CTU Journal of Innovation and Sustainable Development
ISSN 2588-1418 | e-ISSN 2815-6412

DOI:10.22144/ctujoisd.2024.320

Exploring multiple optimization algorithms in transfer learning with EfficientNet models for agricultural insect classification

Hoang-Tu Vo*, Nhon Nguyen Thien, Kheo Chau Mui, Huan Lam Le, and Phuc Pham Tien
Information Technology Department, FPT University, Viet Nam
*Corresponding author (tuvh6@fe.edu.vn)

Article info.

Received 22 Jun 2024
Revised 30 Aug 2024
Accepted 1 Oct 2024

Keywords

Deep learning, explainable AI, grad-cam, insect identification, optimization algorithm, transfer learning

ABSTRACT

Dangerous insects are a significant risk to the global agricultural industry, threatening food security, economic stability, and crop quality. This study investigates the impact of multiple optimization algorithms within transfer learning, employing EfficientNet models for the classification of agricultural insects. The explored optimization algorithms include Adam, Adamax, AdamW, RMSprop, and SGD, while utilizing the EfficientNetB0, EfficientNetB3, EfficientNetB5, and EfficientNetB7 architectures. Experimental results show notable performance differences between optimization algorithms across all EfficiencyNet models in the study. Among the measured metrics are precision, recall, f1-score, accuracy, and loss, the AdamW optimizer consistently demonstrates superior performance compared to other algorithms. The findings underscore the critical influence of optimization algorithms in enhancing classification accuracy and convergence within transfer learning scenarios. Additionally, the study employs various visualization techniques, such as Gradient-weighted Class Activation Mapping (Grad-CAM) to enhance the interpretation of the image classification model's results. By focusing on these methodologies, this research aims to improve the model's performance, optimize its capabilities, and ultimately contribute to effective pest management strategies in agriculture, safeguarding crop yields, farmer livelihoods, and global food security.

1. INTRODUCTION

The global population has expanded significantly compared to the mid-twentieth century. The global human population has grown from approximately 2.5 billion in 1950 to over 8 billion by mid-November 2022. This indicates an increase of one billion people since 2010 and two billion since 1998. Projections suggest that the world's population will continue to rise, nearing a predicted 9.7 billion by 2050 and maybe reaching a peak of 10.4 billion by the mid-2080s, an increase of nearly 2 billion people over the next 30 years (The United Nations, 2023). The rapid increase in population in recent years has led to an increased demand for agricultural products, which in turn has caused a significant expansion of cultivation practices. Agrifood systems provide substantial societal benefits, such as the sustenance we derive from food and employment opportunities that support the livelihoods of more than a billion individuals.

Crop production is always in danger because illnesses and insect pests are a persistent threat (Savary et al., 2012; Shukla et al., 2022). These continuous problems make agriculture productivity vulnerable, necessitating constant watchfulness and creative approaches to managing pests and diseases effectively in order to protect crop yield. Farmers have always relied on chemical treatments and pesticides to keep pests away, but these methods have harmed human health and made damage to groundwater and soil worse (Srivastav, 2020; Zaller & Zaller, 2020; Hossain et al., 2022).

The traditional approach to identifying insect pests depends on the visual inspections of knowledgeable experts; this is a challenging and expertise-dependent process that is not feasible for broad-based agricultural field monitoring. Therefore, there is an immediate need for more efficient alternatives to current approaches that aren't harmful to public health or do not cause environmental harm. The growing accessibility of cutting-edge approaches like modern deep learning techniques and machine learning models is probably going to encourage more study and development in the field of smart farming such as (Park et al., 2023; Parthiban et al., 2023; Sekharamantry et al., 2023; Vo et al., 2023a, 2023b).

This study analyses the influence of various optimization techniques in transfer learning, utilizing EfficientNet models to classify agricultural insects. The examined optimization algorithms encompass Adam, Adamax, AdamW, RMSprop, and SGD, implemented across various EfficientNet architectures such as EfficientNetB0, EfficientNetB3, EfficientNetB5, and EfficientNetB7. This study also utilizes Grad-CAM methodologies to explain and interpret the outcomes derived from the deep learning model.

These visualization techniques offer insightful explanations by highlighting the areas within the model's decision-making process, enhancing the comprehension of how the model arrives at its classifications or predictions. Grad-CAM (Selvaraju et al., 2017) is useful tools that help highlight the image features or regions that are important in helping the model make decisions. This allows for a better understanding of the model's performance and choices.

