the first training pass, only the classifier head of the models is trained. This approach is taken to allow the model to adjust its final layer weights to the new binary classification task while retaining the pretrained features in the other layers. By initially training only the classifier head, we leverage the pretrained features learned from the original multiclass classification task, thereby speeding up convergence and maintaining the beneficial features learned from a larger dataset. In the second training pass, all layers of the models are fine-tuned to further adapt the pretrained weights to the specific dataset. Fine-tuning all layers allows the model to make finer adjustments to the pretrained features to better capture the nuances of the new task. The loss function used is Cross Entropy Loss, with class weights applied to handle class imbalance. Specifically, class-weighted Cross Entropy Loss uses the inverse of class frequency.

$$Class\ Weight_i = \frac{1}{frequency_i} \quad (1)$$

$$frequency_i = \frac{count_i}{count_{total}} \quad (2)$$

The optimizer is Adam, set to a learning rate of 0.001, and a learning rate scheduler with a step size of 7 and gamma of 0.1 is used to adjust the learning rate periodically. However, on the 2nd training pass, the learning rate for the Adam optimizer is reduced to 0.0001 to allow for finer adjustments. This ensures that the updates to the weights are more controlled and precise. These parameters are consistently applied to both the ConvNeXt-Tiny and VGG-16 models to ensure a fair comparison of their performance.

For the SVM model, the following settings are used: the regularization parameter C is set to 1.0, which is the default value. The kernel type is set to radial basis function (rbf). To address class imbalance, the class_weight parameter is set to balanced, which adjusts the weights inversely proportional to class frequencies in the input data.

III. RESULT

A. *Training Model Performance*

The initial training pass focused solely on the classifier head, which converged after approximately 10 epochs (as shown in Fig. 3). During this phase, the selected model attained a peak validation accuracy of 0.8652, with a training loss of 0.2815 and a validation loss of 0.3465. This initial step established a solid baseline performance by effectively tuning the classifier head. The subsequent fine-tuning phase, where the entire model was trained using a smaller learning rate, further enhanced the model's performance. This second training pass also converged after about 10 epochs, resulting in a significantly improved validation accuracy of 0.9478, alongside a training loss of 0.0093 and a validation loss of 0.1615. The loss function graph can be seen in Fig. 3 for all epochs. These results underscore the effectiveness of the two-step transfer learning approach. By initially training the classifier head and then fine-tuning the entire model, substantial improvements in both accuracy and loss metrics were achieved, demonstrating the robustness and efficacy of this method in optimizing model performance.

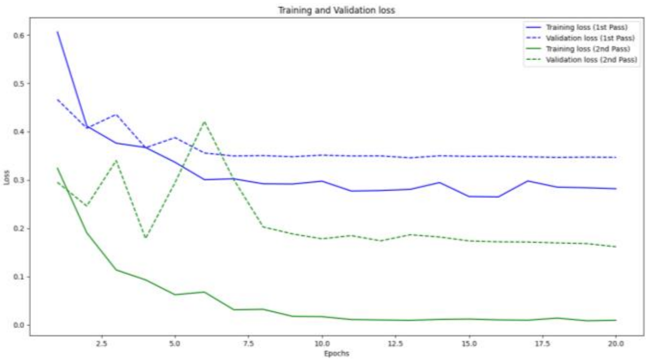

Fig. 3. **Training and validation loss.** This figure presents the training and validation loss for the ConvNeXt-Tiny model on the dataset. The training and validation loss converges around 10 epochs. The second fine-tuning training pass improves the model loss after the first convergence. The lowest loss achieved on the 1st pass is 0.2647 for training and 0.3493 for validation. On the 2nd pass, the lowest loss achieved was 0.0084 for training and 0.1615 for validation.

B. Sample Prediction

Fig. 4 showcases sample predictions from the left-out test set, where each image is annotated with its actual label, predicted label, and the model's predicted probability. To a human observer, the differences between each class are often not clear, underscoring the complexity of the classification task. Despite this challenge, the model demonstrates remarkable performance, accurately distinguishing between classes with a very high accuracy. Two images of misclassification are shown below as an example where the model might be struggling. The result highlights the model's overall effectiveness in correctly predicting the labels for the majority of the test images. However, it is important to note that we cannot guarantee the probabilities generated by the model are representative of the actual probability of being in the class. This limitation arises because the model is trained solely to classify images and not to predict the probability accuracy. The training process utilizes cross-entropy loss and validation accuracy as key metrics, which prioritize correct classification over precise probability calibration.

