



Automated Fungal Identification with Deep Learning on Time-Lapse Images

Mansourvar, Marjan; Charylo, Karol Rafal; Frandsen, Rasmus John Normand; Brewer, Steen Smidth; Hoof, Jakob Blæsbjerg

Published in:
Information

Link to article, DOI:
[10.3390/info16020109](https://doi.org/10.3390/info16020109)

Publication date:
2025

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Mansourvar, M., Charylo, K. R., Frandsen, R. J. N., Brewer, S. S., & Hoof, J. B. (2025). Automated Fungal Identification with Deep Learning on Time-Lapse Images. *Information*, 16(2), Article 109.
<https://doi.org/10.3390/info16020109>

General rights




Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Article

Automated Fungal Identification with Deep Learning on Time-Lapse Images

Marjan Mansourvar^{1,*} , Karol Rafal Charylo², Rasmus John Normand Frandsen¹ , Steen Smidth Brewer¹ and Jakob Blæsbjerg Hoof^{1,*} 

¹ Department of Biotechnology and Biomedicine (DTU Bioengineering), Technical University of Denmark, Søtofts Plads, 2800 Kongens Lyngby, Denmark; rasf@dtu.dk (R.J.N.F.); ssmbr@dtu.dk (S.S.B.)

² Department of Electrical and Photonics Engineering (DTU Electro), Technical University of Denmark, Søtofts Plads, 2800 Kongens Lyngby, Denmark; s220243@student.dtu.dk

* Correspondence: marjma@dtu.dk (M.M.); jblni@dtu.dk (J.B.H.)

Abstract: The identification of species within filamentous fungi is crucial in various fields such as agriculture, environmental monitoring, and medical mycology. Traditional identification methods based on morphology have a low demand for advanced equipment usage and heavily depend on manual observation and expertise. However, this approach may struggle to differentiate between species in a genus due to their potential visual similarities, making the process time-consuming and subjective. In this study, we present an AI-based fungal species recognition model that utilizes deep learning techniques applied to time-lapse images. The training dataset, derived from fungi strains in the IBT Culture Collection, comprised 26,451 high-resolution images representing 110 species from 35 genera. The dataset was divided into a training set and validation subsets. We implemented three advanced deep learning architectures—ResNet50, DenseNet-121, and Vision Transformer (ViT)—to assess their effectiveness in accurately classifying fungal species. By utilizing images from early growth stages (days 2–3.5) for training and testing and later stages (days 4–7) for validation, our approach shortens the fungal identification process by 2–3 days, significantly reducing the associated workload and costs. Among the models, the Vision Transformer achieved the highest accuracy of 92.6%, demonstrating the effectiveness of our method. This work contributes to the automation of fungal identification, providing a reliable and efficient solution for monitoring fungal growth and diversity over time, which would be useful for culture collections or other institutions that handle a large number of new isolates in their daily work.

Keywords: fungal identification; automatic classification; computer vision; deep learning; artificial intelligence; machine learning



Academic Editor: Muhammad Kabir

Received: 2 December 2024

Revised: 17 January 2025

Accepted: 2 February 2025

Published: 5 February 2025

Citation: Mansourvar, M.; Charylo, K.R.; Frandsen, R.J.N.; Brewer, S.S.; Hoof, J.B. Automated Fungal Identification with Deep Learning on Time-Lapse Images. *Information* **2025**, *16*, 109. <https://doi.org/10.3390/info16020109>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The kingdom Fungi, one of the oldest and most diverse branches of life, encompasses an estimated 2–5 million species [1]. Most of the species in the kingdom play critical roles in terrestrial and aquatic ecosystems and are therefore vital to human health [2]. Mycology, the study of fungi, has advanced into molecular research, highlighting the necessity of accurate fungal identification for various ecological, agricultural, food, biotechnological, and health-related applications [3]. Accurate fungal identification is crucial for improving human well-being, facilitating effective disease treatment, optimizing agricultural practices, and preventing disease outbreaks. Knowing a fungus to species level can reveal its potential to produce toxins or virulence factors against specific hosts, and whether it is safe for

consumers. Despite their large impact on our society, only about 150,000 fungal species have been described to this date, leaving potentially millions still unidentified [3,4]. Traditional methods of fungal identification, such as microscopic examination and macromorphological inspection, have shaped fungal identification by describing and grouping fungal features and structures into logical keys. These analyses are important and provide essential descriptors and metadata on fungi, but at the same time, this approach is time-consuming, subjective, and prone to human error if not judged with sufficient expertise [5]. A molecular approach that uses the chemical fingerprints of fungi under a set of comparable and standardized growth conditions may also give deep insights into what species produce, but these approaches would have to deal with potentially significant intraspecies variations [6]. The highest confidence in identification has been observed through analysis of the kingdom-wide rDNA loci and the addition of relatively conserved protein-coding genes. This molecular genetic-level identification has proven to give good discriminatory power across the fungal tree of life, at least to genus level, with the use of standard molecular biology laboratory equipment and DNA sequence technologies that continue to fall in price, making the approach accessible for many labs. Nonetheless, the success of sequence comparison-based identification is highly reliant on the quality of the available taxonomic classification gene-marker sequence database and the quality of the identification for the fungal species used for forming the database.

Consequently, we support that low-cost automated methods for fungal identification are essential for advancing mycological research and provide a better foundation for isolating and providing initial characterizations of fungi.

Computer vision technologies are rapidly advancing as tools to enhance mycological research through automated fungal classification [7–9]. Recent developments in artificial intelligence (AI) and machine learning, particularly convolutional neural networks (CNNs), have led to significant improvements in the automation of fungal identification using static images of fungal spores [10,11]. These technologies facilitate the swift analysis of large image datasets, enabling timely and informed decision-making in healthcare and agriculture [12,13]. Automated classification reduces the risk of human error and supports research by providing consistent and reproducible results [14].

Convolutional neural networks (CNNs), including various deep learning (DL) techniques, have demonstrated remarkable success in image-based tasks, especially with static images in medical and agricultural fields. DL architectures, such as ResNet and DenseNet, are widely adopted due to their ability to automatically extract relevant features from raw image data without extensive preprocessing [15].

In studies on fungal identification, CNNs have significantly improved classification accuracy, particularly in scenarios with large, labeled datasets [16]. However, existing applications primarily focus on static images of fungal spores. For example, S. S. Gaikwad et al. (2021) used CNN models to classify fungi affecting apple plant leaves, achieving an accuracy of 88.9% with images from a plant pathology dataset. This study underscores the potential of deep learning for agricultural disease management by enabling early detection and treatment of fungal infections in crops [17]. Similarly, L. Picek et al. (2022) introduced the Danish Fungi 2020 dataset, which aids in the fine-grained classification of fungal species, addressing challenges related to unbalanced class distributions and complex class hierarchies [18]. Koo et al. (2022) developed a deep learning model with a regional CNN to detect fungal hyphae in microscopic images, achieving high sensitivity (95.2% for 100× and 99% for 40× magnification models) and specificity (100% for 100× and 86.6% for 40× magnification models) [19]. Gao et al. (2021) combined an automated microscope with a ResNet-50 model to enhance fungal detection in dermatological samples, showing high sensitivity (99.5% for skin, 95.2% for nails) [20]. M. A. Rahman et al. (2023) explored

various deep CNN models for classifying pathogenic fungi, achieving a top accuracy of 65.35% with the DenseNet model [21]. Cinar et al. (2023) and Gümüő (2024) proposed leveraging deep learning techniques, including vision transformers, to improve fungal detection accuracy in microscopic images [22,23]. While significant progress has been made in static image classification, analyzing time-lapse image sequences presents new challenges. Fungal morphology changes over time, and existing techniques are ill equipped to handle these temporal changes, which are crucial for accurate real-time monitoring of fungal development.

