Journal of Information System Exploration and Research Vol. 2, No. 2, July 2024, pp. 93-102

SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Information System Exploration and Research

https://shmpublisher.com/index.php/joiser

p-ISSN 2964-1160 | e-ISSN 2963-6361

Inception ResNet v2 for Early Detection of Breast Cancer in Ultrasound Images

Tiara Lailatul Nikmah[1*], Risma Moulidya Syafei[2], Devi Nurul Anisa[3], Elmo Juanara[4], Zohri Mahrus[5]

[1,2,3]Department of Computer Science, Universitas Negeri Semarang, Indonesia
[4]Department of Computer Science, Japan Institute of Science and Technology, Japan
[5]School of Chemistry, University of Melbourne, Australia

DOI: https://doi.org/10.52465/joiser.v2i2.439
Received 28 July 2024; Accepted 30 July 2024; Available online 31 July 2024

Article Info

Keywords:
Breast cancer detection;
Inception ResNet v2;
Data augmentation;
CNN algorithm

Abstract

Breast cancer is one of the leading causes of death in women. Early detection through breast ultrasound images is important and can be improved using machine learning models, which are more accurate and faster than manual methods. Previous research has shown that the use of the CNN (Convolutional Neural Network) algorithm in breast cancer detection still does not achieve high accuracy. This study aims to improve the accuracy of breast cancer detection using the Inception ResNet v2 transfer learning method and data augmentation. The data is divided into training, validation and testing data consisting of 3 classes, namely Benign, Malignant and Normal. The augmentation process includes rotation, zoom, and rescale. The model trained using CNN and Inception ResNet v2 showed good performance by producing the highest accuracy of 89.72% in the training data evaluation data and getting 90% accuracy in the prediction test stage with data testing. This study shows that the combination of data augmentation and the Inception ResNet v2 architecture can improve the accuracy of breast cancer detection in CNN models.

This is an open-access article under the CC BY-SA license.

1. Introduction

Breast cancer is one of the highest causes of death in women. According to the World Health Organization, the death rate from breast cancer in Indonesia is 15.9 per 100,000 women, which means 15.9 out of every 100,000 women die from breast cancer each year [1]. About half of all breast cancers occur in women who have no specific risk factors other than gender and age [2]. Most people will not experience any symptoms when cancer is early, which is why it is important to do early detection.

Early detection of breast cancer is carried out to predict the presence of cancer through the images of breast x-ray images provided. Early detection of breast cancer uses advanced machine learning

* ***Corresponding Author:***

Tiara Lailatul Nikmah,
Department of Computer Science,
Universitas Negeri Semarang,
Semarang, Indonesia.
Email: tiaralaila21@gmail.com

models that can facilitate the detection process. Early detection using machine learning models is considered more effective than manual methods. Machine learning models are trained on very large amounts of data, allowing them to identify patterns and anomalies that humans might miss [3]. This results in a higher level of accuracy in detecting cancer at an early stage. In addition, machine learning models can process data very quickly, allowing cancer detection to be carried out faster. This is very important to increase the chances of treatment success.

Early detection of breast cancer using machine learning models has been widely carried out in previous studies. Many studies use deep learning algorithms, especially the Convolutional Neural Network (CNN) [4], [5]. Deep learning models are able to extract features that are more accurate from images than traditional machine learning models [6]. CNN models are perfect for capturing complex patterns in imagery. In a study conducted by Zeimarani et al. [7] using Breast Lesion Ultrasound Images images, the accuracy was 87.07%. Meanwhile, in another research dataset, Ahmed et al. [8] obtained an accuracy of 88.1% after using the deep convolutional neural network (DCNN) method.

The use of the CNN algorithm alone is considered to lack high accuracy. Therefore, the study aims to improve the accuracy of CNN in the detection of Breast Cancer images using the Inception Resnet v2 transfer learning method. The transfer learning method takes a pretrained model that has been trained in advance to be combined with CNNs. The Inception ResNet-v2 architecture can handle scale variations in imagery data well [9]. With the combination of Inception and ResNet, Inception ResNet-v2 is capable of improving object recognition performance [10]. The Inception module helps in better and more detailed feature extraction, while the residue block deepens the network and prevents gradient vanishing issues, thereby improving the accuracy of object recognition [11]. In a study conducted by Badawy et al. [12] who compared the application of various types of transfer learning on CNN produced the best accuracy in the Inception Resnet v2 transfer learning model with an accuracy of 82.93%. However, this accuracy is still considered too low. Therefore, in this study, data augmentation will be carried out first to process the image so that it is better when processed by the model.