The structure of the paper is as follows: Part 2 presents an extensive literature review. Part 3 presents the methodology, encompassing Data Collection and Preparation, various versions of EfficientNets, Optimization algorithms, Overall Methodology, and Performance Evaluation Measures. Part 4 covers the experimental configuration and the resulting outcomes. Part 5 is an overview of the study's outcomes and final conclusions.

2. RELATED WORKS

Although we used advanced topologies, we struggled with incorrect predictions during the model-building process. Sometimes, building Convolutional Neural Network (CNN) (Wu, 2017) models from scratch on our dataset generates disappointing outcomes, which wastes computational resources and leads to wasteful model training. Transfer learning (TL) (Torrey & Shavlik, 2010) is a technique where a model trained on one task is repurposed for another task. It involves utilizing the knowledge obtained from solving one problem and applying it to a different problem. This approach is particularly significant in deep learning (DL), where models, often pre-trained on huge and diverse datasets, learn generic features that can be adapted to various downstream tasks.

Transfer learning significantly reduces the requirement for huge quantities of labeled data and extensive computational resources for training models from scratch. It expedites model development, enhances model performance, and enables effective deployment in scenarios with limited data availability or computational constraints. By transferring knowledge from one domain to another, transfer learning helps in achieving better generalization, faster convergence, and improved accuracy in various real-world applications.

Transfer learning models are commonly utilized to enhance the accuracy of DL models across a wide range of tasks such as Kathamuthu et al. (2023) investigated various deep TL methods using CNNs to identify the presence of COVID-19 in chest CT photos. Foundation models, including Densenet121, VGG16, VGG19, InceptionV3, Xception, and Resnet50. This study (Gulzar, 2023) uses MobileNetV2 to classify 40 different types of fruits, achieving an accuracy of 99%. This research (Mahmud et al., 2023) includes feature extraction from MRI data using TL. It involves adjusting the weights and utilizing these features to train pre-existing models, eventually constructing a merged classifier. This article (Ikromovich & Mamatkulovich, 2023) delves into employing transfer learning methods to enhance the effectiveness and precision of facial recognition tasks through deep CNNs.

Transfer learning techniques in agriculture also have revolutionized the landscape of agricultural research, offering a potent tool to enhance crop productivity, disease detection, and soil analysis.

By leveraging pre-trained models from domains like computer vision or natural language processing and fine-tuning them with agricultural data, researchers can expedite the development of robust predictive models. In identifying tomato leaf diseases, this study (Saeed et al., 2023) utilized two pre-trained convolutional neural networks (CNNs), Inception V3 and Inception ResNet V2, achieving exceptional accuracy at 99.22%. These models demonstrated the most optimal performance in classifying healthy and unhealthy tomato leaf images. The study (Yonbawi et al., 2023) presents an innovative technique called Modified Metaheuristics with Transfer Learning for Insect Pest Classification in Agricultural Crops (MMTL-IPCAC), showcasing an impressive accuracy rate of 98.73%.

While several research have recently been carried out to increase accuracy though applying Transfer learning models instead of building CNN models from scratch, there are still not many that analyze model accuracy when combining Transfer learning models with optimization algorithms. The significance of optimization algorithms in shaping model performance within Machine Learning and Deep Learning remains a pivotal research focus. In today's world, as deep learning models become more complex, it's crucial to grasp how these models make decisions. Explainable Artificial Intelligence (XAI) (Gunning & Aha, 2017; Samek et al., 2017), steps in to make these decisions clearer and more understandable, bridging the understanding gap between human understanding and the complex workings of AI.

XAI has succeeded in various fields such as medicine (Singh et al., 2020; Schutte et al., 2021; Gaur et al., 2022) agriculture (Quach et al., 2023) and traffic classification (Ahn et al., 2020), making XAI more transparent and accessible in these areas. Moreover, the addition of Grad-CAM techniques has demonstrated significance in explaining the model's decision-making processes.

3. MATERIALS AND METHOD

3.1. Data collection and preparation

The dataset (Jain, 2023) utilized in this study included 1591 photos of 15 distinct insect species that frequently appear in agricultural environments (Figure 1). Additionally, we divided this dataset into three different categories using a ratio of 80:10:10 for training, validation, and testing.

Figure 1. Some images of dangerous insects dataset

3.2. Optimization algorithms

Optimization algorithms serve as the backbone in the training process of deep learning models, playing a pivotal role in enhancing model performance and convergence. At the heart of these algorithms lies the pursuit of minimizing the loss function, aiming to fine-tune the model's parameters iteratively during training.

