A Deep Learning Approach for Identification of Different kind of Fungi

Zeel Jodhani
Northeastern University
Jodhani.z@northeastern.edu

Bhoomi Thakar
Northeastern University
thakar.b@northeastern.edu

Vinas Sukhadiya
Northeastern University
Sukhadiya.v@northeastern.ed

Mihir Kakadiya
Northeastern University
Kakadiya.m@northeastern.ed

Harsh Patel
Northeastern University
Shingala.h@northeastern.edu

Abstract

Traditionally, diagnosis and treatment of fungal infections in humans depend heavily on face-to-face consultations or examinations made by specialized laboratory scientists known as mycologists. This project addresses the challenge of detecting fungi in crops using deep learning techniques. Inspired by recent advancements in convolutional neural networks (CNNs), we propose an efficient and accurate model for automated fungal detection. The dataset used in this study comprises high-resolution images of various fungal infections. To enhance the model's generalization capabilities, a comprehensive data augmentation strategy is employed during training. Our model leverages the transfer learning paradigm, utilizing a pre-trained backbone from the state-of-the-art Efficient Net architecture. The model is fine-tuned on the target dataset to adapt to the specific characteristics of fungal infections. The training process involves a carefully designed loss function and optimization scheme. We employ the cross-entropy loss and the Adam optimizer to ensure effective learning. The training set is split into training and validation subsets, enabling us to monitor the model's performance on unseen data. Additionally, we implement a systematic approach to prevent overfitting and improve robustness. During model evaluation, we employ various metrics such as accuracy and validation loss. The proposed model is tested on a separate test set, and its performance is compared to existing methodologies. Our results demonstrate the effectiveness of the deep learning approach in accurately identifying fungal infections.

1. Introduction

Fungi constituted as an individual kingdom of species from 1969, have been leveraged and ex- ploited in industrial practices such as pharmaceutical production, brewing, baking, and others. The acting of fungi as biological control agents is considered of paramount importance to ecology by contributing to the balance of earth's ecosystems as recyclers and decomposers. The pathogenetic characteristics and diseases caused by some types of fungi have also been recognized by science as an essential matter to humanity. Diagnosis and classification of a fungal infection are made in a laboratory by a specialized biologist known as Mycol-ogist, patient samples such as swabs, blood or scrabs of skin, hair, or nails are processed and cultured in controlled mediums for a range period of 28-31 days (Bosshard, 2011). During the evolutionary process of incubation and growth, the morphological characteristics of the fungi allow Mycologists to suggest a classification diagnosis to medical practitioners such as dermatologists to give early treatment to patients (Pihet et al., 2015). Early treated superficial fungal infections produce less painful and costly treatments to patients. Also, it lowers the percentage of mortality rates, as the 20% increase is related to invasive Candidiasis for patients without early antifungal therapy (Kozel and Wickes (2014).

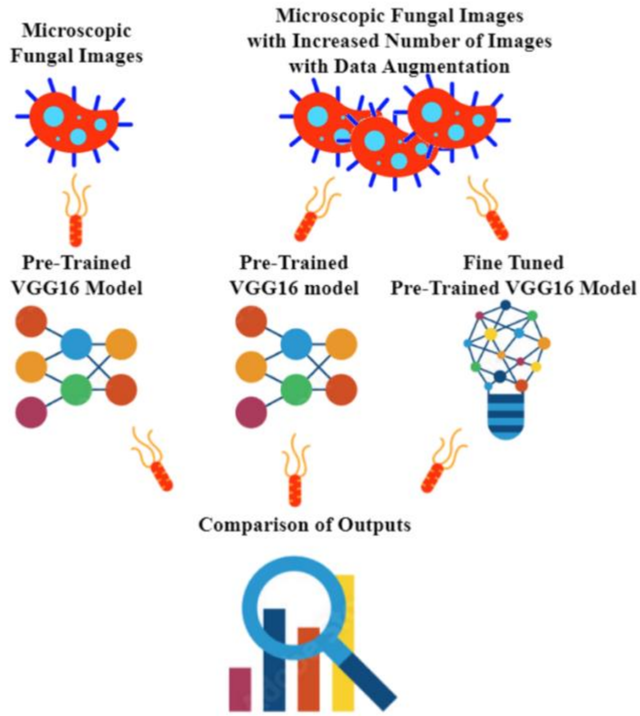
In this project, we explore we explore the use of deep learning techniques to automatically identify type of fungi is present. In our dataset it contains images, and self-generated CVS file refereeing the image name with respective label type.
Thus, to investigate the role of image features in Deep learning, we divide our detection problem into 3 tasks: (1) by using pre-trained model such as efficientb4 where we will use trained weights of the mode and (2) by creating our own CNN [Convolution Neural Network] and (3) at last we will create DNN [Deep Neural Network]