This research contributes to improving the accuracy of the CNN algorithm using the Inception Resnet v2 method, image augmentation to improve the accuracy of the Inception Resnet v2 model. The steps of this research include, data augmentation, data splitting, modelling and evaluation. Model evaluation will use accuracy metrics to measure how accurately the model classifies imagery.

2. Literature Review

In a study conducted by Zeimarani et al. [7], ultrasound images of breast lesions were used to detect breast cancer. They developed a model that utilizes deep learning techniques, specifically by using the Convolutional Neural Network (CNN) algorithm. The model managed to achieve an accuracy rate of 87.07% in classifying images into different categories of breast cancer. These results show that deep learning techniques can provide quite good performance in medical image classification tasks, especially in detecting breast cancer from ultrasound images.

On the other hand, the research conducted by Ahmed et al. [8] used a deep convolutional neural network (DCNN) approach on different datasets. In this study, they also focused on breast cancer detection through medical image analysis. The model developed by Ahmed et al. managed to achieve an accuracy of 88.1%, slightly higher compared to the results obtained by Zeimarani et al. These results indicate that the use of more complex deep learning architectures, such as DCNN, can improve the accuracy of breast cancer detection.

Furthermore, research by Karthik et al. [13] showed more promising results using CNN's Gaussian Dropout Based Stacked Ensemble method. In the dataset used, they managed to achieve an accuracy of 92.15% in breast cancer detection. This Gaussian Dropout method helps in preventing overfitting and improving model generalization, leading to higher accuracy compared to previous approaches.

The three studies show great potential from the use of deep learning algorithms in detecting breast cancer. However, even though the results achieved are quite high, challenges remain in achieving higher and consistent accuracy. Factors such as dataset quality and diversity, data augmentation techniques, as well as model training parameters greatly affect the final result.

3. Method

The research was carried out by carrying out various stages, namely Split Data, Data Augmentation, Training Model, Evaluation and Prediction Test.

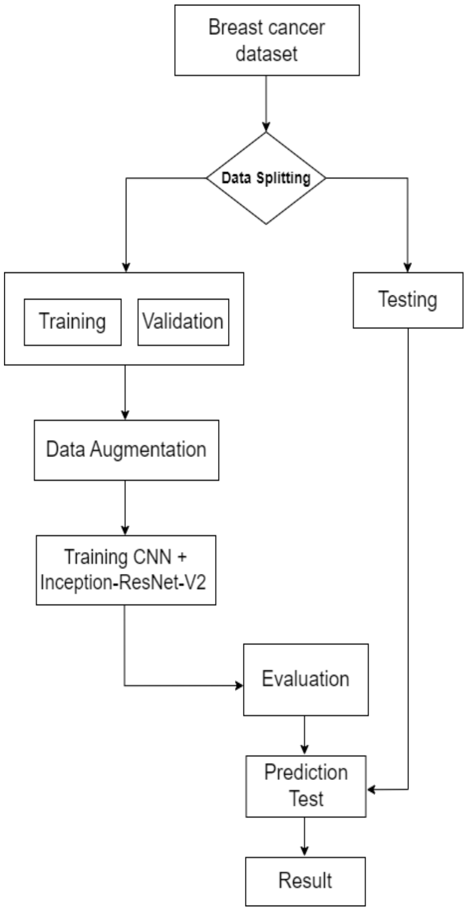

Figure 3. Research Flowchart

Split Data

Data split is the process of dividing a dataset into several subsets, consisting of training data, validation data, and test data. The goal is to ensure that the model can be tested and validated with never-before-seen data, thereby reducing the risk of overfitting and providing more accurate performance estimates. The dataset is divided into 70% training data, 20% validation data, and 10% test data.

