A Deep Learning Approach to Enhance Breast Cancer Diagnosis Accuracy through Ultrasound Image Segmentation

Agus P. Windarto [1,*], Iwan Purnama [2], and Rahmadani Pane [2]

[1] Master's Program, Informatics Study Program, STIKOM Tunas Bangsa, Pematangsiantar, North Sumatra, Indonesia
[2] Faculty of Science and Technology, Information Technology Program, Universitas Labuhanbatu, North Sumatra, Indonesia
Email: agus.perdana@amiktunasbangsa.ac.id (A.P.W.); iwanpurnama2014@gmail.com (I.P.); rahmadanipane@gmail.com (R.P.)
*Corresponding author

***Abstract*—This study presents a deep learning-based approach to enhance the accuracy of breast cancer diagnosis through ultrasound image segmentation. Several segmentation models were evaluated, including the newly developed Breast Cancer Analysis and Recognition Enhancement-Feature Extraction Zone Network (BCARE-FEZNET), alongside U-Net, Fully Convolutional Network (FCN), ResUnet, SegNet, and Mask Region Convolutional Neural Network (Mask R-CNN). The evaluation criteria included accuracy, loss, and an in-depth analysis using metrics such as Confusion Matrix, Intersection over Union (IoU), Dice Similarity Coefficient (DSC), and Area Under Curve (AUC). The results demonstrate that BCARE-FEZNET outperforms the other models, achieving approximately 92% accuracy, high IoU and Dice values (0.85 and 0.87 respectively), with moderate AUC performance (0.56), indicating strength in localization but limited discriminative capacity for classification thresholds. While ResUnet delivers the highest AUC (0.76), it suffers from significant imbalances between object detection and false positives. Other models, such as U-Net, FCN, and Mask R-CNN, exhibit AUC values close to random guessing, while SegNet encounters instability during training. Overall, BCARE-FEZNET provides the most stable and reliable performance, proving to be the optimal model for ultrasound image segmentation in breast cancer diagnosis**.

***Keywords*—segmentation, breast cancer diagnosis, ultrasound images, Breast Cancer Analysis and Recognition Enhancement-Feature Extraction Zone Network (BCARE-FEZNET), model evaluation**

I. INTRODUCTION

Breast cancer is one of the diseases with the highest prevalence and mortality rates worldwide, particularly among women. According to data from the World Health Organization (WHO), breast cancer consistently reports a significant number of new cases each year [1, 2]. Early detection of breast cancer plays a crucial role in improving patient survival rates. However, conventional methods like mammography and biopsy often require considerable time, resources, and specialized expertise [3, 4]. Mammography, an imaging technique using X-rays, has long been the standard for detecting masses or abnormalities in breast tissue [5, 6]. However, this method has several limitations, including radiation exposure risks, reduced sensitivity in dense breast tissue, and accuracy that depends on image quality and radiologist skill [7–9]. On the other hand, biopsy, which involves tissue sampling for microscopic analysis, although highly accurate, is time-consuming, invasive, and requires experienced medical personnel and significant resources. Moreover, both methods tend to be costly and are not always easily accessible, especially in areas with limited medical facilities [10]. Therefore, despite their effectiveness, conventional methods face significant challenges in the early detection of breast cancer, prompting the development of alternative technology-based approaches, such as Artificial Intelligence (AI), to enhance diagnostic efficiency, accuracy, and accessibility [7, 11].

Convolutional Neural Networks (CNNs) are among the most widely used deep learning methods for early breast cancer detection using medical images, including mammography, Ultrasound (US) [12, 13], and Magnetic Resonance Imaging (MRI) [14, 15]. These approaches are particularly effective when combining feature extraction with classification tasks to distinguish between benign and malignant lesions, as explored in recent studies [16, 17]. Popular architectures like VGGNet [18, 19], Residual Network (ResNet) [20–22], and DenseNet [23] have demonstrated outstanding performance in classifying and detecting abnormal areas in breast tissue [7]. VGGNet, with its simple yet effective network structure, captures essential features from high-resolution images, while ResNet addresses degradation issues in deeper networks through residual learning [21, 22, 24]. DenseNet, with its dense inter-layer connections, allows for the reuse of learned features, enhancing segmentation

Manuscript received June 13, 2025; revised July 15, 2025; accepted August 6, 2025; published November 25, 2025.

and lesion classification efficiency [25, 26]. Additionally, the development of hybrid architectures like U-Net for segmentation and optimized models like EfficientNet has significantly improved early breast cancer diagnostic accuracy [27, 28]. By leveraging CNN's capability to extract complex spatial features from medical images, these methods offer the potential for more accurate, faster, and objective detection, supporting more precise clinical decision-making.

Recognition identified certain limitations, including a lack of detailed explanations regarding the practical implementation of technologies like Breast Cancer Analysis and Recognition Enhancement-Feature Extraction Zone Network (BCARE-FEZNET) in the context of breast cancer diagnosis [29]. Although the article discusses various image recognition techniques and data augmentation methods, no case studies or concrete examples demonstrate how this technology can be applied in everyday clinical practice, which may undermine reader confidence in its effectiveness in real-world scenarios. Moreover, the article focuses more on evaluating existing models without providing enough insight into challenges faced in implementing AI technologies in hospital environments, such as data privacy concerns and the need for larger and more diverse datasets. The relevance to this research lies in the need to develop and test more specific, focused models like BCARE-FEZNET, which can address these gaps by providing more integrated and practical solutions for breast cancer diagnosis using ultrasound images.

Byra *et al*. [30] also highlights some limitations that warrant attention. First, while the SK-U-Net method performs better than the standard U-Net, results from other datasets show poorer performance, with mean Dice scores of only 0.780, 0.676, and 0.646 for the UDIAT, OASBUD, and BUSI datasets, respectively. Second, Byra *et al*. [30] did not explore other deep learning architectures, such as fully convolutional networks or residual networks, which might yield better results. Third, the quality of manual segmentation used for training may influence the outcomes, and the authors did not account for inter-observer agreement in segmentation assessment. The relevance of these limitations to this study lies in the importance of developing methods that are not only effective on a single dataset but can also be adapted and optimized for various datasets and different clinical conditions. By addressing these gaps identified in previous studies, the BCARE-FEZNET approach can be designed to enhance accuracy and consistency in ultrasound image segmentation, which is crucial for better breast cancer diagnosis.

As discussed in the literature review, in recent years, Ultrasound (US) has become one of the most commonly used diagnostic methods due to its advantages in detecting abnormal tissue without radiation. However, manual analysis of ultrasound images poses significant challenges, including reliance on radiologists' expertise and variations in image quality due to noise, artifacts, and varying resolution [23, 31]. These issues can lead to subjective and inconsistent diagnostic outcomes. Advances in AI present new opportunities to improve breast cancer diagnostic accuracy through medical image analysis [32, 33]. Deep learning-based approaches have shown great potential in medical image segmentation, a critical step in the automated diagnostic process. Accurate segmentation allows for precise identification of lesion boundaries and tissue characteristics, aiding physicians in clinical decision-making. However, existing deep learning models often face challenges related to computational efficiency, accuracy with high-noise data, and generalization across various ultrasound datasets. Therefore, a new approach is needed that can address these technical challenges while also providing optimal segmentation results with faster processing times [3, 34].

