Pharmacology, Medical Reports, Orthopedic, and Illness Details
https://ojs.transpublika.com/index.php/COMORBID
Online ISSN 2809-9745

https://doi.org/10.55047/comorbid.v5i1.1963

Original Article

Identification of Neisseria gonorrhoeae Bacteria Using a Convolutional Neural Network (CNN) Based on Image Classification

Haris Maulana[1*], Mudyawati Kamaruddin[2], Agus Suyanto[3], Auliyaur Rabban[4]

[1]Master of Science in Medical Laboratory Science, Graduate Program, Universitas Muhammadiyah Semarang, Indonesia
[2-4]Universitas Muhammadiyah Semarang, Indonesia
Email: [1)] harismaulana@umaha.ac.id, [2)] mudyawati@unimus.ac.id, [3)] agussuyanto.kh@unimus.ac.id, [4)] auliyaur.rabbani@umsida.ac.id

Received: 27 September - 2025
Accepted: 23 December - 2025
Published online: 12 January - 2026

Abstract

Neisseria gonorrhoeae (*gonococcus*) is the primary bacterium responsible for the sexually transmitted infection gonorrhea, which is transmitted through sexual contact. Traditional identification methods, such as Polymerase Chain Reaction (PCR), are still widely used but have limitations in terms of cost, time, and the need for multiple reagents. This study aims to develop a faster and more efficient identification method using Artificial Intelligence (AI) through a Convolutional Neural Network (CNN) approach based on the Inception V3 architecture. The dataset used consists of 84 JPEG images, comprising 42 images of Neisseria gonorrhoeae and 42 non-Neisseria images. The model was trained using 50 epochs with an early stopping mechanism, which optimally halted at epoch 25, achieving a training accuracy of 94.74% and a validation accuracy of 100%. The resulting model achieved 96% classification accuracy, correctly identifying all 8 positive and 4 negative test images. These findings indicate that CNN based on Inception V3 is effective in classifying Neisseria gonorrhoeae images and has strong potential as a fast, accurate, and efficient diagnostic alternative.

Keywords: Artificial Intelligence, Convolutional Neural Network (CNN), Image Classification, Inception V3, Neisseria Gonorrhoeae.

1. Introduction

Identification of microorganisms is an essential component of the diagnostic process in healthcare because microorganisms can cause various infectious diseases. Microorganisms are living organisms of very small size that can only be observed using a microscope and are capable of carrying out basic life functions such as growth, energy production, and reproduction at the single-cell level (Badaring et al., 2020; Pratiwi et al., 2020). Microorganisms can be found in various environments, such as air, water, soil, inanimate objects, and within the human body, and may be either pathogenic or non-pathogenic. Because pathogenic and non-pathogenic microorganisms cannot be distinguished with the naked eye, accurate identification becomes critically important for determining appropriate diagnosis and therapy.

In the field of microbiology, conventional methods for identification and classification of microorganisms, such as culture, molecular analysis, and imaging, remain widely used (Franco-Duarte et al., 2019). However, these methods generally require complex laboratory procedures, specialized expertise, and relatively long examination times (Arbefeville et al., 2024). Various challenges such as sample handling, difficulties in the culture process, identification errors, and limitations in antimicrobial susceptibility testing frequently occur and can lead to treatment delays (Franco-Duarte et al., 2019). Errors or delays in pathogen identification pose risks of inappropriate therapy administration and deterioration of patient condition (Kumar et al., 2023).

The development of Artificial Intelligence (AI) technology offers solutions to these various limitations. AI is capable of analyzing large volumes of data, recognizing complex patterns, and accelerating diagnostic processes, thus playing an important role in early disease detection, therapy monitoring, and epidemic surveillance (Bohr & Memarzadeh, 2020; Franiko et al., 2025). Image-based approaches using AI enable automatic recognition and classification of microorganisms through visual characteristics in microscopic images. The application of this technology has the potential to improve identification accuracy and reduce the risk of misdiagnosis in a shorter time compared to conventional methods (Naik et al., 2022).

One AI method widely used in image classification is the Convolutional Neural Network (CNN), a deep learning architecture specifically designed to process grid-structured data such as images. CNNs operate by automatically extracting features from pixel data and have been extensively applied in various fields, including face recognition, document analysis, and medical image classification. In the context of microbiology, CNNs process microscopic images in the form of pixel arrays, both grayscale images and color images with Red-Green-Blue (RGB) channels, to classify microorganisms based on learned morphological patterns.

Although research related to AI-based bacterial image classification continues to develop, there remain limitations in studies focusing on specific pathogens that have high clinical significance and are difficult to culture. One such bacterium is Neisseria gonorrhoeae, a human-specific pathogen causing gonorrhea, a sexually transmitted infection that remains a public health concern. N. gonorrhoeae typically infects cuboidal or columnar epithelium on mucosal membrane surfaces, such as the urethra, vagina, rectum, and pharynx (Leboffe & Pierce, 2011; Quillin & Seifert, 2018). Laboratory detection of this bacterium generally employs nucleic acid amplification methods using Polymerase Chain Reaction (PCR) because this bacterium is difficult to cultivate on conventional culture media. Although it has high sensitivity, the PCR method requires relatively high costs, complex procedures, and the use of various reagents.

Therefore, an alternative approach that is faster, more efficient, and more resource-efficient for detecting N. gonorrhoeae is needed. The application of AI technology using the CNN method has the potential to be used for detecting gonorrhea-causing bacteria directly from microscopic images without requiring culture processes or the use of reagents as in the PCR method. However, research specifically evaluating CNN performance in classifying microscopic images of N. gonorrhoeae remains limited, particularly studies comparing model classification results with manual labeling by microbiology experts as the reference standard (ground truth).