Data Augmentation

Data augmentation is a technique that is done by making variations of existing data through transformations such as rotation, mirroring, cropping, brightness changes, and others. The goal is to increase the amount and diversity of training data to help the model learn more robust features and better generalize to the new data.

Training Model

Model training is carried out using the Convolutional Neural Network (CNN) algorithm, which is then combined with the Inception ResNet v2 model as the base model. At this stage, the pre-trained Inception ResNet v2 model with large datasets such as ImageNet is used to leverage its ability to recognize complex features in images. This process uses various layers including flatten, dense and Inception ResNet v2 as the base model.

Convolutional Neural Network (CNN)

Convolutional Neural network (CNN) is a variation of MultiLayerPerception inspired by human neural networks based on the findings of Hubel and Wiesel who conducted a study of the visual cortex in cats' sense of vision. The study was particularly inspired by the way the visual cortex works in animals that are very powerful in the visual processing system [14]. CNNs are multi-layerperceptrons which are one variant of deep neural networks that are often used as a method to classify digital image data. Today, CNNs are widely applied to various areas of life, including system recommendations, medical image analysis, natural language processing, financial time series analysis [15]. Convolutional Neural network (CNN) is a deep learning method that is widely used in computer vision, such as classification, detection, and segmentation. CNN learns to extract features from imagery by repeating the learning and then generating a feature map [16]. The structure of a CNN contains an input layer, a hidden layer, and an output layer. The commonly used type of CNN is similar to Multilayer Perceptron (MLP) which includes several Convolutional Layers, Pooling Layers, and finally a Fully Connected Layer.

The CNN architecture has two main parts, namely Feature learning and Classification as seen in the following image

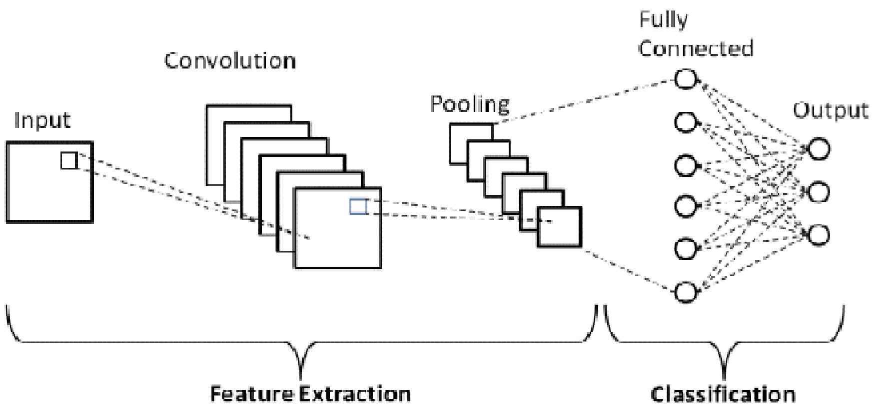

Figure 1. CNN Structure [17]

Inception-ResNet-V2

The advantage of the InceptionResNet-v2 architecture method is its ability to recognize objects with high accuracy and handle unstructured data by extracting relevant features from images using convolution. Where Convolution is the act of image extraction to obtain a model in the form of a kernel matrix. In this process, filtering is carried out that shifts with a certain "step value" on an image input. Or a process to look at the value of a parameter that determines how much the filter shifts in the input image. Furthermore, the results of the convolution become inputs for the fully connected parts for the classification process.

The InceptionResNetV2 architecture combines initial and residual connection models to improve performance [18]. This hybrid approach allows the network to take advantage of the benefits of both models, including faster training times and avoiding disappearing gradient issues. Residual connections also allow the network to pass through multiple layers during training. In addition, InceptionResNetV2 uses a double-sized kernel in a single layer to draw patterns with diverse hierarchies, which further improves the network's ability to capture features of varying complexity [19]. The Inception-ResNet model has several blocks that contain convolutional layers, filter merging, ReLU activation functions, ResNet, and the initial structure. The architectural layout of this block is depicted in Figure 2. The distinctive design of the Inception-ResNetV2 model and the effective use of filters have resulted in impressive results in the task of medical image classification. As a result, we used the Inception-ResNetV2 pre-training model as the foundation for this study.