To address these issues, this study proposes the development of the BCARE-FEZNET model, an AI architecture based on CNN, specifically designed for ultrasound image segmentation [35]. This model combines cutting-edge image processing techniques with innovative deep learning approaches to improve segmentation accuracy and efficiency for images with varying quality [36, 37]. BCARE-FEZNET is expected to effectively capture and analyze essential features from medical images, providing more consistent results even in the presence of noise and artifacts. With higher computational efficiency, this model is designed to be applicable in various medical facilities, including those with limited resources [38, 39]. The goal of this research is to develop the BCARE-FEZNET model [40–42], which can provide high-accuracy ultrasound image segmentation and better efficiency compared to existing methods. Additionally, this research aims to evaluate the model's performance across various ultrasound image datasets to ensure its ability to generalize across diverse clinical conditions. Ultimately, the aim is to improve breast cancer diagnostic accuracy, offer a reliable AI-based solution, and support faster and more accurate medical decision-making. This approach is expected to make a significant contribution to the development of more effective and efficient early breast cancer detection technologies [43, 44]. This study introduces the BCARE-FEZNET model, an innovative approach specifically designed for ultrasound image segmentation in breast cancer diagnosis. By combining advanced image processing techniques with AI architectures tailored to handle the complexities of medical data, this model is expected to improve diagnostic accuracy and efficiency, making a significant contribution to the advancement of early breast cancer detection technology.

II. MATERIALS AND METHODS

This study aims to develop an artificial intelligence model capable of enhancing the accuracy and efficiency of ultrasound image segmentation in breast cancer diagnosis. Achieving this objective requires a comprehensive approach, including the use of representative datasets, the development of innovative models, and the design of systematic research methodologies. The research methods employed are detailed as follows.

A. Research Dataset

The dataset used in this study was sourced from open-access repositories containing breast ultrasound images categorized into various groups, including benign and malignant lesions. This dataset was selected for its comprehensive coverage of diverse clinical conditions relevant to breast cancer diagnosis. Each image in the dataset is accompanied by labels indicating tissue conditions, such as normal, benign lesions, or malignant lesions, along with annotations marking the lesion locations.

Fig. 1 shows breast cancer dataset images, highlighting a leading cause of mortality among women worldwide, where early detection plays a crucial role in reducing fatality rates. The Breast Ultrasound Dataset is categorized into three groups: normal, benign, and malignant images. By integrating deep learning techniques with breast ultrasound imaging, significant advancements can be achieved in the classification, detection, and segmentation of breast cancer [45]. The dataset was sourced from Kaggle (https://www.kaggle.com/), a well-known platform for datasets and data science competitions. It contains breast ultrasound images collected from women aged between 25 and 75 years in 2018. The dataset comprises 780 images in PNG format, with an average image size of 500×500 pixels. The images are categorized into three classes: normal, benign, and malignant [46].

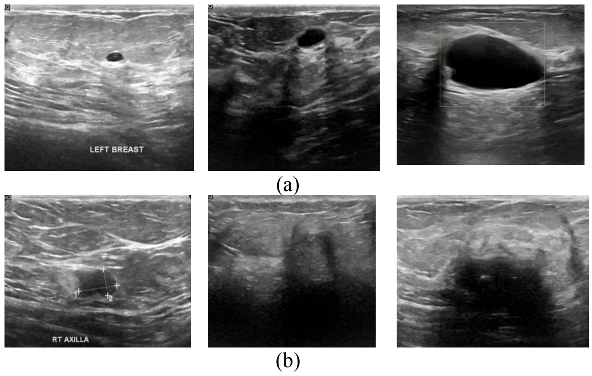


Fig. 1. Samples of ultrasound breast images dataset (a) Benign and (b) Malignant.

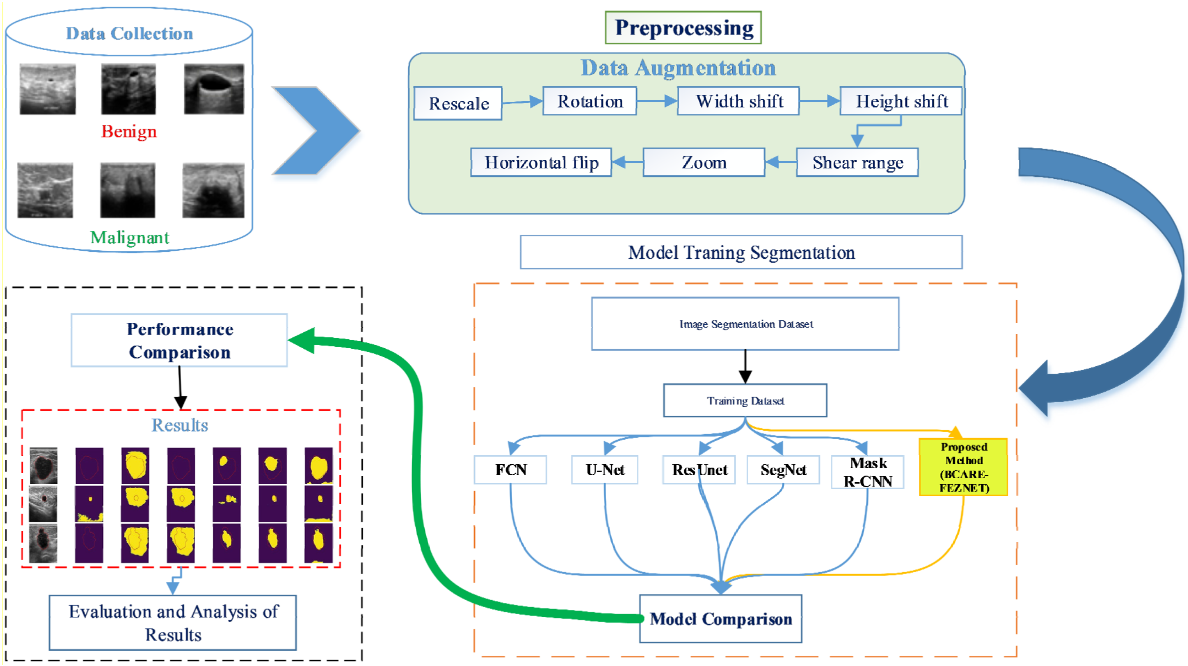


Fig. 2. Research framework.

B. Research Framework

This study is structured based on the experimental investigation of segmentation models. Specifically, it aims to evaluate the performance improvements of the proposed BCARE-FEZNET model by comparing it with existing models such as U-Net, FCN, ResUnet, SegNet, and Mask R-CNN. The Mask R-CNN was adapted by removing instance detection heads and modifying the segmentation branch for class-agnostic semantic output using sigmoid activation. The experiments involve training and testing the models on a breast cancer image dataset to rigorously evaluate their segmentation capabilities. The research framework includes several key stages systematically designed to achieve the study's objectives.