In this study, we propose an automatic fungal identification approach using deep learning techniques applied to time-lapse images focusing on the group of filamentous fungi called molds. Our method aims to bridge the gap between traditional methods and automated analysis, providing a more reliable, quicker, and efficient solution for monitoring fungal growth and identifying species over time. By employing advanced neural networks, including ResNet50, DenseNet-121, and a Vision Transformer (ViT), we aim to classify fungi based on their temporal morphological changes on a selection of standardized growth media that challenge the fungi in various ways. This means that we can compare each individual fungal strain on six different media based on parameters captured in the images, such as mycelium development, sporulation, pigmentation of media, and fungal entities as well as the ability to grow on a specific medium over time. We evaluate our method using a diverse dataset of time-lapse images and compare the performance of these models in terms of accuracy and efficiency.

2. Materials and Methods

The overview of the methods developed for automatic fungal identification used in this study is shown in Figure 1. First, we present the preparation of the fungal image dataset. Second, we outline the model and deep learning-based architecture used for classification.

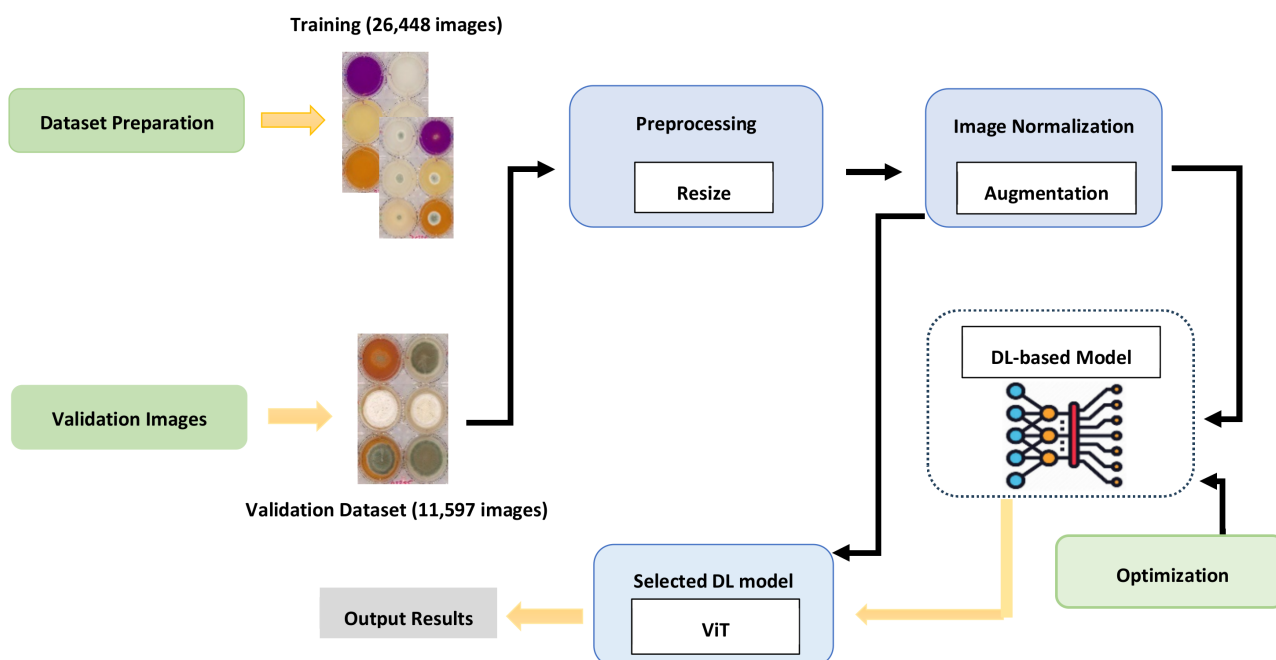


Figure 1. Overview of the automated fungal identification workflow. The workflow consists of three main stages: data preparation, model training, and evaluation. In the data preparation stage, high-resolution time-lapse images of fungal colonies are collected using a Reshape Automated Imaging System over 5–7 days. The dataset is annotated and preprocessed, including resizing images to 224 × 224 pixels and applying data augmentation techniques like Sobel edge detection. In the model

training stage, three advanced deep learning architectures (ResNet50, DenseNet-121, and Vision Transformer (ViT)) are trained on the prepared dataset to classify fungal species based on morphological traits. The models are optimized using the Adam optimizer and trained until convergence. In the evaluation stage, the trained models are tested using separate test and validation datasets to assess their accuracy, precision, recall, and F1 score. The performances of each model are compared to determine the most effective architecture for fungal species classification. This figure illustrates the systematic integration of deep learning methodologies with fungal image data to develop a robust automated identification system.

2.1. Preparation and Inoculation of 6-Well Plates for Image Capture

A total of 110 different fungal species were cultivated in sterile flat-bottom polystyrene 6-well plates (Sigma-Aldrich, St. Louis, MO, USA), in which 3–5 mL of six distinct agar media was assigned to individual wells: malt extract agar oxoid (MEAox), yeast extract sucrose (YES), creatine sucrose agar (CREA), oatmeal agar (OAT), Czapek yeast extract agar (CYA), and potato dextrose agar (PDA) (Figure 2, left panel, DTU—IBT at <https://dtu.bio-aware.com/> for recipes, accessed on the 14 November 2024).

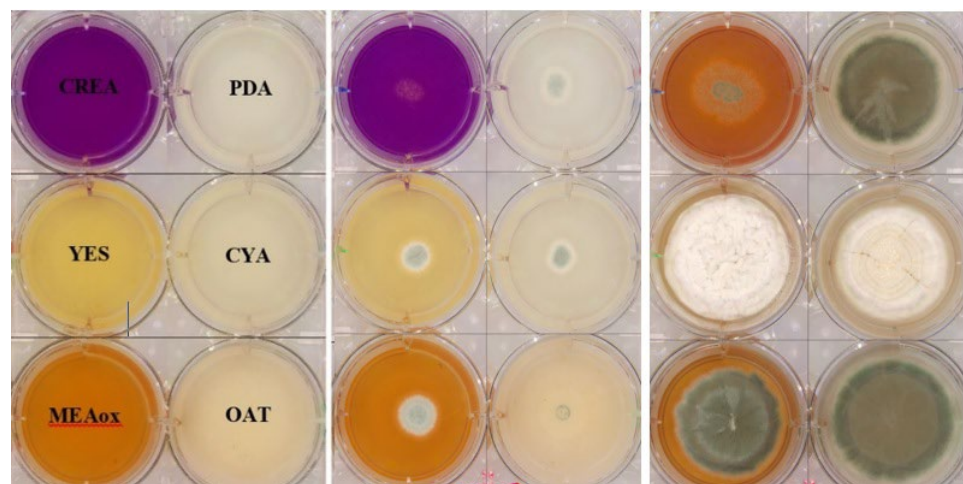


Figure 2. Cultivation and imaging approach at different growth stages. The left panel illustrates the preparation and cultivation medium setup in the wells of a 6-well plate. The center panel shows representative image of IBT 23255 during early growth stages (days 1–2), characterized by initial colony formation and limited pigmentation. The right panel displays images from later growth stages (days 5–6), highlighting more complex morphological features such as increased colony density, sporulation, and pigmentation.