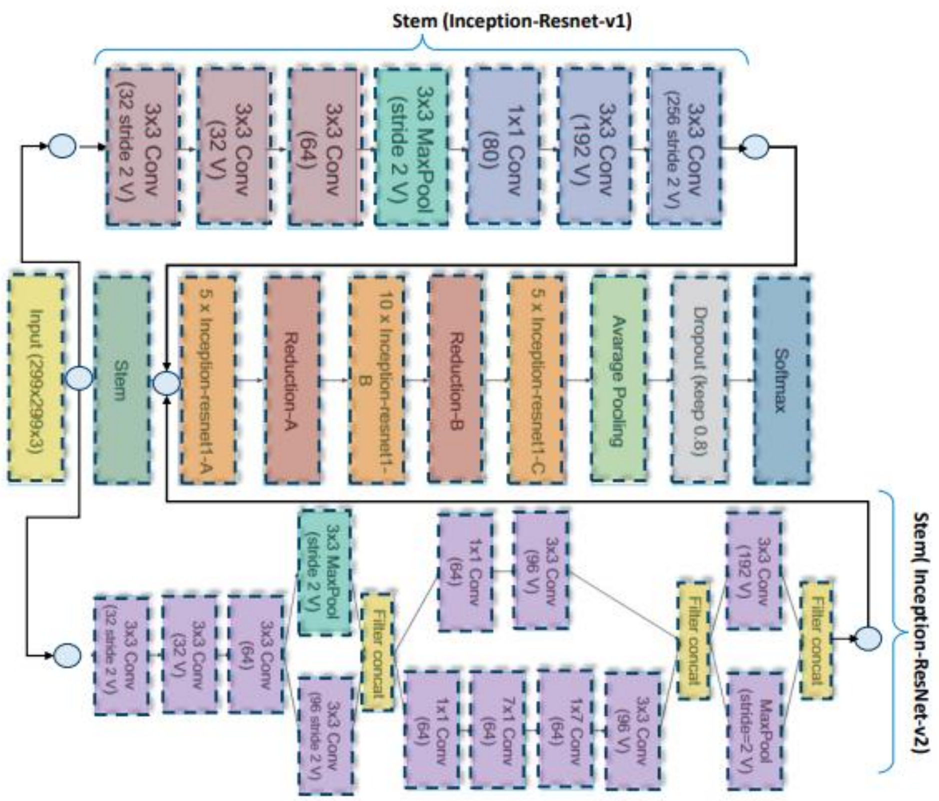

Figure 2. Schematic of Inception-Resnet v2 networks [20]

The Inception-ResNet-v2 approach is often used for CT-scan image classification due to its ability to extract important features from medical images [21]. This deep learning architecture combines two well-known models, namely Inception and ResNet. The Inception model excels at extracting features at various scales, while ResNet effectively addresses the problem of vanishing gradients during training [22]. The Inception-ResNet-v2 approach is designed to overcome the shortcomings of these two models and achieve high accuracy in image classification tasks. In addition, this approach is effective in handling large data sets and reducing computing time.

Evaluation

Evaluation is the process of measuring model performance using validation data that is not included in the training data. The evaluation metric used is accuracy. Accuracy metrics assess how well a model can correctly predict data.

Prediction Test

Prediction tests are carried out using test data to evaluate the model's ability to predict new data. This test data consists of samples that have never been used before in the training or validation process, thus providing an accurate measure of the model's performance on completely new data. The results of these prediction tests are usually summarized in the form of a confusion matrix, which provides a detailed picture of how well the model classifies each category.

4. Results and Discussion

The dataset will be divided into 3 parts, namely data train, test and validation. Train data and validation data will be used to train the model and test data will be used to evaluate the model. The data was divided into 10% test data, 20% validation data and 70% training data.

Training data and validation data will go through the data augmentation process using an imagedatagenerator. The data augmentation process is rotation, zoom, shear, fill mode, and rescale. The data augmentation parameters can be seen in Table 1.

