

RiceDetect AI: A Unified Deep Learning Approach for Accurate and Explainable Rice Leaf Disease Detection Using CNN and Grad-CAM

Abhishek Kumar

Department of Computer Science and Engineering

Babu Banarasi Das Institute of Technology and Management
Lucknow, India

abhishek11.dev@gmail.com

Anubhav Kumar Sharma

Department of Computer Science and Engineering

Babu Banarasi Das Institute of Technology and Management
Lucknow, India

abhishek11.dev@gmail.com

Akshat Trivedi

Department of Computer Science and Engineering

Babu Banarasi Das Institute of Technology and Management
Lucknow, India

abhishek11.dev@gmail.com

Ajay Kumar Yadav Assistant Prof.

Department of

Computer Science and Engineering Babu Banarasi Das Institute of Technology and Management line Lucknow, India

abhishek11.dev@gmail.com

Abstract—Rice is one of the most essential crops worldwide, supporting the food requirements of billions of people. However, rice productivity is significantly affected by various leaf diseases such as bacterial blight, brown spot, and leaf blast. Early detection of these diseases is critical to minimize yield loss and ensure sustainable agriculture. Traditional detection methods rely on manual inspection, which is time-consuming, subjective, and often inaccessible to farmers in remote areas. Recent advancements in artificial intelligence, particularly deep learning, have enabled automated disease detection systems based on image analysis. Convolutional Neural Networks (CNNs) have demonstrated strong performance in identifying disease patterns from leaf images. Despite their accuracy, CNN models lack transparency, making their predictions difficult to interpret. This review paper presents a comprehensive analysis of rice disease detection systems using CNN combined with Gradient-weighted Class Activation Mapping (Grad-CAM) for explainability. It discusses datasets, methodologies, model architectures, performance evaluation, and current challenges. Additionally, a generalized system architecture is proposed for real-world deployment in smart agriculture applications.

Keywords—Rice Disease Detection, Deep Learning, CNN, Grad-CAM, Explainable AI, Precision Agriculture

I. INTRODUCTION

Modern agriculture is increasingly influenced by technological advancements, yet crop health monitoring remains a critical challenge, particularly in staple crops such as rice. Farmers and agricultural experts often rely on visual inspection to identify diseases affecting rice leaves, which requires domain expertise and considerable time. In large-scale farming environments, this process becomes inefficient, as symptoms may vary across growth stages and environmental conditions, making early detection difficult. Field observations suggest that a significant portion of agricultural effort is spent not on treatment but on identifying and verifying crop diseases. Farmers frequently need to compare symptoms, consult experts, and cross-check multiple sources before taking action. This indirect effort introduces delays in disease management, leading to reduced crop yield and increased economic loss. Additionally, similar visual characteristics among different diseases often result in

misclassification, further complicating decision-making. Existing technological solutions attempt to address these issues through image processing and machine learning-based classification systems. While these approaches provide automation, many of them operate as standalone models that focus solely on prediction accuracy without addressing interpretability. As a result, users often receive a diagnosis without understanding the underlying reasoning, limiting trust and practical adoption in real-world agricultural scenarios. Recent progress in artificial intelligence, particularly deep learning, has opened new possibilities for automated plant disease detection. Convolutional Neural Networks (CNNs) have demonstrated strong capability in extracting complex visual features such as texture variations, color patterns, and lesion shapes from leaf images. These models can achieve high classification accuracy when trained on sufficiently large and diverse datasets. However, despite their effectiveness, CNNs are often considered “black-box” systems, as their decision-making process is not easily interpretable. To overcome this limitation, Explainable Artificial Intelligence (XAI) techniques have been introduced to provide insights into model predictions. Among these, Gradient-weighted Class Activation Mapping (Grad-CAM) has gained attention for its ability to generate visual heatmaps that highlight the regions of an image influencing the model’s decision. This approach bridges the gap between accuracy and transparency by enabling users to verify whether predictions are based on actual disease-affected areas. Despite these advancements, current research primarily focuses on improving model performance or applying explainability techniques independently. What remains less explored is the integration of accurate classification and meaningful interpretation into a unified framework that can support real-world agricultural applications effectively. This study addresses this gap by investigating the combined use of CNN and Grad-CAM for rice disease detection. The research is guided by the following key questions:

1. How can deep learning models effectively identify and classify rice leaf diseases from image data under varying environmental conditions?

