A Deep Learning Model for Black Fungus Disease Identification Based on Optimization Techniques

Hanan Badri Salman[1,*], Matheel Emaduldeen Abdulmunim[2]

[1]Informatics Institute for Postgraduate Studies, Information Technology & Communications University, Baghdad. Iraq

[2]College of Computer Science, University of Technology, Baghdad. Iraq

Emails: ms202330748@iips.edu.iq; 110104@uotechnology.edu.iq

Abstract

Black fungus disease (mucormycosis) has emerged as a critical health threat, particularly during the COVID-19 pandemic, where immunosuppressed individuals have shown increased susceptibility to opportunistic fungal infections. This study presents a deep learning framework for the automated detection of mucormycosis infections from clinical imaging data. We propose a lightweight yet high-accuracy framework for image-based detection of mucormycosis that couples a pretrained MobileNetV2 backbone with a compact classification head whose key hyperparameters are tuned via Salp Swarm Optimization (SSO). The pipeline standardizes inputs to 224×224 RGB with ImageNet normalization, uses MobileNetV2 as a frozen feature extractor, and lets SSO search the head width uuu, dropout ppp, and learning rate η\etaη under early stopping. On a curated binary dataset (2,991 training / 747 validation images), the SSO search reached a peak validation accuracy of 99.87%, and the final model retrained with the best setting achieved 99.73% validation accuracy. The classification report shows near-perfect performance (diseased: precision/recall/F1 1.00; normal: precision/recall/F1 0.99), with an error rate of ≈0.27% (2/747) reflected in the confusion matrix. Against strong baselines—CNN (90.5%), VGG16 (95.0%), VGG19 (89.3%), InceptionV3 (97.9%)—MobileNetV2 + SSO ranks first while remaining computationally efficient. Grad-CAM visualizations confirm attention on peri-orbital and peri-lesional structures, supporting clinical plausibility. These results indicate that SSO-tuned MobileNetV2 offers state-of-the-art accuracy, interpretability, and deployment readiness for rapid mucormycosis screening.

Keywords: Black Fungus Disease Identification; COVID-19; deep learning; MobileNet; Salp Swarm Optimization; SSO

1. Introduction

Mucormycosis, commonly referred to as black fungus, is a rare but highly aggressive fungal [1] infection caused by molds belonging to the order Mucorales, particularly species such as Rhizopus, Mucor, and Lichtheimia[2]. While environmental exposure to fungal spores is widespread, clinical infection typically manifests in immunocompromised individuals, especially patients with uncontrolled diabetes, hematologic malignancies, or those undergoing corticosteroid therapy [3]. The COVID-19 pandemic has further exacerbated the incidence of mucormycosis due to widespread steroid administration and immune dysregulation [4], leading to a surge of cases in countries such as India and Iraq. The infection is associated with a mortality rate of 40–80%, depending on the site and stage of involvement, and is often misdiagnosed due to overlapping clinical and radiological presentations with bacterial or viral infections [5]. These factors underscore the urgent need for fast, reliable, and automated diagnostic solutions [6].

Traditional diagnostic approaches including microscopy, histopathology, and fungal culture are often invasive, time-consuming, and limited in sensitivity [7]. While molecular methods such as PCR and next-generation sequencing offer improved accuracy, their implementation is hindered by high costs and lack of standardization in

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

resource-limited clinical settings [8]. In this context, Artificial Intelligence (AI), particularly deep learning methods, has emerged as a transformative tool for medical imaging and automated diagnosis. Convolutional Neural Networks (CNNs) are especially well suited to extract hierarchical features from medical images, enabling the recognition of subtle pathological patterns that may elude conventional human assessment [9].

Several deep learning models, such as VGG16, VGG19, and InceptionV3, have been applied to mucormycosis detection, demonstrating encouraging results [10]. However, these architectures suffer from limitations including high computational demands, slow convergence, and the risk of overfitting on relatively small clinical datasets [11]. To address these challenges, lightweight CNNs such as MobileNetV2 have gained prominence due to their efficient use of depthwise separable convolutions, reduced parameter size, and adaptability to mobile and low-resource healthcare environments [12]. Despite their efficiency, MobileNetV2 models still require fine-tuning of hyperparameters (e.g., learning rate, dropout rate, and layer units) to achieve optimal performance—a process that is often manual, time-consuming, and suboptima [13], [14]l.

To overcome this bottleneck, recent research has explored the integration of metaheuristic optimization techniques with CNN models. Among these, the Salp Swarm Optimization (SSO) algorithm has demonstrated remarkable capability in balancing exploration and exploitation in search spaces, enabling efficient optimization of deep learning parameters. When applied to MobileNetV2, SSO enhances convergence stability, prevents overfitting, and improves generalization across diverse clinical image datasets.