Based on this background, this study aims to analyze the performance of an Inception V3-based CNN model in classifying microscopic images of Neisseria gonorrhoeae bacteria using accuracy, precision, and recall metrics. Furthermore, this study evaluates the concordance between CNN model classification results and manual labeling by microbiology experts. This research also identifies various factors that influence model prediction quality, whether originating from data characteristics, model architecture, or training parameters, and is therefore expected to provide comprehensive understanding regarding the effectiveness of CNNs in classifying bacterial images that present high diagnostic challenges.

2. Literature Review

2.1. Microorganisms and Conventional Methods of Microscopic Identification

Microorganisms are all tiny creatures that are invisible to the naked eye. They are small and simple in structure, and can usually only be seen with a microscope. Environmental Microorganisms (EMs) specifically refer to species of microorganisms that live in natural environments (such as mountains, rivers, and oceans) and artificial environments (such as gardens and fish ponds). The identification of microorganisms is useful for diagnostic processes, especially for infected patients. Generally, classical techniques for classifying and identifying microorganisms are mainly based on biological techniques such as observing the morphology of microbial colonies on Petri dishes containing appropriate selective media, colony counting, Gram staining techniques, biochemical identification tests, antibiotic tests, etc.

Observation under a microscope is a common and important method in microorganism analysis. Stereo scanning electron microscopes are used for microorganism analysis in soil (Gray, 1967). In Daley and Hobbie (1975), a modified epifluorescence technique based on a microscope was used for counting water bacteria. In Collins et al. (1993), environmental scanning electron microscopy was used for microorganism analysis. However, these microscopic methods have several drawbacks. First, there are many types of microorganisms. As estimated by (Locey & Lennon, 2016), earth is inhabited by 10^{11}–10^{12} microbial species. Thus, experts' knowledge is always inadequate. When experts use this method for microorganism analysis, they often have to consult a lot of literature. Second, the training cycle for researchers and the overall detection time are very time-consuming. For example, counting phytoplankton using traditional microscopic methods is very time-consuming. In addition, an operator must have extensive professional knowledge (Embleton et al., 2003). Due to the limitations of microscopic methods, we need more efficient methods for analysing microorganisms. For example, image classification using AI is a suitable method.

2.2. Convolutional Neural Network (CNN) in Image Classification

With the advancement of computational technology and the increasing capacity and intelligence of computers, computational science enables machines to extract information from images automatically for object recognition purposes. One of the most widely used methods in image processing is the Convolutional Neural Network (CNN). CNN is a development of the Multi Layer Perceptron (MLP) and is included in Deep Learning algorithms.

Before CNNs were widely adopted, image classification was generally performed by manually extracting features from images, then inputting them into classification algorithms such as Support Vector Machine (SVM). Several approaches used image pixel values as feature vectors, for instance, training SVM with 784 features representing pixel values from images sized 28×28.

CNNs operate with a hierarchical model that builds networks progressively, resembling a funnel structure, and concludes with a fully connected layer to produce classification output. The CNN concept was first introduced by Yann LeCun in digit classification using a single convolutional layer (LeCun & Bengio, 1998). CNN development accelerated significantly after the introduction of AlexNet in 2012, which employed multiple convolutional layers and demonstrated superior performance on the ImageNet dataset, thereby establishing CNN as the primary method in various image classification challenges.

2.3. Application of Artificial Intelligence in Pathogen Image Classification

Artificial intelligence (AI) has become a growing trend in recent years. One of the tasks that AI can accomplish is computer vision, which is the ability of computers to process and analyse images, with the aim of mimicking human vision. One of the main tasks of computer vision is image classification, which is the process of labelling images into 'classes'. For example, if there are images of several objects, and these images need to be categorised into 'classes', such as 'Neisseria gonorrhoeae (gonococci)', that is image classification.

In today's healthcare industry, the use of artificial intelligence for pathogen image classification is very important. Artificial Intelligence (AI) can help identify and classify pathogens quickly and accurately, saving time and resources. This technology enables faster and more accurate diagnosis of diseases, reducing the time and cost required for conventional diagnosis, and can also help address medical staff shortages, especially in hard-to-reach locations. This can help doctors and medical staff make better decisions about diagnosis and treatment of diseases.

Therefore, this research has significant potential to improve the efficiency and accuracy of pathogen image classification, as well as provide great benefits to the healthcare industry as a whole (Kumar et al., 2022, Pillai et al., 2022).

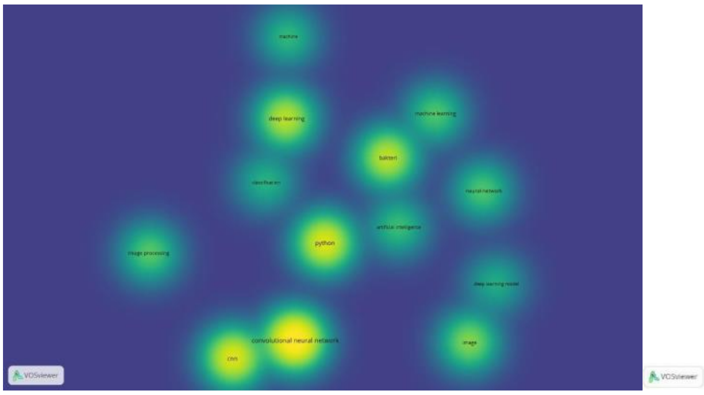

Figure 1. Research authenticity using VOSviewer

This study was conducted to classify images using the CNN method. Several topics from previous studies can be used as a basis for the thought process in the design and development, as well as in its application, in line with the needs and developments of relevant current technology. The hypotheses taken before this study were as follows:

1) The Convolutional Neural Network (CNN) model based on the Inception V3 architecture is capable of classifying microscopic images of Neisseria gonorrhoeae bacteria with an accuracy, precision, and recall rate of ≥ 85%.
2) The image classification results by the Inception V3-based CNN model are consistent with the ground truth labelling by microbiology experts.
3) Image quality and bacterial morphological variation significantly affect the success and failure of CNN model classification.

