

## PARKINSONS DISEASE DETECTION USING INCEPTION AND X-CEPTION WITH ATTENTION MECHANISM

Eka Rahma Agustina<sup>1\*</sup>; Hendra Marcos<sup>1</sup>

Informatics<sup>1</sup>

Universitas Amikom Purwokerto, Purwokerto, Indonesia<sup>1</sup>

www.amikompurwokerto.ac.id<sup>1</sup>

ekarahmaagustina284@gmail.com\*; hendra.marcos@amikompurwokerto.ac.id

(\*) Corresponding Author

(Responsible for the Quality of Paper Content)



The creation is distributed under the Creative Commons Attribution-Non Commercial 4.0 International License.

**Abstract**—Parkinson's disease is one of the global health challenges that requires early detection to slow the progression of symptoms. This study proposes an automatic detection system based on deep learning using the InceptionV3 and Xception architectures combined with a multi-head awareness mechanism. The dataset used consists of 72 handwritten spiral images, comprehensively distributed between the Healthy and Parkinson's categories. The process includes preprocessing in the form of normalization and image resizing, as well as model training using the Adam algorithm and the binary cross-entropy loss function. The results show that the model is able to classify both categories with high accuracy. The use of the attention mechanism provides a performance increase of 4.2% on InceptionV3 and 3.1% on Xception compared to the version without attention. In data testing, the InceptionV3 model with attention achieved 100% accuracy, 100% precision, 100% recall, and 100% F1-score. Meanwhile, the Xception model with attention achieved 88% accuracy, 90% precision, 88% recall, and 87% F1-score. The attention mechanism also helps the model in capturing important features such as vibration and irregularity of the spiral pattern. This research makes an important contribution to the development of an artificial intelligence-based automated early diagnosis system to detect Parkinson's disease more accurately and responsively.

**Keywords:** attention mechanism, deep learning, inceptionv3 and Xception, parkinson's disease detection, transfer learning.

**Intisari**— Penyakit Parkinson merupakan salah satu tantangan kesehatan global yang memerlukan deteksi dini untuk memperlambat perkembangan gejala. Penelitian ini mengusulkan sistem deteksi otomatis berbasis deep learning dengan menggunakan arsitektur InceptionV3 dan Xception yang dipadukan dengan mekanisme multi-head attention. Dataset yang digunakan terdiri dari 72 gambar spiral tulisan tangan, terbagi secara merata antara kategori Sehat dan Parkinson. Proses melibatkan pra-proses berupa normalisasi dan perubahan ukuran gambar, serta pelatihan model menggunakan algoritma Adam dan fungsi loss binary cross-entropy. Hasil penelitian menunjukkan bahwa model mampu mengklasifikasikan kedua kategori dengan akurasi tinggi. Penggunaan mekanisme attention memberikan peningkatan performa sebesar 4,2% pada InceptionV3 dan 3,1% pada Xception dibanding versi tanpa attention. Pada data uji, model InceptionV3 dengan attention mencapai akurasi 100%, presisi 100%, recall 100%, dan F1-score 100%. Sementara itu, model Xception dengan attention memperoleh akurasi 88%, presisi 90%, recall 88%, dan F1-score 87%. Mekanisme attention juga membantu model dalam menangkap fitur penting seperti tremor dan ketidakraturan pola spiral. Penelitian ini memberikan kontribusi penting dalam pengembangan sistem diagnosis dini otomatis berbasis kecerdasan buatan untuk mendeteksi penyakit Parkinson secara lebih akurat dan responsif.

**Kata Kunci:** mekanisme atensi, pembelajaran mendalam, inceptionv3 dan Xception, deteksi penyakit parkinson, pembelajaran transfer.