Table 1. Data Augmentation Parameters

Function	**Parameters**
Rotation_range	20
Zoom_range	0.2
Shear_range	0.2
Fill_mode	'nearest'
rescale	1./255
Validation_split	0.2
Target_size	150, 150
Class_mode	'categorical'

A comparison between the original sample and the sample from the data augmentation results can be seen in Figure 4.

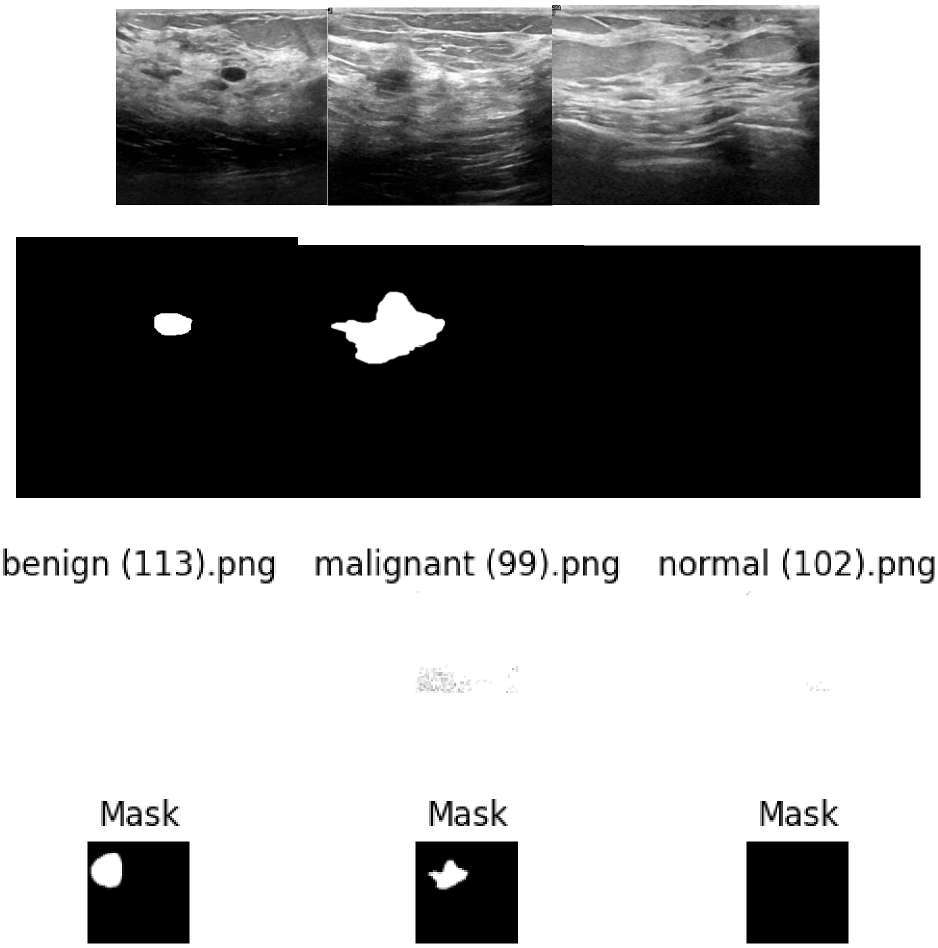

Figure 4. Comparison between the original sample and the data augmentation sample

This image displays side-by-side examples of original data with variations generated through the augmentation process. Through this comparison, we can clearly see how data augmentation adds diversity to the training dataset, by applying various transformations such as rotation, rescale, brightness changes, and contrast adjustments.

The training process is carried out using the pre-trained Inception ResNet v2 model which has been trained with ImageNet big data. This model was chosen for its ability to recognize complex patterns in images, given its deep and sophisticated architecture. As a model that has been trained with a very large and diverse dataset, Inception ResNet v2 is able to provide more accurate and efficient results than models trained from scratch. The Inception ResNet v2 model will be used as the Base Model for the CNN architecture model that will be built later. In more detail, the model architecture can be seen in Table 2.