3. Methods

3.1. Type of Research

This type of research falls under the category of quantitative experimental research with a development and evaluation study approach. The main objective is to develop and test the performance of a Convolutional Neural Network (CNN) model based on the Inception V3 architecture in classifying microscopic images of Neisseria gonorrhoeae bacteria.

3.2. Population, Sample Size, and Sampling Technique

Based on the flow shown in the figure above, the population source for this study comes from a collection of microscopic images of bacteria taken from two main databases, namely Digital Image of Bacterial Species (DiBaS) and Kaggle. The DiBaS dataset includes 330 images divided into 20 different bacterial categories, while the Kaggle dataset consists of 235 images grouped as non-Neisseria gonorrhoeae. From this population, the researchers screened the images to select samples that were suitable for the research objectives. From the DiBas dataset, one specific category was selected containing 42 images of Neisseria gonorrhoeae. From this population, researchers screened samples to select those that were suitable for the research objectives. From the DiBas dataset, one specific category was selected containing 42 images of Neisseria gonorrhoeae. On the other hand, for comparison, 42 images of non-Neisseria gonorrhoeae bacteria were taken from a combination of non-Neisseria images found on Kaggle.

In this way, the total number of samples used in this study was 84 images consisting of 42 positive images (Neisseria gonorrhoeae) and 42 negative images (non-Neisseria gonorrhoeae). The sampling method applied was purposive sampling, which is the deliberate selection of samples based on certain criteria appropriate for the research objectives. In this case, the criteria applied included label clarity, image quality, and suitability for the positive or negative category. This method was chosen to ensure that the final data set truly represented the two groups that would be used in training and testing the Convolutional Neural Network (CNN) model.

3.3. Research Variables

The variables in this study were the CNN model, accuracy, precision, and F1-Score. The independent variable in this study was the CNN model, while the dependent variables were accuracy, precision, and F1-score.

To ensure consistency and clarity, this study establishes the following operational definitions: Image Classification refers to the process by which the model automatically identifies and categorises microscopic images into two classes, namely Neisseria gonorrhoeae or non-Neisseria. The Convolutional Neural Network (CNN) model is an artificial intelligence system specifically designed to process visual data. The Inception V3 architecture is a specific configuration of CNN that optimises feature extraction. Accuracy is measured as the ratio of successful model predictions to the entire test data. Training Data is a set of images used to train the model, while Test Data is a separate set of images used to evaluate the performance of the trained model.

3.4. Research Tools and Materials

In developing this system, researchers used the following tools:

a. Hardware

A set of computers with the following specifications:

Processor : Intel(R) Core(TM) i5-9300HF CPU @ 2.40GHz

Memory : 16 GB DDR4

Hardisk : SSD 256 GB
OS : Windows 11 Pro 64-bit

b. Software

1) Jupyter Notebook

Jupyter Notebook is an open-source web application that allows users to create and share documents containing interactive code, equations, visualisations, and narrative text. Its uses include data cleaning and transformation, numerical simulation, statistical modelling, data visualisation, machine learning, and much more. Jupyter Notebook is a spin-off project from the Python project, which previously had its own Python Notebook project. The name Jupyter comes from the core programming languages it supports: Julia, Python, and R. Jupyter is equipped with an Ipython kernel.

2) Python

Python is a high-level, interpreted, object-oriented programming language with dynamic semantics. It has high-level built-in data structures, combined with dynamic typing and dynamic binding, making it very attractive for rapid application development and as a scripting or glue language for connecting existing components.

Python's simple and easy-to-learn syntax emphasises readability, thereby reducing programme maintenance costs. Python supports modules and packages, which encourage programme modularity and code reuse. The Python interpreter and extensive standard libraries are provided in source or binary code, available on all major platforms, and distributed free of charge. The materials used in this research are image data obtained through DiBas and Kaggle open sources.

3.5. Data Analysis

The dataset used in this study comprises microscopic images of Neisseria gonorrhoeae bacteria and a control group consisting of non-Neisseria samples, obtained from two primary sources: the Digital Images of Bacterial Species (DIBaS) dataset and curated image searches via Google Images. The DIBaS dataset serves as a widely recognized reference collection in machine learning-based microorganism classification research, providing high-quality digital images of various bacterial species. To complement this, additional images from Kaggle and Google Images were included to introduce a broader range of non-Neisseria samples with visual characteristics resembling those found in clinical environments. Before these microscopic images are used for training Convolutional Neural Network (CNN) models, a series of pre-processing and augmentation steps are applied to ensure that the model learns from clean, standardized, and diverse data representations. These steps are crucial for enhancing the model's generalization capability, minimizing the risk of overfitting, and improving its robustness in handling image variations that may occur in real-world diagnostic scenarios.

3.6. CNN Experiment Design with Inception V3

The CNN model used in this study adapts the Inception V3 architecture. Inception V3 is a CNN architecture that can perform image recognition processes so that it can be used for image classification processes. Inception V3 combines various convolutions with different kernel sizes, as it can extract features efficiently. Inception V3 also uses regulation and reduction techniques to prevent overfitting. The Inception V3 architecture has a deep and complex structure (Nugraini, 2024). Inception V3 is a model trained to perform the classification process of nisseriae bacteria.

