

**AI for Pig Disease Classification**

Ai in society



November 27, 2024

Lucas Weijs

Contents

[1. Introduction 1](#_Toc183596031)

[2. Sustainable Context and Societal Impact 1](#_Toc183596032)

[3. Critical Analysis of the AI Project 1](#_Toc183596033)

[4. Data Collection and Preparation 1](#_Toc183596034)

[5. Model Training and Testing 1](#_Toc183596035)

[6. Visualization and Communication 2](#_Toc183596036)

[7. Project Reporting and Reflection 2](#_Toc183596037)

[8. Entrepreneurial Insights and Professional Development 2](#_Toc183596038)

[9. Conclusion and Future Work 2](#_Toc183596039)

# Abstract

Pig disease classification is a pivotal challenge in the agricultural and veterinary sectors, impacting animal health, farm productivity, and food security. Leveraging Artificial Intelligence (AI), this project aims to develop a robust classification system to detect and diagnose pig diseases with high accuracy. The research integrates advanced machine learning techniques, comprehensive data preprocessing, and model optimization to address the complexities of real-world applications.

The study emphasizes sustainability, ethical considerations, and societal impacts while adhering to data legislation. Challenges such as data quality, imbalance, and model interpretability are critically analyzed, with proposed solutions to mitigate these issues. Additionally, interactive visualizations are employed to effectively communicate findings to stakeholders, including farmers and industry professionals.

The results demonstrate the potential of AI to revolutionize disease management in agriculture, contributing to sustainable practices and enhanced animal welfare. This research sets the stage for future advancements in AI-driven veterinary diagnostics, aligning with broader goals of intelligent and responsible farming innovations.

# 1. Sustainable Context and Societal Impact

The use of AI for pig disease classification addresses several sustainability and societal challenges, particularly in agriculture and veterinary medicine. By leveraging AI to identify diseases earlier and more accurately, farmers and veterinarians can focus on specific issues sooner, allowing for targeted interventions. This early detection minimizes the spread of diseases within herds, reduces the need for broad-spectrum treatments like antibiotics, and promotes animal welfare.

From a sustainability perspective, improved disease management contributes to more efficient farming practices by reducing livestock losses and optimizing resource utilization. Healthier animals require fewer resources, such as feed and water, and produce higher yields, supporting sustainable food production systems. This aligns with global goals of reducing environmental footprints in agriculture.

Ethically, the project emphasizes the responsible use of AI by adhering to data legislation and ensuring transparency in its applications. It also helps farmers and veterinarians make informed decisions based on reliable, data-driven insights rather than relying solely on manual assessments, which can be error-prone and time-consuming.

The societal impact extends beyond farms. By mitigating the risks of disease outbreaks in livestock, the project supports public health by reducing the potential transmission of zoonotic diseases. Additionally, fostering trust in AI-driven solutions among stakeholders promotes the adoption of innovative practices in farming, bridging the gap between traditional agriculture and modern technology.

# 2. Critical Analysis of the AI Project

This chapter identifies the challenges encountered in the AI project, such as data scarcity and model generalizability. Solutions like data augmentation and transfer learning are discussed, highlighting their effectiveness in addressing these issues.

# 3. Data Collection and Preparation

Data collection forms the backbone of the project. This chapter details the sourcing of high-quality data, its preprocessing, and adjustments made to enhance usability for AI models. Processes such as cleaning, normalization, and annotation are explained.

Data collection formed the foundation of this project, emphasizing the importance of high-quality and balanced datasets for effective AI model performance. To ensure the dataset's usability, an extensive manual review of images was conducted to remove poor-quality samples that could hinder model training. Multiple datasets were considered during this phase; however, the chosen dataset was selected based on its prior use in another AI project, which had already demonstrated a baseline accuracy of 87%. This provided confidence in its reliability and relevance to the task.

The scarcity of openly available data on pig diseases posed a significant challenge. To address this, the dataset was expanded to include diverse visual diseases rather than limiting the classification to binary categories like "healthy" and "ill." This approach not only enhanced the dataset’s depth but also better reflected real-world complexities. Preprocessing steps such as cleaning, normalization, and annotation were rigorously applied, while tools like ImageDataGenerator were employed to automate data preparation. The dataset was ultimately split into training (80%) and validation (20%) subsets to ensure balanced and effective learning.

# 4. Model Training and Testing

This section covers the selection of AI models, training methodologies, and testing strategies. Evaluation metrics such as accuracy, precision, and recall are used to assess model performance. Hyperparameter tuning and cross-validation techniques are also discussed.

The VGG16 model architecture was selected due to its well-documented success in image classification tasks. Transfer learning techniques were employed, leveraging VGG16’s pre-trained layers while adding custom layers tailored to the pig disease classification task. Training started with an initial accuracy of 15.4% in the first epoch, gradually improving as the model learned to distinguish between diverse disease categories.

This project explored various models, including custom Convolutional Neural Networks (CNNs) and fully connected neural networks. While these models showed promise, they failed to match the performance of VGG16, which consistently achieved the highest accuracy. The inclusion of multiple disease categories instead of a simple binary classification played a critical role in enhancing the model's predictive power.

Advanced training methodologies were applied, including hyperparameter tuning and cross-validation, to optimize model performance. Regularization techniques such as dropout were also employed to prevent overfitting. The results reaffirmed that VGG16, combined with a thoughtfully curated and diverse dataset, was the most effective solution for this task.

# 5. Visualization and Communication

Creating engaging and informative visualizations is key to communicating findings. This chapter presents examples of data storytelling tailored to stakeholders, including farmers and researchers.

# 6. Project Reporting and Reflection

A methodologically sound approach to documentation is essential for transparency and reproducibility. This section outlines the reporting of the project’s progress, outcomes, and lessons learned.

# 7. Entrepreneurial Insights and Professional Development

The project demonstrates an entrepreneurial mindset by addressing real-world challenges in animal health. This chapter reflects on personal growth and aligns the project with future professional ambitions.

# 8. Conclusion and Future Work

This paper concludes by summarizing the outcomes of the AI-based pig disease classification project. Future work aims to refine model accuracy and expand its applicability across different contexts. The long-term vision includes integrating AI solutions into broader veterinary practices.