International Journal of Computer Applications (0975 – 8887)
Volume 187 – No.76, January 2026

FAWINSTARNet: A Lightweight MobileNetV2 Model for Early Instar Fall Armyworm Detection in Maize

Monica Shinde
Research Scholar
D Y Patil Institute of MCA and Management, Akurdi, Pune; Savitribai Phule Pune University, Maharashtra, India

Kavita Suryawanshi, PhD
Professor, D Y Patil Institute of MCA and Management, Akurdi, Pune; Savitribai Phule Pune University, Maharashtra, India

ABSTRACT
The Fall Armyworm (Spodoptera frugiperda) has emerged as a major constraint on maize cultivation throughout warm-climate agricultural zones. Management practices are most effective during the earliest larval stages, making precise recognition of first- and second-instar caterpillars essential for minimizing crop damage and limiting indiscriminate pesticide application. In response to this requirement, the present work proposes FAWINSTARNet, a computationally efficient deep-learning framework derived from the MobileNetV2 family and tailored for six-category instar discrimination. An initial image repository containing 12,169 samples validated by entomological experts was systematically enlarged to 187,152 images through controlled augmentation to enhance feature variability. A group of ten pretrained convolutional neural networks was evaluated to determine an appropriate trade-off between predictive performance and resource demand. The selected FAWINSTARNet configuration attained an accuracy near 97% and was sufficiently lightweight for execution on mobile hardware, thereby supporting on-site pest surveillance for growers. The study offers a full account of dataset development, experimental procedures, architectural design, and comparative assessment of competing models.

General Terms
Deep Learning, Image Processing, Pattern Recognition, Computer Vision

Keywords
Fall Armyworm; FAWINSTARNet; instar classification; maize pests; convolutional neural networks.

1. INTRODUCTION
A rapid geographical spread and severe foliar feeding have positioned the Fall Armyworm (Spodoptera frugiperda) as a high-risk pest in global agriculture. Following its establishment in African and Asian production zones, the species has contributed to pronounced yield reductions and financial losses in areas dependent on maize cultivation [1,3]. Timely recognition of larval development is critical, because chemical and cultural interventions are most successful during the initial two instars, whereas later stages inflict substantial vegetative and reproductive injury [2]. In practice, visual differentiation of instars is difficult—early larvae are small, display only minor morphological variation, and frequently evade accurate field assessment—limiting the reliability of manual scouting efforts.

Artificial intelligence has expanded the technological options available to agricultural management, with deep learning architectures now routinely applied to tasks such as crop disease categorization, insect identification, and visual symptom assessment on foliage [5,6]. Compact CNN designs suitable for mobile hardware have further made it possible to conduct diagnostic analysis on handheld devices rather than relying on laboratory environments. Within this class of efficient networks, MobileNetV2 is notable for its balanced computational footprint and its capacity to derive discriminative representations through the use of inverted residual structures and linear bottleneck layers [10].

Although recent progress in agricultural artificial intelligence has been considerable, detailed categorization of FAW larval instars has received limited attention. Prior work has concentrated primarily on identifying pest species or evaluating feeding injury on host plants, yet effective in-field decision-making depends on discriminating specific instars so that control measures can be executed at the most responsive stages. FAWINSTARNet is designed to meet this operational need by leveraging an extensive expert-annotated image corpus, systematic augmentation to increase visual diversity, and a refined MobileNetV2-based configuration engineered for rapid inference on mobile platforms.

2. LITERATURE REVIEW
The worldwide establishment of the Fall Armyworm has prompted extensive work on surveillance and early-alert mechanisms. Global assessments consistently highlight the need for fast detection to preserve crop productivity and reinforce integrated pest management programs [1]. Conventional diagnostic practice depends on morphological traits, but reliably separating early instars demands specialized entomological skill and considerable time investment. Molecular tools, including PCR and LAMP assays, can verify species identity with high precision; however, their dependence on laboratory infrastructure constrains their suitability for routine field deployment [11].

Large-area surveillance using remote sensing platforms and unmanned aerial vehicles has been investigated extensively for tracking FAW outbreaks. Although these technologies are useful for mapping stress signatures associated with infestation, their spatial granularity is insufficient for detecting single caterpillars or differentiating between larval stages [4]. As a result, close-range imaging solutions remain indispensable for timely and accurate instar-level identification.