4. Results and Discussion

4.1. Research Results

4.1.1. Model Performance Evaluation Results

The performance evaluation of the model obtained from the initial epoch value was established by the researcher at 50, which means that the model was allowed to undergo the training process for up to 50 full cycles of training data. The researcher also applied an early stopping parameter to halt the training process once the model had reached optimal performance, as indicated by a low validation loss value and a high accuracy value. However, the training process stopped at epoch 25 because the best performance had already been achieved at epoch 25. Early stopping also plays a role in preventing futile training and saving computation. Thus, overfitting was successfully prevented because training was stopped before the model began to 'memorise' the training data excessively. The model is ready to be used or tested further, with high accuracy and low loss values, as shown in Table 1 below.

Table 1. Model Performance Evaluation (epoch)

Epoch	Training Accuracy	Training Loss	Val Accuracy	Val Loss
21/50	97.37%	0.1057	100%	0.1177
22/50	97.37%	0.1146	100%	0.0416
23/50	97.37%	0.1291	100%	0.0666
24/50	96.05%	0.1399	100%	0.0634
25/50	94.74%	0.1167	100%	0.0560

Based on Table 1, the model demonstrates high and relatively stable training accuracy, ranging from 94.74% to 97.37%, with low training loss values at each observed epoch. Furthermore, validation accuracy reaches 100% with consistently low and stable validation loss values. This consistency between training and validation performance indicates that the model is capable of learning data patterns effectively without significant performance differences between training data and validation data. The pattern of decline and stabilization in loss values indicates that the model has reached a convergent state and possesses good generalization capability toward previously unseen data.

4.1.2. Training and Validation Results

Based on Table 1 Convolutional Neural Network (CNN) Model – Inception V3, the data was trained with 50 iterations (epochs) using an early stopping mechanism. At epoch 25, the results were 3s/step – accuracy: 0.9474 – loss: 0.1167 – val_accuracy: 1.0000 – val_loss: 0.0560. Training automatically stopped at epoch 25 due to no significant changes in the validation metrics, indicating that the model had reached optimal convergence, as shown in Table 1.

Training accuracy was stable in the range of 94.74%–97.37%, with the highest value of 97.37% occurring between epochs 21 and 25. The training loss varied between 0.1057–0.1399, indicating minor differences in training. The validation accuracy consistently reached 100% from epoch 21 to epoch 25, indicating that the model was very good at processing new data. The validation loss (val_loss) decreased significantly from 0.1177 (epoch 21) to 0.0560 (epoch 25), proving that the model is becoming more precise. Although the training accuracy decreased slightly in the final epoch (from 97.37% to 94.74%), the validation accuracy remained perfect (100%). This indicates that the model is not overfitting and is able to adapt to data not seen during training. Stopping at epoch 25 proves that adding epochs is unnecessary to save computation time without sacrificing performance.

a) Evaluation metrics

The evaluation metric in the form of the validation process for Neisseria gonorrhoeae bacterial image labels in the Digital Image of Bacterial Species (DIBaS) dataset involves a series of rigorous steps carried out by microbiologists to ensure data accuracy and reliability. The following is a detailed explanation:

1. Bacterial Sample Preparation

Prior to image validation, microbiologists ensure that:

a) The bacterial culture used is indeed Neisseria gonorrhoeae (confirmed through biochemical tests such as positive oxidase and glucose fermentation).

b) Gram staining is performed correctly (should show kidney-shaped Gram-negative diplococci).

2. Microscopic Image Acquisition

a) Images are taken using a microscope with high magnification, typically 100x with immersion oil.

b) Lighting and focus parameters are standardised for consistency.

3. Validation by Microbiologists

a) Morphological Examination

The expert verifies that the image shows the characteristic features of N. gonorrhoeae:

(1) Shape: Diplococci (kidney-shaped pairs of cells).

(2) Colour: Pink (Gram-negative) on Gram staining.

(3) Cell arrangement: Usually in pairs, sometimes in small groups.

b) Elimination of artefacts or contamination

Ensuring the image does not contain:

(1) Dust or other microscopic particles.

(2) Other bacteria (e.g., Neisseria meningitidis, which is morphologically similar but clinically distinct).

c) Cross-Check with Laboratory Data

(1) Image labels are matched with laboratory test results (PCR, MALDI-TOF MS) to confirm genetic identity.

4. Labelling and Quality Control (QC)

(1) Multi-rater review: Several independent microbiologists assess the same image. If there are differences, a discussion or retest is conducted.

(2) Consensus criteria: Images are only accepted if ≥90% of experts agree with the *N. gonorrhoeae label.

5. Documentation and Metadata

Each image is accompanied by metadata that includes:

(1) Staining method (Gram, methylene blue).

(2) Culture conditions (Thayer-Martin medium, incubation temperature).

(3) Name of the expert who validated it.

6. Validation Example in DIBaS Research

In the study by Maltby et al. (2018) gonorrhoeae images in DIBaS were validated by 3 microbiology experts. Labelling accuracy reached 98% after the QC process, as shown in Figure 2 below.