At last, we will compare all results of our 3 different created model with each other to see how pre-trained and our created model differ from each other.

NOTE: We have only one pre-trained model because it is very computational expensive.

2. Materials and Methods

In this section, details regarding the dataset, convolutional neural networks, transfer learning and fine tuning, and the experimental setup are elucidated. The graphical illustration of the study can be found in Fig.1.



A. Dataset Details:

The data used in the study was obtained from 'UCI Machine Learning Repository' . The images in the dataset are stated to be taken from superficial fungal infections caused by yeasts, molds, or dermatophyte fungi. In addition, it was stated that the images were manually divided into 5classes and edited with the help of the subject expert assistance. The dataset consists of 9114 images in 5 different categories. Detailed information about the dataset is given in Table I. Examples of images in the dataset are given in Fig.2.

Table I

Category	Number of Images
H1	4404
H2	2334
H3	819
H5	818
H6	739
Total	**9114**

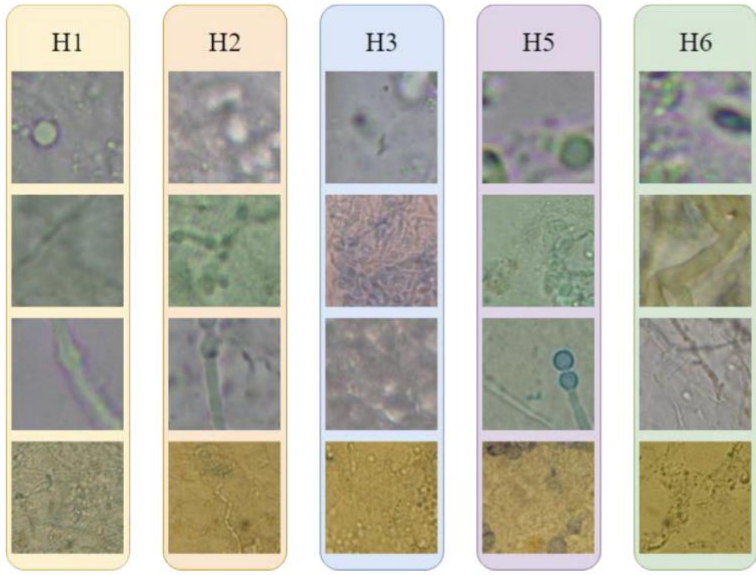


Fig. 2 Examples of images in the dataset

B. Data Augmentation

Data augmentation was performed to balance the number of images of the classes in the existing dataset. For data augmentation, images were subjected to a series of operations such as horizontal, vertical, “+” and “-” 45-degree rotation. After the data augmentation, the total number of images reached 21,691. Detailed information about the data set after data augmentation is given in Table II, and examples of data augmentation are given in Fig.3.

Table II

Category	Number of Images After Data Augmentation
H1	4404
H2	4668
H3	4095
H5	4090
H6	4434
Total	**21,691**