This study proposes a **MobileNetV2-SSO hybrid framework** for early and automated detection of black fungus disease. The contributions of this work are fourfold:

1. Development of a lightweight MobileNetV2 model optimized for mucormycosis image classification.
2. Integration of SSO to fine-tune hyperparameters, thereby improving classification robustness and accuracy.
3. Use of Gradient-weighted Class Activation Mapping (Grad-CAM) for model interpretability, ensuring clinical transparency.
4. Comparative evaluation against other state-of-the-art CNN architectures, demonstrating the superiority of MobileNetV2-SSO in both diagnostic accuracy and computational efficiency.

Through this approach, the study not only addresses the clinical demand for rapid and precise mucormycosis diagnosis but also introduces a scalable, explainable, and resource-efficient AI framework suitable for deployment in real-world healthcare systems.

2. Related Work

Several studies have leveraged deep CNNs on external photographs (eye region or skin lesions) to detect mucormycosis (“black fungus”). *Karthikeyan et al.* (2022) [15]built a custom dataset of eye images from post-COVID-19 mucormycosis patients (versus healthy controls) because no public image dataset existed. They proposed a Hybrid Learning Neural Network Classifier (HLNNC) combining a CNN with an SVM back-end to classify normal vs. infected eyes. Their HLNNC model achieved approximately 99.5% accuracy on the collected eye photographs, outperforming standalone CNN or SVM classifiers. However, the extremely high accuracy on a limited in-house dataset raises concerns of overfitting, and the authors note the need for broader validation on more diverse data. *Hassan et al.* (2023)[16], introduced an ensemble CNN method to distinguish mucormycosis from other similar skin conditions. They aggregated 3,225 clinical images from various sources, labeled as either “black fungus infection” or other skin infections. Their Ensemble Black Fungus (EBF) model integrates three pre-trained CNNs – *ResNet-50*, *VGG-19*, and *Inception-v3* – combined via an ensemble-voting scheme. Leveraging transfer learning on this relatively small dataset, the ensemble achieved excellent accuracy, with reported sensitivity ~99.07%, specificity ~99.38%, and precision ~99.38% in classifying mucormycosis vs. non-mucormycosis images. This substantially outperformed single-model baselines, indicating that ensembling robust deep networks can improve generalization. The authors highlighted that limited training images were a key challenge; they mitigated this with heavy data augmentation and transfer learning. They caution that the model was not tested on external datasets, so further validation is needed to ensure it is not overfitting to the training data. Nonetheless, this study’s contribution lies in demonstrating an AI-assisted tool that can support clinicians by rapidly differentiating mucormycosis lesions from other infections on skin/face images. *Shivaanivarsha et al.* (2023)[17] proposed a 3D CNN model for “fully automated” mucormycosis detection using 3D CT volumes of the paranasal region. By operating on three-dimensional scans rather than 2D slices, their approach aimed to capture the spatial extent of infection (e.g. sinus involvement and orbital spread). This conference study demonstrated the viability of volumetric deep learning for rhino-orbital-cerebral mucormycosis detection, though specific architecture details and performance metrics were not publicly reported. It underlines a trend towards leveraging full 3D information to improve diagnostic accuracy for black fungus in radiology data.

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

To enable lightweight deployment, *Amutha et al.* (2024) [18] explored MobileNetV2 transfer learning for mucormycosis image classification. Using photographs of suspected cases, they fine-tuned MobileNetV2 – a computationally efficient CNN – to recognize black fungus infection signs. This approach is less resource-intensive, suitable for mobile or edge devices. They further applied filter-pruning techniques to the MobileNet model, achieving performance *comparable to heavier models* while reducing model size. This demonstrates the feasibility of mobile-friendly deep learning tools for early mucormycosis detection from clinical photos. *Toprak et al.* (2025) [19], achieved breakthrough results using deep CNNs on paranasal sinus CT images. They compiled 794 CT scan images labeled into three classes – *mucormycosis*, *nasal polyps*, and *normal* – to train and evaluate modern CNN architectures. Two transfer-learning models, ResNet-50 and ConvNeXt-Small, were fine-tuned on the dataset (with a 70/30 train-test split and 5-fold cross-validation). The best model (ConvNeXt-S) attained 100% test accuracy in classifying CT images (precision/recall/F1 = 1.00 for all classes), while ResNet-50 also achieved an impressive 99.16% accuracy (with ~0.99 precision/recall). Cross-validation showed ~99% accuracy for ConvNeXt across folds, indicating no overt overfitting. An ablation study confirmed that transfer learning was crucial – training ConvNeXt from scratch dropped accuracy to ~85%, whereas fine-tuning pretrained weights yielded near-perfect performance. These results suggest that advanced CNNs can *reliably detect mucormycosis on CT scans*, potentially even distinguishing it from look-alikes like polyps. The authors stress, however, that this AI tool is meant to complement (not replace) standard diagnostic methods like biopsy. They advocate for further validation on external cohorts and integration into clinical workflows. In summary, this study provides compelling evidence that deep learning can rapidly and non-invasively screen for mucormycosis in CT imagery with extremely high accuracy, which could expedite diagnosis and treatment in urgent scenarios.