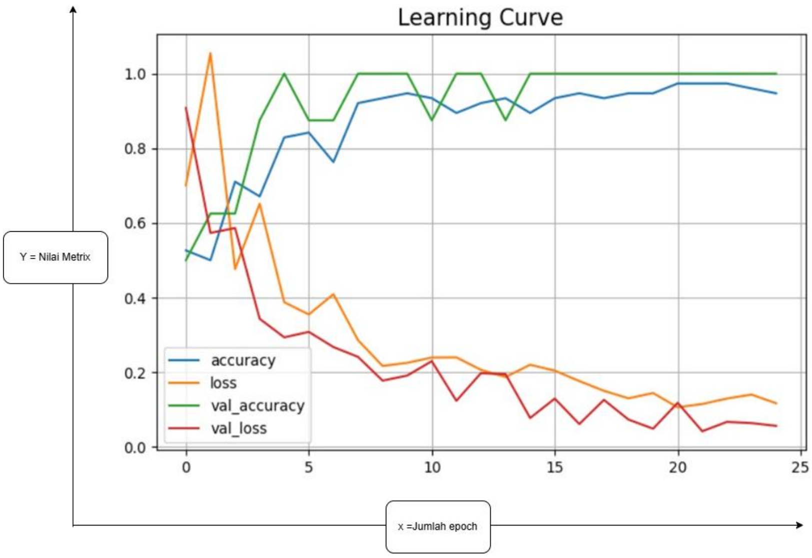

Figure 2. Learning Curve Graph

The results of the researchers who conducted comprehensive retraining from epoch 1 to epoch 25. First, regarding accuracy (training accuracy), the graph results do not show a significant difference with validation accuracy. In epochs 1–10, the model was still unstable in terms of generalisation, as training accuracy ranged from 60% to 90%, while validation accuracy fluctuated between 70% and 85%. In epochs 11–20, the difference was less than 3%, indicating the effectiveness of the data argument in preventing overfitting, as training accuracy reached 93%–96%, while validation accuracy reached 95%–98%. In epochs 21–25, the validation accuracy percentage reached 100%, higher than the training accuracy, which ranged from 94% to 97%, so it can be concluded that the model showed ideal results. Second, regarding loss (data reading errors), the graph results were not much different from the validation loss. In epochs 1–14, initial convergence was successful because the training loss and validation loss dropped dramatically from 1.0 to below 0.3. In epochs 14–25, the decline continued dramatically and consistently to 0.056, proving that the model became more accurate.

4.1.3. Test Data Classification Results

The classification of Neisseria and non-Neisseria test data is shown in Table 2.

Table 2. Table of Neisseria and non-Neisseria image test classes

Test Class	Number of Images	Correct Predictions	Accuracy
Yes Neisseria	8	8	100%
Non Neisseria	4	4	100%
Total	12	12	100%

The classification results of Neisseria and non-Neisseria test data are presented in Table 2. A total of 12 test images were used, consisting of 8 Neisseria images and 4 non-Neisseria images. The model was able to correctly classify all test images, thus achieving an accuracy rate of 100% for each class. Specifically, all Neisseria and non-Neisseria images were successfully predicted according to their actual classes. These results demonstrate that the model possesses excellent classification performance in distinguishing Neisseria and non-Neisseria images on the test data used.

Results of image classification model testing using the CNN method to detect the presence of Neisseria and non-Neisseria bacteria. The images consist of 12 classified images, where 8 bacterial images are Neisseria and 4 are non-Neisseria bacteria.

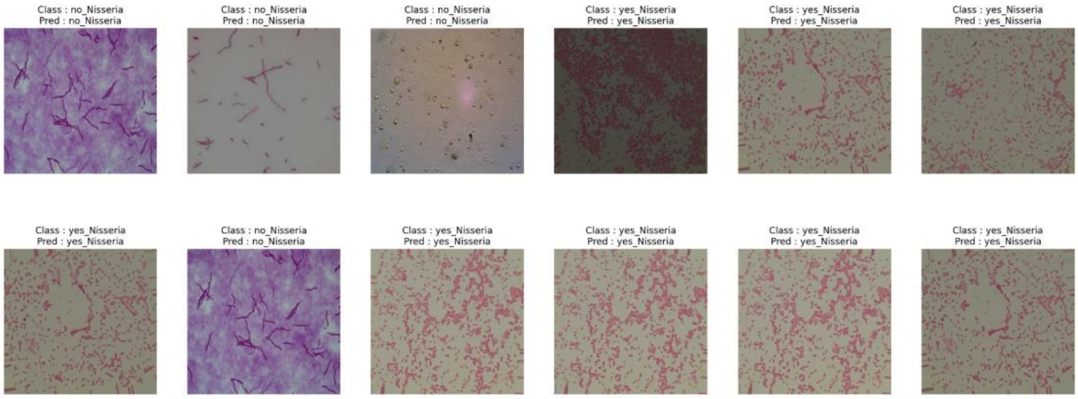

Figure 3. Neisseria and non-Neisseria testing data

The results of the CNN model prediction with the inceptionV3 architecture and 96% accuracy showed perfect performance in identifying Neisseria gonorrhoeae. All 8 positive images were correctly identified, and the 4 negative images were classified accurately.

Based on the visualization of classification results in Figure 3, it can be observed that the CNN model with InceptionV3 architecture is capable of consistently distinguishing Neisseria and non-Neisseria images. Each test image displays concordance between the actual class label and the model's prediction results, as indicated by the absence of classification errors across all test data. This demonstrates that the morphological features of bacteria present in microscopic images have been successfully learned by the model. The model's success in accurately classifying all test images indicates that the InceptionV3 architecture possesses strong feature representation capability for detecting the presence of Neisseria gonorrhoeae bacteria. Nevertheless, considering the relatively limited number of test data, further testing with a larger and more diverse dataset remains necessary to ensure the consistency and generalization of model performance.