Advances in deep learning have reshaped the analysis of agricultural imagery, making it possible to perform reliable pest and pathogen recognition under heterogeneous field environments. Prior investigations indicate that architectures such as ResNet, DenseNet, and Inception deliver high accuracy across a range of crop-related classification problems [12]. Streamlined networks, including MobileNet and comparable derivatives, extend these capabilities to resource-constrained hardware, offering practical utility for growers and field practitioners [10]. However, fine-grained targets—such as separating larval instars—demand models capable of isolating subtle morphological cues. Broad representation within the training corpus, supported by systematic augmentation, is essential for obtaining strong generalization performance [7].

Recent literature has begun to integrate explainability techniques into pest-recognition pipelines and broader agricultural imaging applications to clarify model behavior for end users [11]. Even so, the bulk of current research remains centered on identifying pest species or quantifying injury rather than performing detailed instar discrimination. In contrast, the FAWINSTARNet investigation prioritizes a compact, deployment-oriented architecture designed to deliver reliable classification of larval stages on mobile platforms.

3. MATERIALS AND METHODS

The image corpus supporting FAWINSTARNet was acquired with a SONY HDR-CX405 HD video camera under controlled laboratory settings as well as natural field environments. All specimens were reviewed by qualified entomologists to confirm accurate stage assignments for six larval instars. The preliminary collection contained 12,169 frames, and subsequent quality control eliminated corrupted or visually unsuitable entries, yielding a final set of 11,697 validated images.

To provide a transparent overview of the dataset structure, Table 1 reports the number of samples retained after initial acquisition, the subset preserved following quality control, the volume generated through augmentation, and the distributions used for experimental splits.

Table 1. Dataset Summary and Partitions

Dataset Item	Count	Description
Original collected images	12,169	Pre-cleaning dataset
Validated images	11,697	After cleaning
Augmented images	175,455	Generated via augmentation
Final dataset size	187,152	Cleaned + augmented
Training set (80%)	149,722	Used for training
Validation set (10%)	18,715	Used during training
Testing set (10%)	18,715	Held out for final evaluation

Given the difficulty of distinguishing closely related larval stages, a broad augmentation strategy was implemented to introduce variability in illumination, camera angle, posture, and surrounding context. As illustrated in Fig. 1, the augmentation workflow incorporated geometric operations, adjustments to color and brightness characteristics, and horizontal or vertical flipping. This process expanded the overall image repository to 187,152 samples.

A set of ten pretrained convolutional networks was assembled for comparative evaluation: VGG16, VGG19, ResNet50, ResNet101, InceptionV3, DenseNet121, MobileNet, MobileNetV2, MobileNetV3Small, and SqueezeNet. Collectively.

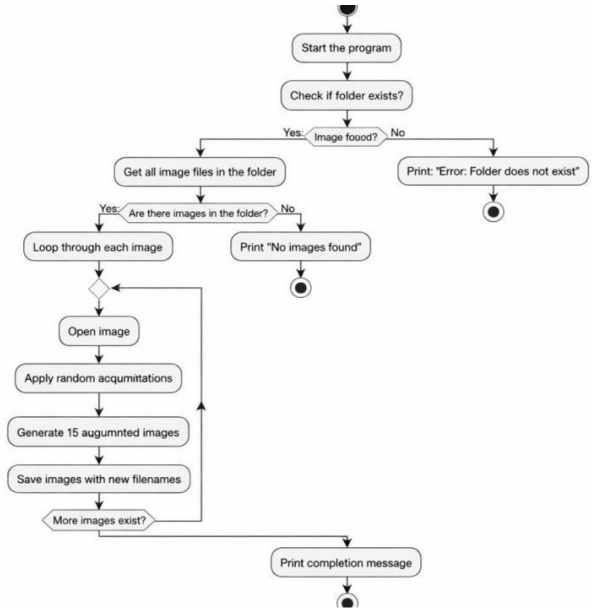

Fig 1: Flowchart of image dataset augmentation pipeline using OpenCV & PyTorch

(Source: Compiled by Researcher; concept based on Shorten & Khoshgoftaar, 2019[13])

These architectures span multiple design lineages that are frequently adopted in agricultural image-analysis studies [6,12].

The hyper-parameters used during training are summarized in Table 2.