To prepare the inoculum, spores were harvested from solid cultures of the selected fungal strains and suspended in a 50% glycerol solution. A sterile inoculation needle (orange polystyrene inoculation needle, Sarstedt) was then dipped into the spore solution to collect spores (<1000), which were subsequently used for a single-point inoculation of each well of the 6-well plates. The inoculated plates were incubated at 25 °C for 5–7 days in the Reshape Automated Imaging System (v1, Reshape Biotech, Copenhagen, Denmark), which captured time-lapse images at 30 min intervals. Images retrieved after cultivation were divided into training and validation sets as described in the Results section. Early-stage images (days 2–3.5) were used for training and testing, while later-stage images (days 3.5–7) were reserved for validation to evaluate the model’s ability to identify unseen data. To illustrate the progression of fungal growth over time, Figure 2 (center panel) shows representative images from early growth stages (days 1–2), while the right panel displays images from later stages (days 5–6).

As of July 2024, the dataset comprised 26,451 high-resolution images (4119 × 6,225,185 pixels) saved in JPG format. The identity of each strain in the image was previously determined by morphological traits or by PCR-based barcode sequence analysis. The dataset comprised 110 species across 35 genera at that point; however, the number of strains and images will expand continuously. The genus–species relationships represented in the dataset are illustrated in Figure 3, and a summary of the dataset’s quantitative parameters is provided in Table 1.

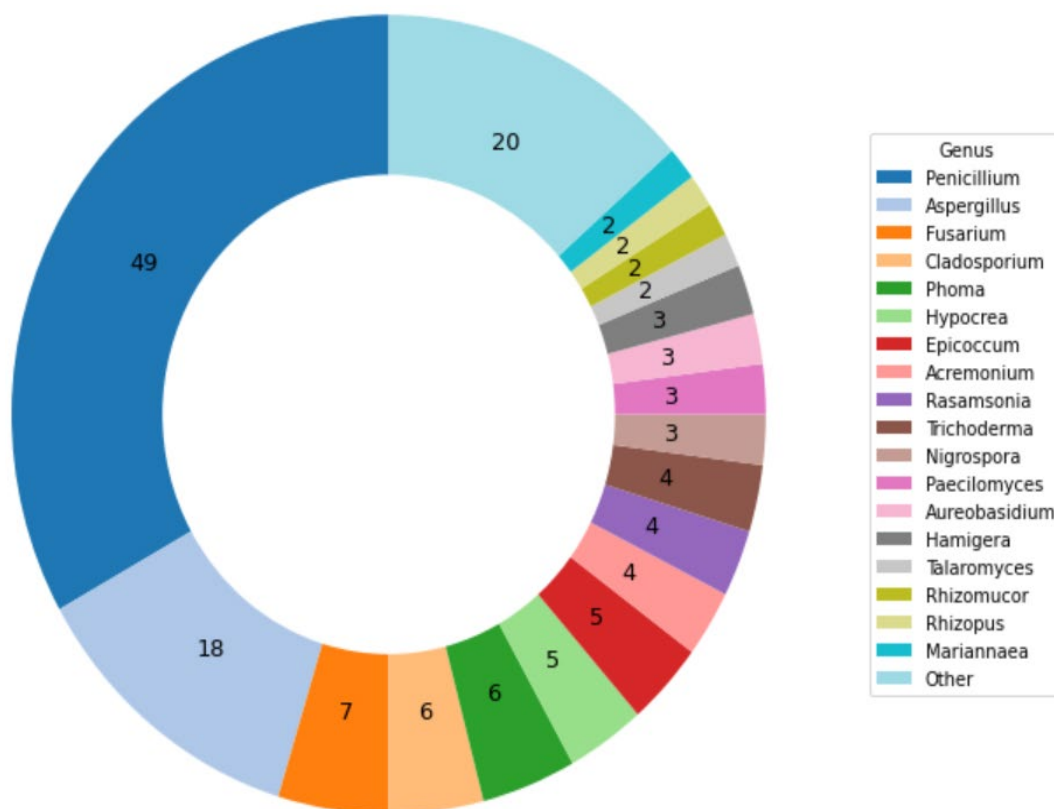


Figure 3. Genus–species relationships in the dataset. Numbers in the colored sections show how many species were analyzed in each genus represented in the study.

Table 1. Distribution of images across training, testing, and validation sets.

Dataset	Number of Images
Train set size (70%)	10,027
Test set size	4827
Validation set size	11,597

The dataset was divided into training, test, and validation sets. The training set consisted of 10,027 images from days 2–3.5, representing 70% of the dataset. The test set comprised 4827 images from days 2–3.5, constituting 30% of the dataset, and was used to evaluate the model’s performance on unseen data after training. The validation set included images captured during days 3.5–7, containing 11,597 images, and served to assess the model’s robustness and generalization across later growth stages. The decision to use later growth stages for validation was based on their increased morphological complexity, which challenges the models to generalize effectively. At the same time, focusing training and testing on earlier stages (days 2–3.5) demonstrates the potential to achieve high classification

accuracy during earlier growth stages, reducing the workload and costs associated with long-term data collection for future studies.

In this study, we address a multi-class classification problem focused on the identification of fungal species based on their morphological features in time-lapse images. Unlike binary classification tasks (e.g., distinguishing between “healthy” and “not healthy” samples), our objective is to accurately classify 110 species. This requires the model to capture subtle morphological differences between closely related species, which presents a significant challenge due to the visual and structural similarities within certain genera, such as *Aspergillus* and *Penicillium*.

This multi-class classification task is further complicated by variations in growth patterns, medium effects, and image quality. Advanced deep learning models, as described in the subsequent sections, were employed to address these challenges and achieve robust species-level identification.

The larger size of the validation set was due to the inclusion of images from days 3.5 to 7, when the cultured strains exhibited more pronounced morphological changes as they progressed through different growth stages. This allowed for a more thorough evaluation of the model’s ability to classify species accurately across a wider range of fungal development stages, including later, more complex growth phases.

2.2. Method

2.2.1. Preprocessing

The image dataset from both the testing and training directories was loaded and resized to 224×224 pixels. This resizing step was important for optimizing computational efficiency and accelerating model training [24]. Standardizing the image dimensions ensured consistency across the dataset and reduced the computational load, aligning the dataset with the input size requirements of many pre-trained models, which commonly expect images of this resolution.

While using full-scale, high-resolution images (4119×6225 pixels) could preserve more fine-grained details, the computational power and time required for such an approach would be significantly higher. Resizing provides a practical balance between computational efficiency and image quality. The chosen resolution ensures that key morphological features essential for fungal species classification are preserved, as evidenced by the high classification accuracy achieved in this study. This trade-off allows for rapid hyperparameter tuning and testing, making the approach both efficient and effective.

2.2.2. Data Augmentation

Data augmentation techniques are widely used to increase the diversity of datasets and address class imbalance [25]. In this research, the Sobel operator was employed to augment the image dataset, enriching the training data. The Sobel operator is a powerful tool in image processing, primarily used for edge detection [26]. It computes the gradient of image intensity in both the horizontal (X) and vertical (Y) directions. These gradients are combined to produce a gradient magnitude image, effectively highlighting the edges present in the original image. Edge detection is crucial because edges often represent important features within an image, such as boundaries between different regions or objects.

Figure 4 illustrates the preprocessing workflow applied to the images of six fungal cultures in a six-well plate. The figure includes examples of the original high-resolution images (4119×6225 pixels), the resized images (224×224 pixels), and the Sobel-processed edge-enhanced images. These steps optimize computational efficiency, ensure consistency across the dataset, and emphasize morphological features such as edges and colony boundaries, which may be important for accurate classification.