4.1.4. The potential of conventional screening tools/methods with the CNN method

1. Primary Source on DIBaS and Validation by Experts
   a. Original DIBaS Article
   (1) Title: 'Deep learning approach to bacterial colony classification'
   (2) Authors: Zieliński et al. (2017)
   (3) Link: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0184554
   (4) Contents:
   Explains the creation of the DIBaS dataset, including validation standards by microbiology experts. The labelling process involved several experts to verify the morphology of the bacteria.
   b. Multi-Expert Validation Study
   (1) Title: 'Quality assurance for antimicrobial susceptibility testing of Neisseria gonorrhoeae in Latin American and Caribbean countries, 2013-2015'
   (2) Authors: Sawatzky et al. (2018)
   (3) Link: https://sti.bmj.com/content/94/7/479.abstract

(4) Contents:
This study evaluates inter-laboratory agreement in antimicrobial susceptibility testing of Neisseria gonorrhoeae across multiple countries, using MIC values as a standardized reference. The results demonstrate generally high agreement levels (>90%) among laboratories, while also revealing variability for certain antibiotics. Although not focused on image-based labelling, this study provides empirical evidence that expert agreement in microbiological assessment can vary even under standardized protocols, underscoring the need for careful validation when dealing with organisms such as Neisseria gonorrhoeae that present classification challenges.

2. Specific Validation Procedures for Neisseria
a. CDC Guidelines for the Identification of N gonorrhoeae
(1) Source: Centres for Disease Control and Prevention (CDC)
(2) Link: https://www.cdc.gov/gonorrhea/php/laboratories/index.html
(3) Contents:
Standard protocol for identifying N. gonorrhoeae through Gram staining, biochemical tests, and PCR, used as a reference in the validation of microbiology datasets.
b. Study on Labelling Errors
(1) Title: 'The laboratory diagnosis of Neisseria gonorrhoeae: current testing and future demands'
(2) Author: Meyer & Buder (2020)
(3) Link: https://www.mdpi.com/2076-0817/9/2/91
(4) Contents:
Diagnosis of Neisseria gonorrhoeae using only microscopy or conventional methods carries a risk of misidentification because non-pathogenic Neisseria can appear similar in Gram preparations. Therefore, additional confirmation using molecular techniques such as PCR is necessary to improve specificity and reduce diagnostic errors.

3. Examples of Validation Implementation in AI
a. Paper on Bacterial Classification with DIBaS
(1) Title: 'Small-Scale Depthwise Separable Convolutional Neural Networks for Bacteria Classification'
(2) Authors: Mai & Ishibashi (2021)
(3) Link: https://www.mdpi.com/2079-9292/10/23/3005
(4) Content:
Uses expert-validated DIBaS images to train a CNN for classifying 33 bacterial species; includes preprocessing and data augmentation to improve data quality, though it does not explicitly mention criteria for excluding ambiguous image.
b. Validation for Explainable AI (XAI) Models
(1) Title: 'Explainable Artificial Intelligence for Human Decision-Support System in Medical Domain'
(2) Author: Knapič et al. (2021)
(3) Link: https://www.mdpi.com/2504-4990/3/3/37
(4) Contents:
Emphasises the importance of expert-validated ground truth in AI development.

4. Additional Sources

Book: 'Manual of Clinical Microbiology' (ASM Press)

(1) The chapter on Neisseria identification explains the microscopic validation procedure. (Carroll et al., 2024)

(2) Link: https://www.who.int/publications/i/item/978924001 534

(3) Content: To detect gonorrhoea

Table 3. Table of Neisseria and non-Neisseria images

Method	Detection Time	Accuracy	Cost	Equipment Requirements	Sensitivity	Specificity	Reference
Culture	24-72 hours	70-90%	Low-Medium	Incubator, Culture Media	60-75%	85-95%	CDC, 2022
PCR	2-6 hours	95-99%	High	Thermocycler, Reagen PCR	95-99%	98-99%	WHO,2023
Microscopy (Gram)	15-30 minutes	50-70%	Low	Microscope, Gram Staining	50-65%	70-85%	Clin. Microbiol Rev, 2021
CNN (InceptionV3)	1 second-5 minutes	90-98%	Medium*	Computer/GPU, AI Software	92-97%	94-95%	Research Results

As shown in table 3, it can be seen that the identification of Neisseria bacteria using the CNN method shows faster and more efficient results than the identification of Neisseria bacteria using conventional methods. It also has the potential to be an aid/screening tool to improve research on gonorrhoea bacteria, especially Neisseria gonorrhoeae. A comparison of Neisseria gonorrhoeae detection methods shows that traditional culture methods, although specific (85-95%), are slow (24-72 hours) and less sensitive (60-75%), while PCR is more accurate (95-99%) but expensive and requires complex infrastructure. Gram staining is fast (15-30 minutes) and inexpensive, but highly dependent on operator expertise with limited accuracy (50-70%). The CNN InceptionV3-based solution offers the highest speed (1-5 seconds) with accuracy comparable to PCR (90-98%), although it requires initial model training and further clinical validation. This AI approach has the potential to advance diagnostics through automation, reducing reliance on experts while lowering long-term operational costs. This model not only demonstrates the technical feasibility of AI-based Neisseria gonorrhoeae identification but also opens the door to digital transformation in microbiological diagnostics. This approach has the potential to become the new standard in sexually transmitted infection management.

4.2. Discussion

The results of this study demonstrate that the InceptionV3-based CNN model possesses excellent performance in classifying Neisseria gonorrhoeae images. The consistency between high training accuracy and validation accuracy reaching 100% indicates that the model is capable of learning bacterial morphological patterns without experiencing overfitting. The early stopping mechanism proved effective in terminating training at the optimal convergence point, thereby accelerating the training process without sacrificing accuracy (Nurtanio et al., 2022). The model's success in accurately classifying all test images demonstrates that CNN can detect important bacterial morphological features accurately, in alignment with manual labeling by microbiology experts and laboratory validation standards (Zieliński et al., 2017; Maltby et al., 2018).