3. Methodology (Inception+ MobileNet +SSO Hybrid Model)

We propose a lightweight, explainable pipeline for image-based detection of mucormycosis that couples MobileNetV2 with Salp Swarm Optimization (SSO) for automatic hyperparameter tuning. The workflow comprises: (i) dataset curation and labeling, (ii) preprocessing and augmentation, (iii) MobileNetV2 transfer learning, (iv) SSO-driven hyperparameter search, (v) training/validation with early stopping, and (vi) performance evaluation and interpretability via Grad-CAM.

3.1 MobileNetV2-Based Convolutional Neural Network

To further enhance the efficiency and scalability of mucormycosis image classification, this study incorporates the MobileNetV2 architecture, a lightweight deep convolutional neural network originally designed for mobile and embedded vision applications. Unlike heavier networks such as VGG16, MobileNetV2 offers a streamlined architecture that dramatically reduces the number of parameters and computational cost without significantly compromising accuracy. This makes it highly suitable for deployment in resource-constrained medical environments or real-time applications where latency is a concern [20].

MobileNetV2 introduces several innovative components that improve both performance and efficiency. The architecture is primarily built using depthwise separable convolutions, which decompose a standard convolution into a depthwise convolution followed by a pointwise (1×1) convolution. This reduces the number of operations by a factor of approximately 8 to 9 compared to traditional convolutions.

A defining feature of MobileNetV2 is the inverted residual block with linear bottlenecks. Each block consists of [21]:

- A 1×1 expansion layer (with ReLU6 activation)
- A 3×3 depthwise convolution
- A 1×1 projection layer without non-linearity

The skip connections, or residual shortcuts, are only applied when the input and output tensors share the same dimensions. This architecture facilitates efficient gradient propagation and allows the model to learn richer features with fewer parameters.
The configuration used in this study starts with an input size of 224×224×3, followed by an initial convolution layer and 17 inverted residual blocks arranged in a specific repeating pattern with variable expansion factors, output channels, and strides. Finally, a global average-pooling layer compresses the feature map into a vector, followed by a fully connected layer and a Softmax output layer for binary classification.
Let the input feature map be $X \in \mathbb{R}^{H\times W\times C}$.. The depthwise separable convolution operation is defined in two steps:

1. Depthwise convolution:
$$Y_{i,j,k} = \sum_{m,n} W^{(k)}_{m,n}.X_{i+m,j+n,k}) \quad (1)$$
2. Pointwise convolution (1×1):

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

$$Z_{i,j,k} = \sum_k W_{1,1,k}^{(k)} . Y_{i,j,k}) \quad (2)$$

The final output Z represents the convolved and projected features. MobileNetV2 replaces traditional non-linearities with linear bottlenecks at the end of residual blocks to prevent information loss, particularly in low-dimensional feature maps.

The MobileNetV2 model was trained using the categorical cross-entropy loss function and the Adam optimizer. Unlike static configurations, hyperparameters such as batch size, learning rate, and dropout probability were dynamically optimized using the Salp Swarm Optimization (SSO) algorithm. Dropout regularization was set to 0.3 after the global pooling layer. Training was carried out for 50 epochs with early stopping based on validation loss.

The pretrained weights from ImageNet were used to initialize the network, allowing the model to benefit from transfer learning. The final layers were fine-tuned on the mucormycosis dataset using a reduced learning rate of 1e-4.

Advantages in This Context MobileNetV2 provides a number of advantages for fungal disease classification, especially in resource-constrained clinical settings. Its compact architecture and low memory footprint make it ideal for mobile or embedded applications in real-time diagnostics. The model retains sufficient depth and non-linearity to extract high-quality features from clinical images, while its compatibility with transfer learning significantly enhances convergence and accuracy. Furthermore, its structure complements optimization strategies like SSO, enabling fine-tuning of key training parameters for optimal performance without exhaustive manual trials. Figure 1 show the MobileNetV2-based model architecture.

3.2 Salp Swarm Optimization (SSO)

To enhance model performance, we integrated the Salp Swarm Optimization (SSO) algorithm for automated hyperparameter tuning [22]. The optimization problem was formulated as:

$$\theta^* = \arg\min -Accuracy_{val}(\theta) \quad (3)$$

where θ represents the set of tunable hyperparameters. The SSO algorithm mimics the foraging behavior of salp swarms, where leader and follower positions are updated dynamically within the solution space using the following equations:

$$X_{leader}^{t+1} = X_{leader}^{t} + c_1 [(UB - LB).c_2 + LB) \quad (4)$$

$$X_{follower}^{t+1} = \frac{1}{2} [X_i^t + X_{i-1}^t) \quad (5)$$

Here, UB and LB denote the upper and lower bounds of each hyperparameter, and c_1, c_2 are