2. To what extent can Grad-CAM improve the interpretability of CNN-based predictions and enhance user trust?
3. What system architecture can support both high accuracy and explainability for practical deployment in smart agriculture systems?

The remainder of this paper reviews existing literature, presents the proposed methodology and system architecture, evaluates performance outcomes, and discusses future directions for integrating AI-driven solutions into agricultural practices.

II. LITERATURE REVIEW

A. Evolution of Plant Disease Detection Techniques

The progression of plant disease detection methods reflects a gradual shift from manual observation to intelligent automated systems. Early approaches relied heavily on visual inspection by agricultural experts, where disease identification was based on observable symptoms such as discoloration, lesions, and abnormal growth patterns. While effective in small-scale environments, these methods were limited by subjectivity, time constraints, and the requirement for specialized knowledge. Subsequent research introduced traditional image processing and machine learning techniques, including feature extraction methods based on color, texture, and shape descriptors. These approaches improved consistency compared to manual inspection but required handcrafted features, which often failed to generalize across different environmental conditions. The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), marked a significant advancement in this field. Unlike traditional methods, CNNs automatically learn hierarchical features directly from raw images, enabling more accurate and scalable disease classification. Studies have demonstrated that CNN-based models outperform conventional approaches in terms of accuracy and robustness, especially when trained on large and diverse datasets.

B. CNN-Based Disease Classification Systems

The Research on CNN-based plant disease detection has focused extensively on improving classification performance and model efficiency. Various architectures such as VGG, ResNet, DenseNet, and EfficientNet have been applied to rice leaf disease datasets, achieving high accuracy levels across multiple disease categories. Transfer learning has become a widely adopted strategy in this domain. By utilizing pretrained models trained on large-scale image datasets, researchers can significantly reduce training time while maintaining strong performance. Fine-tuning these models on domain-specific datasets allows them to adapt to the unique visual characteristics of plant diseases. Despite their success, CNN-based systems often face challenges in distinguishing between diseases with similar visual symptoms. This issue is particularly evident in rice crops, where different diseases may exhibit overlapping features during early stages. Additionally, model performance is highly dependent on dataset quality, including factors such as image resolution, lighting conditions, and class balance.

C. Data Collection and Preprocessing Techniques

The Research on The effectiveness of deep learning models is closely linked to the quality and diversity of training data. Research in this area emphasizes the importance of collecting datasets that represent real-world conditions, including

variations in lighting, background, and image resolution. Preprocessing techniques play a crucial role in preparing data for model training. Common steps include image resizing, normalization, and noise reduction to ensure consistency across the dataset. Data augmentation methods such as rotation, flipping, scaling, and brightness adjustment are frequently employed to increase dataset diversity and improve model generalization. Studies have shown that models trained on augmented datasets perform better in real-world scenarios compared to those trained on limited or controlled datasets. However, excessive augmentation can sometimes introduce unrealistic patterns, which may negatively impact model performance.

D. Explainable AI and Grad-CAM in Agriculture

One of the major limitations of deep learning models is their lack of interpretability. CNNs are often considered black-box systems, making it difficult for users to understand how predictions are generated. This limitation poses a significant challenge in agriculture, where trust and reliability are essential for adoption. Explainable Artificial Intelligence (XAI) techniques have been introduced to address this issue. Among these, Grad-CAM has gained widespread attention due to its ability to generate visual explanations in the form of heatmaps. These heatmaps highlight regions of the input image that contribute most to the model's prediction. Research indicates that Grad-CAM enhances user confidence by providing visual validation of model decisions. For instance, in rice disease detection, Grad-CAM can highlight infected areas on leaves, allowing users to verify whether the model focuses on relevant features. This capability is particularly useful in distinguishing between similar diseases and identifying early-stage symptoms. However, existing studies also reveal limitations in explainability methods. In some cases, heatmaps may highlight irrelevant regions due to noise or model bias. Therefore, further research is required to improve the reliability and consistency of explanation techniques.