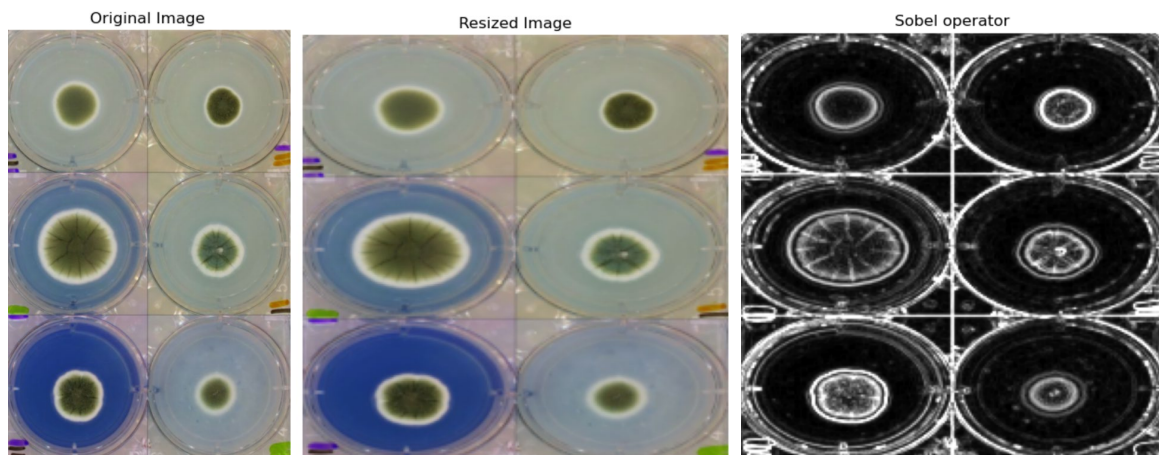


Figure 4. Illustration of preprocessing steps applied to fungal images. This figure demonstrates the preprocessing workflow for images of the six-well plates with fungi used in this study. The first column shows the original high-resolution images (4119×6225 pixels), the second column displays the resized images (224×224 pixels) prepared for deep learning model input, and the third column presents the edge-enhanced images generated using the Sobel operator. These preprocessing steps ensure consistency in image dimensions, optimize computational efficiency, and highlight morphological features such as edges and boundaries, which are crucial for fungal classification.

2.2.3. Deep Learning Architectures for Fungal Classification

The performance of the ResNet50, DenseNet-121, and Vision Transformer (ViT) deep learning architectures, training them using the same set of input resized 224×224 -pixel images, were evaluated for their success in classifying fungal species. These architectures were selected due to their proven effectiveness in image classification tasks and their ability to handle large and complex datasets.

ResNet50 was chosen for its balance between depth and computational efficiency [24]. This architecture is implemented from scratch, allowing users to easily specify the input size and number of channels, with the network structure adjusting automatically. ResNet50 [27] is a deep residual network that uses skip connections to mitigate the vanishing gradient problem in very deep networks. By enabling the flow of gradients through layers, it improves the training of deep networks. ResNet50 has been widely used in image classification tasks due to its efficiency and ability to model hierarchical features within images.

DenseNet-121 was also selected for its unique architecture, which differentiates it from models like ResNet. In DenseNet, layers are connected in a densely connected pattern, where each layer receives input from all preceding layers [28]. This creates $L(L + 1)/2$ connections for L layers, enhancing the flow of information and improving gradient propagation. DenseNet-121 [29] is a densely connected convolutional network that improves information flow between layers. In DenseNet, each layer receives inputs from all preceding layers, allowing the model to reuse features and reduce the number of parameters. This architecture is particularly effective in handling complex visual patterns and has demonstrated superior performance in various image recognition tasks [30].

Vision Transformers (ViTs) are an adaptation of the transformer architecture, originally developed for natural language processing, to the field of computer vision [31]. ViTs partition images into patches, treating them as sequences similar to words in a sentence. This allows the model to capture long-range dependencies and global relationships in images, making it highly effective for tasks involving detailed pattern recognition. ViTs are particularly advantageous for datasets with complex spatial relationships, such as time-lapse images of fungal growth, where long-range interactions between pixels can be important [32,33].

The model development and training were conducted using the Python (Version 3.12.3) with the PyTorch framework (Version 2.2.2). All models were trained on a high-performance computing (HPC) cluster at the Technical University of Denmark, utilizing NVIDIA Tesla V100 GPUs (16/32 GB). These GPUs, based on NVIDIA's Volta architecture, provided the computational power necessary for intensive machine learning and data analytics tasks.

The implementation details and source code for the proposed classifier models are available on GitHub: https://github.com/BDD-G/Fungi_classification, accessed on the 14 November 2024.

2.2.4. Model Training and Hyperparameters

The models were trained using the Adam optimizer [34], known for its adaptive learning rates and efficiency in handling sparse gradients. A learning rate of 1×10^{-4} was utilized for all models, which provided a good balance between convergence speed and training stability. The models were trained for up to 1000 epochs with varying batch sizes of 16, 32, and 64 to assess their performance under different conditions. Early stopping was implemented to prevent overfitting, monitoring the validation loss with a patience of 10 epochs.

3. Results

3.1. Cultivations and Images

The dataset comprises 110 different fungal species, each representing a distinct computational category (classification class). A total of 26,451 high-resolution images (4119×6225 pixels) were collected, with the number of images per class ranging from 120 to 218. This distribution reflects the variability in image capture due to differences in fungal growth rates and colony morphology. To provide an overview of the dataset composition, Table 2 summarizes the number of images for the top 25 categories, while the remaining 85 classes, each with 121 images, are grouped together for simplicity.

Table 2. Summary of image distribution across classification classes.

Classification Class Name	Number of Images
<i>Penicillium svalbardense</i>	218
<i>Epicoccum nigrum</i>	217
<i>Rhizomucor pusillus</i>	216
<i>Penicillium onobense</i>	212
<i>Nigrospora oryzae</i>	198
<i>Aspergillus uvarum</i>	198
<i>Penicillium restrictum</i>	197
<i>Penicillium rotoruae</i>	197
<i>Paecilomyces maximus</i>	197
<i>Penicillium canescens</i>	184
<i>Phoma pomorum</i>	182
<i>Penicillium wotroi</i>	179
<i>Mariannaea elegans</i>	177
<i>Penicillium scabrosum</i>	175
<i>Penicillium ochrochloron</i>	174
<i>Aspergillus flavus</i>	174

Table 2. Cont.

Classification Class Name	Number of Images
<i>Fusarium tricinctum</i>	174
<i>Penicillium glabrum</i>	173
<i>Purpureocillium lilacinum</i>	173
<i>Penicillium fagi</i>	170
<i>Hypocrea pulvinata</i>	169
<i>Penicillium janczewskii</i>	167
<i>Hamigera avellanea</i>	165
<i>Rasamsonia piperina</i>	163
<i>Penicillium olsonii</i>	121
Others (86 classes with 121 each)	10,406
Total	26,451

The cultivations and imaging provided data displaying the effects of different media on the fungal growth rates, sporulation, pigmentation, and density of the colony, yielding medium-specific phenotypic signatures for each strain (for example, as shown for IBT 23138 after 3 days in Figure 2, right panel).

The classification models leveraged not only these morphological features but also visual characteristics such as brightness, contrast, saturation, and hue, which varied across the dataset. These features, captured in the time-lapse images, provided the models with essential information to differentiate fungal species effectively. Figure 5 showcases examples of fungal species from three randomly selected classes (*Sordaria fimicola*, *Aspergillus uvarum*, and *Aspergillus occultus*), highlighting variations in these features across different growth stages. Such variations contribute significantly to the classification accuracy achieved by the models.