Several factors influence model prediction quality. Data characteristics, including image quality and bacterial morphological variation, affect the model's capability in recognizing patterns, where datasets that are less representative or have imbalanced class distributions can decrease accuracy on new data (Ahmed et al., 2025). The use of InceptionV3 architecture enables complex multi-scale feature extraction, and transfer learning from pre-trained models enhances performance on limited datasets, allowing the model to distinguish N. gonorrhoeae from other similar species (Kim et al., 2022). Training strategies such as data augmentation and early stopping also play important roles in improving generalization and preventing overfitting (Khovidhunkit et al., 2025). Model evaluation demonstrates clinical potential as a rapid screening tool that can improve laboratory efficiency and reduce dependence on expert personnel, although its use should remain combined with standard diagnostic methods such as PCR (Soe et al., 2024). The combination of these factors confirms that the model possesses significant technical effectiveness and clinical relevance, supporting the development of image-based bacterial detection systems in healthcare laboratories. Overall, these findings affirm that InceptionV3-based CNN is capable of providing rapid, accurate, and consistent bacterial image classification, thereby supporting the potential adoption of AI methods in the digital transformation of microbiological diagnostics.

These findings have important implications for the development of digital diagnostic systems. The InceptionV3-based CNN model has the potential to be integrated into telemedicine systems, enabling medical personnel to detect and classify Neisseria gonorrhoeae bacteria rapidly using portable devices. This approach is particularly beneficial for medical personnel in remote areas with limited access to microbiology laboratories, as it can accelerate screening and support clinical decision-making.

Still, this study is without limitations. Image validation remains a crucial factor, as labeling errors in the dataset can affect model accuracy. Furthermore, the amount of data used remains limited, so the model needs to be tested with additional, more diverse datasets to improve prediction robustness and reliability. Technical variations, such as Gram staining thickness, as well as morphological similarities between N. gonorrhoeae and other similar species, can complicate both visual interpretation and automatic classification. Therefore, further development requires improvement in dataset quality, diversification of data sources, and cross-laboratory validation so that the model can be applied consistently in broader clinical contexts.

5. Conclusion

This study aims to develop a Convolutional Neural Network (CNN) model based on InceptionV3 architecture for binary classification of Neisseria gonorrhoeae images. The model trained on a limited dataset demonstrates high performance, with an accuracy of 94.74%, precision of 100%, and validation accuracy of 100%, thus capable of distinguishing Neisseria and non-Neisseria bacteria on the dataset used. Model predictions were also compared with reference images from the Digital Images of Bacterial Species (DIBaS) and expert-verified Google images, showing consistent concordance. However, due to the small dataset size, the risk of overfitting remains, and the model's ability to generalize to broader clinical datasets cannot yet be ascertained. For future research, the primary priority is the collection of a larger and clinically representative image dataset to validate model robustness and its potential use as a screening tool. Additionally, exploration of additional parameter settings or combinations, as well as utilization of pretrained InceptionV3 weights, can support further model development and improve transfer learning efficiency.

6. References

Ahmed, S. R., Befano, B., Egemen, D., Rodriguez, A. C., Desai, K. T., Jeronimo, J., Ajenifuja, K. O., Clark, C., Perkins, R., & Campos, N. G. (2025). Generalizable deep neural networks for image quality classification of cervical images. *Scientific Reports*, *15*(1), 6312. https://doi.org/10.1038/s41598-025-90024-0

Arbefeville, S. S., Timbrook, T. T., & Garner, C. D. (2024). Evolving strategies in microbe identification—A comprehensive review of biochemical, MALDI-TOF MS and molecular testing methods. *Journal of Antimicrobial Chemotherapy*, *79*(Supplement_1), i2–i8. https://doi.org/10.1093/jac/dkae275

Badaring, D. R., Fiqriansyah, M., & Bahri, A. (2020). Identifikasi Morfologi Mikroba pada Ruangan Water Closet Jurusan Biologi Universitas Negeri Makassar. *Prosiding Seminar Nasional Biologi FMIPA UNM. Makassar: Program Studi Biologi, Universitas Negeri Makassar*, 161–168.

Bohr, A., & Memarzadeh, K. (2020). The rise of artificial intelligence in healthcare applications. In *Artificial Intelligence in healthcare* (pp. 25–60). Elsevier. https://doi.org/10.1016/B978-0-12-818438-7.00002-2

Carroll, K. C., Pfaller, M. A., Karlowsky, J., Landry, M. L., McAdam, A. J., Patel, R., & Pritt, B. S. (2024). *Manual of clinical microbiology, 4 Volume Set*. John Wiley & Sons.

Collins, S. P., Pope, R. K., Scheetz, R. W., Ray, R. I., Wagner, P. A., & Little, B. J. (1993). Advantages of environmental scanning electron microscopy in studies of microorganisms. *Microscopy Research and Technique*, *25*(5-6), 398–405.

Daley, R. J., & Hobbie, J. E. (1975). Direct counts of aquatic bacteria by a modified epifluorescence technique 1. *Limnology and Oceanography*, *20*(5), 875–882. https://doi.org/10.4319/lo.1975.20.5.0875

Embleton, K. V, Gibson, C. E., & Heaney, S. I. (2003). Automated counting of phytoplankton by pattern recognition: a comparison with a manual counting method. *Journal of Plankton Research*, *25*(6), 669–681. https://doi.org/10.1093/plankt/25.6.669

Franco-Duarte, R., Černáková, L., Kadam, S., S. Kaushik, K., Salehi, B., Bevilacqua, A., Corbo, M. R., Antolak, H., Dybka-Stępień, K., & Leszczewicz, M. (2019). Advances in chemical and biological methods to identify microorganisms—from past to present. *Microorganisms*, *7*(5), 130. https://doi.org/10.3390/microorganisms7050130

Franiko, R., Suhendra, E. J., & Thevany, T. (2025). Innovation in Urology Telerobotic Surgery in Improving Treatment Effectiveness and Safety through a Systematic Journal Review. *PHARMACOLOGY, MEDICAL REPORTS, ORTHOPEDIC, AND ILLNESS DETAILS*, *4*(1), 41–53. https://doi.org/10.55047/comorbid.v4i1.1658

Gray, T. R. G. (1967). Stereoscan electron microscopy of soil microorganisms. *Science*, *155*(3770), 1668–1670.