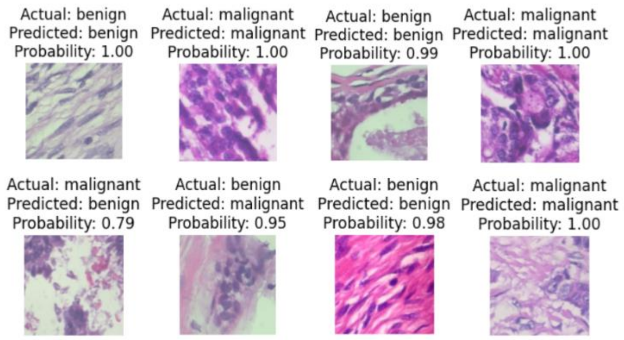

Fig. 4. **Predicted class and probability for image in test cases.** This figure presents the actual label, predicted label, and the associated probability for eight images obtained from the left-out test set. Each subfigure includes the test image along with the classification obtained from the ConvNeXt-Tiny model. The two images in the bottom left demonstrate instances of misclassification, whereas the remaining images show correct predictions.

C. Comparison with Other Machine Learning Models

As depicted in Fig. 5, we utilized five performance indicators to measure and evaluate SVM, VGG16 and ConvNext proficiency. Our approach demonstrates a significant advantage in all five measurements. Specifically, as we can see there is a very wide gap between ConvNext and VGG16 for AUC-ROC which is 0.0394. And the biggest gap between ConvNext and SVM is for F1-score which is 0.1216. ConvNeXt is able to achieve superior performance due to its efficient architecture design, which allows training and inference to occur more quickly. Through the use of convolution operations, ConvNeXt extracts meaningful features from input data, resulting in an increase in accuracy and efficiency. Moreover, ConvNeXt's extensive training data, resulting from its widespread adoption, further enhances system performance.

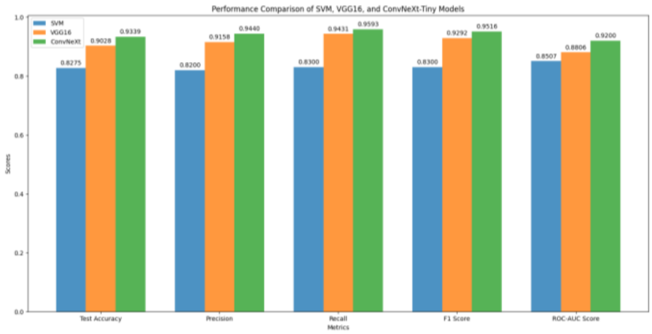

Fig. 5. **Model results comparison.** The chart shows the performance comparison of Support Vector Machines (SVM), VGG16, and the ConvNeXt-TINY model. Performance is evaluated based on not only accuracy scores but also precision, recall metrics, F1-score, and ROC AUC Score. There are several reasons for the differences in the performance results of the models. The most basic reason is the differences in model complexity and structure. SVM is the simplest machine learning approach compared to two others in these experiments. While VGG16 is a straightforward design of CNN that uses small convolutional filters and a large number of layers. And ConvNeXt-tiny is a modernized version of CNN.

IV. CONCLUSION

In this work, we developed an approach for breast cancer image classification using the ConvNeXt-Tiny model. Our model selection demonstrated superior performance compared to conventional models like VGG16 and SVM. We detailed the training process involving data augmentation, transfer learning, and fine-tuning, which significantly improved model accuracy and F1-Score. Our extensive experiments confirmed that the ConvNeXt-Tiny model with classifier head modification achieved the highest F1-Score of 0.9516, outperforming VGG16 and SVM. This research highlights the potential of the ConvNeXt architecture in providing efficient and accurate breast cancer detection, paving the way for more reliable AI-assisted diagnostic tools in medical imaging. While our approach shows promising results, there are several areas for future improvement. One limitation of our current model is the lack of clear explanations for its predictions. Future work could address this by incorporating saliency maps, which would allow domain experts to see which parts of the image the model focuses on, enhancing interpretability. Additionally, the probabilities generated by our model are not calibrated, which can affect the reliability of its confidence scores. Implementing probability calibration techniques could further refine the model's performance. Lastly, we aim to explore the implementation of our model on mobile devices, making it more accessible and easier to use in various healthcare settings. This would facilitate remote screening and monitoring, providing valuable support in the fight against breast cancer.