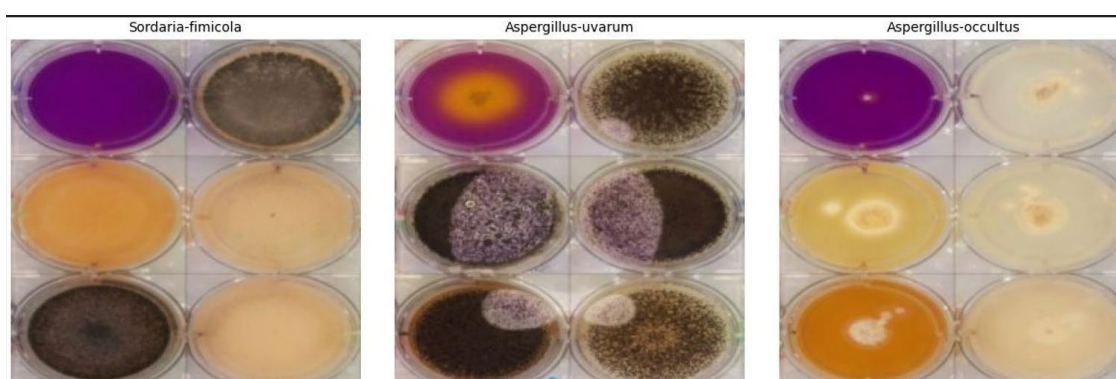


Figure 5. Representative examples of fungal species from three classification classes. Images of *S. fimicola*, *A. uvarum*, and *A. occultus* highlight variations in brightness, contrast, saturation, and hue, which play a critical role in the classification process.

For some species, the mycelium (fungus) grew to cover the entire well between days 4 and 7, but variations as a function of genera and species were observed. Manual inspections of the experiments were implemented to approve or reject species identifications regardless of the computed score. This step was critical to ensure that the final identification was according to the taxon determined for the presumed species selected for the analysis. The validation process involved manually reviewing images captured at different growth stages

over several days to confirm the identity of a particular fungal species, as illustrated in Figure 6. Variations in the incubation time needed for successful validation arose from the differing complexities of identifying fungal species, which depend on their growth stages and the morphological characteristics exhibited at those stages.

Distribution of the incubation time needed of correct classification

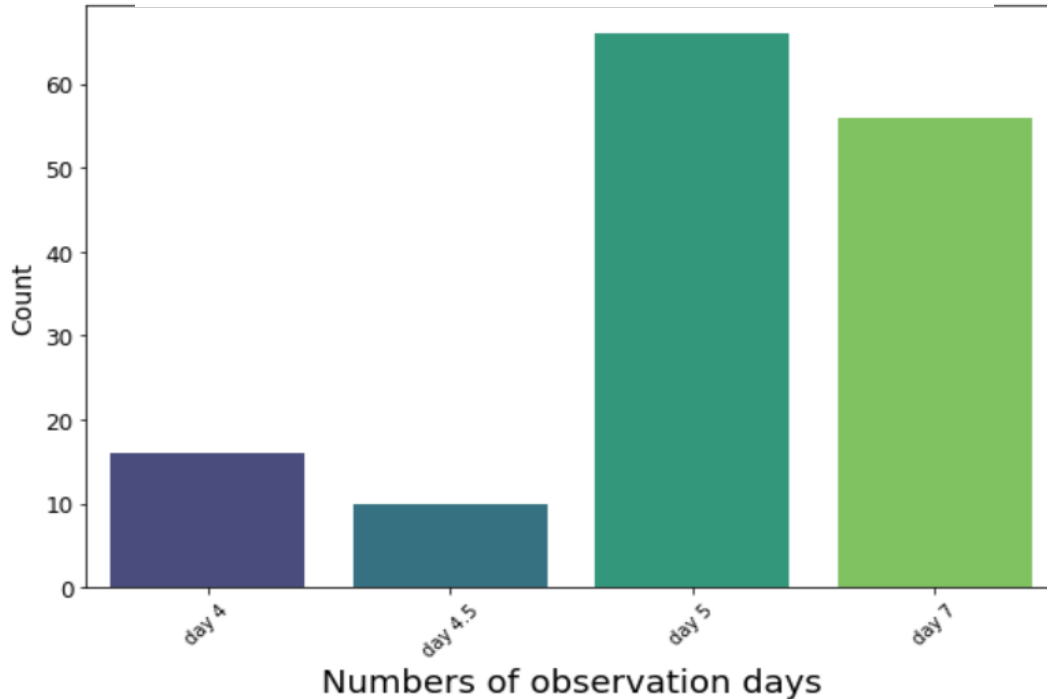


Figure 6. Validation time across different days of fungal growth. The graph illustrates the number of image validations performed on days 4 through 7, highlighting the variability in growth completion across different fungal species. Early validations occur as some species reach maturity sooner, while others require longer incubation periods for accurate identification.

3.2. Evaluation and Performance of Models

For the deep learning classifiers, the dataset was divided into three subsamples: 70% of the images were used to train the model, 30% were reserved for testing, and 11,597 unseen images, selected from the last 4 days of fungal growth, were utilized for evaluation to assess the accuracy of the selected models.

The performance of the three deep learning architectures—ResNet50, DenseNet-121, and Vision Transformer (ViT)—was evaluated using the following metrics:

Accuracy: Accuracy measures the overall correctness of the model by calculating the ratio of correctly predicted instances to the total number of instances [35], as shown in Equation (1).

$$Accuracy(Acc) = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + false\ Negative} \quad (1)$$

Precision: Precision is the ratio of true positive predictions to the total number of positive predictions, indicating the model's ability to correctly identify relevant instances. This is expressed in Equation (2).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

Recall: Recall, the ratio of true positive predictions to the total number of actual positives, measures the model's effectiveness in capturing all relevant instances, as shown in Equation (3).

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (3)$$

F1 score: The F1 score, the harmonic mean of precision and recall, provides a balanced measure of the model's performance, especially when dealing with imbalanced datasets [36]. This is represented in Equation (4).

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

The models were trained on a high-performance computing (HPC) cluster, with ResNet50 taking 1.11 h, DenseNet-121 taking 3.61 h, and ViT taking 1.18 h to complete training. Table 3 presents a comparison of the results obtained from cross-validation for the three deep learning architectures. Figure 7 displays histogram plots illustrating the metrics for the three models. Among the three models, the Vision Transformer (ViT) demonstrated the best overall performance, with the highest accuracy, precision, recall, and F1 score. This indicates that ViT is particularly effective at identifying fungal species with minimal misclassification, capturing relevant morphological details across the dataset.

Table 3. Performance comparison of ResNet50, DenseNet-121, and Vision Transformer (ViT) architectures.