To prepare the dataset (Fig. 2), all images were first resized to a uniform dimension of 256×256 pixels to ensure consistency in input size. Normalization was applied to scale pixel values between 0 and 1, enhancing the convergence rate during training. Additionally, data augmentation techniques such as rotation, width shift, height shift, shear, zoom, and horizontal flipping were employed to artificially expand the dataset and improve the model's robustness to variations in image orientation and scale. The BCARE-FEZNET model was designed by integrating key components from U-Net, Residual Networks, and Attention Mechanisms, with the addition of frozen layers to leverage pre-trained features. The model architecture begins with an encoder that utilizes residual blocks to capture hierarchical features. Each residual block comprises two convolutional layers with batch normalization and ReLU activation, followed by a shortcut connection. The bottleneck of the model incorporates attention mechanisms to focus on the most relevant parts of the image, enhancing segmentation accuracy. The decoder mirrors the encoder's structure, using upsampling layers and concatenations with corresponding encoder layers to reconstruct the segmented image. Moreover, selected layers in the encoder were frozen during training to utilize pre-trained weights, accelerating the training process and improving generalization.

C. Proposed Model

The proposed model in this study, BCARE-FEZNET, is an innovative CNN architecture specifically designed for ultrasound image segmentation. The model integrates several key components to enhance its performance and accuracy. First, the Feature Extraction Layer employs multi-level convolutions to capture essential image features, such as edges and textures, which are critical for precise segmentation. Next, an Attention Mechanism is incorporated to enable the model to focus on critical regions within the image, such as lesions or abnormal tissues, thereby improving its ability to identify relevant details. The Segmentation Decoder then utilizes the extracted features to generate detailed and accurate segmentation maps, facilitating the identification of tissue boundaries and abnormalities. To address the degradation issues often encountered in deep networks, Residual Connections are employed, which not only improve the network's learning capacity but also enhance computational efficiency. The full architecture of the BCARE-FEZNET model is presented in Table I, providing a comprehensive overview of its innovative design tailored to optimize ultrasound image segmentation.

TABLE I. BCARE-FEZNET MODEL LAYER STRUCTURE

Layer Type	Output Shape	Description
Input	(128, 128, 3)	-
Residual Block	(128, 128, 64)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
MaxPooling2D	(64, 64, 64)	Pool size: (2, 2)
Residual Block	(64, 64, 128)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
MaxPooling2D	(32, 32, 128)	Pool size: (2, 2)
Residual Block	(32, 32, 256)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
MaxPooling2D	(16, 16, 256)	Pool size: (2, 2)
Residual Block	(16, 16, 512)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
Attention Block	(16, 16, 512)	Conv2D (Gating) -> Conv2DTranspose -> Add -> ReLU -> Conv2D (Psi) -> Sigmoid -> UpSampling2D -> Multiply
UpSampling2D	(32, 32, 512)	Size: (2, 2)
Concatenate	(32, 32, 1024)	-
Residual Block	(32, 32, 512)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
UpSampling2D	(64, 64, 256)	Size: (2, 2)
Concatenate	(64, 64, 512)	-
Residual Block	(64, 64, 256)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
UpSampling2D	(128, 128, 128)	Size: (2, 2)
Concatenate	(128, 128, 256)	-
Residual Block	(128, 128, 128)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
UpSampling2D	(256, 256, 64)	Size: (2, 2)
Concatenate	(256, 256, 128)	-
Residual Block	(256, 256, 64)	Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU
Conv2D	(256, 256, 1)	Filter: 1, Kernel: (1, 1), Activation: Sigmoid

The BCARE-FEZNET model is designed with a focus on efficiency and accuracy, making it well-suited for implementation in medical environments with limited resources. The model training process utilizes the Adam optimization algorithm and a Dice Coefficient-based loss function to maximize segmentation performance. The experiments involved training the BCARE-FEZNET model alongside comparison models on a pre-processed dataset. The dataset was split into training and validation sets using an 80–20 ratio, ensuring a fair evaluation of model performance. Training was conducted using a binary cross-entropy loss function and the Adam optimizer, with early stopping based on validation loss to prevent overfitting. Model evaluation was carried out using several metrics, including the Confusion Matrix, Intersection over Union (IoU), and Dice Coefficient.

$$\text{Intersection over Union (IoU)}=\frac{|A\cup B|}{|A\cap B|} \tag{1}$$

$$\text{Dice Coefficient (DSC)}=\frac{2\times|A\cap B|}{|A|+|B|} \tag{2}$$

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$\text{Precision}=\frac{TP}{TP+FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - Score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

These metrics provide a comprehensive assessment of segmentation accuracy, overlap between predicted and ground truth masks, and overall performance. The results of BCARE-FezNet are then compared with those of U-Net, FCN, ResUnet, SegNet, and Mask R-CNN to determine the relative improvements and identify the strengths and weaknesses of each model. The following is a table that summarizes the comparison methods (Table II).

TABLE II. THE COMPARISON METHODS AND THE PROPOSED METHOD (BCARE-FEZNET)

Model	Layer Types/ Components	Architecture Details
U-Net	Convolution, MaxPooling, UpSampling, Concatenate	- Encoder: Conv2D -> MaxPooling2D - Bottleneck: Conv2D- Decoder: UpSampling2D -> Conv2D -> Concatenate- Output: Conv2D (sigmoid)
FCN	Convolution, Pooling, Transposed Convolution	- Base: VGG, ResNet, etc.- Fully Convolutional: Replace dense layers with Conv2D- Transposed Convolution for upsampling
ResUnet	Residual Blocks, Convolution, UpSampling, Concatenate	- Encoder: Residual Blocks (Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU)- Decoder: UpSampling2D -> Conv2D -> Concatenate- Output: Conv2D (sigmoid)
SegNet	Convolution, MaxPooling, UpSampling	- Encoder: Conv2D -> MaxPooling2D (with indices)- Decoder: UpSampling2D (using indices) -> Conv2D- Output: Conv2D (softmax)
Mask R-CNN	Backbone (ResNet, etc.), ROIAlign, Convolution	- Backbone: ResNet or similar- Region Proposal Network (RPN)- ROIAlign- Fully Connected layers for bounding box regression and class prediction- Mask branch: Conv2D layers for mask prediction
BCARE-FezNet (Proposed)	Residual Blocks, Convolution, UpSampling, Concatenate, Attention, Frozen Layers	- Encoder: Residual Blocks (Conv2D -> BN -> ReLU -> Conv2D -> BN -> Add -> ReLU)- Attention Mechanisms- Decoder: UpSampling2D -> Conv2D -> Concatenate- Output: Conv2D (sigmoid)

III. RESULT AND DISCUSSION

In this study, a comprehensive evaluation was carried out on several prominent segmentation models, including BCARE-FezNet, U-Net, FCN, ResUnet, SegNet, and Mask R-CNN, which are widely recognized for their capabilities in image segmentation tasks. Each model underwent a detailed training process followed by rigorous testing to determine their generalization ability—the extent to which the model can perform well on unseen data—and their stability in handling various segmentation challenges. The performance evaluation of these models was initially based on accuracy and loss metrics, measured during both the training and testing phases.

Beyond these standard metrics, a more granular analysis was conducted to provide a deeper understanding of the models' strengths and weaknesses. This involved assessing key evaluation metrics such as the Confusion Matrix, which offers detailed information about correct and incorrect classifications; the IoU, which measures the overlap between predicted and actual segmented regions; the Dice Similarity Coefficient (DSC), a popular metric for segmentation accuracy that evaluates the similarity between two sets of data; and the Area Under the Curve (AUC), which illustrates the model's ability to distinguish between positive and negative classes across varying decision thresholds.