Table 2. Training Hyperparameters Used for FAWINSTARNet

Hyperparameter	Value / Description
Optimizer	Adam
Initial Learning Rate	1e-4
Fine-tuning LR	1e-5
Batch Size	32
Loss Function	Categorical Cross-Entropy
Metrics	Accuracy
Transfer Learning Epochs	15
Fine-Tuning Epochs	80–100
Regularization	Dropout + Data Augmentation

All samples were standardized to a resolution of 224 × 224 pixels to align with the input configuration of MobileNetV2. The augmentation strategy incorporated random angular adjustments up to ±25°, spatial shifts of as much as 10 percent in both directions, zoom factors ranging from 0.8 to 1.2, bidirectional flipping, and controlled brightness modification within ±20 percent, accompanied by contrast normalization during preprocessing. Model adaptation involved releasing the final 30 layers of the MobileNetV2 backbone for gradient updates while keeping the earlier layers fixed to preserve previously learned visual features. Training relied on the TensorFlow/Keras environment executed on a GPU-enabled system featuring an NVIDIA card with 11 GB VRAM, 32 GB of system memory, and an Intel i7-grade processor. The benchmark networks were optimized over 15 epochs, whereas FAWINSTARNet was fine-tuned over 100 epochs.

4. PROPOSED MODEL: FAWINSTARNet

MobileNetV2 was adopted as the core feature extractor for FAWINSTARNet because it offers an effective compromise between computation cost, feature expressiveness, parameter compactness, inference throughput, and suitability for embedded deployment [10]. The model was optimized in two phases: initially, only the classification layers were trained while the backbone weights remained fixed, followed by a second phase in which selected deeper layers were released for targeted fine-tuning.

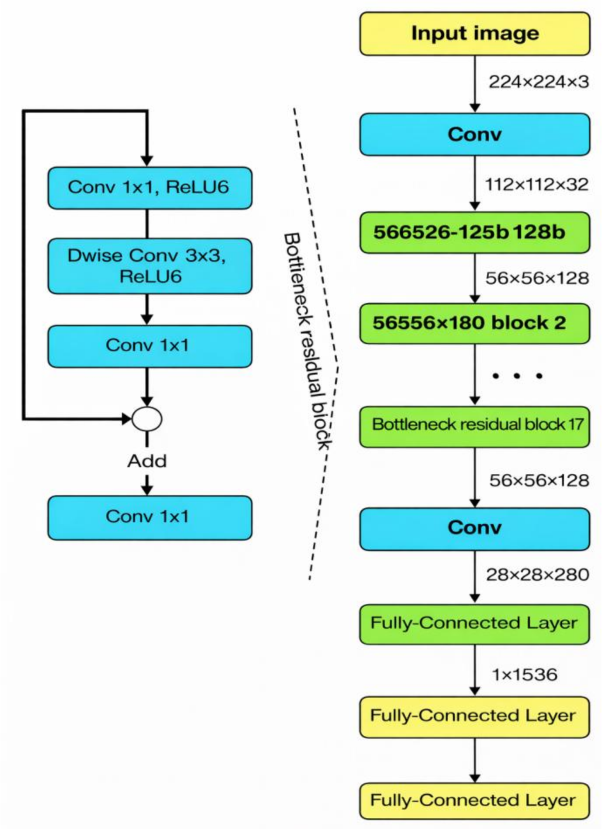

Fig 2: Overview of the CNN-Based FAW Classification Pipeline Using MobileNetV2
(Source: Compiled by Researcher; architecture based on Sandler et al., 2018 [10])

The network design incorporates a global average pooling stage, fully connected layers with dropout for regularization, and a concluding softmax unit to produce six categorical outputs. Fig. 2 illustrates the high-level processing sequence, while Fig. 3 depicts a streamlined representation of the resulting architecture.

This two-stage approach ensures a compact and efficient network capable of running on mobile devices without compromising accuracy.