Khovidhunkit, S. P., Phosri, K., Thanathornwong, B., Rungraungrayabkul, D., Poomrittigul, S., & Treebupachatsakul, T. (2025). Performance of deep learning models for the classification and object detection of different oral white lesions using photographic images. *Scientific Reports*, *15*(1), 30834. https://doi.org/10.1038/s41598-025-14450-w

Kim, H. E., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, M. E., & Ganslandt, T. (2022). Transfer learning for medical image classification: a literature review. *BMC Medical Imaging*, *22*(1), 69. https://doi.org/10.1186/s12880-022-00793-7

Knapič, S., Malhi, A., Saluja, R., & Främling, K. (2021). Explainable artificial intelligence for human decision support system in the medical domain. *Machine Learning and Knowledge Extraction*, *3*(3), 740–770. https://doi.org/10.3390/make3030037

Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. (2023). Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *Journal of Ambient Intelligence and Humanized Computing*, *14*(7), 8459–8486. https://doi.org/10.1007/s12652-021-03612-z

Leboffe, M. J., & Pierce, B. E. (2011). Photographic Atlas for the microbiology laboratory. 4th editio. *USA: Douglas N. Morton*.

LeCun, Y., & Bengio, Y. (1998). Convolutional networks for images, speech, and time series. *The Handbook of Brain Theory and Neural Networks*.

Locey, K. J., & Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences*, *113*(21), 5970–5975. https://doi.org/10.1073/pnas.1521291113

Mai, D.-T., & Ishibashi, K. (2021). Small-scale depthwise separable convolutional neural networks for bacteria classification. *Electronics*, *10*(23), 3005.

Maltby, V. E., Lea, R. A., Graves, M. C., Sanders, K. A., Benton, M. C., Tajouri, L., Scott, R. J., & Lechner-Scott, J. (2018). Genome-wide DNA methylation changes in CD19+ B cells from relapsing-remitting multiple sclerosis patients. *Scientific Reports*, *8*(1), 17418. https://doi.org/10.1038/s41598-018-35603-0

Meyer, T., & Buder, S. (2020). The laboratory diagnosis of Neisseria gonorrhoeae: current testing and future demands. *Pathogens*, *9*(2), 91. https://doi.org/10.3390/pathogens9020091

Naik, N., Hameed, B. M., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Ibrahim, S., Patil, V., & Smriti, K. (2022). Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? *Frontiers in Surgery*, *9*, 862322.

Nugraini, V. F. (2024). *Klasifikasi Penyakit Kulit Nevus Menggunakan Metode Convolutional Neural Network Dengan Arsitektur Inception V3*. Universitas Islam Sultan Agung Semarang.

Nurtanio, I., Bustamin, A., Yohannes, C., & Handoyo, A. T. (2022). Multi Classification of Bacterial Microscopic Images Using Inception V3. *ILKOM Jurnal Ilmiah*, *14*(1), 80–90. https://doi.org/10.33096/ilkom.v14i1.1121.80-90

Pillai, N., Ramkumar, M., & Nanduri, B. (2022). Artificial intelligence models for zoonotic pathogens: a survey. *Microorganisms*, *10*(10), 1911. https://doi.org/10.3390/microorganisms10101911

Pratiwi, K. N. F., Mastra, N., & Sofi Yanty, J. (2020). *Identifikasi Bentuk Bakteri Dari Swab Tangan Petugas Pengangkut Sampah Di Desa Dalung, Kecamatan Kuta Utara, Kabupaten Badung*. Poltekkes Denpasar Jurusan Teknologi Laboratorium Medis.

Quillin, S. J., & Seifert, H. S. (2018). Neisseria gonorrhoeae host adaptation and pathogenesis. *Nature Reviews Microbiology*, *16*(4), 226–240. https://doi.org/10.1038/nrmicro.2017.169

Sawatzky, P., Martin, I., Galarza, P., Carvallo, M. E. T., Rodriguez, P. A., Cruz, O. M. S., Hernandez, A. L., Martinez, M. F., Borthagaray, G., & Payares, D. (2018). Quality assurance for antimicrobial susceptibility testing of Neisseria gonorrhoeae in Latin American and Caribbean countries, 2013–2015. *Sexually Transmitted Infections*, *94*(7), 479–482. https://doi.org/10.1136/sextrans-2017-053502

Soe, N. N., Yu, Z., Latt, P. M., Lee, D., Ong, J. J., Ge, Z., Fairley, C. K., & Zhang, L. (2024). Evaluation of artificial intelligence-powered screening for sexually transmitted infections-related skin lesions using clinical images and metadata. *BMC Medicine*, *22*(1), 296. https://doi.org/10.1186/s12916-024-03512-x

Zieliński, B., Plichta, A., Misztal, K., Spurek, P., Brzychczy-Włoch, M., & Ochońska, D. (2017). Deep learning approach to bacterial colony classification. *PloS One*, *12*(9), e0184554. https://doi.org/10.1371/journal.pone.0184554