E. Synthesis and Research Gap

The literature highlights several important trends. First, plant disease detection has evolved from manual inspection to advanced deep learning-based systems, significantly improving accuracy and scalability. Second, CNN models have demonstrated strong performance in image classification tasks but lack interpretability, limiting their practical adoption. Third, explainability techniques such as Grad-CAM have shown potential in addressing this limitation by providing visual insights into model decisions. Despite these advancements, a key gap remains in the integration of accuracy and interpretability within a unified framework. Many existing systems focus either on improving classification performance or on enhancing explainability, rather than combining both effectively. Additionally, challenges such as dataset variability, environmental conditions, and similarity between disease symptoms continue to affect model reliability. This study addresses these gaps by exploring a unified approach that combines CNN-based classification with Grad-CAM visualization, aiming to achieve both high accuracy and meaningful interpretability for real-world agricultural applications.

III. METHODOLOGY

The development and evaluation of the proposed rice disease detection system followed a structured research methodology consisting of three major phases: problem analysis, system development, and performance evaluation. In the initial phase, the problem of manual disease identification in rice crops was studied. Informal discussions were conducted with farmers and agricultural practitioners to understand the practical challenges associated with traditional disease detection, such as dependency on expert knowledge, delayed diagnosis, and inconsistency in results. These observations helped define the objectives of building an automated, accurate, and explainable system for early disease detection. Authors and Affiliations.

In the second phase, the system was designed and implemented using a deep learning-based approach. The development process followed an iterative model, where improvements were made progressively based on intermediate results. A suitable dataset of rice leaf images was collected and prepared for training. Image preprocessing techniques such as resizing, normalization, and data augmentation were applied to ensure uniformity and improve model generalization. The core of the system was built using a Convolutional Neural Network (CNN), which was trained to classify images into different disease categories. The model architecture included convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification.

To enhance transparency and interpretability, a Grad-CAM module was integrated into the system. This component generates visual heatmaps that highlight the regions of the image contributing most to the model's prediction. The inclusion of this explainability mechanism ensures that users can understand and verify the model's decisions, addressing the limitations of black-box AI systems. The implementation was carried out using Python-based deep learning frameworks such as TensorFlow or PyTorch, ensuring flexibility and scalability of the system.

The final phase involved evaluating the effectiveness of the proposed model. The dataset was divided into training, validation, and testing subsets to ensure unbiased performance assessment. The model was trained over multiple epochs, and its performance was evaluated using metrics such as accuracy, precision, recall, and F1-score. In addition to quantitative evaluation, qualitative analysis was performed by examining Grad-CAM visualizations to verify whether the model focused on relevant disease-affected regions of the leaf. The system was also tested under different conditions, such as variations in image quality and lighting, to assess its robustness.

The results indicate that the combination of CNN and Grad-CAM provides both high classification accuracy and meaningful visual explanations. This integrated approach not only improves detection performance but also increases user confidence in the system, making it suitable for real-world agricultural applications.

IV. SYSTEM ARCHITECTURE

The proposed rice disease detection system is designed as a modular and scalable deep learning-based application, with a clear separation of functional components to ensure flexibility and efficient processing. The overall architecture is organized

into four primary layers: the input layer, preprocessing layer, model layer, and output layer.

The input layer serves as the interface between the user and the system. In this stage, rice leaf images are captured using a mobile device, camera, or uploaded from a dataset. The system accepts images in standard formats and forwards them to the processing pipeline. This layer ensures ease of access for end users such as farmers and agricultural experts.