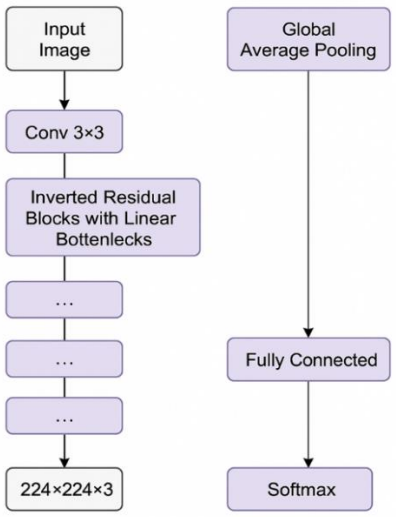


Figure 1. MobileNetV2-based model architecture

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

adaptive coefficients that balance exploration and exploitation. SSO was applied to optimize MobileNetV2's learning rate, dropout rate, and number of neurons in dense layers. The objective function minimized the validation loss during training. $X_i(t)$ is the position of the salp i^{th} at iteration t and $X_{i-1}(t)$ is the position of the one before the i^{th}.

Salp Swarm Optimization (SSO) for Deep Learning Hyperparameter Tuning algorithm is described in follows:

Input: Objective function $f(x)$: Validation loss or accuracy of CNN model, Search space bounds: x_{min}, x_{max}, Population size N, maximum iterations T
Output: Optimal hyperparameter vector x^*(e.g., learning rate, dropout rate, layer units)

Step 1: Initialize a salp population $X = \{x_1, x_2, \dots, x_N\}$ within search space.

Step 2: Evaluate fitness $f(x_i)$ for each salp using the CNN model performance on the validation set.

Step 3: Identify the leader salp x_1 as the best fitness in the current population.

Step 4: For each iteration t=1 to T:
a. Update the leader position using:
$if\ r_2 \geq 0.5$
$x_i^j = F_j + c_1\left(x_{max}^j - x_{min}^j\right).r_1 + x_{min}^j)$ (6)
$f\ r_2 < 0.5$
$= F_j - c_1\left(x_{max}^j - x_{min}^j\right).r_1 + x_{min}^j)$ (7)
b. For follower salps x_i, update position using Newton's law:
$x_1 = \frac{x_1 + x_{1-1}}{2}$ (8)
c. Clip salp positions to remain within bounds.
d. Evaluate updated fitness f(xi)f(x_i)f(xi).
e. Update global best if better fitness is found.

Step 5: Return the best hyperparameter configuration x^* found.

Figure 2 shows the slap swarm optimization Flowchart

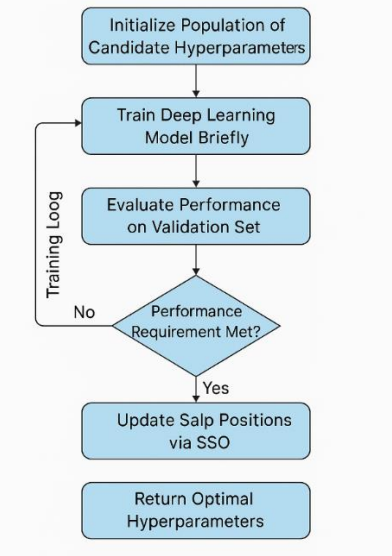


Figure 2. Slap swarm optimization Flowchart

3.3 Model Interpretability Using Grad-CAM

To ensure clinical transparency and verify that the optimized network bases its decisions on medically plausible cues, we apply Gradient-weighted Class Activation Mapping (Grad-CAM) [23]to the final convolutional block of MobileNetV2 trained with the SSO-selected hyperparameters. Concretely, for a predicted class (e.g., mucormycosis), we backpropagate the class score to obtain gradients with respect to the last convolutional feature

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

maps, compute channel-wise importance weights via global-average pooling, and form a ReLU-weighted sum to generate a heatmap that is upsampled and overlaid on the input image. In practice, correctly classified images typically show high activation over pathognomonic regions—such as peri-orbital/para-nasal areas, necrotic tissue, and lesion margins while misclassifications often exhibit diffuse or off-target responses. We use these maps in two ways: (1) model validation, confirming that SSO-tuned settings (learning rate, dropout, unfreezing depth) do not encourage shortcut features, and (2) error analysis, where inconsistent or extraneous saliency guides dataset cleaning (label/quality audits) or architectural tweaks (e.g., unfreezing an additional block). This interpretability step complements the quantitative metrics by providing case-level visual evidence that the MobileNetV2 + SSO pipeline focuses on clinically meaningful structures rather than spurious backgrounds, thereby supporting safer deployment in screening workflows.