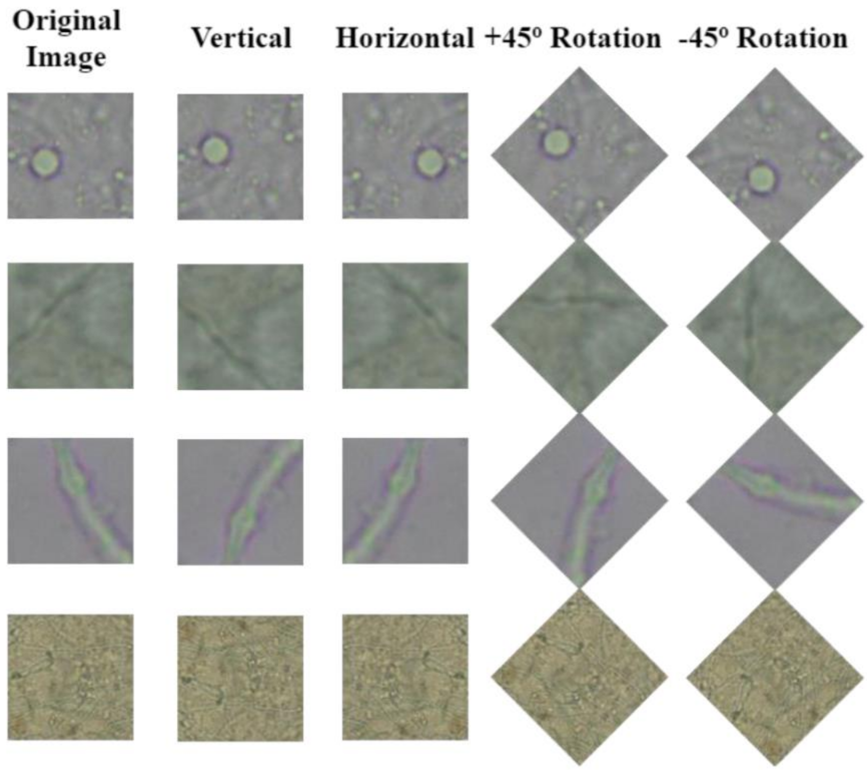
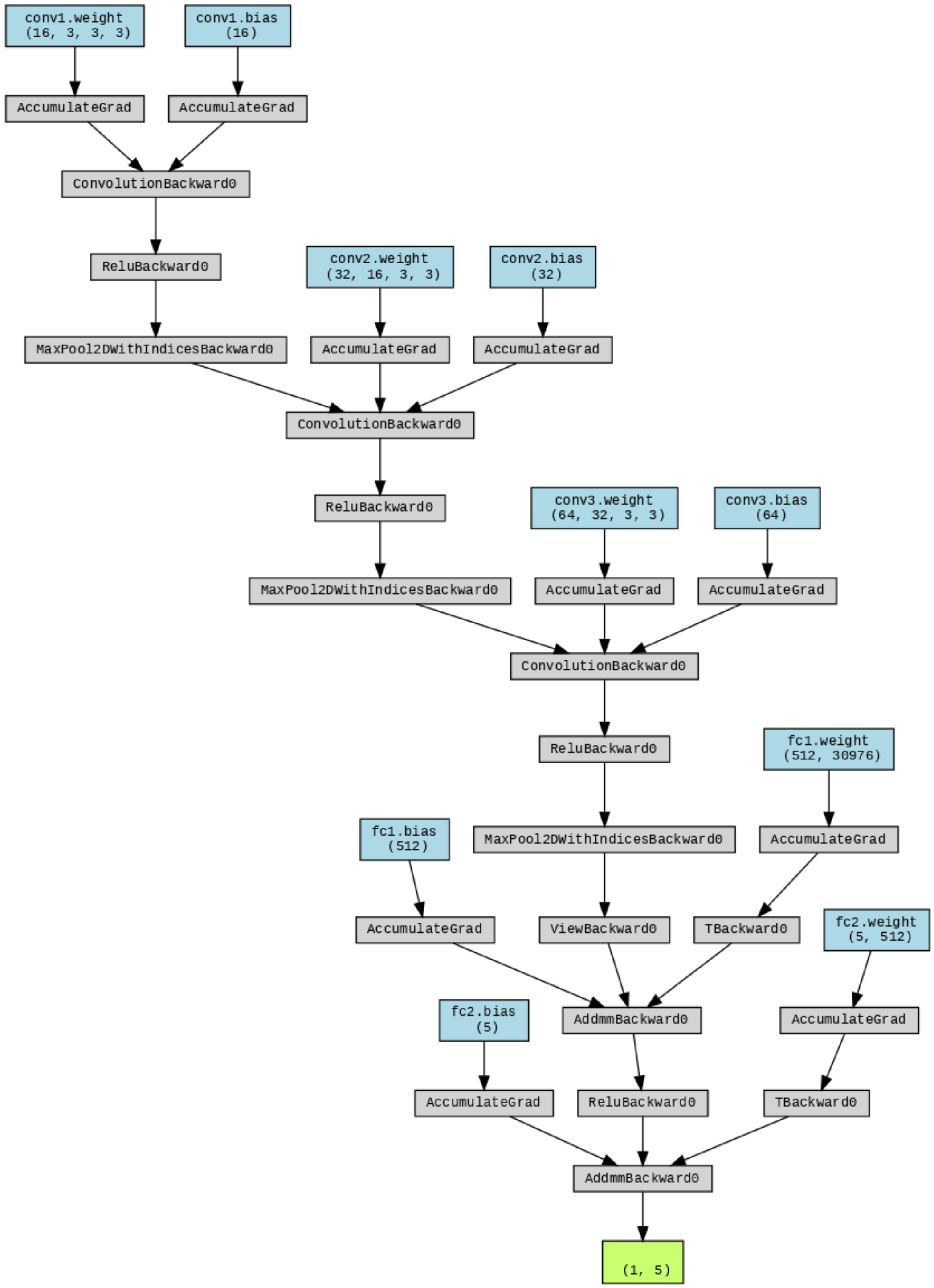