This section will theoretically introduce several optimization algorithms employed in the research such as Adaptive Moment Estimation (Adam) (Kingma, 2014), Root Mean Squared Propagation (RMSprop) (Zou et al., 2019), AdaMax (Kingma, 2014), Adam with decoupled weight decay (AdamW) (Loshchilov & Hutter, 2017), and Stochastic Gradient Descent (SGD) (Bottou, 2010).

3.3. Overall methodology

In this study, through the integration of optimization techniques and transfer learning models, we propose a new model to address the classification of dangerous insects in agriculture.

Transfer learning is a machine learning approach that utilizes a model developed for one specific task to be applied or modified for another distinct task. Instead of training a model from scratch, transfer learning uses the knowledge or features learned from one problem domain and applies it to another domain. Initially, pre-trained models such as EfficientNetB0, EfficientNetB3, EfficientNetB5, and EfficientNetB7 are employed for feature extraction.

The fine-tuning of the last layers of these models serves as the classifier, incorporating additional layers such as BatchNormalization, Dense, Dropout, and Output Layer.

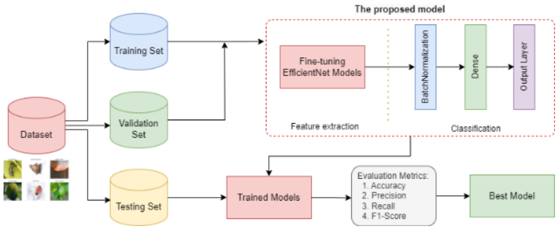

Figure 2. Diagram illustrating the recommended deep learning methodology for training

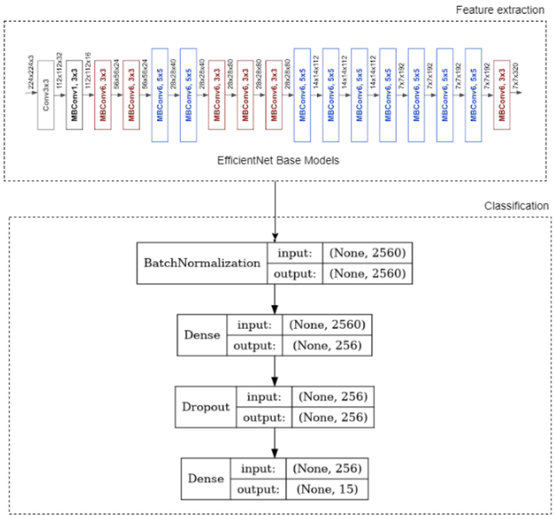

Figure 3. The architecture of the proposed model is based on the EfficientNet model

Details of the proposed deep learning method for training are presented in Figure 2 and Figure 3. Each model has five optimization methods installed before it is trained: Adam, Adamax, AdamW, RMSprop, and SGD. The study then analyzes the test dataset using Grad-CAM and HiresCAM approaches with the objective provide explanatory insights into the evaluation model. Figure 4 shows the insect's classification and provides an explanation of the model's observations.

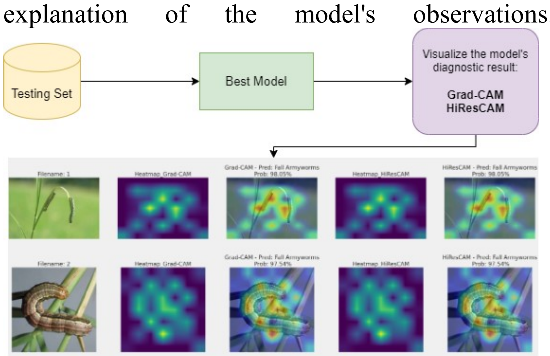

Figure 4. Classification process and explain the model's outcomes

4. RESULTS AND DISCUSSION

4.1. Environmental settings

The experimental results were obtained from trials conducted on the Kaggle platform. The hardware used for these trials included 13GB of RAM and a P100 GPU featuring 16GB of memory. The models were trained for a total of 40 epochs, utilizing a batch size of 32 consistently throughout the process.

4.2. Performance comparison of DL models

The results displayed in Table 1 showcase the performance of different optimization algorithms across various EfficientNet models for the given classification task. Overall, Adam, AdamW, and RMSprop consistently demonstrate better performance compared to Adamax and SGD in most models. Particularly, AdamW consistently exhibits strong performance across all EfficientNet models, maintaining high scores in precision, recall, F1-score, accuracy, and lower loss values.