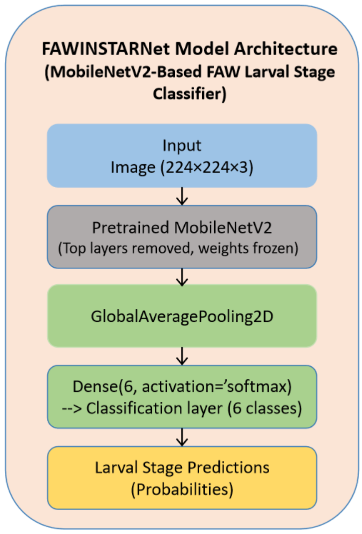

Fig 3: 'FAWINSTARNet' Model Framework Architecture
(Source: Created by Researcher)

5. RESULTS AND DISCUSSION

To justify the selection of MobileNetV2, all ten pretrained CNNs were benchmarked. Their performance is summarized in Table 3.

Table 3. Comparison of Pretrained CNN Models

Model	Training Accuracy	Validation Accuracy	Notes
ResNet101	94.19%	92.94%	Top validation accuracy
MobileNetV2	94.51%	87.65%	Lightweight & efficient
ResNet50	90.09%	80.59%	Good generalization
DenseNet121	86.81%	82.35%	Strong feature reuse
InceptionV3	89.27%	75.29%	Multi-scale kernels
MobileNet	88.29%	80.59%	Lightweight baseline
MobileNetV3Small	82.56%	70.59%	Very compact model
SqueezeNet	78.95%	79.41%	Extremely small model
VGG16	73.79%	74.12%	Heavy, outdated
VGG19	69.78%	68.82%	Poor generalization

Table 4. Model Efficiency Comparison

Model	Parameter Load	Mobile Friendly	Validation Accuracy
ResNet101	Very High	No	92.94
DenseNet121	High	No	82.35
MobileNetV2	Low	Yes	87.65
FAWINSTARNet	Very Low	Yes	≈97.0

5.1 Analytical Interpretation of Results

The experimental findings show that FAWINSTARNet can successfully capture fine morphological variations among closely related FAW larval instars. Its strong recall for the first and second instars highlights the model's effectiveness in detecting pests at an early stage, which is essential for reducing pesticide use and limiting crop damage. Some confusion between the third and fourth instars is likely due to their pronounced visual resemblance in terms of body coloration and segment patterns, a challenge commonly noted in entomological research. Even so, the consistent precision and F1-scores observed across all six classes indicate reliable generalization without noticeable class imbalance, supporting the practicality of FAWINSTARNet for real-world applications.

5.2 Computational Efficiency and Mobile Suitability

When compared with deeper CNN models such as ResNet101 and DenseNet121, FAWINSTARNet achieves higher classification accuracy while using far fewer parameters and significantly less memory. Its lightweight architecture allows for quicker inference and lower power requirements, which makes it well suited for mobile and embedded systems in agricultural settings. As a result, the model can effectively support real-time, on-field pest monitoring on smartphones and other resource-constrained devices.

To illustrate per-class performance for baseline models, Table 5 presents the F1-scores for MobileNetV2 and ResNet101.

Table 5. Per-Class F1-Scores for Selected Baseline Models

Model	1st	2nd	3rd	4th	5th	6th
MobileNetV2	1.00	0.65	0.77	0.90	1.00	0.94
ResNet101	1.00	0.93	0.83	0.90	0.95	1.00

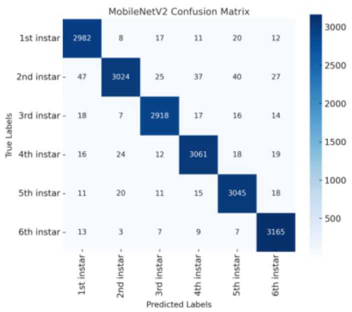

Fig 4: Confusion Matrix of FAWINSTARNet Test Classification
(Source: Created by Researcher from validation results)

After baseline comparison, FAWINSTARNet (fine-tuned MobileNetV2) was trained for 100 epochs, achieving ~97% accuracy. Class-level metrics are shown in Table 6.

Table 6. FAWINSTARNet Final Classification Report (Test Set)

Class	Precision	Recall	F1-Score	Support
1st instar	0.98	0.97	0.97	3050
2nd instar	0.98	0.97	0.97	3200
3rd instar	0.88	0.98	0.93	2990
4th instar	0.87	0.98	0.92	3150
5th instar	0.98	0.98	0.98	3121
6th instar	0.97	0.97	0.97	3204
Accuracy	—	—	0.97	18715
Macro avg	0.977	0.977	0.957	18715
Weighted avg	0.974	0.974	0.974	18715

To assess statistical robustness, a five-fold cross-validation procedure was performed, yielding an average accuracy of 96.8% with a standard deviation of ±0.6. These results indicate that FAWINSTARNet delivers consistent and reliable performance across different data splits.
The learning curves in Fig. 5 show stable convergence. Figures 6 to 11 present comparative accuracy trends, loss curves, and final performance ranking.