Table 2. Model Architecture

Layers	Parameters
Flatten	Last_output
Dense	256, activation = 'relu'
Dropout	0.2
Dense	3, activation = 'softmax'
Base model	InceptionResnetV2

The training process uses 10 epochs with 143 steps per epoch. The Loss function used is 'categorical_crossentropy' and the optimizer 'Adam' with a learning_rate of 0.001. As for the evaluation metric, it is 'accuracy'. The results of the model training can be seen in Table 3.

Table 3. The results of the model training

Epoch	Accuracy	Val_accuracy
1	0.8620	0.7809
2	0.8524	0.7951
3	08620	0.8269
4	0.8972	0.8481
5	0.8972	0.8799
6	0.8032	0.4417
7	0.8049	0.8481
8	0.8462	0.7350
9	0.8787	0.8339
10	0.8928	0.8975

Based on the results seen in Table 3, the training process of the CNN model combined with Inception ResNet v2 shows good performance. The highest accuracy value was recorded in the 5th epoch with an accuracy value of 89.72%. This shows that the model with this architecture is able to learn quickly and achieve optimal performance in a relatively short time. In addition, the achievement of high accuracy in the initial epoch indicates that the model is able to generalize patterns and features from the training data effectively. This stable and high accuracy value is important to ensure that the model is not only good at remembering training data but is also able to provide accurate predictions on validation data. This achievement confirms that the use of Inception ResNet v2, with its complexity and depth, has successfully improved the model's ability to recognize and classify various complex features in data. Further evaluation and detailed analysis of the model's performance at each epoch can be seen in Table 3, which provides a complete picture of the accuracy development during the training process.

To see the visualization of the training results, it can be seen in Figure 5 which presents a graph of changes in accuracy and loss values in both training and validation data. This graph provides a clear picture of how the model's performance evolves as the epoch increases.

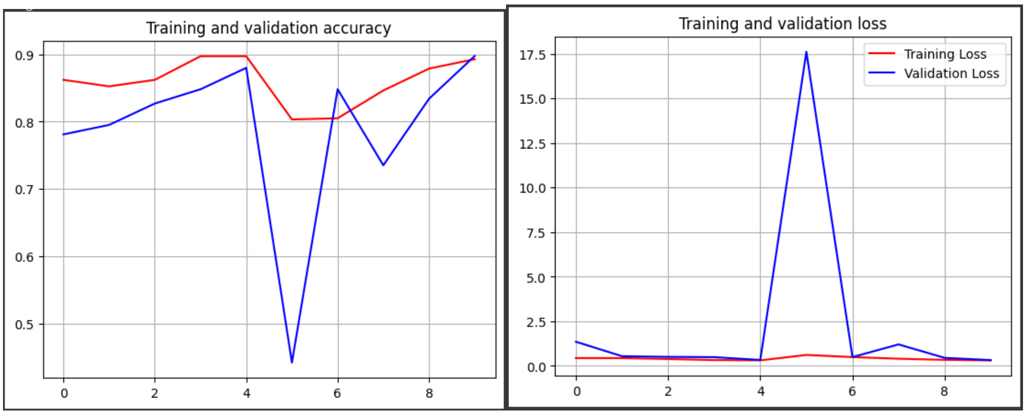

Figure 5. Visualization of model training results

In the accuracy graph, we can observe the trend of increasing accuracy values in training and validation data over time. Although there was a significant decline, the accuracy value dropped again and did not interfere with the model's performance. Continuously increasing accuracy values indicate that the model is increasingly able to recognize important patterns and features in the training data, while the accuracy trends in the validation data give an indication that the model is able to generalize that knowledge to data that has never been seen before.

In contrast, the loss graph shows a decrease in the loss value in the training and validation data. Although there was a significant increase, the loss value fell again and did not interfere with the model's performance. A decreasing loss value indicates that the model is getting better at minimizing prediction errors. The decrease in the loss value in the training data shows that the model has successfully learned from the data effectively, while the decrease in the loss value in the validation data shows that the model does not overfitting and still maintains its prediction ability on new data.

After training the model, the best model will be obtained, namely the model with the highest accuracy results, then prepare the testing data for model testing. The results of the model test obtained 90% accuracy.

Furthermore, the model that has been trained will be subjected to a prediction test process using test data. The results of this prediction test can be seen in Figure 6.