## INTRODUCTION

Parkinson's disease is a neurodegenerative condition that gradually deteriorates the quality of life of patients [1]. Until 2022, Parkinson's disease (PD) has attacked around 90 million people in the world [2]. The estimated global population of people with Parkinson's disease doubled from 1990 to 2016, from 2.5 million to 6.1 million [3]. This disease is characterized by decreased movement control, tremors, muscle strength, and slowing body activity [4]. This disease is one of the global health challenges. Early detection is critical, considering the increasing number of sufferers, and is expected to slow the rate of this disease.

Decision-making in this cutting-edge era can be done more quickly, efficiently, and accurately thanks to computers for information processing, such as AI technology [5]. In recent years, artificial intelligence technology has opened up many new opportunities, including diagnosing various diseases, including Parkinson's. Several analyses have shown that PD sufferers have shorter maximum phonation times and decreased susceptibility [6]. Deep learning is a branch of artificial neural network (ANN)-based AI that allows doctors to automatically assist in the diagnosis process [7].

Convolutional Neural Networks (CNN), which are popular in analyzing complex image data, have architectures such as Inception and Xception. InceptionV3 uses a multi-branch convolution approach that allows for feature extraction at multiple scales in parallel, making it efficient in capturing spatial patterns from medical images [8]. Meanwhile, Xception is a development of the Inception architecture that replaces standard convolution with depthwise separable convolution, resulting in a lighter model but still able to extract deep features efficiently.

Both models are able to extract relevant features in images, but still have limitations in identifying specific areas that are important, such as in the process of diagnosing diseases. With a unique structure that supports multi-scale analysis processes on medical images, both models are designed to recognize certain patterns that are relevant in early detection of Parkinson's disease [9].

However, to improve the model's ability to detect Parkinson's more precisely, an additional mechanism is needed that is able to direct the model's focus to important areas in the image related to the disease. Attention mechanism is a technique that gives more weight to relevant parts of the image. In this study, multi-head attention was

used, namely an architecture that divides the model's focus into several parts in parallel, so that it is able to recognize visual patterns more accurately [10]. The model will be able to concentrate more on specific and relevant aspects of the diagnosis process through the attention mechanism, such as tremors and movement disorders [11]. It is expected that by implementing this mechanism, it can produce more accurate and responsive detection.

Previous research has been conducted by Ahmad Irwansyah and colleagues with the title Application of Expert System Technology with Bayes Theorem Method for Early Detection of Parkinson's Disease [12]. Previous studies used probability rules to determine the level of certainty of diagnosis. As an update, this study proposes a deep learning approach based on InceptionV3 and Xception architectures that is able to recognize complex visual patterns in spiral images. With the support of multi-head attention mechanism, the model can focus attention on important areas of the image, thereby improving the classification accuracy and detecting tremor features that are difficult to recognize by statistical methods.

The selection of the method in this study is based on the effectiveness of the deep learning approach in analyzing medical images. The CNN model with inception and X-caption architecture was chosen because of its ability to extract relevant features from medical images, while the attention mechanism was added to increase the model's focus on more significant areas in Parkinson's diagnosis [13].

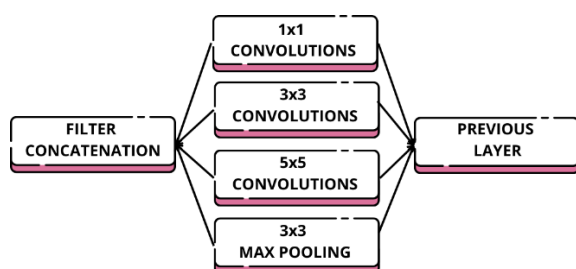
The selection of algorithms in this study is based on the effectiveness of the deep learning analysis model in medical images. The InceptionV3 and Xception architectures were chosen because of their ability to extract complex features from patient spiral images, while the attention mechanism was applied to increase the model's focus on relevant areas in Parkinson's diagnosis. Compared to probability-based approaches such as Bayes' Theorem, the CNN method is superior in recognizing complex visual patterns, so it is expected to improve the accuracy of detecting this disease.