On the other hand, SGD consistently demonstrates comparatively lower performance metrics across all models. This comparison highlights the impact of optimization algorithms on model performance, indicating the potential superiority of Adam, AdamW, and RMSprop in this classification context, with a particular emphasis on AdamW.

The results highlight the supremacy of EfficientNetB3 combined with the AdamW optimizer, achieving the highest accuracy of 95.5%. This combination stands out among the various models and optimization algorithms tested, emphasizing the critical role of model architecture and optimizer choice in achieving superior performance in classification tasks (Figure 5).

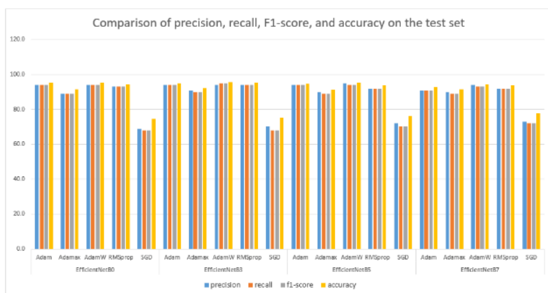

Figure 5. Comparison of precision, recall, F1-score, and accuracy on the test set

Table 1. Table of results for the fine-tuning model on the test set

Model	Optimizer	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)	Loss
EfficientNetB0	Adam	94.0	94.0	94.0	95.2	0.48
	Adamax	89.0	89.0	89.0	91.5	1.73
	AdamW	94.0	94.0	94.0	95.2	0.47
	RMSprop	93.0	93.0	93.0	94.5	0.41
	SGD	69.0	68.0	68.0	74.6	7.96
EfficientNetB3	Adam	94.0	94.0	94.0	95.0	0.5
	Adamax	91.0	90.0	90.0	92.3	1.76
	AdamW	94.0	95.0	95.0	95.5	0.5
	RMSprop	94.0	94.0	94.0	95.2	0.43
	SGD	70.0	68.0	68.0	75.3	8.05
EfficientNetB5	Adam	94.0	94.0	94.0	94.8	0.51
	Adamax	90.0	89.0	89.0	91.2	1.56
	AdamW	95.0	94.0	94.0	95.3	0.52
	RMSprop	92.0	92.0	92.0	93.6	0.46
	SGD	72.0	70.0	70.0	76.0	8.24
EfficientNetB7	Adam	91.0	91.0	91.0	92.9	0.57
	Adamax	90.0	89.0	89.0	91.6	1.5
	AdamW	94.0	93.0	93.0	94.4	0.56
	RMSprop	92.0	92.0	92.0	93.7	0.48
	SGD	73.0	72.0	72.0	77.6	8.36

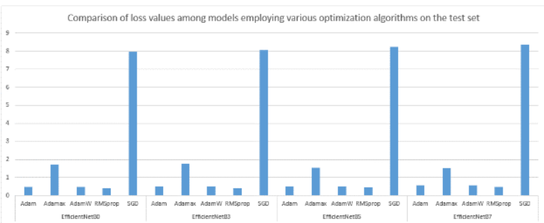

Figure 6. Comparison of loss values among models employing various optimization algorithms on the test set

Moreover, concerning the loss metric across different models and optimizers, notable variations in the loss values are apparent. Typically, models trained with Adam, AdamW, and RMSprop optimizers consistently exhibit lower loss values in comparison to those trained with Adamax and SGD as displayed in Figure 6.

Additionally, Figure 7 provides an overview of the performance metrics, encompassing both loss and accuracy, evaluated during both the training and validation phases of EfficientNetB3 with the AdamW optimizer and Figure 8 illustrates the confusion matrix.

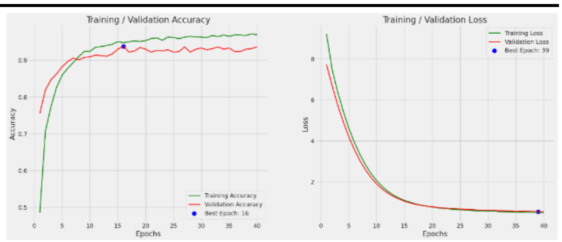

Figure 7. Loss and accuracy plots of the model EfficientNetB3 with the AdamW optimizer