Fig. 3 Examples of data augmentation operations

C. Convolutional Neural Networks (CNN)

In the landscape of modern data analysis demands, Convolutional Neural Networks (CNNs) stand as a revolutionary leap in computer vision. This architecture employs intricate mathematical operations to hierarchically extract and comprehend features within data. This specialized breed of artificial neural networks has achieved remarkable feats, especially in domains like image recognition, object detection, and classification. What fuels the prowess of CNNs is their inherent ability to automatically discern local patterns in data, rendering them invaluable in tasks such as texture analysis, facial recognition, medical imaging, and more. Amidst the existing body of literature, the adaptability and learning capacity of CNNs have transformed them into an exhilarating realm of exploration across a wide spectrum of applications.

NOTE: Below image is self-created CNN model.

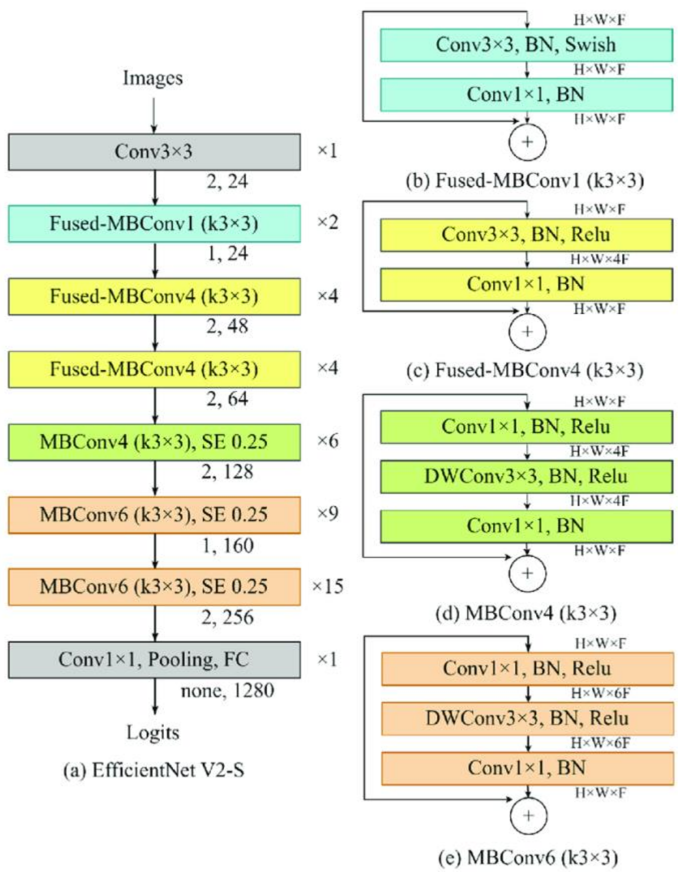


conv1.weight
(16, 3, 3, 3)
conv1.bias
(16)
AccumulateGrad
AccumulateGrad
ConvolutionBackward0
ReluBackward0
conv2.weight
(32, 16, 3, 3)
conv2.bias
(32)
MaxPool2DWithIndicesBackward0
AccumulateGrad
AccumulateGrad
ConvolutionBackward0
ReluBackward0
conv3.weight
(64, 32, 3, 3)
conv3.bias
(64)
MaxPool2DWithIndicesBackward0
AccumulateGrad
AccumulateGrad
ConvolutionBackward0