This study aims to develop a Parkinson's disease detection model by implementing deep learning, combining invention and exception, and equipped with an attention algorithm. By implementing technology, it is expected that the resulting system will be able to provide a faster and more accurate diagnosis.

## MATERIALS AND METHODS

A deep learning algorithm called CNN (Convolutional Neural Network) is useful for spatial data image processing [14]. This algorithm is very popular and is often used for image pattern recognition processing, object detection, and facial recognition [15]. CNN uses image characteristic readings as input and is re-described as a tensor using the relationship between each neuron and the correlated perceptron [16].

The InceptionV3 is a convolutional neural network architecture that was first presented by Google researchers in 2014. It was designed to solve efficiency problems, particularly in deep learning image classification models [17]. It is made up of several layers: convolutional layers, pooling layers, and additional classifiers. The inception model consists of 4 parallel layers of different sizes (1x1, 3x3, 5x5), as shown in Figure 1.



Source: (maddala, 2024)

Figure 1. Model Inception

The X-ception architecture differs from inception in that the initiation module is replaced by convolution and can be separated in depth. X-ception has the same parameters as inception V3, namely 36 convolution layers [18].

In this study, the dataset used consists of handwritten spiral images divided into two categories: healthy and parkinson, with each category consisting of 36 images, so the total dataset used is 72 images. This data is taken from the Kaggle Parkinson's Drawings source and organized into two separate folders for each category. <https://www.kaggle.com/datasets/kmader/parkinsons-drawings> This is the source link of the dataset. After the data is collected, a preprocessing process is carried out to prepare the images so that the deep learning model can use them. The images are resized to 224x224 pixels and normalized using ImageDataGenerator to adjust the input format required by the model.

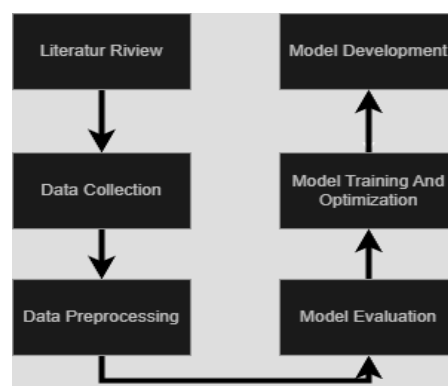
Two transfer learning-based models are used, namely InceptionV3 and Xception. Both models are pre-trained with the ImageNet dataset

as the central feature extractor. InceptionV3 consists of several parallel layers with different filter sizes (1x1, 3x3, 5x5), which allows the model to capture various spatial features in the image. On the other hand, To enable the model to absorb more complex spatial patterns, Xception replaces the initial InceptionV3 module with depthwise separable convolutions. Both architectures are used as the basis for feature extraction, which is then equipped with additional layers to improve accuracy and prevent overfitting.

One of the essential additions in the model is the attention mechanism, specifically multi-head attention, which is used to improve the model's ability to focus attention on more relevant parts of the image, such as patterns or details that distinguish between healthy and Parkinson's conditions. This allows the model to emphasize more important features, which improves the classification quality and the model's ability to identify relationships between image elements [18]. In addition, Gaussian noise and dropout layers are also added to reduce overfitting, keep the model from being too dependent on the training data, and enhance the model's accessibility to test data.

Model training was performed using Adam optimization with binary\_crossentropy loss function for two-class classification, namely Healthy and Parkinson. The model was trained on data divided into three subsets: train, validation, and test. The training process was performed using Google Colab as a programming environment with Python language, which allows automatic evaluation by using the sci-kit-learn library to calculate evaluation metrics such as accuracy, precision, recall, and F1 scores. The results of this evaluation are used to measure the model's performance in detecting Parkinson's disease from processed spiral images.