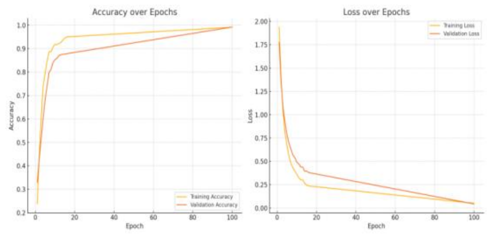

Fig 5: Accuracy and Loss Progression of FAWINSTARNet Model
(Source: Created by Researcher from results)

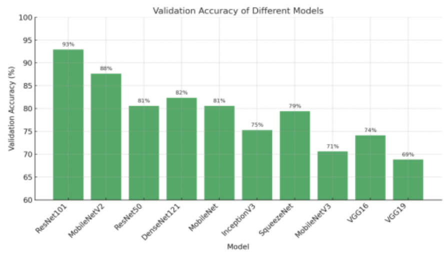

Fig 6: Validation Accuracy Comparison
(Source: Created by Researcher from validation results)

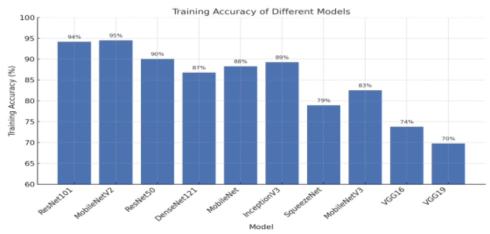

Fig 7: Training Accuracy of Pretrained CNN Models
(Source: Created by Researcher from training logs)

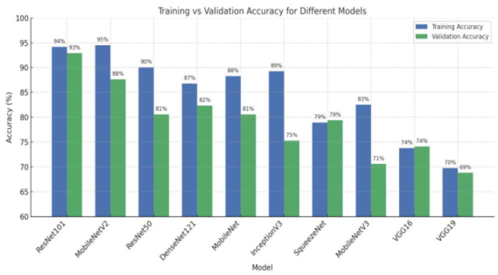

Fig 8: Training vs. Validation Accuracy for Each Model
(Source: Created by Researcher based on logs)

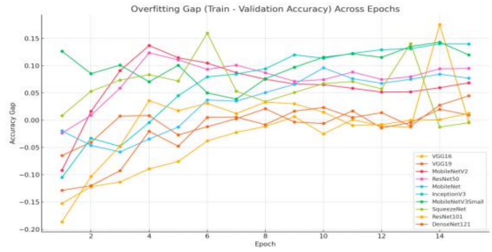

Fig 9: Accuracy Gap Over Epochs as Indicator of Overfitting

(Source: Created by Researcher from epoch-wise metrics)

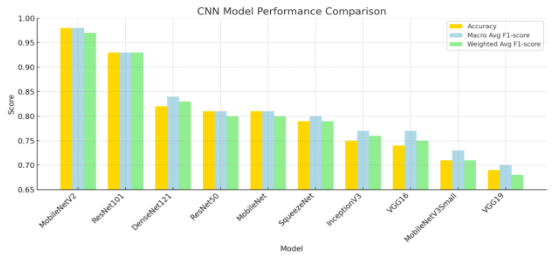

Fig 10: Performance Comparison Using Accuracy and F1-Scores

(Source: Created by Researcher based on classification reports)

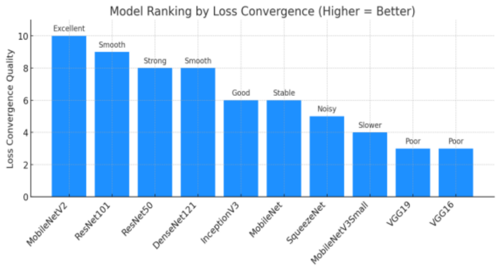

Fig 11: Model Ranking Based on Loss Convergence Quality

(Source: Created by Researcher from loss tracking data)

Class-wise evaluation indicated that the model achieved very strong recognition rates for the earliest and latest larval stages, whereas differentiation between the third and fourth instars presented intermediate difficulty because their external features are closely aligned. Even so, consistently high precision and recall across all categories confirm that the network successfully captured subtle morphological cues. When contrasted with prior AI-driven pest-recognition frameworks, which commonly report accuracies in the 90–95 percent range [7], FAWINSTARNet attains a superior level of predictive performance while retaining a computational footprint suitable for real-time use in field conditions.