ReluBackward0
fc1.weight
(512, 30976)
fc1.bias
(512)
MaxPool2DWithIndicesBackward0
AccumulateGrad
AccumulateGrad
ViewBackward0
TBackward0
fc2.weight
(5, 512)
fc2.bias
(5)
AddmmBackward0
AccumulateGrad
AccumulateGrad
ReluBackward0
TBackward0
AddmmBackward0
(1, 5)

D. *EfficientNetV2-S Model:*

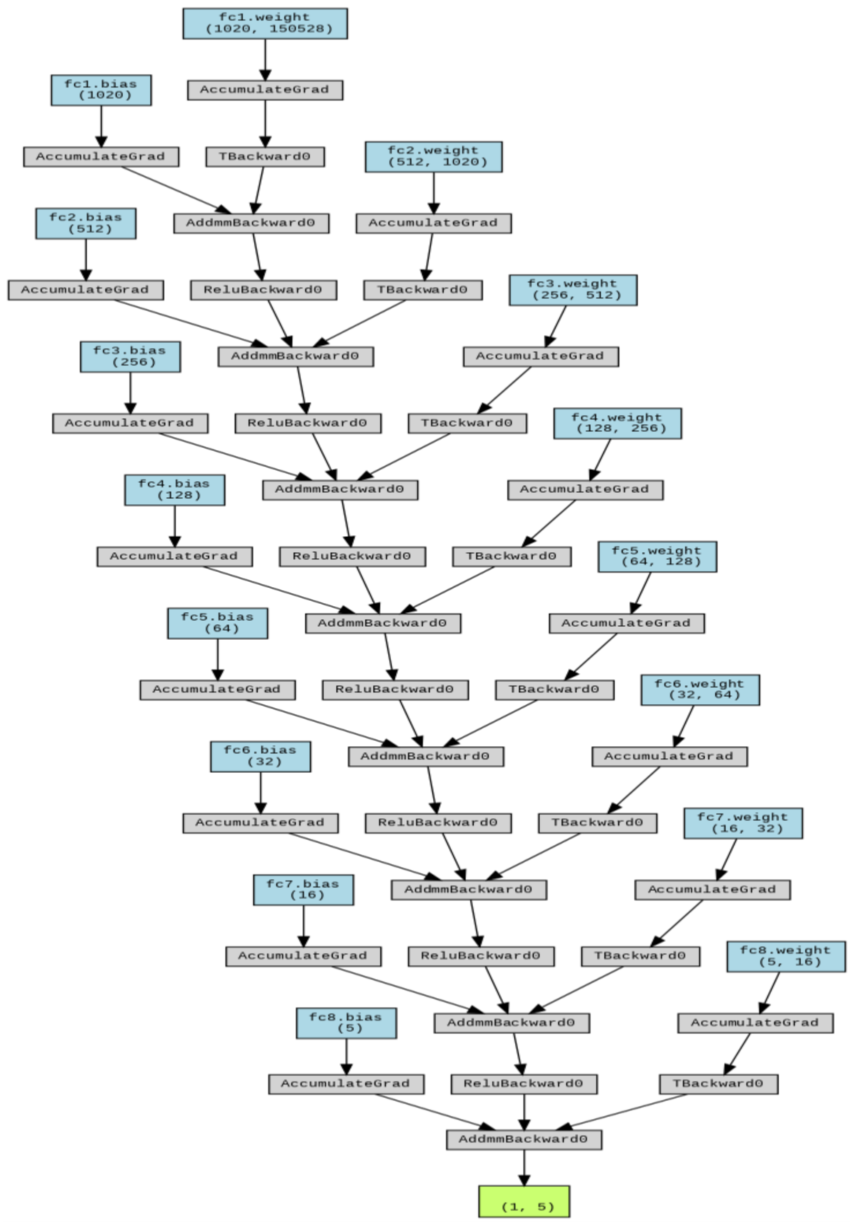
Efficient Net utilizes a compound scaling method, balancing the model's depth, width, and resolution simultaneously. The scaling coefficients determine the size of the network, with a formula that involves the baseline architecture's parameters (depth, width, and resolution). Specifically, for a given scaling factor. This compound scaling approach ensures that all dimensions are proportionally adjusted, optimizing the trade-off between model accuracy and computational efficiency. Additionally, efficient building blocks like inverted residuals with linear bottlenecks and squeeze-and-excitation blocks contribute to the model's efficiency. The EfficientNetV2-S model, a state-of-the-art deep learning architecture, was employed for image classification. The model underwent training for 50 epochs, utilizing the Adam optimizer with a learning rate of 0.0001. A multi-step learning rate scheduler was employed, adjusting the learning rate at predefined milestones. This adaptive learning rate approach helps in fine-tuning the model's performance over time.