This research was conducted with steps as shown in the following image:



Source: (Research results, 2024)

Figure 2. Research Flow

The research flow is shown in Figure 2 a structured methodology consisting of the following steps:

1. Literatur Riwiew: Conducting an in-depth study of previous research related to Parkinson's disease detection using spiral images.
2. Data Collection: Acquiring handwritten spiral images from the Kaggle Parkinson's Drawings dataset and categorizing them into two groups: Healthy and Parkinson's.
3. Data Preprocessing: Preparing the images by resizing them to 224x224 pixels, normalizing them using ImageDataGenerator, and performing augmentation techniques to enhance model generalization.
4. Model Development: Implementing two deep learning architectures (InceptionV3 and Xception) as feature extractors and integrating an attention mechanism to enhance focus on relevant image regions.
5. Model Training And Optimization: The dataset is split into 80% for training, 10% for validation, and 10% for testing utilizing a stratified division (random\_state=42). The model employs the Adam optimizer with a learning rate of  $1 \times 10^{-4}$  and utilizes the binary\_crossentropy loss function for training. Training occurs over 10 epochs using a batch size of 16. Regularization methods like Gaussian noise ( $\sigma=0.25$ ) and dropout layers (rate=0.25) are utilized to avoid overfitting. Early stopping with a patience of 5 is implemented to retain the best model throughout training.
6. Model Evaluation: Measuring model performance based on accuracy, precision, recall, and F1-score using the scikit-learn library. To strengthen result reliability, multiple evaluation runs are performed using k-fold cross-validation.

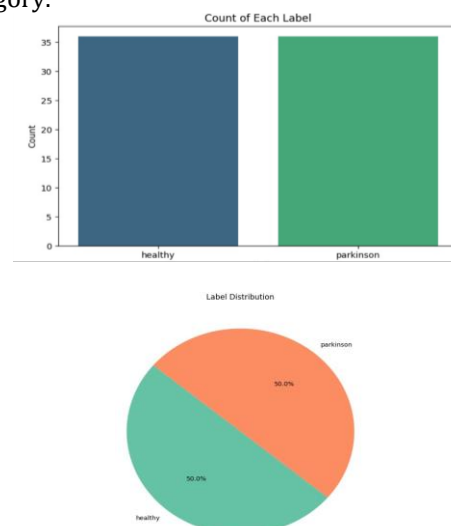
The Result Analysis And Interpretation: Analyzing the model's classification results to assess its effectiveness in distinguishing Parkinson's patients from healthy individuals.

## RESULTS AND DISCUSSION

This study will present the results of Parkinson's disease detection through spiral image analysis. Two transfer learning-based deep learning models, namely InceptionV3 and Xception, are used in this study, which are combined with attention mechanisms to improve the model's ability to recognize visual patterns that distinguish healthy

individuals from Parkinson's patients. The dataset is divided into training (80%), validation (10%), and testing (10%) using a stratified split. Evaluation metrics include accuracy, precision, recall, and F1-score.

The dataset used in this study consists of handwritten spiral images, which are divided into two categories: Healthy and Parkinson (Parkinson sufferers). This dataset is taken from a file directory containing images in PNG and JPG formats. Figure 3 provides a sample visualization of the number of datasets for the Healthy category, which has 36 images, while the Parkinson category has 36 images. The total data used is 72 images. The label distribution shows that the dataset is balanced, visualized in bar graphs and pie charts, illustrating the comparison of the number of data in each category.



Source: (Research results, 2024)

Figure 3. Dataset Visualization

After that, a visualization of sample images from the dataset used for model training was performed. The results show the first five images for each category (healthy and Parkinson) in the form of subplots consisting of two rows and five columns. The first row displays healthy images, while the second row displays Parkinson images. Figure 4 below is a visualization of sample datasets from two categories: healthy and Parkinson.