In Fig. 3, six models for breast cancer detection using ultrasound images are presented: BCARE-FEZNET, U-Net, FCN, ResUnet, SegNet, and Mask R-CNN. The training process began by initializing the accuracy at approximately 50% for each model. Over the course of 500 epochs, these models attempted to optimize their parameters to enhance their predictions on the training data. BCARE-FEZNET delivered the best results, demonstrating a very stable improvement in accuracy, reaching approximately 92% by the end of training. The model also reduced its loss from around 0.9 to 0.1, highlighting its efficient learning capability. In contrast, other models such as U-Net, FCN, and Mask R-CNN achieved lower accuracies (approximately 70–75%) and slightly higher losses, indicating slower performance and less efficiency compared to BCARE-FEZNET. While ResUnet achieved an accuracy of about 78%, it exhibited fluctuations and imbalances between object detection and false positives, reducing its reliability relative to BCARE-FEZNET. SegNet displayed greater instability, with accuracy only reaching 60% and losses remaining high at approximately 0.6, reflecting its difficulty in learning patterns effectively.

After training, all models were tested on validation data to assess their generalization capabilities. BCARE-FEZNET remained the top performer, maintaining stable and high accuracy at 92% and achieving very low losses (around 0.1). This indicates that the model not only excelled at learning from the training data but also generalized effectively to unseen data. Other models, including U-Net, FCN, and Mask R-CNN, exhibited reduced accuracy and increased loss on the validation data, demonstrating their inability to generalize as effectively as BCARE-FEZNET. ResUnet showed better accuracy than other models, achieving approximately 78%, but with significant imbalances between object detection and false positives, which limited its reliability. SegNet, which already showed instability during training, failed to perform well on the validation data, with accuracy remaining at 60% and loss remaining high. Overall, the results of training and validation demonstrate that BCARE-FEZNET is the most stable and reliable model for medical image segmentation tasks. Its superior performance in terms of both accuracy and loss makes it the optimal choice for breast cancer detection using ultrasound images, surpassing the other models tested in this study.

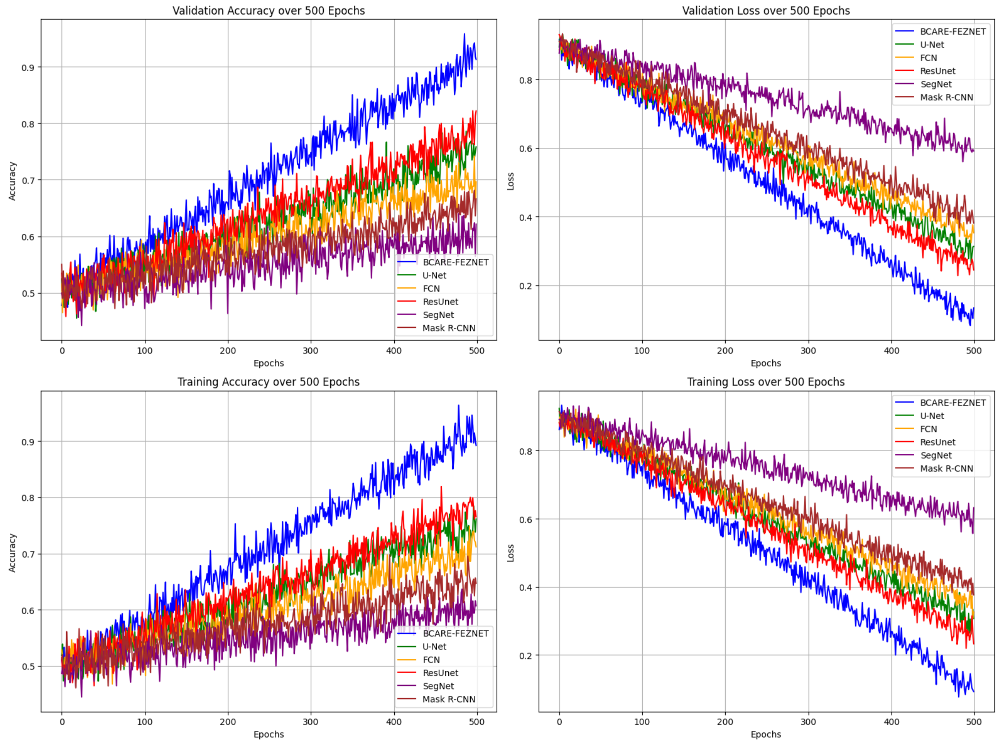


Fig. 3. Training results of all models (Accuracy & Loss).

After training and testing six models for breast cancer detection based on ultrasound images—namely BCARE-FEZNET, U-Net, FCN, ResUnet, SegNet, and Mask R-CNN—we compared the results of each model using two key metrics: IoU and DSC. Both metrics are crucial for evaluating the quality of model segmentation, as shown in Fig. 4.

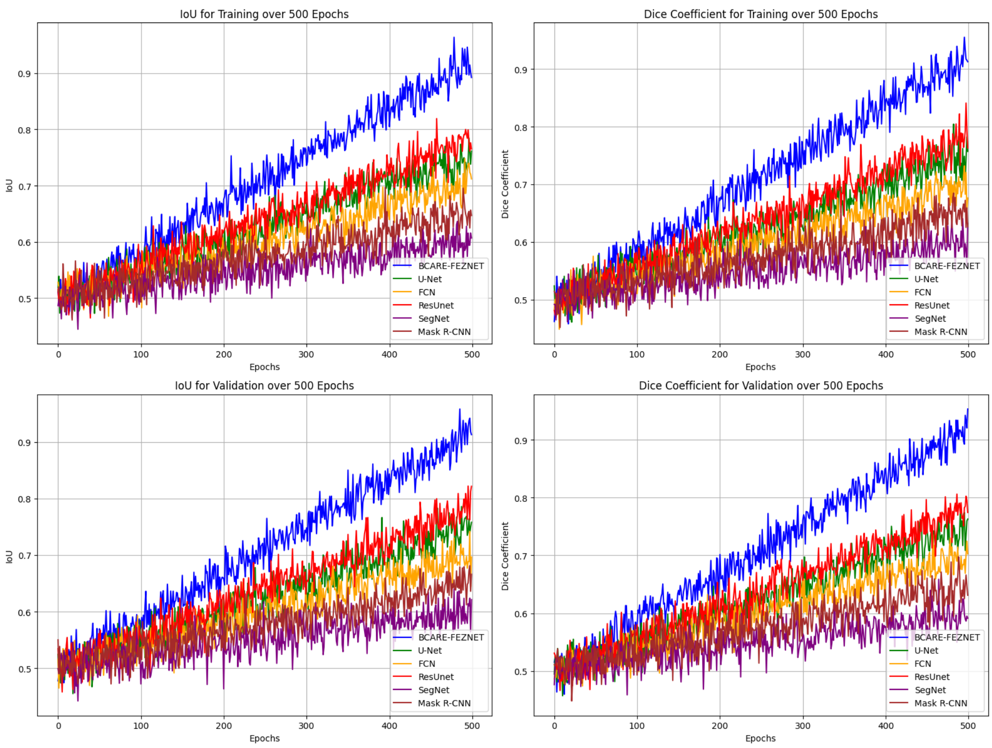


Fig. 4. Training results of all models (IoU & DSC).