REFERENCES

[1] B. Smolarz, A. Z. Nowak, and H. Romanowicz, “Breast Cancer—Epidemiology, Classification, Pathogenesis and Treatment (Review of Literature),” *Cancers*, vol. 14, no. 10, p. 2569, May 2022, doi: 10.3390/cancers14102569.

[2] N. Yao *et al.*, “Clinicopathologic characteristics and prognosis for male breast cancer compared to female breast cancer,” *Sci Rep*, vol. 12, no. 1, p. 220, Jan. 2022, doi: 10.1038/s41598-021-04342-0.

[3] E. J. Cathcart-Rake, K. J. Ruddy, A. Bleyer, and R. H. Johnson, “Breast Cancer in Adolescent and Young Adult Women Under the Age of 40 Years,” *JCO Oncology Practice*, vol. 17, no. 6, pp. 305–313, Jun. 2021, doi: 10.1200/OP.20.00793.

[4] N. Bilani, E. C. Zabor, L. Elson, E. B. Elimimian, and Z. Nahleh, “Breast Cancer in the United States: A Cross-Sectional Overview,” *Journal of Cancer Epidemiology*, vol. 2020, pp. 1–8, Oct. 2020, doi: 10.1155/2020/6387378.

[5] S. S. Panigoro *et al.*, “Intronic Variant of MUTYH Gene Exhibits A Strong Association with Early Onset of Breast Cancer Susceptibility in Indonesian Women Population,” *Asian Pac J Cancer Prev*, vol. 22, no. 12, pp. 3985–3991, Dec. 2021, doi: 10.31557/APJCP.2021.22.12.3985.

[6] M. Park, D. Kim, S. Ko, A. Kim, K. Mo, and H. Yoon, “Breast Cancer Metastasis: Mechanisms and Therapeutic Implications,” *IJMS*, vol. 23, no. 12, p. 6806, Jun. 2022, doi: 10.3390/ijms23126806.

[7] L. Wang, S. Zhang, and X. Wang, “The Metabolic Mechanisms of Breast Cancer Metastasis,” *Front. Oncol.*, vol. 10, p. 602416, Jan. 2021, doi: 10.3389/fonc.2020.602416.

[8] M. Hayati *et al.*, “Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning,” *Procedia Computer Science*, vol. 216, pp. 57–66, 2023, doi: 10.1016/j.procs.2022.12.111.

[9] A. Y. Paulindino, B. Pardamean, and G. N. Elwirehardja, “A Systematic Literature Review of CNN Approaches in Classifying Brain Tumor,” in *2023 6th International Conference of Computer and Informatics Engineering (IC2IE)*, Lombok, Indonesia: IEEE, Sep. 2023, pp. 226–231. doi: 10.1109/IC2IE60547.2023.10331371.

[10] A. A. Hidayat, J. P. Trinugroho, R. Nirwantono, D. Sudigyo, and B. Pardamean, “Utilizing Semi-supervised Method in Predicting BRCA1 Pathogenicity Variants,” *Procedia Computer Science*, vol. 227, pp. 36–45, 2023, doi: 10.1016/j.procs.2023.10.500.

[11] R. M. Mann, R. Hooley, R. G. Barr, and L. Moy, “Novel Approaches to Screening for Breast Cancer,” *Radiology*, vol. 297, no. 2, pp. 266–285, Nov. 2020, doi: 10.1148/radiol.2020200172.

[12] O. Ginsburg *et al.*, “Breast cancer early detection: A phased approach to implementation,” *Cancer*, vol. 126, no. S10, pp. 2379–2393, May 2020, doi: 10.1002/cncr.32887.

[13] B. Mahesworo, A. Budiarto, A. A. Hidayat, and B. Pardamean, “Cancer Risk Score Prediction Based on a Single-Nucleotide Polymorphism Network,” *Healthc Inform Res*, vol. 28, no. 3, pp. 247–255, Jul. 2022, doi: 10.4258/hir.2022.28.3.247.