Source: (Research results, 2024)

Figure 4. Visualization of sample images



The models used involve two transfer learning-based architectures, InceptionV3 and Xception, equipped with attention mechanisms to improve classification performance. In the InceptionV3 architecture, pre-trained base layers trained on the ImageNet dataset are used as feature extractors, with these layers frozen during training to prevent weight updates. After feature extraction, a multi-head attention mechanism is applied to capture more complex spatial patterns. Next, Gaussian noise layers, global average pooling, dense layers with 512 neurons, and batch normalization and dropout are used to prevent overfitting. In both architectures, pre-trained ImageNet models serve as feature extractors, followed by frozen base layers, attention mechanisms, Gaussian noise layers, global average pooling, dense layers, batch normalization, and dropout to prevent overfitting. Models are trained using Adam optimizer (learning rate = 0.0001) with binary cross-entropy loss for 10 epochs. Training convergence is monitored with early stopping (patience = 5).

```
Epoch 8/10
4/4 ----- 1s 123ms/step - accuracy: 0.9888 - loss:
0.0700 - val_accuracy: 0.8571 - val_loss: 0.3443
Epoch 9/10
4/4 ----- 1s 137ms/step - accuracy: 1.0000 - loss:
0.0243 - val_accuracy: 0.8571 - val_loss: 0.3370
Epoch 10/10
4/4 ----- 1s 391ms/step - accuracy: 1.0000 - loss:
0.0195 - val_accuracy: 0.8571 - val_loss: 0.1380
```

Source: (Research results, 2024)

Figure 5. Data preprocessing

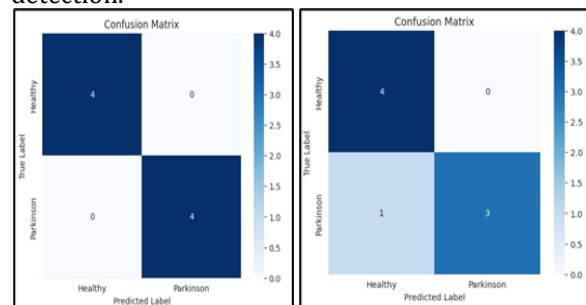
Experimental results show that both models perform very well despite differences in some evaluation metrics. InceptionV3 achieves a training accuracy of 1.00 with a very low training loss (0.0195), but on the validation data, the accuracy drops to 0.8571 with a loss of 0.1380. This indicates the possibility of overfitting, even though the model recognizes patterns in the training data very well. Visualization of accuracy and loss per epoch metrics also shows stable convergence, achieving high accuracy in a relatively short training time. The attention mechanism in InceptionV3 plays a key role in helping the model recognize complex patterns in the spiral data, distinguishing images from healthy individuals and Parkinson's patients.

	precision	recall	f1-score	support
healthy	1.00	1.00	1.00	4
parkinson	1.00	1.00	1.00	4
accuracy			1.00	8
macro avg	1.00	1.00	1.00	8
weighted avg	1.00	1.00	1.00	8

Source: (Research results, 2024)

Figure 6. Evaluation results

The evaluation results show that the developed model has excellent performance in classifying spiral images between healthy individuals and Parkinson's patients, with perfect precision, recall, and F1-score (1.00). The main factor affecting model performance is dataset quality, where data balance and the presence of noise in the image, such as variations in hand pressure when drawing, can affect accuracy [19]. The attention mechanism plays an important role in capturing more complex visual patterns, improving the model's ability to distinguish between the two categories. However, the model still depends on the quality of the provided image, so exploring data augmentation methods and integrating temporal features from the spiral drawing process can be an important step to improve the generalization and reliability of the model in real-world Parkinson's detection.



Source: (Research results, 2024)

Figure 7. Confusion matrix comparison

Figure 7 shows the confusion matrix, and classification reports show that InceptionV3 achieved perfect scores (precision, recall, F1 = 1.00) on the test set, while Xception scored slightly lower (macro average F1 = 0.87). This suggests that InceptionV3 better captures image-level features, likely due to its filter diversity and aggressive multi-scale feature extraction. Meanwhile, Xception's depthwise separable convolutions may reduce redundancy but also limit initial learning capacity with small datasets.