The optimizer with a weight decay of 0.1 was used, and a multi-step learning rate scheduler adjusted the learning rate at predefined milestones to enhance convergence. The training process involved continuous monitoring of both training and validation metrics, including loss and accuracy.



E. Simple DNN

A CNN is a kind of network architecture for deep learning algorithms specifically used for image recognition and tasks involving pixel data processing data. Simple CNN model was implemented as an alternative to the EfficientNetV2-S model. This model comprises three convolutional layers and two fully connected layers. It was trained for 50 epochs using the Adam optimizer with a learning rate of 0.0001. The model's architecture is defined with three convolutional layers followed by two fully connected layers. CNNs have proven highly effective in tasks like image classification, object detection, and image segmentation. The training process and monitoring of training and validation metrics were executed similarly to the EfficientNetV2-S model.



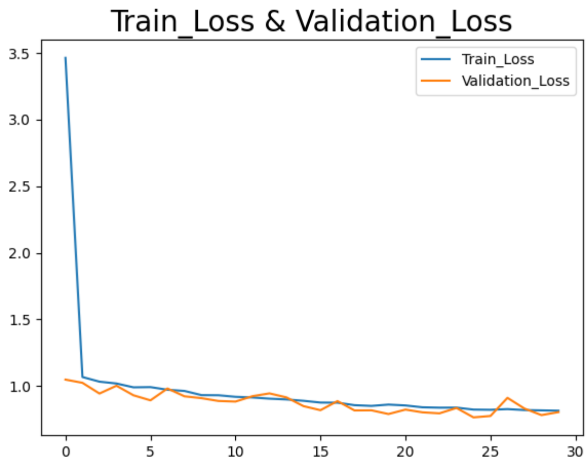
F. Experimental Setup

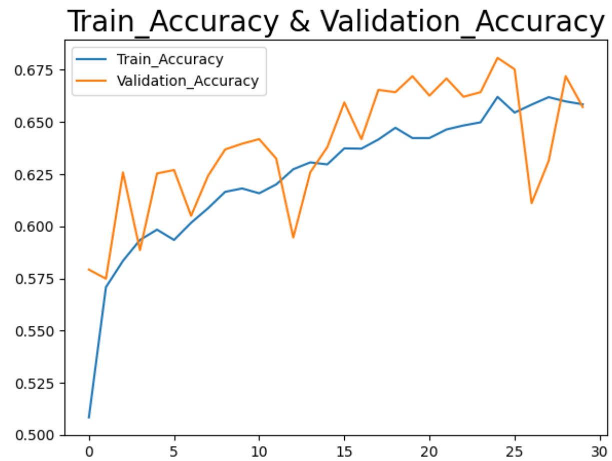
For training model, we used training options specified in the below Table

SOLVER	INITIAL LEARN RATE	VALIDATION FREQUENCY	EPOCH	MINI BATCH SIZE	Weight Decay	MOMENTUM
CNN	***0.00001***	***0.2***	***50***	***8***	***0.1***	***0.9***