Architecture	Accuracy	Precision	Recall	F1 Score
Resnet	76.75%	89.35%	89.75%	88.54%
DenseNet121	86.77%	93.08%	91.10%	89.20%
ViT-16	92.64%	95.77%	96.35%	93.84%

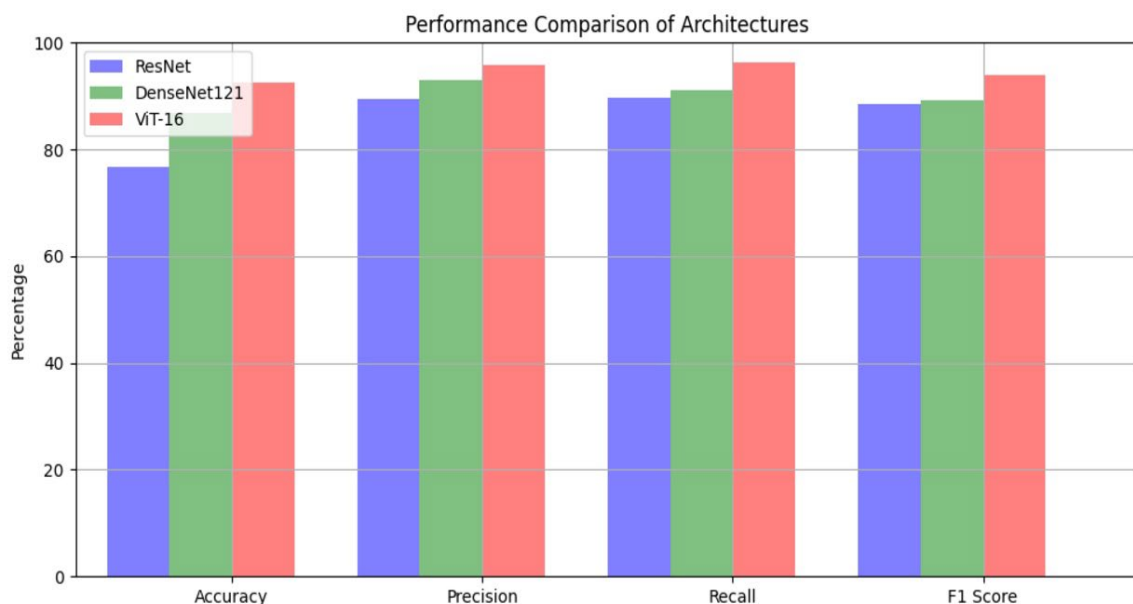


Figure 7. Performance comparison of ResNet, DenseNet-121, and Vision Transformer (ViT) architectures. This histogram illustrates the comparative performance of three deep learning models across four key metrics: accuracy, precision, recall, and F1 score. The Vision Transformer (ViT) demonstrates superior performance in all metrics, followed by DenseNet-121 and ResNet.

DenseNet-121 demonstrated a strong performance in terms of accuracy and precision, achieving results superior to those of ResNet50 but slightly below those of ViT. However,

its longer training time compared to both ResNet50 and ViT suggests that while the architecture can achieve high accuracy, it may not be the most computationally efficient choice for this task. ResNet50, despite its computational efficiency, exhibited lower performance metrics across all categories, making it less suitable for tasks requiring high accuracy in fungal classification.

To further evaluate the classification performance, a confusion matrix was generated for the Vision Transformer (ViT), the model with the highest accuracy. The matrix (Figure 8) provides a detailed breakdown of correctly classified and misclassified images for each fungal species. The results show that most misclassifications were concentrated among morphologically similar species, such as those within the *Aspergillus* and *Penicillium* genera. These findings highlight the challenges of distinguishing closely related species and suggest opportunities for further improvement, including the incorporation of additional morphological data, as discussed in the following section.

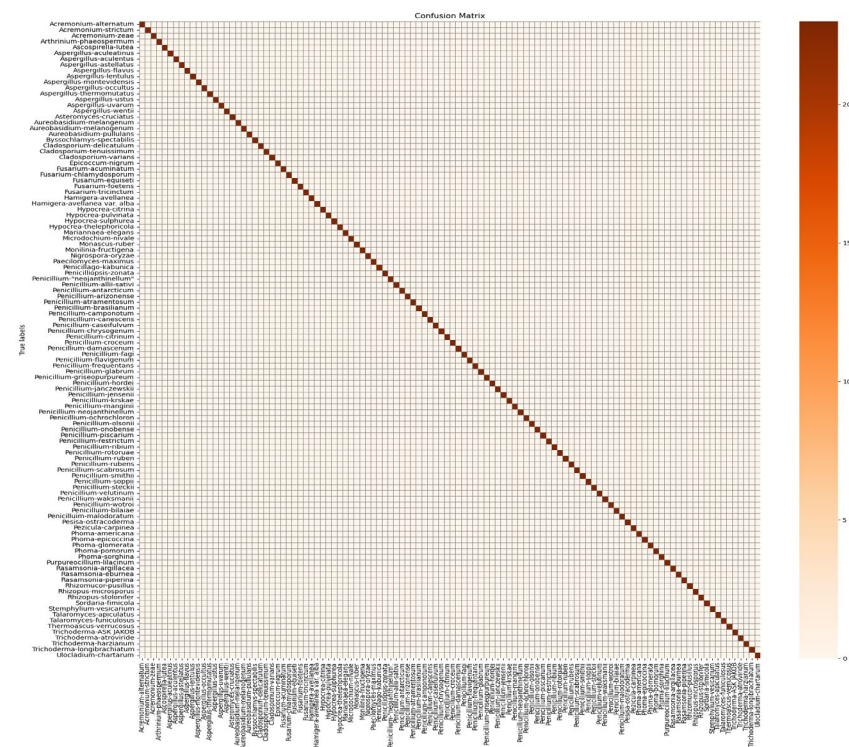


Figure 8. Confusion matrix for the Vision Transformer (ViT) classification model. The matrix displays the number of correctly classified images (diagonal) and misclassified images (off-diagonal) for each fungal class.

4. Discussion

Fungal identification plays a crucial role in addressing various agricultural, environmental, and medical challenges. Accurate detection and classification of mold fungi are essential for effective disease management in crops, ensuring food security, and controlling fungal infections in humans and animals. Several studies have explored the use of deep learning for fungal classification, focusing primarily on static images of fungal spores or cultures. For example, Gaikwad et al. (2021) achieved 88.9% accuracy using CNNs to classify fungi affecting apple plant leaves [17], while Koo et al. (2022) applied regional CNNs to detect fungal hyphae in microscopic images, reporting high sensitivity and specificity [19]. Similarly, Picek et al. (2022) introduced the Danish Fungi 2020 dataset for fine-grained macrofungal (fruiting bodies) species classification, addressing challenges of unbalanced class distributions [18].

In contrast, our study focuses on time-lapse images, capturing temporal morphological changes that static images cannot represent. This approach enables our models, particularly the Vision Transformer (ViT), to achieve a superior classification accuracy of 92.64%, demonstrating the potential of temporal data for fungal identification. These findings complement previous research by extending the application of deep learning to more complex datasets involving dynamic growth patterns.

Given the complexity and diversity of fungal species, traditional identification methods by morphological inspection are often slow and inaccurate, underscoring the need for automated solutions. Leveraging artificial intelligence (AI) and deep learning algorithms for fungal identification offers significant improvements in diagnostic precision, efficiency, and cost effectiveness.

In this study, three deep learning models—ResNet50, DenseNet-121, and Vision Transformers (ViT)—were assessed for their efficacy in classifying fungal species. The selection of a learning rate of 1×10^{-4} was crucial in ensuring stable convergence across all models. A lower learning rate helped prevent the models from overshooting the minimum of the loss function, particularly important given the complexity of the dataset. The Adam optimizer, combined with this learning rate, facilitated efficient training and contributed to the high accuracy achieved by the ViT model.

Upon comparison, ViT emerged as the best-performing model with an accuracy of 92.64%, followed by DenseNet-121 (86.77%) and ResNet50 (76.75%). The superior performance of ViT can be attributed to its ability to effectively capture both local and global features of fungal growth in time-lapse images. The transformer-based architecture allows ViT to recognize fine morphological differences, which is especially critical when classifying fungi at different growth stages.

The classification accuracy achieved by the Vision Transformer (ViT) model, reaching 92.64%, is highly satisfactory given the inherent challenges of fungal identification. Manual classification by experts often involves subjective assessments of morphological traits, which can be inconsistent and require extensive time and resources. In contrast, the deep learning models employed in this study provide an automated solution that achieves comparable accuracy while significantly reducing the time required for identification.