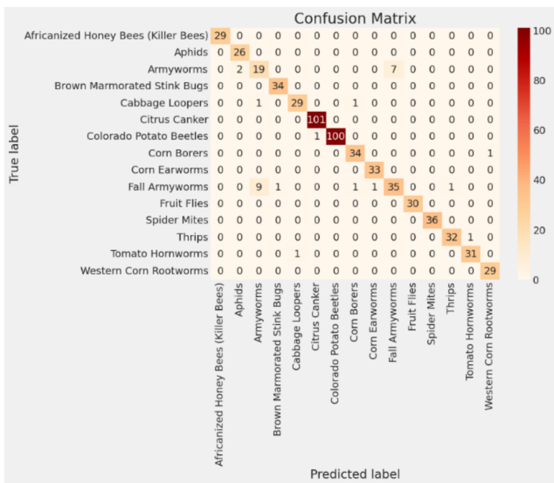

Figure 8. EfficientNetB3 model confusion matrix with the AdamW optimizer

4.3. Using Grad-CAM to visualize and interpret model predictions

The study investigates the influence of optimization algorithms on model performance and identifies pivotal image areas affecting predictions. Using Grad-CAM and HiResCAM techniques, the research visualizes crucial image regions, offering insights into the model's focus during prediction. Grad-CAM generates heatmaps highlighting significant areas, while HiResCAM provides further comparison for a deeper understanding of visual elements impacting the model's attention.

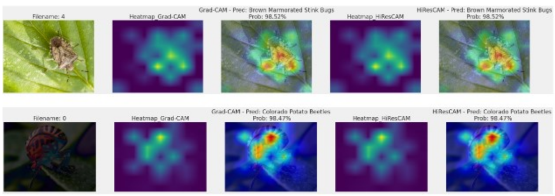

Figure 9. Examples of the brown marmorated stink bugs explained using Grad-CAM and HiResCAM

Figure 9 clearly displays specific regions essential for accurate classification by the model, unique to each type. Moreover, the heatmap identifies crucial image regions corresponding to the model's attention during the prediction phase, demonstrating a strong correlation with its predictions.

5. CONCLUSION

The study underscores the crucial role of optimization algorithms, particularly Adam, AdamW, and RMSprop, in enhancing the performance of EfficientNet models for classification tasks, consistently outperforming Adamax and SGD in accuracy, precision, recall, and F1-score. Notably, AdamW stands out with consistently superior performance metrics and lower loss values across all evaluated measures, while SGD consistently trails behind.

Additionally, the integration of Grad-CAM and HiResCAM techniques not only validates the model's predictions but also provides valuable insights into its decision-making process, ultimately enhancing the model's comprehensibility and reliability for practical applications.

REFERENCES

Ahn, S., Kim, J., Park, S. Y., & Cho, S. (2020). Explaining deep learning-based traffic classification using a genetic algorithm. *IEEE Access, 9*, 4738-4751.

Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. *In Proceedings of COMPSTAT'2010: 19th International Conference on Computational StatisticsParis France, August 22-27, 2010 Keynote, Invited and Contributed Papers* (pp. 177-186). Physica-Verlag HD.

Gaur, L., Bhandari, M., Razdan, T., Mallik, S., & Zhao, Z. (2022). Explanation-driven deep learning model for prediction of brain tumour status using MRI image data. *Frontiers in Genetics*, *13*, 822666.

Gulzar, Y. (2023). Fruit image classification model based on MobileNetV2 with deep transfer learning technique. *Sustainability*, *15*(3), 1906.

Gunning, D., & Aha, D. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI magazine*, *40*(2), 44-58.

Hossain, M. E., Shahrukh, S., & Hossain, S. A. (2022). Chemical fertilizers and pesticides: impacts on soil degradation, groundwater, and human health in Bangladesh. *In Environmental Degradation: Challenges and Strategies for Mitigation* (pp. 63-92). Cham: Springer International Publishing.

Ikromovich, H. O., & Mamatkulovich, B. B. (2023). Facial recognition using transfer learning in the deep CNN. *Open Access Repository*, *4*(3), 502-507.

Jain, T. R. (2023). *Dangerous farm insects dataset*. https://www.kaggle.com/datasets/tarundalal/dangerous-insects-dataset.

Kathamuthu, N. D., Subramaniam, S., Le, Q. H., Muthusamy, S., Panchal, H., Sundararajan, S. C. M., ... & Zahra, M. M. A. (2023). A deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications. *Advances in Engineering Software*, *175*, 103317.