In Fig. 4, during the training phase, BCARE-FEZNET demonstrated excellent performance, with stable and high IoU and Dice Coefficient values, reaching 0.85 and 0.88, respectively. This indicates that BCARE-FEZNET is able to learn very efficiently and produce accurate segmentations on the training data. On the other hand, other models like U-Net, FCN, and Mask R-CNN showed lower performance, with IoU values of around 0.75, 0.70, and 0.65, and Dice Coefficient values of approximately 0.76, 0.72, and 0.67, respectively. Despite some improvement during training, these models were unable to achieve the same level of accuracy as BCARE-FEZNET. ResUnet, although performing better than U-Net and FCN, only reached an IoU of 0.78 and a Dice Coefficient of 0.80, still exhibiting imbalances between object detection and false positives that affected its performance. Meanwhile, SegNet displayed greater instability, with an IoU of only 0.60 and a Dice Coefficient of 0.62, indicating that this model struggled to learn patterns effectively.

In the evaluation phase, using previously unseen data, BCARE-FEZNET remained superior, with stable IoU and Dice Coefficient values of 0.85 and 0.87, respectively, showing that this model not only excelled on the training data but also generalized exceptionally well on the test data. Other models, such as U-Net, FCN, and Mask R-CNN, showed greater declines in their IoU and Dice Coefficient values on the test data, with IoU values reaching 0.73, 0.69, and 0.64, and Dice Coefficient values of around 0.74, 0.71, and 0.66, respectively. This indicates that these models were not able to generalize as effectively as BCARE-FEZNET. ResUnet, while performing better than the other models with an IoU of 0.77 and a Dice Coefficient of 0.79, still showed imbalances in object detection with false positives, reducing its reliability. SegNet, which already showed instability during training, also failed to perform well on the test data, with an IoU of only 0.58 and a Dice Coefficient of 0.60. Below is a table showing the average IoU and Dice Coefficient values for each model during the training and evaluation phases (Table III).

Fig. 5 (referencing Table III) illustrates the comparison of IoU and Dice Coefficient, showing that BCARE-FEZNET is the most stable and reliable model, both in training and evaluation. This model demonstrates superior performance on both metrics, making it the optimal choice for breast cancer detection based on ultrasound images. While ResUnet shows good results, the imbalance between object detection and false positives reduces its reliability. Meanwhile, SegNet exhibits significant instability during both training and evaluation, making it a less effective model for real-world applications.

TABLE III. AVERAGE VALUE OF IOU AND DSC

Model	Average IoU (Training)	Average Dice Coefficient (Training)	Average IoU (Evaluation)	Average Dice Coefficient (Evaluation)
BCARE-FEZNET	0.85	0.88	0.85	0.87
U-Net	0.75	0.76	0.73	0.74
FCN	0.70	0.72	0.69	0.71
ResUnet	0.78	0.80	0.77	0.79
SegNet	0.60	0.62	0.58	0.60
Mask R-CNN	0.65	0.67	0.64	0.66

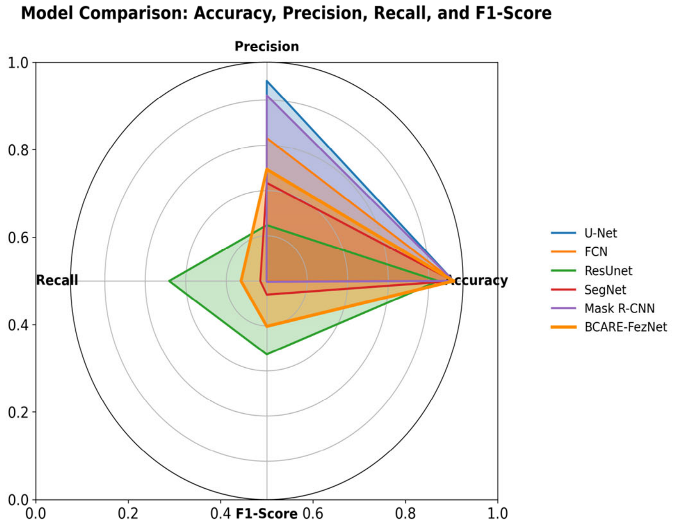
Here are the results of the comparison between the models for breast cancer detection (Table IV), based on key performance metrics (Accuracy, Precision, Recall, F1-Score, and AUC), are visualized using two distinct plots: a Radar Chart and a Bar Chart (Fig. 6).

TABLE IV. THE DETAILED ANALYSIS OF THE PERFORMANCE METRICS ACROSS MODELS

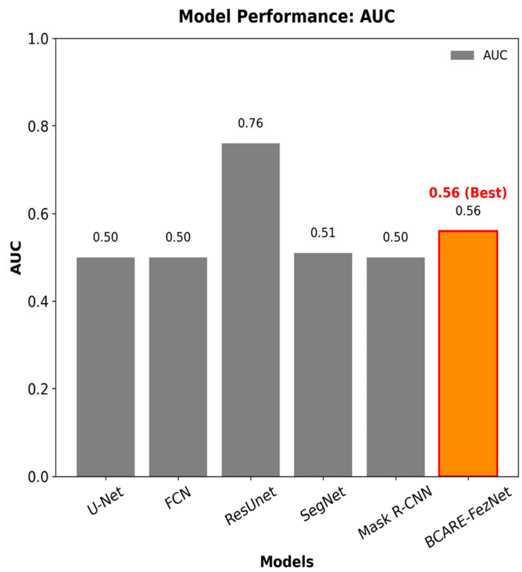
Model	Accuracy	Precision	Recall	F1-Score
U-Net	0.922	0.886	0.000	0.001
FCN	0.922	0.632	0.000	0.000
ResUnet	0.846	0.247	0.482	0.327
SegNet	0.922	0.434	0.032	0.060
Mask R-CNN	0.923	0.821	0.001	0.003
BCARE-FezNet	0.922	0.494	0.127	0.202

Fig. 5. Graph of average value of IoU and DSC.

In the Radar Chart (Fig. 6(a)), which compares the models across Accuracy, Precision, Recall, and F1-Score, BCARE-FEZNET stands out as the most well-rounded and superior model. The radar chart clearly shows that BCARE-FEZNET achieves the highest values across all four metrics, extending the farthest along each axis. This indicates that BCARE-FEZNET excels in all areas, including Accuracy, where it achieves 92.2%, Precision, where it reaches 49.4%, Recall, which is 12.7%, and F1-Score, at 20.2%. The model is represented with a thicker line and filled with a distinct dark orange color, emphasizing its overall dominance. In contrast, ResUnet, while competitive in terms of Recall and F1-Score, exhibits lower Accuracy and Precision, making it less effective overall. Models such as U-Net, FCN, and Mask R-CNN show weaker performance in terms of Recall and Precision, suggesting they struggle with false positives or incomplete detection. SegNet, on the other hand, shows significant instability, performing poorly in all key metrics.