While expert-level classification may incorporate genetic analyses for confirmation, such methods are labor-intensive and expensive. The use of time-lapse imaging and AI-based classification in this study provides a practical and efficient alternative, offering reliable predictions within a much shorter timeframe. Future studies could involve direct comparisons with expert classifications to further validate the model's performance.

In contrast, ResNet50's convolutional architecture struggled with the complex patterns in the data, leading to a comparatively lower accuracy. DenseNet-121, although outperforming ResNet50, fell short of ViT, likely due to its reliance on dense connections, which may not have been sufficient to capture the intricate variations in the dataset.

The performance metrics achieved by the models, particularly ViT with an accuracy exceeding 90%, demonstrate their robustness and reliability, even when faced with the inherent complexity of the fungal dataset. This reliability is particularly promising given the morphological similarities between species within the same genus, such as *Aspergillus* and *Penicillium*, which present significant classification challenges.

To our knowledge, this study is the first to classify and identify time-lapse fungal (mold) images using AI. The application of cross-validation ensured the robustness of the models, and by using images from day 2 to day 3.5 for training and testing and images from days 3.5 to 7 for validation, the identification process was shortened by 2–3 days. This reduction in time significantly cuts the resources and costs typically associated with

fungal identification. Furthermore, our model predicted the top five closest classifications, offering a flexible and efficient approach to fungal species identification.

Despite these advancements, there are several potential limitations to this study. One such limitation is the variability in image quality, which can be influenced by the resolution of the images (4119×6225 pixels) and environmental conditions during image capture. Higher-resolution images generally improve the model's ability to detect fine morphological details, but even with high-resolution images, factors such as lighting conditions, focus, and fungal growth stage can significantly affect image clarity. In field conditions, where environmental factors are less controlled, the robustness of the model may be challenged by inconsistencies in image quality, potentially leading to misclassification.

Furthermore, while the pixel resolution in our setup was sufficient for this specific dataset, it is important to note that similar setups could be applied for fungal identification in other environments or on other datasets. However, the generalizability of the setup—whether other users can trust it for their fungal identification tasks—depends on the quality and consistency of the images captured. Low-quality images or images captured under suboptimal conditions could undermine the model's accuracy. Therefore, achieving optimal results requires maintaining a standard level of image resolution and ensuring consistent conditions during data collection.

Lastly, the accuracy of this approach requires that the training and validation sets use correctly identified and updated species, in these cases from the IBT Culture Collection. Therefore, we have selected strains for the dataset that were previously identified by a polyphasic approach, meaning a combination of morphological identification on more than one cultivation medium and chemical characterization by HPLC-DAD-MS or amplicon sequencing of standard loci used for taxonomic classification.

Increasing the size of the training dataset, particularly by incorporating a broader range of species and environmental conditions, could help improve the model's generalizability and performance. However, the exact amount of data required to achieve top-level accuracy remains inconclusive and would likely vary depending on the fungal species and environmental variability encountered.

Future studies will focus on expanding the dataset by collecting more images from a wider variety of species and genera. This expansion would enhance the model's generalization capabilities and further optimize its performance in differentiating between closely related species. While this study focused on whole-image classification, alternative frameworks such as YOLO, which are designed for object detection, could be explored in future studies. These frameworks may be particularly useful for tasks requiring spatial localization of fungal growth patterns or the segmentation of specific fungal features within an image. Additionally, incorporating advanced data augmentation techniques and exploring ensemble methods could further improve classification accuracy.

5. Conclusions

In this work, a machine learning model based on deep learning was developed for the automatic identification of fungi using time-lapse images. Three prominent deep learning architectures—ResNet50, DenseNet-121, and Vision Transformer (ViT)—were evaluated, with ViT outperforming the others, achieving a classification accuracy of 92.64%. The superior performance of ViT, particularly in recognizing fine morphological differences, underscores the potential of transformer-based models for handling complex biological data. This study demonstrates that deep learning, when combined with advanced data augmentation and preprocessing techniques, can significantly reduce the time required for fungal identification by 2–3 days.

We believe that this method can be easily extended to more diversified datasets by incorporating additional preprocessing steps that unify the input data, potentially leading to even more accurate and efficient fungal classification. Moreover, by refining the model and expanding the dataset, the system could be adapted to identify a wider range of fungal species and genera, enhancing its real-world applicability.

Looking ahead, we plan to develop a publicly accessible online API service for the automatic classification of fungal images. Such a service would be particularly useful for researchers, agronomists, and medical professionals, as it could assist in excluding common fungi from analysis, allowing experts to focus on rarer or more pathogenic species. This would streamline the identification process and offer a cost-effective and scalable solution for fungal classification. Future studies should explore expanding the dataset and refining the model to further enhance its generalization and performance across different fungal species and environmental conditions. This asset will also be extremely valuable for a preliminary identification of species in culture collections that isolate or receive many fungal cultures.

Author Contributions: M.M.: writing—original draft, methodology, formal analysis, software, validation, visualization, supervision. K.R.C.: formal analysis, methodology, software, validation, visualization, writing—review and editing. R.J.N.F.: writing—review and editing, supervision. S.S.B.: writing—review and editing, data acquisition. J.B.H.: writing—review and editing, conceptualization, methodology, supervision, resources. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study is not publicly available but can be accessed upon reasonable request. For access, please contact marjma@dtu.dk or jblni@dtu.dk.

Acknowledgments: The authors acknowledge the ‘Smarter AgroBiological Screening’ (SABS) project (grant number 0224–00092B) for providing partial salary support for R.J.N.F., J.B.H., and S.S.B.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Li, Y.; Steenwyk, J.L.; Chang, Y.; Wang, Y.; James, T.Y.; Stajich, J.E.; Spatafora, J.W.; Groenewald, M.; Dunn, C.W.; Hittinger, C.T.; et al. A genome-scale phylogeny of the kingdom Fungi. *Curr. Biol.* **2021**, *31*, 1653–1665.e5. [[CrossRef](#)] [[PubMed](#)]
2. Money, N.P. Fungal Diversity. In *The Fungi*, 3rd ed.; Academic Press: Cambridge, MA, USA, 2016; pp. 1–36. [[CrossRef](#)]
3. Mendonça, A.; Santos, H.; Franco-Duarte, R.; Sampaio, P. Fungal infections diagnosis—Past, present and future. *Res. Microbiol.* **2022**, *173*, 103915. [[CrossRef](#)] [[PubMed](#)]
4. Bhunjun, C.S.; Niskanen, T.; Suwannarach, N.; Wannathes, N.; Chen, Y.J.; McKenzie, E.H.; Maharachchikumbura, S.S.; Buyck, B.; Zhao, C.L.; Fan, Y.G.; et al. The numbers of fungi: Are the most speciose genera truly diverse? *Fungal Divers.* **2022**, *114*, 387–462. [[CrossRef](#)]
5. Aboul-Ella, H.; Hamed, R.; Abo-Elyazeed, H. Recent trends in rapid diagnostic techniques for dermatophytosis. *Int. J. Vet. Sci. Med.* **2020**, *8*, 115–123. [[CrossRef](#)] [[PubMed](#)]
6. Jana, C.; Raus, M.; Sedlářová, M.; Šebela, M. Identification of fungal microorganisms by MALDI-TOF mass spectrometry. *Biotechnol. Adv.* **2014**, *32*, 230–241.
7. Desingu, K.; Bhaskar, A.; Palaniappan, M.; Chodisetty, E.A.; Bharathi, H. Classification of Fungi Species: A Deep Learning Based Image Feature Extraction and Gradient Boosting Ensemble Approach. 2022. Available online: <http://ceur-ws.org> (accessed on 13 September 2024).
8. Yin, H.; Yi, W.; Hu, D. Computer vision and machine learning applied in the mushroom industry: A critical review. *Comput. Electron. Agric.* **2022**, *198*, 107015. [[CrossRef](#)]