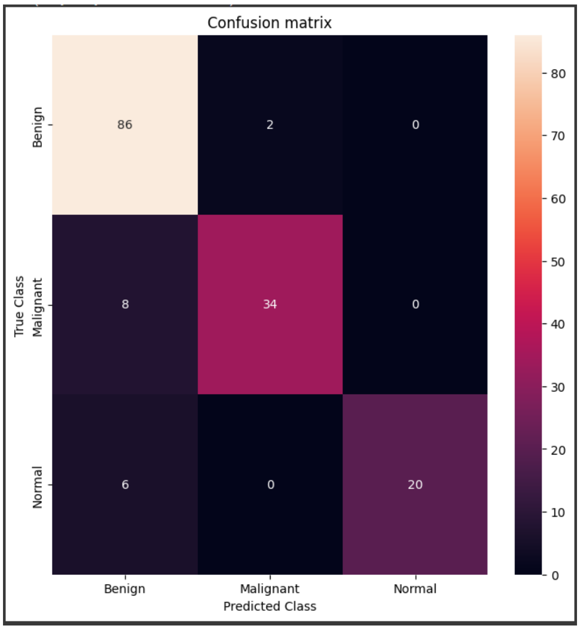

Figure 6. Confusion Matrix of Prediction Test

Figure 6 is a confusion matrix table of prediction test results that provides a detailed overview of the model's performance in classifying various data categories. In Figure 6, it can be seen that the model has done a good job of predicting, which is shown by most of the images that have been successfully classified correctly. This shows that the model has a high level of accuracy and is able to identify relevant patterns in the data.

However, there are still some prediction errors that need to be noted. There are 2 images that should be included in the Benign category but are predicted to be Malignant. This error may be due to the similarity of visual features between the two categories. In addition, there are 8 Malignant images that are predicted to be Benign, which shows that the model still needs refinement in detecting the distinctive features of the Malignant category. Another error noted was 6 images that should have been classified as Normal but were predicted to be Benign, which could be due to variations in the visual representation of the Normal imagery. Overall, however, the model has shown good performance in classifying data with sufficient accuracy.

Based on the research that has been conducted, it is found that the use of data augmentation for the Inception-ResNet-V2 architecture gets better performance results compared to previous research. The research conducted by Badawy et al. obtained an accuracy of 82.93% using Inception Resnet v2 while this study obtained greater accuracy of 90%, with a difference of 7.07%.

5. Conclusion

This study discusses the process of developing and evaluating a CNN model that uses the Inception ResNet v2 architecture to classify medical images. The data augmentation process is carried out to produce data variations to help the model recognize patterns better. The model is trained using CNN algorithms and Inception ResNet v2 which provides a solid foundation for complex pattern recognition. The highest accuracy value was recorded in the 5th epoch with a score of 89.72%, indicating that the model was able to learn and generalize patterns from the training data effectively. After training, the model was tested using test data and the results showed 90% accuracy. Overall, this study shows that the use of data augmentation and the Inception ResNet v2 architecture provide good performance. The use of other pre-trained models can be done to obtain higher accuracy in future studies.

References

[1] Kemenkes RI, “Jika Tidak Dikendalikan 26 Juta Orang Di Dunia Menderita Kanker,” Kemenkes RI. Accessed: Jul. 27, 2024. [Online]. Available: http://www.kemenkes-ri.htm/

[2] Etinosa Yvonne, “Breast cancer,” World Health Organization (WHO). Accessed: Jul. 27, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[3] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, “Machine learning for anomaly detection: A systematic review,” *Ieee Access*, vol. 9, pp. 78658–78700, 2021.

[4] M. Desai and M. Shah, “An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN),” *Clinical eHealth*, vol. 4, pp. 1–11, 2021, doi: https://doi.org/10.1016/j.ceh.2020.11.002.

[5] C. B. Gonçalves, J. R. Souza, and H. Fernandes, “Classification of static infrared images using pre-trained CNN for breast cancer detection,” in *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)*, 2021, pp. 101–106. doi: 10.1109/CBMS52027.2021.00094.