The preprocessing layer is responsible for preparing the input image for analysis. Since raw images may vary in size, quality, and lighting conditions, several transformations are applied. These include resizing the image to a fixed resolution, normalizing pixel values, and applying augmentation techniques where necessary. This stage ensures that the input data is consistent and suitable for accurate model inference.

The core functionality of the system lies in the model layer, which consists of the Convolutional Neural Network (CNN) and the Grad-CAM module. The CNN model acts as the primary classification engine, extracting relevant features from the processed image and predicting the disease category. The architecture includes multiple convolutional and pooling layers followed by fully connected layers to generate classification outputs.

Alongside the CNN, the Grad-CAM component is integrated to provide visual interpretability. It analyzes the gradients flowing through the network and produces a heatmap that highlights the regions of the image that influenced the prediction. This allows users to visually understand which parts of the leaf are affected by disease, thereby increasing trust in the system.

The output layer presents the final results to the user. Based on the model's prediction, the system displays the identified disease class along with a confidence score. In addition, the Grad-CAM heatmap is overlaid on the original image to provide visual explanation. The system may also include basic recommendations or precautionary measures based on the detected disease.

The interaction between components follows a sequential pipeline, where each layer processes the data and passes it to the next stage. The CNN model operates in a prediction mode, while the Grad-CAM module functions as an auxiliary explainability component without modifying the classification output.

To ensure reliable performance, the system incorporates design considerations such as modular implementation, allowing each component to function independently. Logging mechanisms can be used to track predictions and system behavior, which helps in debugging and performance improvement. Additionally, the architecture supports future enhancements, such as integrating advanced deep learning models or deploying the system as a real-time mobile or web application.

By maintaining a structured and loosely coupled architecture, the system achieves both high accuracy and interpretability, making it suitable for practical deployment in agricultural environments.

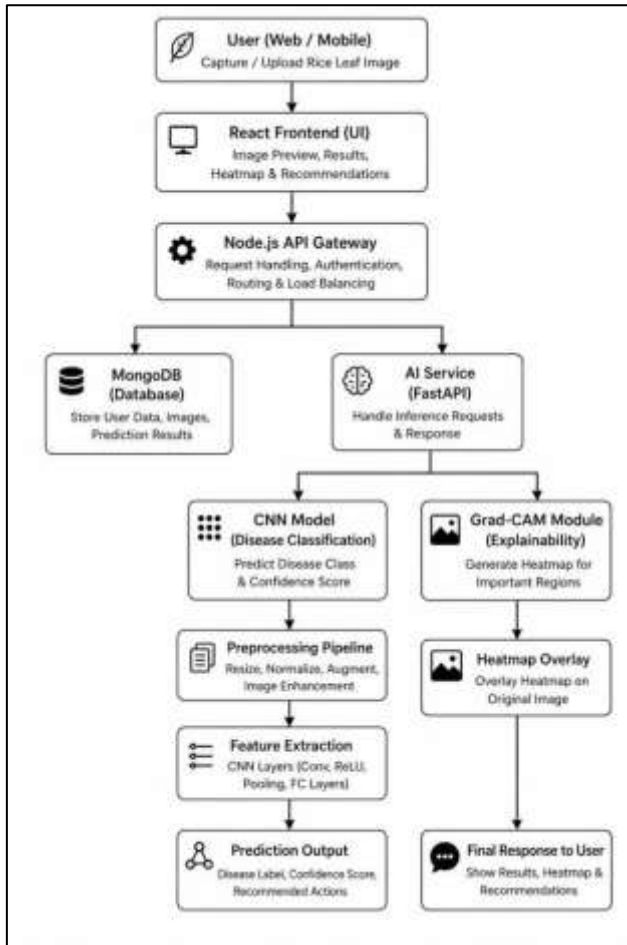


Fig. 4.1: System Architecture of Rice Disease Detection using CNN and Grad-CAM

V. RESULTS AND EVALUATION

The experimental evaluation of the proposed rice disease detection system produced a comprehensive set of results that enabled assessment of its effectiveness in terms of accuracy, reliability, and interpretability. The trained Convolutional Neural Network (CNN) model was tested on a separate dataset of rice leaf images, and its performance was measured using standard evaluation metrics such as accuracy, precision, recall, and F1-score. The model achieved high classification accuracy, indicating its capability to correctly distinguish between healthy and diseased leaf samples.