Kingma, D. P. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Mahmud, T., Barua, K., Barua, A., Das, S., Basnin, N., Hossain, M. S., ... & Sharmen, N. (2023, August). Exploring deep transfer learning ensemble for improved diagnosis and classification of Alzheimer's disease. *In International Conference on Brain Informatics* (pp. 109-120). Cham: Springer Nature Switzerland.

Loshchilov, I., & Hutter, F. (2017). Fixing weight decay regularization in adam. *arXiv preprint arXiv:1711.05101, 5*.

Parthiban, S., Moorthy, S., Sabanayagam, S., Shanmugasundaram, S., Naganathan, A., Annamalai, M., & Balasubramanian, S. (2023). Deep learning based recognition of plant diseases. *In Computer Vision and Machine Intelligence Paradigms for SDGs: Select Proceedings of ICRTAC-CVMIP 2021* (pp. 83-93). Singapore: Springer Nature Singapore.

Park, Y. H., Choi, S. H., Kwon, Y. J., Kwon, S. W., Kang, Y. J., & Jun, T. H. (2023). Detection of soybean insect pest and a forecasting platform using deep learning with unmanned ground vehicles. *Agronomy*, *13*(2), 477.

Quach, L. D., Nguyen, K. Q., Nguyen, A. Q., Thai-Nghe, N., & Nguyen, T. G. (2023). Explainable deep learning models with gradient-weighted class activation mapping for smart agriculture. *IEEE Access*, *11*(August), 83752-83762.

Saeed, A., Abdel-Aziz, A. A., Mossad, A., Abdelhamid, M. A., Alkhaled, A. Y., & Mayhoub, M. (2023). Smart detection of tomato leaf diseases using transfer learning-based convolutional neural networks. *Agriculture*, *13*(1), 139.

Samek, W. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint arXiv:1708.08296.*

Savary, S., Ficke, A., Aubertot, J. N., & Hollier, C. (2012). Crop losses due to diseases and their implications for global food production losses and food security. *Food Security*, *4*(4), 519-537.

Schutte, K., Moindrot, O., Hérent, P., Schiratti, J. B., & Jégou, S. (2021). Using stylegan for visual interpretability of deep learning models on medical images. *arXiv preprint arXiv:2101.07563*.

Sekharamantry, P. K., Melgani, F., & Malacarne, J. (2023). Deep learning-based apple detection with attention module and improved loss function in YOLO. *Remote Sensing*, *15*(6), 1516.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. *In Proceedings of the IEEE international conference on computer vision* (pp. 618-626).

Singh, A., Sengupta, S., & Lakshminarayanan, V. (2020). Explainable deep learning models in medical image analysis. *Journal of Imaging*, *6*(6), 52.

Shukla, S., Upadhyay, D., Mishra, A., Jindal, T., & Shukla, K. (2022). Challenges faced by farmers in crops production due to fungal pathogens and their effect on Indian economy. *In Fungal diversity, ecology and control management* (pp. 495-505). Singapore: Springer Nature Singapore.

Srivastav, A. L. (2020). Chemical fertilizers and pesticides: role in groundwater contamination. *In Agrochemicals detection, treatment and remediation* (pp. 143-159). Butterworth-Heinemann.

The United Nations. (2023). *Population*. https://www.un.org/en/global-issues/population.

Torrey, L., & Shavlik, J. (2010). Transfer learning. *In Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (pp. 242-264). IGI Global.

Vo, H. T., Quach, L. D., & Hoang, T. N. (2023a). Ensemble of deep learning models for multi-plant disease classification in smart farming. *International Journal of Advanced Computer Science and Applications, 14*(5).

Vo, H. T., Thien, N. N., & Mui, K. C. (2023b). Tomato disease recognition: Advancing accuracy through xception and bilinear pooling fusion. *International Journal of Advanced Computer Science and Applications*, *14*(8).

Wu, J. (2017). Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, *5*(23), 495.

Yonbawi, S., Alahmari, S., Daniel, R., Lydia, E. L., Ishak, M. K., Alkahtani, H. K., ... & Mostafa, S. M. (2023). Modified Metaheuristics with Transfer Learning Based Insect Pest Classification for Agricultural Crops. *Computer Systems Science & Engineering*, *46*(3).

Zaller, J. G., & Zaller, J. G. (2020). Pesticide impacts on the environment and humans. *Daily Poison: Pesticides-An Underestimated Danger*, 127-221.

Zou, F., Shen, L., Jie, Z., Zhang, W., & Liu, W. (2019). A sufficient condition for convergences of Adam and RMSProp. *In Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition* (pp. 11127-11135).