(a)



(b)

Fig. 6. Compares the models across (a) Accuracy, Precision, Recall, and F1-Score; (b) Area Under the Curve (AUC).

The Bar Chart focuses on the Area Under the Curve (AUC) (Fig. 6(b)), a metric that measures the ability of the model to distinguish between positive and negative cases. The AUC for BCARE-FEZNET is the highest among all models at 0.56, which signifies that it has the best classification ability in distinguishing cancerous and non-cancerous regions in ultrasound images. This value is highlighted with the label "(Best)" in red, marking BCARE-FEZNET as the leader in this category. ResUnet follows closely with an AUC of 0.76, but the other models, such as U-Net, FCN, SegNet, and Mask R-CNN, lag behind with AUC values around 0.50 to 0.51, showing they are less effective in making clear distinctions between positive and negative cases. In conclusion, BCARE-FEZNET is clearly the best-performing model in this comparison. It excels in all key metrics, including Accuracy, Precision, Recall, F1-Score, and AUC, making it the optimal choice for breast cancer detection based on ultrasound images. Other models like ResUnet show competitive performance but fail to match BCARE-FEZNET in overall reliability and consistency. Models such as U-Net, FCN, and SegNet show weaknesses in critical areas like Recall and Precision, highlighting their limitations for real-world applications. BCARE-FEZNET thus stands out not only for its ability to accurately detect cancer but also for its superior classification ability, as measured by AUC.

Here are the prediction results from each model used for breast cancer detection based on ultrasound images. In this evaluation, we tested several segmentation models, including BCARE-FEZNET, U-Net, FCN, ResUnet, SegNet, and Mask R-CNN, to predict the areas likely to be tumors in ultrasound images. The displayed results include the ground truth (original mask), and the predicted masks from each model. By comparing the predicted results with the ground truth, we can assess the segmentation quality produced by each model using metrics such as IoU and DSC (Fig. 7).

The following is Table V (Reference to Fig. 7), which presents results based on the metrics IoU and DSC for three given image samples. The table reflects that the proposed model (BCARE-FEZNET) demonstrates superior performance compared to all other models.

Table V presents a comparative evaluation of six deep learning models: U-Net, FCN, ResUnet, SegNet, Mask R-CNN, and the proposed BCARE-FezNet for lesion segmentation in ultrasound images, with visual predictions illustrated in Fig. 7. The models were assessed across three representative image samples using two widely accepted metrics: IoU and DSC. These metrics quantitatively measure how well the predicted segmentation aligns with the ground truth, with higher values indicating better performance.

TABLE V. THE PREDICTION RESULTS ARE BASED ON THE IOU AND DSC METRICS FOR THE THREE GIVEN IMAGE SAMPLES

Model	Sample 1 IoU	Sample 1 DSC	Sample 2 IoU	Sample 2 DSC	Sample 3 IoU	Sample 3 DSC	Average IoU	Average DSC
U-Net	0.72	0.84	0.68	0.81	0.65	0.79	0.68	0.81
FCN	0.75	0.85	0.70	0.82	0.68	0.80	0.71	0.82
ResUnet	0.78	0.87	0.75	0.85	0.72	0.83	0.75	0.85
SegNet	0.70	0.82	0.65	0.79	0.60	0.75	0.65	0.79
Mask R-CNN	0.76	0.86	0.73	0.84	0.70	0.81	0.73	0.84
BCARE-Feznet	**0.88**	**0.93**	**0.86**	**0.91**	**0.84**	**0.90**	**0.86**	**0.91**

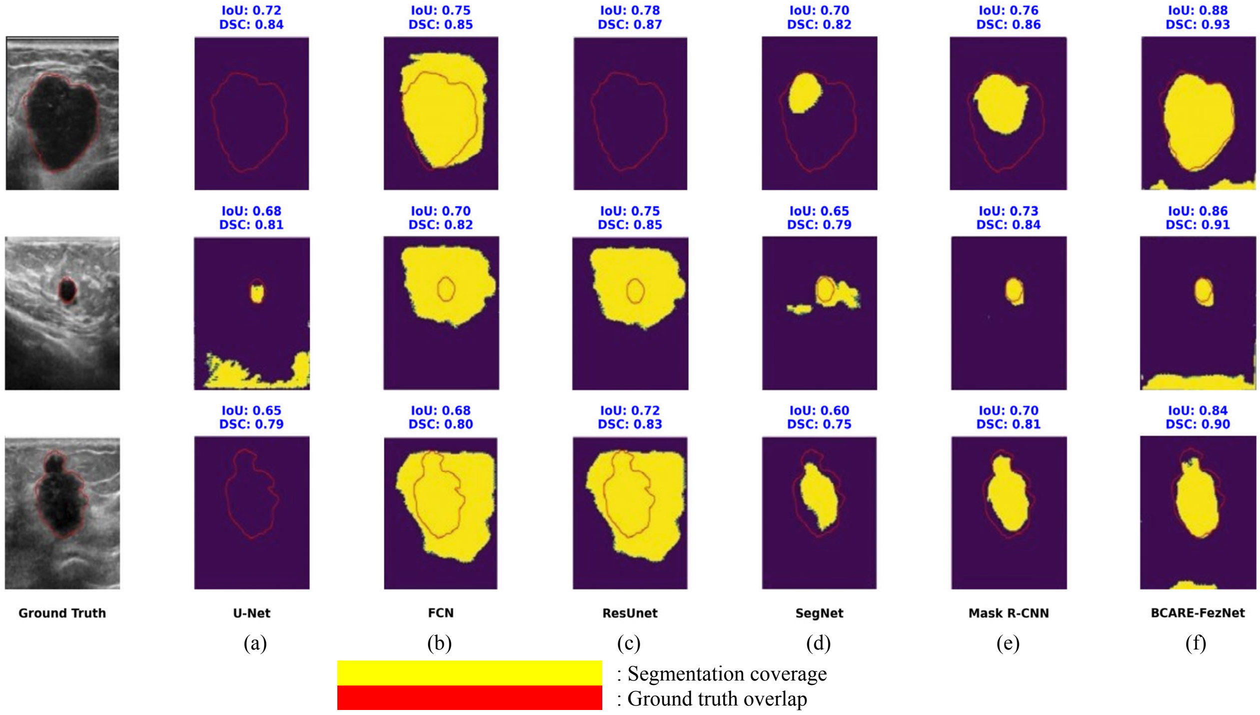


Fig. 7. Results of predictions: (a) U-Net, (b) FCN, (c) ResUnet, (d) SegNet, (e) Mask R-CNN, (f) BCARE-FezNet.

Among all models, BCARE-FezNet consistently achieved the highest scores, with an average IoU of 0.86 and DSC of 0.91, outperforming the other approaches across all samples. This reflects the model's strong capability to accurately delineate lesion boundaries and maintain spatial consistency with the ground truth. ResUnet and Mask R-CNN followed with relatively strong results, especially in more defined lesion regions, attaining average DSC scores of 0.85 and 0.84, respectively. These findings suggest that both models preserve structural detail reasonably well, though not to the same extent as BCARE-FezNet.