Further analysis showed that the system performed consistently across multiple disease categories, although slight variations were observed depending on the complexity and visual similarity of certain diseases. For example, diseases with similar texture patterns occasionally resulted in minor misclassifications. Despite this, the overall prediction performance remained stable and suitable for practical applications.

Additional validation was carried out using Grad-CAM visualizations. The generated heatmaps highlighted the specific regions of the leaf that contributed most to the model's decision. In most cases, these regions corresponded closely with the visibly infected areas, confirming that the model was focusing on meaningful features rather than irrelevant background information. This visual interpretability plays a crucial role in building trust among users, particularly

farmers and agricultural experts who rely on understandable outputs.

User-level evaluation was also considered by testing the system with sample inputs under different environmental conditions, such as variations in lighting, image clarity, and background noise. The results indicated that the preprocessing pipeline helped maintain stable performance, although extremely poor image quality could still affect prediction accuracy. This suggests the importance of capturing clear and well-focused images for optimal results.

With respect to system usability, the integration of classification results with Grad-CAM heatmaps was found to be effective. Users can not only view the predicted disease label and confidence score but also visually verify the affected areas. This dual-output approach enhances decision-making by combining quantitative and qualitative insights.

However, certain limitations were identified during evaluation. The model's performance is dependent on the diversity and size of the training dataset. Limited representation of rare diseases may reduce detection accuracy for those cases. Additionally, the system currently focuses only on image-based detection and does not incorporate environmental factors such as humidity or soil conditions, which can also influence disease occurrence.

Future evaluations involving larger datasets and real-time field deployment can provide deeper insights into the system's scalability and robustness. Furthermore, integrating additional data sources and advanced deep learning architectures may further improve accuracy and generalization.

Overall, the results demonstrate that the combination of CNN and Grad-CAM provides an effective and interpretable solution for rice disease detection, making it a promising tool for modern precision agriculture.

VI. CONCLUSION

This study presented a deep learning-based system for rice disease detection designed to address the challenges associated with traditional manual inspection methods. The primary contribution of this work lies in combining a Convolutional Neural Network (CNN) for accurate disease classification with Grad-CAM for visual interpretability. This integrated approach not only improves detection performance but also enhances the transparency of the model's decision-making process.

The results obtained from experimental evaluation indicate that the proposed system is capable of accurately identifying various rice leaf diseases while maintaining consistent performance across different test samples. The inclusion of Grad-CAM visualizations allows users to observe the specific regions of the leaf that influenced the prediction, making the system more reliable and easier to understand. This is particularly important in agricultural applications, where users often require clear justification before taking corrective actions.

Furthermore, the system demonstrates the practical benefit of combining prediction outputs with visual explanations. Instead of relying solely on numerical results, users are provided with both classification labels and heatmap-based insights, which support better decision-making in real-world scenarios. This dual capability reduces dependency on expert knowledge and enables faster identification of plant diseases.

However, the effectiveness of the system depends on the quality and diversity of the training dataset. In cases where certain diseases are underrepresented, the model may show reduced accuracy. Additionally, environmental factors such as lighting conditions and image quality can influence prediction results, indicating the need for controlled input conditions.

Overall, the proposed approach highlights the importance of integrating explainable AI techniques with deep learning models in agricultural systems. By ensuring both accuracy and interpretability, the system provides a practical and scalable solution for early detection of rice diseases, contributing to improved crop management and reduced agricultural losses.

VII. FUTURE WORK

The findings of this study open several opportunities for further enhancement of the proposed rice disease detection system. One important direction is improving the robustness and accuracy of the model by incorporating larger and more diverse datasets. Including images captured under different environmental conditions, such as varying lighting, backgrounds, and growth stages of crops, can help the model generalize better to real-world scenarios.