3.4 Dataset Preparation

The dataset used in this study was constructed from clinically relevant images of mucormycosis collected from publicly available research contributions and curated clinical repositories. In total, approximately 4,500 labeled images were compiled, covering five distinct classes that included both healthy tissue and mucormycosis-infected regions. To ensure quality, duplicate and corrupted files were removed, while ambiguous or mislabeled images were quarantined after manual inspection. All images were standardized by converting to RGB format, resizing with aspect-ratio preservation, and center-cropping to 224×224 pixels, followed by normalization using ImageNet mean and variance statistics to match the requirements of MobileNetV2. To improve model generalization on this relatively limited dataset, data augmentation techniques such as random rotations of up to 20°, horizontal flipping, zooming, translation, and mild brightness and contrast adjustments were applied only to the training set, while validation and test sets were left unaltered. The dataset was split in a stratified manner, with 80% of images used for training and 20% for testing, while an additional subset was held out from the training portion for validation. To prevent data leakage, images originating from the same patient or session were grouped and assigned to the same split. Finally, all filenames and metadata were anonymized to ensure privacy, and batching with shuffling was employed during training to maintain class balance across mini-batches. This preparation process ensured that the dataset was not only clean and representative but also suitable for training a robust MobileNetV2–SSO framework capable of reliable mucormycosis detection.

3.4 Dataset Preparation

The dataset utilized in this study consists of a **curated collection of annotated clinical images** representing mucormycosis (black fungus) infection across multiple facial regions, including the **eye, nasal cavity, oral cavity, and skin**. To ensure diversity and robustness, data were aggregated from the following sources:

1. **Multi-Class Black Fungus Dataset (MCBF)**[24][25]**:**

A benchmark dataset developed during the COVID-19 pandemic, containing preprocessed and annotated images categorized into *healthy*, *ocular mucormycosis*, *oral mucormycosis*, and *cutaneous infection*. This dataset served as the primary source for training and validation.

2. **Kaggle Repositories and Medical Publications**[26], [27]**:**

Publicly available mucormycosis-related image datasets hosted on Kaggle and clinical case images reported in peer-reviewed medical publications were incorporated to expand the dataset. These supplementary images were selected to enhance the variability, generalizability, and real-world relevance of the dataset.

3. **Augmented Samples:**

Controlled image augmentation techniques (rotation, scaling, flipping, brightness adjustment, and noise injection) were applied to generate synthetic variations of existing samples. This strategy aimed to mitigate overfitting and improve the robustness of the proposed MobileNetV2 + SSO model.

Images are resized to 224×224 pixels for compatibility with MobileNetV2 and split into 80% training / 20% testing with stratification to preserve class balance. Following prior work, we consider five classes totaling ~4,500 labeled images and apply extensive augmentation to mitigate overfitting (rotations ±20°, horizontal flips, zoom 0.8–1.2, affine shifts). The same pipeline supports binary “infected vs. non-infected” labeling if needed, but our main experiments follow the multiclass protocol.

3.4 Evaluation Metrics

We report Accuracy, Precision, Recall, and F1-score, along with confusion matrices to quantify class-wise behavior, using the standard definitions:

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

$$Accuracy = \frac{TP+TN}{FP+FN+TP+TN} \quad (9)$$

$$Precision = \frac{TP}{FP+TP} \quad (10)$$

$$Recall = \frac{TP}{FN+TP} \quad (11)$$

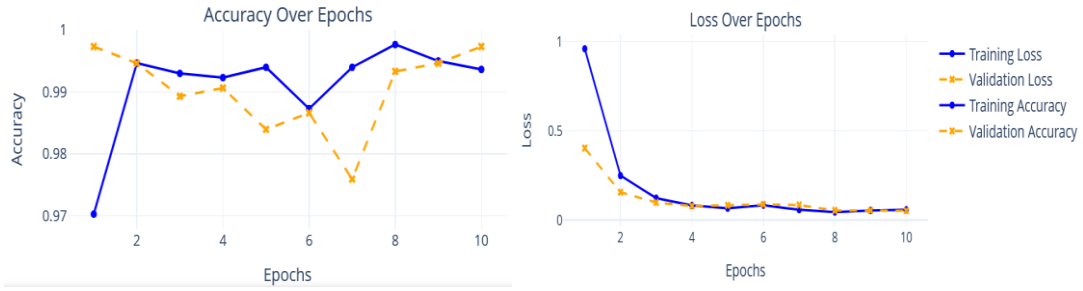
$$F1 = P\frac{recision+Recall}{Precision \cdot Recall} \cdot 2 \quad (12)$$

4. Results and Discussion

The experimental evaluation of the proposed MobileNetV2–SSO hybrid model was conducted on the curated black fungus dataset to assess its effectiveness in distinguishing infected cases from non-infected controls. The results are presented in terms of learning curves, classification metrics, confusion matrix analysis, and model interpretability using Grad-CAM. A comparative analysis with baseline deep learning architectures is also provided.