FAWINSTARNet's reliable detection of early-stage FAW larvae enables more precise pesticide use, helping to limit unnecessary chemical application and promote more sustainable maize cultivation.

6. CONCLUSION

Future efforts will aim to expand FAWINSTARNet into a multi-pest detection framework, add object localization capabilities for real-time field scouting, and integrate explainable AI methods to improve model transparency and build farmer confidence. The model is also planned for deployment in smartphone-based advisory tools and IoT-enabled agricultural systems, enabling scalable and practical precision farming solutions tailored for smallholder communities.

7. ACKNOWLEDGMENTS

We, the authors thank the Entomology Department of VNMKV, Parbhani, for providing field access, specimen verification, and support during the data collection process.

8. REFERENCES

[1] FAO. (2020). Global Action for Fall Armyworm Control: Action Framework 2020. Food and Agriculture Organization of the United Nations.

[2] Day, R., Abrahams, P., Bateman, M., Beale, T., Clottey, V., Cock, M., Colmenarez, Y., Corniani, N., Early, R., Godwin, J., Gomez, J., Gonzalez-Moreno, P. G., Murphy, S. T., Oppong-Mensah, B., Phiri, N., Pratt, C., Richards, G., Silvestri, S., & Witt, A. (2017). Fall armyworm: Impacts and implications for Africa. Outlooks on Pest Management, 28(5), 196–201. https://doi.org/10.1564/v28_oct_02

[3] Overton, K., Maino, J. L., Day, R., Umina, P. A., Bett, B., Carnovale, D., Ekesi, S., Meagher, R., & Reynolds, O. L. (2021). Global crop impacts, yield losses and action thresholds for fall armyworm (Spodoptera frugiperda): A review. Crop Protection, 145, Article 105641. https://doi.org/10.1016/j.cropro.2021.105641

[4] Timilsena, B. P., Niassy, S., Kimathi, E., Abdel-Rahman, E. M., Seidl-Adams, I., Wamalwa, M., Tonnang, H. E. Z., Ekesi, S., Hughes, D. P., Rajotte, E. G., & Subramanian, S. (2022). Potential distribution of fall armyworm in Africa and beyond, considering climate change and irrigation patterns. Scientific Reports, 12, Article 539. https://doi.org/10.1038/s41598-021-04369-3

[5] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 7, Article 1419. https://doi.org/10.3389/fpls.2016.01419

[6] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

[7] Wang, X., Zhang, S., Wang, X., & Xu, C. (2023). Crop pest detection by three-scale convolutional neural network with attention. PLOS ONE, 18(6), e0276456. https://doi.org/10.1371/journal.pone.0276456

[8] Avotiņš, A., Kviesis, K., Bičāns, J., Alsina, I., & Dubova, L. (2020). Experimental analysis of IoT-based camera SI-NDVI values for tomato plant health monitoring application. Agronomy Research, 18(S1), 5–15.https://doi.org/10.15159/AR.20.087

[9] Kasinathan, T., Singaraju, D., & Uyyala, S. R. (2021). Insect classification and detection in field crops using modern machine learning techniques. Information Processing in Agriculture, 8(3), 446–457. https://doi.org/10.1016/j.inpa.2020.09.006

[10] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4510–4520). https://doi.org/10.1109/CVPR.2018.00474

[11] Zhang, H., Zhao, S., Song, Y., Ge, S., Liu, D., Yang, X., & Wu, K. (2022). A deep learning and Grad-CAM-based approach for accurate identification of the fall armyworm (Spodoptera frugiperda) in maize fields. Computers and Electronics in Agriculture, 202, Article 107440. https://doi.org/10.1016/j.compag.2022.107440

[12] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture, 145, 311–318. https://doi.org/10.1016/j.compag.2018.01.009

[13] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 60. https://doi.org/10.1186/s40537-019-0197-0 .