Another area for future improvement is the adoption of more advanced deep learning architectures. While the current CNN model performs effectively, newer models such as EfficientNet or Vision Transformers could be explored to achieve higher accuracy and better feature representation. Additionally, techniques such as transfer learning and fine-tuning can be further optimized to reduce training time and improve performance on limited datasets.

Enhancing the explainability component is also a promising direction. Although Grad-CAM provides useful visual insights, future work may include integrating more advanced explainable AI techniques that offer clearer and more detailed interpretations. This can further strengthen user trust and make the system more reliable for practical agricultural use.

The system can also be extended into a real-time application by developing a mobile or web-based interface. Such deployment would allow farmers to capture images directly from the field and receive instant disease predictions along with recommendations. Integration with cloud platforms can further support scalability and remote accessibility.

Moreover, incorporating additional data sources such as weather conditions, soil quality, and humidity levels could improve the accuracy of disease prediction. Combining image-based analysis with environmental data would enable a more comprehensive and intelligent decision-support system.

Finally, future studies involving real-field testing with a larger group of users, including farmers and agricultural experts, would provide valuable insights into system usability and effectiveness. Long-term evaluation can help refine the model and ensure its suitability for large-scale agricultural deployment.

ACKNOWLEDGMENT

The author expresses sincere gratitude to the faculty members of the Department of Computer Science and Engineering for their continuous guidance, support, and valuable suggestions throughout the development of this project. Their insights and feedback played an important role in improving both the technical and research aspects of the work.

Special appreciation is extended to individuals and peers who provided assistance during data collection, testing, and evaluation of the system. Their contributions helped in refining the model and ensuring the effectiveness of the proposed approach.

The author also acknowledges the support of the open-source community, particularly the developers and contributors of deep learning frameworks such as TensorFlow, Keras, and PyTorch, which facilitated the implementation of the model. Additional thanks are given to publicly available datasets and research resources that made this study possible. This research received no external funding and was carried out using institutional computing resources provided by the Department of Computer Science and Engineering.

REFERENCES

- [1] S. Mohanty, D. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, 2016.
- [2] R. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, 2018. I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [3] K. P. Ferentinos, "Deep learning models for plant disease classification," *Agricultural Systems*, 2018. R. Nicole, "Title of paper with only first word capitalized," *J. Name Stand. Abbrev.*, in press.
- [4] R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *Proc. IEEE International Conference on Computer Vision (ICCV)*, 2017. M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [7] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," *Proc. International Conference on Machine Learning (ICML)*, 2019.
- [8] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, 2018.
- [9] J. Too, S. Yujian, N. Njuki, and L. Yingchun, "A comparative study of fine-tuned deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, 2019.
- [10] H. Hughes and M. Salathé, "An open access repository of plant disease images to enable the development of mobile disease diagnostics," *arXiv preprint arXiv:1511.08060*, 2015.
- [11] S. Sladojevic et al., "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, 2016.
- [12] P. Picon, A. Alvarez-Gila, A. Seitz, et al., "Deep convolutional neural networks for mobile capture device-based crop disease classification," *Computers and Electronics in Agriculture*, 2019.
- [13] G. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization," *Applied Artificial Intelligence*, 2017.
- [14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, 2015.
- [15] B. Zoph, V. Vasudevan, J. Shlens, and Q. Le, "Learning transferable architectures for scalable image recognition," *Proc. CVPR*, 2018.
- [16] D. Ghosh, A. Singh, and R. Kumar, "Rice leaf disease detection using deep learning techniques," *International Journal of Advanced Computer Science and Applications*, 2021.
- [17] N. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Computers and Electronics in Agriculture*, 2018.

- [18] X. Zhang, Y. Qiao, F. Meng, et al., "Identification of maize leaf diseases using improved deep convolutional neural networks," IEEE Access, 2018.
- [19] S. Wang, D. Wang, and Y. Zhang, "Plant disease detection using CNN and visualization techniques for explainable AI," IEEE Access, 2020.
- [20] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems (NeurIPS), 2012.