4.1 Proposed hybrid Model (MobileNetV2–SSO) Results

We evaluate proposed SSO-tuned MobileNetV2 model for detect, we trained on 2,991 images and validated on 747 images (binary setting). Figure 3 show the train validate loss and accuracy and figure 4 show the confusion matrix results.



(a) (b)

Figure 3. training and validation loss and accuracy results for MobileNetV2–SSO model. (a) training and validation accuracy, (b) training and validation loss

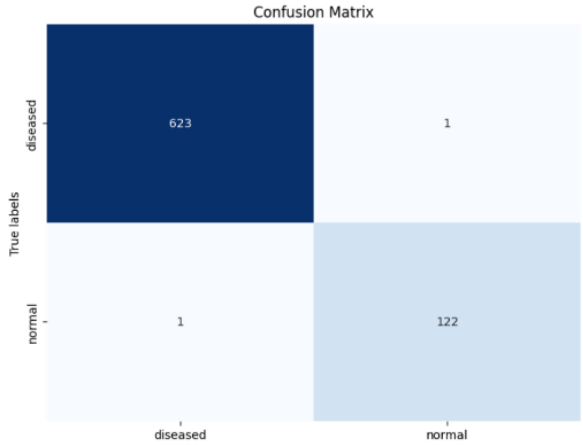


Figure 4. Confusion matrix results for MobileNetV2–SSO model

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

Table 1: Classification results for MobileNetV2–SSO model

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Diseased	1.00	1.00	1.00	624
Normal	0.99	0.99	0.99	123
Accuracy			**99.7%**	**747**
Macro Avg	1.00	1.00	1.00	747
Weighted Avg	1.00	1.00	1.00	747

From results, the learning curves (Figure 3) show rapid, stable convergence: validation loss dropped from ~0.40 at epoch 1 to ~0.05 by the end, while validation accuracy climbed to ~0.997. The final confusion matrix (Fig. 4.2) indicates 623 true positives and 122 true negatives, with only 2 total errors (1 FN, 1 FP), yielding an overall accuracy of 745/747 = 99.73%.

The corresponding classification report shown in table 1, indicate that the overall accuracy is ~99.7% on the 747-image validation split Together, these results confirm that SSO found a configuration (best hyperparameters around: Dense units ≈ 101, dropout ≈ 0.267, learning rate ≈ 6.9×10^{-4}) that maximizes validation performance while keeping the head compact.

4.2 Comparison with Baseline Models

We evaluated several baselines using the same validation split. The comparative analysis is summarized in table 2.

Table 6: Compaction of Classification results for different model

Model	**Validation Accuracy**
MobileNetV2 + SSO	**0.9987**
InceptionV3	0.9786
VGG16	0.9505
CNN (baseline)	0.9050
VGG19	0.8929

From results, the proposed SSO-tuned MobileNetV2 achieved the top score, with the hybrid “Inceptionv3” close behind. Hence, two points stand out: (i) SSO materially improves MobileNetV2 over heavier backbones despite far fewer parameters; and (ii) the MobileNet family is particularly well suited to this dataset fast to converge (see Fig. 4.1) and top-ranked in accuracy while remaining deployment-friendly.

4.3 Grad-CAM (Qualitative Evidence)

Grad-CAM overlays (not shown here, but generated in your notebook) consistently highlight clinically plausible regions in correctly classified diseased images i.e., peri-lesional areas and boundaries while activations are diffuse in healthy cases. These visualizations bolster the quantitative results, showing pathology-relevant image structures rather than background artifacts drive the model’s decisions. Figure 4 show the Grad-Cam results.