To benchmark the proposed method, we compared it to prior research using conventional CNNs, SVMs, and Naive Bayes classifiers [20]. Traditional machine learning models often require hand-crafted features and achieve lower accuracy on similar datasets (typically <85%). Nevertheless, this study has limitations. The dataset is relatively small (n = 72) and lacks diversity in handwriting style, stroke pressure, and drawing speed. Additionally, temporal features from the drawing process (e.g., direction, speed) are not captured, which could significantly improve the model's ability to distinguish motor control patterns. Future

studies should explore the use of dynamic datasets and multi-modal inputs, such as pen pressure and time-series data, to enhance model generalization.

Finally, k-fold cross-validation and external dataset testing demonstrate that the models are reliable and robust, showing consistent accuracy without major overfitting. This confirms the potential of using InceptionV3 and Xception with attention mechanisms for early-stage Parkinson's detection.

### CONCLUSION

Based on the results obtained, the deep learning models based on InceptionV3 and Xception with attention mechanism proved effective in detecting Parkinson's disease through spiral image analysis. Both models achieved 100% accuracy in classifying images into Healthy and Parkinson's. The multi-head attention mechanism improved the performance by focusing on important elements in the image, even amidst noise. The risk of overfitting was also reduced by layers such as Gaussian noise and dropout, improving the generalization ability. Although the results are promising, limitations such as the small dataset size need to be considered. Further research is recommended to use larger and more varied datasets and advanced image preprocessing techniques. Testing with real-world data will help evaluate the potential implementation of this system in clinical practice. Overall, the combination of transfer learning and attention mechanism improves the accuracy of Parkinson's detection and has the potential for handwriting-based early screening.

### REFERENCE

- [1] J. Zhang, "Mining imaging and clinical data with machine learning approaches for the diagnosis and early detection of Parkinson's disease," *npj Park. Dis.*, vol. 8, no. 1, p. 13, 2022, doi: 10.1038/s41531-021-00266-8.
- [2] J. Mei, C. Desrosiers, and J. Frasnelli, "Machine learning for the diagnosis of Parkinson's disease: a review of literature," *Front. Aging Neurosci.*, vol. 13, p. 633752, 2021, doi: 10.3389/fnagi.2021.633752.
- [3] H. Marcos, B. Warsito, and A. Wibowo, "Feature Selection and Classification Algorithm Approaches for Parkinson's Disease Detection," *Proc. - 2021 IEEE 5th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. Appl. Data Sci. Artif. Intell. Technol. Glob. Challenges Dur. Pandemic Era, ICITISEE 2021*, pp. 95-99, 2021, doi: 10.1109/ICITISEE53823.2021.9655774.
- [4] R. A. Fadillah and I. A. Nachrowi, "Implementasi Sistem Pakar Untuk Mendiagnosa Penyakit Paru-Paru Menggunakan Metode Forward Chaining," *J. Inform. Multi*, vol. 1, no. 4, pp. 524-528, 2023.
- [5] C. Quan, K. Ren, and Z. Luo, "A Deep Learning Based Method for Parkinson's Disease Detection Using Dynamic Features of Speech," *IEEE Access*, vol. 9, pp. 10239-10252, 2021, doi: 10.1109/ACCESS.2021.3051432.
- [6] S. N. Nugraha, R. Pebrianto, and E. Fitri, "Penerapan Deep Learning Pada Klasifikasi Tanaman Paprika Berdasarkan Citra Daun Menggunakan Metode CNN," *Inf. Syst. Educ. Prof. J. Inf. Syst.*, vol. 8, no. 2, pp. 133-142, 2023, doi: 10.51211/isbi.v8i2.2671.
- [7] R. Kurniawan, P. B. Wintoro, Y. Mulyani, and M. Komarudin, "Implementasi Arsitektur Xception Pada Model Machine Learning Klasifikasi Sampah Anorganik," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 2, pp. 233-236, 2023, doi: 10.23960/jitet.v11i2.3034.
- [8] C. Wibisono and S. Budi, "Form Recognition dan Character Mapping Menggunakan Image Segmentation dan Optical Character Recognition," *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 1, pp. 154-164, 2021, doi: 10.28932/jutisi.v7i1.3340.
- [9] R. Rina, T. Hidayat, and D. Uki Eka Saputri, "Analisis Percepatan Pemulihan Ekonomi Indonesia Pasca Pandemi Dengan Big Data Dan Deep Learning," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 3, pp. 3244-3252, 2024, doi: 10.36040/jati.v8i3.9420.
- [10] A. Taufiqur and D. H. Fudholi, "Image Captioning For Mobile Application," *Automata*, vol. 3, no. 2, 2022.
- [11] R. A. Kurniawan, A. Sunyoto, and A. Nasiri, "Pengaruh Arsitektur Convolutional Neural Network Untuk Klasifikasi Penyakit Daun Tomat," *Tek. Teknol. Inf. dan Multimed.*, vol. 4, no. 2, pp. 126-131, 2023, doi: 10.46764/teknimedia.v4i2.111.
- [12] F. Ridho and M. Syahputra, "Perancangan Sistem Informasi atau Aplikasi Monitoring Absensi Karyawan pada PT. Socfindo Menggunakan QR Code Berbasis Web," *SIKOM J. Sist. Inf. Komput.*, vol. 1, no. 1, pp. 37-50, 2024.
- [13] M. Boulaabi, T. B. A. Gader, A. K. Echi, and S. Mbarek, "Advanced Segmentation of Diabetic Retinopathy Lesions Using