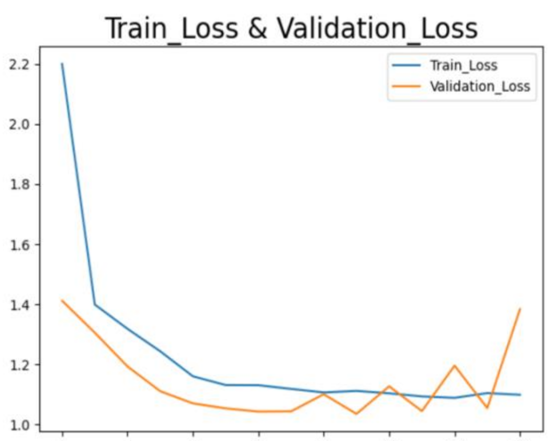
EfficientNetV2-S MODEL

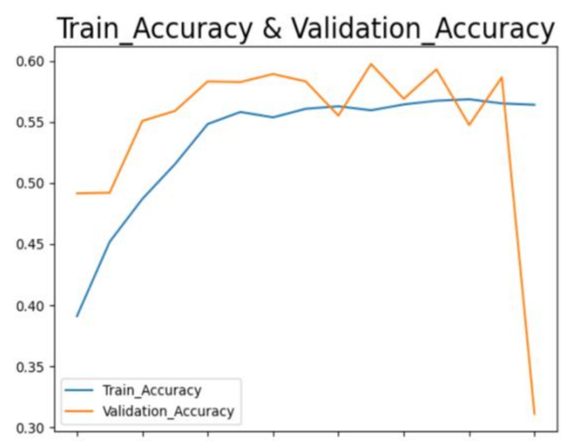
Convolutional Neural Networks (CNNs):





Deep Neural Networks (DNNs):





3. Why Deep Learning and not tradition Machine learning Algorithm?

1. Feature Learning:
   - Deep Learning (DL): Deep neural networks, especially convolutional neural networks (CNNs), automatically learn hierarchical features. In the case of images, lower layers might capture basic features like edges, while deeper layers capture more complex patterns like shapes and textures.
   - Traditional ML: Traditional methods, such as SVM or logistic regression, rely on predefined features. Engineers often manually extract features, and this process may be less effective for large and complex image datasets.

2. End-to-End Learning:
   - DL: Deep learning models can learn end-to-end mappings from raw input to output. They automatically discover relevant features and representations during training, reducing the need for manual preprocessing.
   - Traditional ML: Traditional methods often involve manual feature engineering and preprocessing steps. For image datasets, this might include extracting edges, textures, or other relevant features.

3. Scale and Complexity:
   - DL: Deep learning models can handle the scale and complexity of large datasets. As the dataset size increases, deep learning models have the capacity to learn more intricate and nuanced representations.
   - Traditional ML: Traditional algorithms may struggle to capture the complexity of large datasets without extensive feature engineering. They might not automatically adapt to the increasing richness of the data.

4. Conclusion

MODEL	VALIDATION ACCURACY	VALIDATION LOSS
EfficientNetV2-S MODEL	**0.923**	**0.2052**
Convolutional Neural Networks (CNNs):	**0.657**	**0.8047**

Deep Neural Networks (DNNs):	**0.311**	**1.3836**

In the pursuit of robust image classification for superficial fungal diseases, the performance of three distinct models—EfficientNetV2-S, Convolutional Neural Networks (CNNs), and Deep Neural Networks (DNNs)—was rigorously evaluated based on validation metrics. The EfficientNetV2-S model demonstrated remarkable accuracy, achieving a validation accuracy of 92.3% with a minimal validation loss of 0.2052. In contrast, CNNs displayed a respectable but comparatively lower validation accuracy of 65.7%, accompanied by a validation loss of 0.8047. Meanwhile, DNNs exhibited a lower validation accuracy of 31.1%, coupled with a higher validation loss of 1.3836. These results highlight the superior performance of the EfficientNetV2-S model in accurately classifying microscopic images of superficial fungal diseases, showcasing its potential as a leading choice for automated diagnostic applications.