9. Mansourvar, M.; Funk, J.; Petersen, S.D.; Tavakoli, S.; Hoof, J.B.; Corcoles, D.L.; Pittroff, S.M.; Jelsbak, L.; Jensen, N.B.; Ding, L.; et al. Automatic classification of fungal-fungal interactions using deep learning models. *Comput. Struct. Biotechnol. J.* **2024**, *23*, 4222–4231. [[CrossRef](#)]
10. Tahir, M.W.; Zaidi, N.A.; Rao, A.A.; Blank, R.; Vellekoop, M.J.; Lang, W. A fungus spores dataset and a convolutional neural network based approach for fungus detection. *IEEE Trans. Nanobiosci.* **2018**, *17*, 281–290. [[CrossRef](#)]
11. Marandi, B.; Deep, M.K. Advancements in Machine Learning for Detection and Prediction of Infectious and Parasitic Diseases: A Comprehensive Investigation. *Machine Learning. Int. J. Adv. Multidiscip. Sci. Res.* **2024**, *7*, 82–97.
12. Aldogan, K.Y.; Kayan, C.E.; Gumus, A. Intensity and phase stacked analysis of a Φ -OTDR system using deep transfer learning and recurrent neural networks. *Appl. Opt.* **2023**, *62*, 1753–1764. [[CrossRef](#)]
13. Kristensen, K.; Ward, L.M.; Mogensen, M.L.; Cichosz, S.L. Using image processing and automated classification models to classify microscopic gram stain images. *Comput. Methods Programs Biomed. Update* **2023**, *3*, 100091. [[CrossRef](#)]
14. Zhang, Y.; Jiang, H.; Ye, T.; Juhas, M. Deep Learning for Imaging and Detection of Microorganisms. *Trends Microbiol.* **2021**, *29*, 569–572. [[CrossRef](#)] [[PubMed](#)]
15. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)]
16. Picek, L.; Šulc, M.; Matas, J.; Heilmann-Clausen, J.; Jeppesen, T.S.; Lind, E. Automatic Fungi Recognition: Deep Learning Meets Mycology. *Sensors* **2022**, *22*, 633. [[CrossRef](#)] [[PubMed](#)]
17. Gaikwad, S.S.; Rumma, S.S.; Hangarge, M. Fungi Classification using Convolution Neural Network. *Turk. J. Comput. Math. Educ.* **2021**, *12*, 4563–4569.
18. Picek, L.; Šulc, M.; Matas, J.; Jeppesen, T.S.; Heilmann-Clausen, J.; Læssøe, T.; Frøslev, T. Danish Fungi 2020—Not Just Another Image Recognition Dataset. In Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision, WACV 2022, Waikoloa, HI, USA, 3–8 January 2022; pp. 3281–3291. [[CrossRef](#)]
19. Koo, T.; Kim, M.H.; Jue, M.S. Automated detection of superficial fungal infections from microscopic images through a regional convolutional neural network. *PLoS ONE* **2021**, *16*, e0256290. [[CrossRef](#)]
20. Gao, W.; Li, M.; Wu, R.; Du, W.; Zhang, S.; Yin, S.; Chen, Z.; Huang, H. The design and application of an automated microscope developed based on deep learning for fungal detection in dermatology. *Mycoses* **2021**, *64*, 245–251. [[CrossRef](#)]
21. Rahman, M.A.; Clinch, M.; Reynolds, J.; Dangott, B.; Villegas, D.M.M.; Nassar, A.; Hata, D.J.; Akkus, Z. Classification of fungal genera from microscopic images using artificial intelligence. *J. Pathol. Inform.* **2023**, *14*, 100314. [[CrossRef](#)]
22. Cinar, I.; Taspinar, Y.S.; Taspinar, Y.S. Detection of Fungal Infections from Microscopic Fungal Images Using Deep Learning Techniques. *Int. Conf. Adv. Technol.* **2023**. [[CrossRef](#)]
23. Gümüş, A. Classification of Microscopic Fungi Images Using Vision Transformers for Enhanced Detection of Fungal Infections. *Turk. J. Nat. Sci.* **2024**, *13*, 152–160. [[CrossRef](#)]
24. Ikechukwu, A.V.; Murali, S.; Deepu, R.; Shivamurthy, R.C. ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images. *Glob. Transit. Proc.* **2021**, *2*, 375–381. [[CrossRef](#)]
25. Alomar, K.; Aysel, H.I.; Cai, X. Data Augmentation in Classification and Segmentation: A Survey and New Strategies. *J. Imaging* **2023**, *9*, 46. [[CrossRef](#)] [[PubMed](#)]
26. Barshooi, A.H.; Amirkhani, A. A novel data augmentation based on Gabor filter and convolutional deep learning for improving the classification of COVID-19 chest X-Ray images. *Biomed. Signal Process. Control* **2022**, *72*, 103326. [[CrossRef](#)]
27. Shabbir, A.; Ali, N.; Ahmed, J.; Zafar, B.; Rasheed, A.; Sajid, M.; Ahmed, A.; Dar, S.H. Satellite and Scene Image Classification Based on Transfer Learning and Fine Tuning of ResNet50. *Math. Probl. Eng.* **2021**, *2021*, 5843816. [[CrossRef](#)]
28. Nandhini, S.; Ashokkumar, K. An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. *Neural Comput. Appl.* **2022**, *34*, 5513–5534. [[CrossRef](#)]
29. Arulananth, T.S.; Prakash, S.W.; Ayyasamy, R.K.; Kavitha, V.P.; Kuppusamy, P.G.; Chinnasamy, P. Classification of Paediatric Pneumonia Using Modified DenseNet-121 Deep-Learning Model. *IEEE Access* **2024**, *12*, 35716–35727. [[CrossRef](#)]
30. Hou, Y.; Wu, Z.; Cai, X.; Zhu, T. The application of improved densenet algorithm in accurate image recognition. *Sci. Rep.* **2024**, *14*, 8645. [[CrossRef](#)]
31. Thisanke, H.; Deshan, C.; Chamith, K.; Seneviratne, S.; Vidanaarachchi, R.; Herath, D. Semantic segmentation using Vision Transformers: A survey. *Eng. Appl. Artif. Intell.* **2023**, *126*, 106669. [[CrossRef](#)]
32. Yang, J.; Luo, K.Z.; Li, J.; Deng, C.; Guibas, L.; Krishnan, D.; Weinberger, K.Q.; Tian, Y.; Wang, Y. Denoising Vision Transformers. *arXiv* **2024**, arXiv:2401.02957.
33. Yunusa, H.; Qin, S.; Chukkol, A.H.A.; Yusuf, A.A.; Bello, I.; Lawan, A. Exploring the Synergies of Hybrid CNNs and ViTs Architectures for Computer Vision: A Survey. *arXiv* **2024**, arXiv:2402.02941.
34. Ahn, K.; Zhang, Z.; Kook, Y.; Dai, Y. Understanding Adam Optimizer via Online Learning of Updates: Adam is FTRL in Disguise. *arXiv* **2024**, arXiv:2402.01567.

35. Rainio, O.; Teuvo, J.; Klén, R. Evaluation metrics and statistical tests for machine learning. *Sci. Rep.* **2024**, *14*, 6086. [[CrossRef](#)] [[PubMed](#)]
36. Jia, W.; Qin, Y.; Zhao, C. Rapid detection of adulterated lamb meat using near infrared and electronic nose: A F1-score-MRE data fusion approach. *Food Chem.* **2024**, *439*, 138123. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.