- DeepLabv3+,” in *2024 IEEE/ACS 21st International Conference on Computer Systems and Applications (AICCSA)*, IEEE, pp. 1–8, 2024, doi: 10.1109/AICCSA63423.2024.10912539,
- [14] T. S. Winanto, C. Rozikin, and A. Jamaludin, “Analisa Performa Arsitektur Transfer Learning Untuk Mengidentifikasi Penyakit Daun Pada Tanaman Pangan,” *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 68–81, 2023, doi: 10.30871/jaic.v7i1.5991.
- [15] E. D. Sefrila, B. Rahmat, and A. N. Sihananto, “Implementasi Arsitektur Inception V3 Dengan Optimasi Adam, SGD dan RMSP Pada Klasifikasi Penyakit Malaria,” *Bridg. J. Publ. Sist. Inf. dan Telekomun.*, vol. 2, no. 2, pp. 69–84, 2024, doi: 10.62951/bridge.v2i2.62.
- [16] H. Maddala, “Building an Image Classifier Model Using Pretrained InceptionV3,” *juni 3*, 2024.
- [17] T. A. Tresi and M. Salim, “Analisa Perbandingan Inception dan Xception Berbasis CNN untuk Klasifikasi Wajah Hewan,” *J. Tek. Inform. dan Desain Komun. Vis.*, vol. 3, no. 1, pp. 34–41, 2024.
- [18] A. S. Atmaja, N. F. Muzakki, and Z. D. Oktavian, “Peramalan Tinggi Muka Air Menggunakan Long-Short Term Memory dengan Mekanisme Multi-Head Attention,” in *Seminar Nasional Official Statistics*, 2024, pp. 625–636.
- [19] C. E. Sukmawati, A. R. Pratama, and N. Nurjanah, “Algoritma Naïve Bayes, Random Forest Dan Support Vector Machine Untuk Klasifikasi Penyakit Sapi,” *J. Ilmu Komput. dan Teknol. Inf.*, vol. 9, no. 1, pp. 44–50, 2024.
- [20] X. Wang *et al.*, “A recognition method of ancient architectures based on the improved inception v3 model,” *Symmetry (Basel)*, vol. 14, no. 12, p. 2679, 2022, doi: 10.3390/sym14122679.