[6] P. Wang, E. Fan, and P. Wang, “Comparative analysis of image classification algorithms based on traditional machine learning and deep learning,” *Pattern Recognit Lett*, vol. 141, pp. 61–67, 2021.

[7] B. Zeimarani, M. G. F. Costa, N. Z. Nurani, S. R. Bianco, W. C. D. A. Pereira, and C. F. F. Costa Filho, “Breast lesion classification in ultrasound images using deep convolutional neural network,” *IEEE Access*, vol. 8, pp. 133349–133359, 2020.

[8] M. Ahmed, A. AlZoubi, and H. Du, “Improving Generalization of ENAS-Based CNN Models for Breast Lesion Classification from Ultrasound Images,” in *Medical Image Understanding and Analysis*, B. W. Papież, M. Yaqub, J. Jiao, A. I. L. Namburete, and J. A. Noble, Eds., Cham: Springer International Publishing, 2021, pp. 438–453.

[9] C. Peng, Y. Liu, X. Yuan, and Q. Chen, “Research of image recognition method based on enhanced inception-ResNet-V2,” *Multimed Tools Appl*, vol. 81, no. 24, pp. 34345–34365, 2022.

[10] J. Wang, X. He, S. Faming, G. Lu, H. Cong, and Q. Jiang, "A real-time bridge crack detection method based on an improved inception-resnet-v2 structure," *IEEE Access*, vol. 9, pp. 93209–93223, 2021.
[11] C. Chen and F. Qi, "Single image super-resolution using deep CNN with dense skip connections and inception-resnet," in *2018 9th international conference on information technology in medicine and education (ITME)*, IEEE, 2018, pp. 999–1003.
[12] S. M. Badawy, A. E.-N. A. Mohamed, A. A. Hefnawy, H. E. Zidan, M. T. GadAllah, and G. M. El-Banby, "Classification of Breast Ultrasound Images Based on Convolutional Neural Networks-A Comparative Study," in *2021 International Telecommunications Conference (ITC-Egypt)*, IEEE, 2021, pp. 1–8.
[13] R. Karthik, R. Menaka, G. S. Kathiresan, M. Anirudh, and M. Nagharjun, "Gaussian dropout based stacked ensemble CNN for classification of breast tumor in ultrasound images," *Irbm*, vol. 43, no. 6, pp. 715–733, 2022.
[14] A. Anton, N. F. Nissa, A. Janiati, N. Cahya, and P. Astuti, "Application of deep learning using convolutional neural network (CNN) method for Womenâ€™ s skin classification," *Scientific Journal of Informatics*, vol. 8, no. 1, pp. 144–153, 2021.
[15] R. Sayuti, *Convolutional Neural Networks Untuk Visi Komputer*, 1 ed. deepublish, 2021.
[16] M. Jogin, M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva, "Feature extraction using convolution neural networks (CNN) and deep learning," in *2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT)*, IEEE, 2018, pp. 2319–2323.
[17] F. Indriyani and I. R. Rahadjeng, "Klasifikasi Tumor Otak Berbasis Magnetic Resonance Imaging Menggunakan Algoritma Convolutional Neural Network," *Digital Transformation Technology*, vol. 3, no. 2, pp. 918–924, 2023.
[18] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the AAAI conference on artificial intelligence*, 2017.
[19] S. Kumar and A. Mallik, "Covid-19 detection from chest x-rays using trained output based transfer learning approach," *Neural Process Lett*, vol. 55, no. 3, pp. 2405–2428, 2023.
[20] M. Neshat, M. Ahmed, H. Askari, M. Thilakaratne, and S. Mirjalili, "Hybrid Inception Architecture with Residual Connection: Fine-tuned Inception-ResNet Deep Learning Model for Lung Inflammation Diagnosis from Chest Radiographs," *Procedia Comput Sci*, vol. 235, pp. 1841–1850, 2024.
[21] C. Do and L. Vu, "An approach for recognizing COVID-19 cases using convolutional neural networks applied to CT scan images," in *Applications of Digital Image Processing XLIII*, SPIE, 2020, pp. 719–727.
[22] M. HeK and Q. RenS, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.