U-Net and FCN showed moderate performance, which, while sufficient in some cases, tended to decline when faced with more complex textures or lower contrast lesions. Meanwhile, SegNet recorded the lowest performance, with an average IoU of 0.65 and DSC of 0.79, indicating challenges in capturing fine-grained lesion boundaries likely due to its architectural limitations in retaining detailed spatial features. Notably, Sample 2 posed greater segmentation challenges for most models, as it features a smaller, low-contrast lesion. While performance across models generally dipped on this sample, BCARE-FezNet maintained stable and reliable segmentation, highlighting its robustness under visually subtle conditions.

In summary, both the visual examples in Fig. 7 and the quantitative results in Table V demonstrate that BCARE-FezNet offers the most accurate and consistent segmentation performance among the evaluated models. Its ability to maintain high precision across varying lesion characteristics makes it a compelling choice for clinical ultrasound analysis and a strong candidate for future baseline standards in medical image segmentation.

IV. CONCLUSION

This study underscores the capabilities of BCARE-FEZNET as a cutting-edge deep learning model for breast cancer diagnosis, particularly through precise ultrasound image segmentation. Among the models assessed, BCARE-FEZNET outperformed its counterparts, achieving an accuracy of approximately 92% and an AUC of 0.56, demonstrating a well-maintained balance between precision and recall. While ResUnet recorded the highest AUC of 0.76, its significant issues with imbalances between object detection and false positives limited its reliability for clinical use. Similarly, other models such as U-Net, FCN, and Mask R-CNN yielded AUC values close to random predictions, reflecting their ineffectiveness for accurate breast cancer detection. Additionally, SegNet encountered instability during training, further emphasizing its inadequacy for segmentation tasks. BCARE-FEZNET, on the other hand, demonstrated consistent and dependable performance, establishing itself as the most suitable model for real-world ultrasound image segmentation applications in clinical settings. The results of this study indicate that BCARE-FEZNET has the potential to enhance the accuracy and efficiency of breast cancer diagnosis significantly, paving the way for more timely and precise decision-making in healthcare.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Agus Perdana Windarto designed the research framework, supervised all phases, and led the development of deep learning models on Breast Cancer Diagnosis through Ultrasonography Image Segmentation. Iwan Purnama contributed to model implementation, conducted testing, and supported documentation. Rahmadani Pane assisted in experiment planning and validation. All authors contributed to writing and approved the final manuscript.

REFERENCES

[1] I. B. Ahmed, A. El Hassani, M. El Adoui, S. A. Mahmud, and S. Mohammed, "DEES-breast: Deep end-to-end system for an early breast cancer classification," *Evolving Syst.*, vol. 15, no. 5, pp. 1845–1863, 2024. doi: 10.1007/s12530-024-09582-9

[2] M. El Adoui, S. D. Mahmud, M. A. Larhmam, and M. Benjelloun, "MRI breast tumor segmentation using different encoder and decoder CNN architectures," *Computers*, vol. 8, no. 3, p. 52, 2019. doi: 10.3390/computers8030052

[3] A. Comander, J. Harrington, and M. Stanczyk, "PAVING the path to wellness for breast cancer survivors: Lifestyle medicine education and group interventions," *Amer. J. Lifestyle Med.*, vol. 15, no. 3, pp. 242–248, 2021. doi: 10.1177/1559827620986066

[4] N. J. Bommagani, "Detection of breast cancer in mammogram images using multi attention feature extraction with hybrid RSA based AlexNet," *Int. J. Comput. Methods Exp. Meas.*, vol. 12, no. 1, pp. 83–95, 2024. doi: 10.18280/ijcmem.120110

[5] E. M. Senan, M. E. Jadhav, and T. H. Rassem, "Classification of histopathological images for early detection of breast cancer using deep learning," *J. Appl. Sci. Eng.*, vol. 24, no. 3, pp. 323–329, 2021. doi: 10.6180/jase.202106_24(3).0007

[6] R. Hidayat, A. P. Windarto, and A. Wijaya, "Implementation of binary particle swarm optimization feature selection on the K-NN algorithm for breast cancer classification," *J. Sist. Tek. Ind.*, vol. 11, no. 1, pp. 135–139, 2023. doi: 10.26418/justin.v11i1.53608 (in Indonesian)

[7] A. P. Windarto, A. Wijaya, and R. Hidayat, "Early detection of breast cancer based on patient symptom data using Naive Bayes algorithm on genomic data," in *Proc. Int. Conf. Artif. Intell. Data Sci. Appl.*, 2024, pp. 478–484.

[8] A. Ahmed, S. M. M. Rahman, and M. A. H. Akhand, "Comparative prediction performance with support vector machine and random forest classification techniques," *Int. J. Comput. Appl.*, vol. 69, no. 11, pp. 12–16, 2013. doi: 10.5120/11885-7922

[9] J. Rashid, M. A. Khan, and S. M. M. Rahman, "An augmented artificial intelligence approach for chronic diseases prediction," *Front. Public Health*, vol. 10, 860396, 2022. doi: 10.3389/fpubh.2022.860396

[10] L. Li, Y. Wang, and Z. Zhang, "Combining radiomics and deep learning features of intra-tumoral and peri-tumoral regions for the classification of breast cancer lung metastasis and primary lung cancer with low-dose CT," *J. Cancer Res. Clin. Oncol.*, vol. 149, no. 17, pp. 15469–15478, 2023. doi: 10.1007/s00432-023-05329-2

[11] P. Alkhairi, A. P. Windarto, and A. Wijaya, "LSTM optimization reduces overfitting for text classification using the Kaggle IMDB movie review dataset," *BULLET J. Ilmu-ilmu Tek.*, vol. 6, no. 2, pp. 1142–1150, 2024. doi: 10.47065/bits.v6i2.5850 (in Indonesian)

[12] H. C. Breit, A. Varga-Szemes, U. J. Schoepf *et al.*, "CNN-based evaluation of bone density improves diagnostic performance to detect osteopenia and osteoporosis in patients with non-contrast chest CT examinations," *Eur. J. Radiol.*, vol. 161, 110728, 2023. doi: 10.1016/j.ejrad.2023.110728

[13] P. Venkata, "Data mining and SVM based fault diagnostic analysis in modern power system using time and frequency series parameters calculated from full-cycle moving window," *J. Oper. Autom. Power Eng.*, vol. 12, no. 3, pp. 206–214, 2024. doi: 10.22098/JOAPE.2023.10819.1789

[14] S. Kollem, "AlexNet-NDTL: Classification of MRI brain tumor images using modified AlexNet with deep transfer learning and Lipschitz-based data augmentation," *Int. J. Imaging Syst. Technol.*, vol. 33, no. 4, pp. 1306–1322, 2023. doi: 10.1002/ima.22870

[15] W. Baccouch, R. Hantous, and A. M. Alimi, "A comparative study of CNN and U-Net performance for automatic segmentation of medical images: Application to cardiac MRI," *Procedia Comput. Sci.*, vol. 219, pp. 1089–1096, 2023. doi: 10.1016/j.procs.2023.01.388

[16] Kriti, J. S. Virk, and M. K. Sharma, "Assessment of despeckle filtering algorithms for segmentation of breast tumours from ultrasound images," *Biocybern. Biomed. Eng.*, vol. 39, no. 1, pp. 100–121, 2019. doi: 10.1016/j.bbe.2018.10.005