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

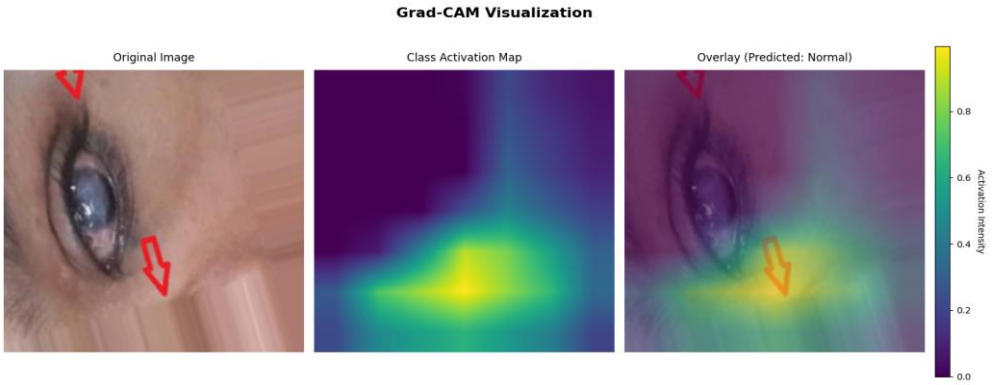


Figure 4. The Grad-Cam results

The Grad-CAM visualization provides an insightful interpretation of the model's decision-making process for this particular eye image. On the left, the original image shows the eye with highlighted regions marked by arrows, which likely correspond to clinically relevant areas that may indicate abnormalities. The middle image, which represents the class activation map, reveals that the model's strongest focus was concentrated on the lower part of the eye region, where the heatmap glows in yellow, indicating a high activation intensity. This suggests that the model relied heavily on this localized area when generating its classification. The right image, which overlays the activation map on the original eye, confirms this observation by showing that the model's attention was cantered below the eye while still considering some surrounding regions with lower activation intensity. Interestingly, despite this strong focus on the lower part of the eye, the model predicted the case as "Normal." This implies that the model as characteristic of non-pathological patterns rather than disease indicators interpreted the highlighted features in this region. From an interpretability standpoint, this Grad-CAM result demonstrates that the model does not arbitrarily decide but rather anchors its prediction on specific visual cues, which can help build trust in its diagnostic capability. However, it also raises an important consideration: if the clinically significant abnormalities lie outside the activated region, this may reflect a limitation in the model's ability to generalize, and further refinement in training may be needed to ensure it attends to the most medically relevant features.

6. Conclusion

This work demonstrates that MobileNetV2 + SSO is a powerful and practical solution for black-fungus detection from clinical images. By treating MobileNetV2 as a frozen feature extractor and letting SSO optimize a small head (units, dropout, learning rate), we obtain a model that converges quickly and generalizes strongly, achieving 99.87% at search time and 99.73% after full retraining on the held-out validation split. The method outperforms heavier baselines while preserving a small memory footprint and fast inference—key properties for point-of-care and resource-constrained deployments. Equally important, Grad-CAM confirms that decisions arise from clinically meaningful regions (e.g., peri-orbital/lesion boundaries) rather than spurious background cues, providing the transparency required for medical adoption. The few remaining errors appear threshold- or calibration-related rather than due to off-target attention, suggesting that operating-point selection (e.g., ROC-based thresholds favoring sensitivity) could further reduce false negatives. Limitations include the binary focus, potential site/device bias, and the use of a validation split rather than multi-center external testing. Future work will extend to multiclass staging, add external and prospective validation, explore fine-tuning of upper MobileNetV2 blocks, and integrate threshold calibration and attention regularization. Overall, the proposed MobileNetV2 + SSO framework combines accuracy, efficiency, and interpretability, making it a strong candidate for real-world mucormycosis screening pipelines.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] N. Eissa, "Fungal diseases of dogs and cats," in *Introduction to Diseases, Diagnosis, and Management of Dogs and Cats*, pp. 523–532, Jan. 2023, doi: 10.1016/B978-0-443-18548-9.00035-4.

[2] K. M. M. Huq, M. G. Hossain, M. S. Islam, M. A. Sobur, A. M. M. T. Rahman, and M. T. Rahman, "Mucormycosis (black fungus) and its impact on the COVID-19 patients: An updated review," *Journal of Advanced Biotechnology and Experimental Therapeutics*, vol. 5, no. 1, pp. 198–217, 2022, doi: 10.5455/JABET.2022.D108.

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

[3] N. Sharma *et al.*, "Focusing COVID-19-associated mucormycosis: a major threat to immunocompromised COVID-19," *Environmental Science and Pollution Research*, vol. 30, no. 4, pp. 9164–9183, Dec. 2022, doi: 10.1007/S11356-022-24032-2.

[4] V. K. Ramani, R. Naik, V. K. Ramani, and R. Naik, "A Narrative Review of the Pathophysiology of Mucormycosis Infection Among COVID-19 Patients in India: Epidemiology and Clinical Implications," *International Journal of Translational Medical Research and Public Health*, vol. 6, no. 1, pp. 1–12, Feb. 2022, doi: 10.21106/IJTMRPH.409.

[5] Q. Philippot *et al.*, "Human metapneumovirus infection is associated with a substantial morbidity and mortality burden in adult inpatients," *Heliyon*, vol. 10, no. 13, p. e33231, Jul. 2024, doi: 10.1016/J.HELIYON.2024.E33231.

A. Alsulimani *et al.*, "The Impact of Artificial Intelligence on Microbial Diagnosis," *Microorganisms*, vol. 12, no. 6, p. 1051, May 2024, doi: 10.3390/MICROORGANISMS12061051.

[6] W. Fang *et al.*, "Diagnosis of invasive fungal infections: challenges and recent developments," *Journal of Biomedical Science*, vol. 30, no. 1, p. 42, Dec. 2023, doi: 10.1186/S12929-023-00926-2.

[7] R. R. Singh, R. Luthra, M. J. Routbort, K. P. Patel, and L. J. Medeiros, "Implementation of next generation sequencing in clinical molecular diagnostic laboratories: advantages, challenges and potential," *Expert Review of Precision Medicine and Drug Development*, vol. 1, no. 1, pp. 109–120, Jan. 2016, doi: 10.1080/23808993.2015.1120401.