[14] I. Sechopoulos, J. Teuwen, and R. Mann, “Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art,” *Seminars in Cancer Biology*, vol. 72, pp. 214–225, Jul. 2021, doi: 10.1016/j.semcancer.2020.06.002.

[15] K. Freeman *et al.*, “Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy,” *BMJ*, p. n1872, Sep. 2021, doi: 10.1136/bmj.n1872.

[16] R. E. Caraka, N. T. Nugroho, S. K. Tai, R. C. Chen, T. Toharudin, and B. Pardamean, “Feature importance of the aortic anatomy on endovascular aneurysm repair (Evar) using boruta and bayesian mcmc,” *Communications in Mathematical Biology and Neuroscience*, vol. 2020, no. 22, 2020, doi: 10.28919/cmbn/4584.

[17] T. W. Cenggoro, B. Mahesworo, A. Budiarto, J. Baurley, T. Suparyanto, and B. Pardamean, “Features Importance in Classification Models for Colorectal Cancer Cases Phenotype in Indonesia,” *Procedia Computer Science*, vol. 157, pp. 313–320, 2019, doi: 10.1016/j.procs.2019.08.172.

[18] I. Intan and Y. J. W. Soetikno, “Classification of Thyroid Carcinoma using Sobel Edge Detection and Adaptive Neuro Fuzzy Inference System Approach,” in *2019 7th International Conference on Cyber and IT Service Management (CITSM)*, Nov. 2019, pp. 1–6. doi: 10.1109/CITSM47753.2019.8965403.

[19] B. Koonce, “ResNet 50,” in *Convolutional Neural Networks with Swift for Tensorflow*, Berkeley, CA: Apress, 2021, pp. 63–72. Accessed: Jul. 07, 2024. [Online]. Available: http://link.springer.com/10.1007/978-1-4842-6168-2_6

[20] W. Xu, Y.-L. Fu, and D. Zhu, “ResNet and its application to medical image processing: Research progress and challenges,” *Computer Methods and Programs in Biomedicine*, vol. 240, p. 107660, Oct. 2023, doi: 10.1016/j.cmpb.2023.107660.

[21] Z. Liu *et al.*, “Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows,” presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 10012–10022. Accessed: Jul. 07, 2024. [Online]. Available: https://openaccess.thecvf.com/content/ICCV2021/html/Liu_Swin_Transformer_Hierarchical_Vision_Transformer_Using_Shifted_Windows_ICCV_2021_paper

[22] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “A ConvNet for the 2020s,” Mar. 2022, doi: 10.48550/arXiv.2201.03545.

[23] Z. Wu and X. Jiang, “Extraction of Pine Wilt Disease Regions Using UAV RGB Imagery and Improved Mask R-CNN Models Fused with ConvNeXt,” *Forests*, vol. 14, no. 8, p. 1672, Aug. 2023, doi: 10.3390/f14081672.

[24] S. Cheng, E. Sun, W. Qian, and Y. Han, “A Breast Cancer Detection Model Based on Modified Convnext v2,” in *Proceedings of the 2023 3rd Guangdong-Hong Kong-Macao Greater Bay Area Artificial Intelligence and Big Data Forum*, Guangzhou China: ACM, Sep. 2023, pp. 556–559. doi: 10.1145/3660395.3660493.

[25] V. K. Chauhan, K. Dahiya, and A. Sharma, “Problem formulations and solvers in linear SVM: a review,” *Artif Intell Rev*, vol. 52, no. 2, pp. 803–855, Aug. 2019, doi: 10.1007/s10462-018-9614-6.

[26] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition.” arXiv, Apr. 10, 2015. Accessed: Jul. 11, 2024. [Online]. Available: http://arxiv.org/abs/1409.1556

[27] P. Agarwal, A. Yadav, and P. Mathur, “Breast Cancer Prediction on BreakHis Dataset Using Deep CNN and Transfer Learning Model,” in *Data Engineering for Smart Systems*, vol. 238, P. Nanda, V. K. Verma, S. Srivastava, R. K. Gupta, and A. P. Mazumdar, Eds., Singapore: Springer Singapore, 2022, pp. 77–88. Accessed: Jul. 07, 2024. [Online]. Available: https://link.springer.com/10.1007/978-981-16-2641-8_8