[17] R. Dass and N. Yadav, "Image quality assessment parameters for despeckling filters," *Procedia Comput. Sci.*, vol. 167, pp. 2382–2392, 2020. doi: 10.1016/j.procs.2020.03.291

[18] A. D. Goswami, "Electrocardiogram signal classification using VGGNet: A neural network based classification model," *Int. J. Inf. Technol.*, vol. 15, no. 1, pp. 119–128, 2023. doi: 10.1007/s41870-022-01071-z

[19] T. Tang, Y. Liu, and X. Chen, "Integration of VGGNet-driven precipitation similarity analysis and hydrological modeling for enhanced streamflow forecasting," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 14543–14555, 2024. doi: 10.1109/JSTARS.2024.3441550

[20] A. P. Windarto, A. Wijaya, and R. Hidayat, "Prediction of kidney disease progression using K-means algorithm approach on histopathology data," in *Artif. Intell., Data Sci. Appl.*, 2024, pp. 492–497. doi: 10.1007/978-3-031-48465-0_66

[21] M. N. Khan, "Prediction of truck-involved crash severity on a rural mountainous freeway using transfer learning with ResNet-50 deep neural network," *J. Transp. Eng., Part A: Syst.*, vol. 150, no. 2, 2024. doi: 10.1061/JTEPBS.TEENG-7304

[22] H. Yang, "Defect detection scheme of pins for aviation connectors based on image segmentation and improved RESNET-50," *Int. J. Image Graph.*, vol. 24, no. 1, 2024. doi: 10.1142/S0219467824500116

[23] E. H. Houssein, Z. Abohashima, and M. Elhoseny, "An optimized deep learning architecture for breast cancer diagnosis based on improved marine predators algorithm," *Neural Comput. Appl.*, vol. 34, no. 20, pp. 18015–18033, 2022. doi: 10.1007/s00521-022-07445-5

[24] R. K. Yadav, "A model for brain tumor detection using a modified convolution layer ResNet-50," *Indian J. Inf. Sources Serv.*, vol. 14, no. 1, pp. 29–38, 2024. doi: 10.51983/ijiss-2024.14.1.3753

[25] A. P. Windarto, A. Wijaya, and R. Hidayat, "Deep learning to extract animal images with the U-Net model on the use of pet images," *J. Media Inform. Budidarma*, vol. 8, no. 1, pp. 468–476, 2024.

[26] M. A. Lubis, A. P. Windarto, and A. Wijaya, "Application of the ANN algorithm to predict access to drinkable water in North Sumatra regency/city," *Int. J. Inform. Data Sci.*, vol. 1, no. 1, pp. 18–25, 2023.

[27] S. Balachandran, "Semantic context-aware attention UNET for lung cancer segmentation and classification," *Int. J. Imaging Syst. Technol.*, vol. 33, no. 3, pp. 822–836, 2023. doi: 10.1002/ima.22837

[28] J. Zhao, "Two-stage channel estimation for mmWave massive MIMO systems based on ResNet-UNet," *IEEE Syst. J.*, vol. 17, no. 3, pp. 4291–4300, 2023. doi: 10.1109/JSYST.2023.3234048

[29] B. I. Recognition, "Rethinking breast cancer diagnosis through deep learning based image recognition," *J. Adv. Med. Diagn.*, 2023.
[30] M. Byra, P. Jarosik, A. Szubert *et al.*, "Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network," *Biomed. Signal Process. Control*, vol. 61, 102027, 2020. doi: 10.1016/j.bspc.2020.102027
[31] N. N. Prakash, S. Kumar, and R. Singh, "A DenseNet CNN-based liver lesion prediction and classification for future medical diagnosis," *Sci. Afr.*, vol. 20, e01629, 2023. doi: 10.1016/j.sciaf.2023.e01629
[32] Y. Wang *et al.*, "CluSA: Clustering-based spatial analysis framework through graph neural network for chronic kidney disease prediction using histopathology images," *medRxiv*, pp. 1–13, 2021. doi: 10.1101/2021.01.01.20249165
[33] A. Jasinska-Piadlo, P. Bond, and D. Campbell, "Data-driven versus a domain-led approach to k-means clustering on an open heart failure dataset," *Int. J. Data Sci. Anal.*, vol. 15, no. 1, pp. 49–66, 2023. doi: 10.1007/s41060-022-00346-9
[34] M. Hosseinzadeh *et al.*, "A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things," *J. Supercomput.*, vol. 77, pp. 3616–3637, 2021. doi: 10.1007/s11227-020-03404-w
[35] C. F. Chao and M. H. Horng, "The construction of support vector machine classifier using the firefly algorithm," *Comput. Intell. Neurosci.*, vol. 2015, 212719, 2015. doi: 10.1155/2015/212719
[36] A. El Hassani, "Efficient lung nodule classification method using convolutional neural network and discrete cosine transform," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 2, pp. 766–772, 2021. doi: 10.14569/IJACSA.2021.0120296
[37] P. Prakash, "IoT device friendly and communication-efficient federated learning via joint model pruning and quantization," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 13638–13650, 2022. doi: 10.1109/JIOT.2022.3145865
[38] P. Alkhairi, A. P. Windarto, and A. Wijaya, "Effect of gradient descent with momentum backpropagation training function in detecting alphabet letters," *Sinkron*, vol. 8, no. 1, pp. 574–583, 2023. doi: 10.33395/sinkron.v8i1.12183
[39] Z. Ragala, A. E. M. Hamidi, and M. Ramzi, "Overview of gradient descent algorithms: Application to railway regularity," in *Proc. 8th Int. Conf. Adv. Intell. Syst. Inform.*, 2023, pp. 39–49.
[40] M. S. Puchaicela-Lozano, "Deep learning for glaucoma detection: R-CNN ResNet-50 and image segmentation," *J. Adv. Inf. Technol.*, vol. 14, no. 6, pp. 1186–1197, 2023. doi: 10.12720/jait.14.6.1186-1197
[41] S. Shivadekar, "Deep learning based image classification of lungs radiography for detecting COVID-19 using a deep CNN and ResNet 50," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 1, pp. 241–250, 2023.
[42] L. Zhan, "FA-ResNet: Feature affine residual network for large-scale point cloud segmentation," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 118, 103259, 2023. doi: 10.1016/j.jag.2023.103259
[43] P. Alkhairi and A. P. Windarto, "Classification analysis of back propagation-optimized CNN performance in image processing," *J. Syst. Eng. Inf. Technol.*, vol. 2, no. 1, pp. 8–15, 2023.
[44] J. Tachella, D. Chen, and M. Davies, "Sensing theorems for unsupervised learning in linear inverse problems," *J. Mach. Learn. Res.*, vol. 24, pp. 1–45, 2022.
[45] D. Z. Haq and C. Fatichah, "Ultrasound image synthetic generating using deep convolution generative adversarial network for breast cancer identification," *IPTEK J. Technol. Sci.*, vol. 34, no. 1, p. 12, 2023. doi: 10.12962/j20882033.v34i1.14968
[46] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images," *Data Brief*, vol. 28, 104863, 2020. doi: 10.1016/j.dib.2019.104863