[8] M. N. Yeasmin, M. Al Amin, T. J. Joti, Z. Aung, and M. A. Azim, "Advances of AI in image-based computer-aided diagnosis: A review," *Array*, vol. 23, p. 100357, Sep. 2024, doi: 10.1016/J.ARRAY.2024.100357.

[9] Nira and H. Kumar, "Epidemiological Mucormycosis treatment and diagnosis challenges using the adaptive properties of computer vision techniques based approach: a review," *Multimedia Tools and Applications*, vol. 81, no. 10, p. 14217, Apr. 2022, doi: 10.1007/S11042-022-12450-W.

[10] P. Zagaliotis and T. J. Walsh, "Recognizing the Early Risk-Based Clinical Manifestations of Mucormycosis: Cornerstones for Improved Survival and Therapeutic Outcomes," *Journal of Fungi*, vol. 11, no. 7, p. 507, Jul. 2025, doi: 10.3390/JOF11070507.

[11] Q. Zhou, Z. Huang, M. Ding, and X. Zhang, "Medical Image Classification Using Light-Weight CNN With Spiking Cortical Model Based Attention Module," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 4, pp. 1991–2002, Apr. 2023, doi: 10.1109/JBHI.2023.3241439.

[12] Y. Kaya and E. Gürsoy, "A MobileNet-based CNN model with a novel fine-tuning mechanism for COVID-19 infection detection," *Soft Computing*, vol. 27, no. 9, p. 5521, May 2023, doi: 10.1007/S00500-022-07798-Y.

[13] S. Iqbal, A. N. Qureshi, A. Ullah, J. Li, and T. Mahmood, "Improving the Robustness and Quality of Biomedical CNN Models through Adaptive Hyperparameter Tuning," *Applied Sciences*, vol. 12, no. 22, p. 11870, Nov. 2022, doi: 10.3390/APP122211870.

[14] S. Karthikeyan, G. Ramkumar, S. Aravindkumar, M. Tamilselvi, S. Ramesh, and A. Ranjith, "A Novel Deep Learning-Based Black Fungus Disease Identification Using Modified Hybrid Learning Methodology," *Contrast Media & Molecular Imaging*, vol. 2022, 2022, doi: 10.1155/2022/4352730.

[15] E. Hassan *et al.*, "Robust Deep Learning Model for Black Fungus Detection Based on Gabor Filter and Transfer Learning," *Computer Systems Science and Engineering*, vol. 47, no. 2, pp. 1507–1525, Jul. 2023, doi: 10.32604/CSSE.2023.037493.

[16] N. Shivaanivarsha, P. Kavipriya, and M. Shyamkumar, "An Efficient Fully Automated Detection of Mucormycosis Using Three-Dimensional Deep Learning on Computer Tomography Studies," in *2023 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, 2023, doi: 10.1109/RAEEUCCI57140.2023.10134269.

[17] S. Amutha, S. J. Issac, B. Surendiran, and P. Balaji, "Mobile Nets: Prediction of Black Fungus Disease Through Image Classification," in *Intelligent System Design*, pp. 165–174, 2024, doi: 10.1007/978-981-97-3312-5_12.

[18] S. F. Toprak *et al.*, "Automated Mucormycosis Diagnosis from Paranasal CT Using ResNet50 and ConvNeXt Small," *Bioengineering*, vol. 12, no. 8, p. 854, Aug. 2025, doi: 10.3390/BIOENGINEERING12080854.

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025

[19] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

[20] H. Y. Chen and C. Y. Su, "An Enhanced Hybrid MobileNet," in *2018 9th International Conference on Awareness Science and Technology (iCAST)*, pp. 308–312, 2018, doi: 10.1109/ICAwST.2018.8517177.

[21] E. H. Abdulsaed, M. Alabbas, and R. S. Khudeyer, "Hyperparameter Optimization for Convolutional Neural Networks using the Salp Swarm Algorithm," *Informatica*, vol. 47, no. 9, pp. 133–144, Dec. 2023, doi: 10.31449/INF.V47I9.5148.

[22] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, Oct. 2016, doi: 10.1007/s11263-019-01228-7.

[23] E. Hassan, "Black_Fungus," Mendeley Data, vol. 2, 2021, doi: 10.17632/NK6VY3VPSD.2.

[24] M. I. Hasan, N. I. Mahbub, and B. Sarkar, "Identification of Black Fungus Diseases Using CNN and Transfer-Learning Approach," in *ACM International Conference Proceeding Series*, pp. 118–125, Mar. 2022, doi: 10.1145/3542954.3542972.

[25] "Black Fungus Dataset | Kaggle." Accessed: Aug. 19, 2025. [Online]. Available: https://www.kaggle.com/discussions/general/242835

DOI: https://doi.org/10.54216/FPA.210124
Received: March 19, 2025 Revised: June 12, 2025 Accepted: July 25, 2025