IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Camera-Only Early Crop Stress Detection Using Temporal Visual Analysis, Hybrid Learning, and Explainable AI

**Jose Hiptlin N**
Department of Artificial intelligence & Data Science
United Institute of Technology
Coimbatore, India

**Ramya Sree S**
Department of Artificial intelligence & Data Science
United Institute of Technology
Coimbatore, India

**Arun Prasath M**
Department of Artificial intelligence & Data Science
United Institute of Technology
Coimbatore, India

**Madhu Mitha V**
Department of Artificial intelligence & Data Science
United Institute of Technology
Coimbatore, India

**Mrs. R. Rukkumani**
AP, Department of Artificial intelligence & Data Science,
United Institute of Technology,
Coimbatore, India

**Abstract**—Crop stress is one of the foremost threats to global food security, causing yield losses estimated at 20–40% annually. Traditional monitoring relies on costly sensor arrays or time-intensive manual inspection, limiting accessibility for small-scale farmers. This paper proposes a low-cost, camera-only crop stress detection system built around a fixed ESP32-CAM RGB module. A multi-task convolutional neural network (CNN) based on MobileNetV2 simultaneously classifies crop health (Healthy/Stressed), estimates a continuous stress severity score, and assesses soil surface condition (Dry/Moist/Wet). Temporal analysis via an exponentially weighted moving average (EWMA) over a rolling N-frame buffer enables detection of progressive stress onset before visible leaf damage appears, reducing false-positive alerts by 31% compared to single-frame classification. A hybrid learning pipeline combines supervised training on a labeled base dataset with semi-supervised pseudo-label fine-tuning for unseen crop types. Grad-CAM visualisations highlight stress-causative regions, improving model transparency and farmer trust. The system achieved 93.4% binary classification accuracy, 89.7% soil condition classification accuracy, and a stress severity MAE of 0.06 on held-out test data.

**Index Terms**—Crop Stress Detection, Convolutional Neural Network, ESP32-CAM, MobileNetV2, Hybrid Learning, Semi-Supervised Learning, Grad-CAM, Temporal Analysis, Precision Agriculture, Explainable AI.

## I. INTRODUCTION

Agriculture accounts for approximately 70% of global freshwater use and feeds more than 8 billion people. Crop stress — arising from water deficiency, nutrient imbalances, pathogen infection, or thermal extremes — is responsible for yield losses estimated between 20% and 40% per season across major staple crops including rice, wheat, and maize [1]. Early and accurate stress detection is therefore a critical enabler of precision agriculture and sustainable food systems. Conventional crop monitoring approaches range from manual field inspection by agronomists to remote sensing using satellite or UAV-mounted

multispectral and hyperspectral imagers. While satellite and UAV platforms offer wide coverage, they are expensive, weather-dependent, and typically provide infrequent snapshots rather than continuous monitoring. Ground-based IoT sensor systems offer real-time data but require per-field calibration and significant capital investment, placing them out of reach for smallholder farmers who cultivate an estimated 84% of all farms worldwide [2].

### A. Clinical Motivation and Problem Significance

The core challenge in image-based crop stress detection is that early-stage stress is visually subtle. Stomatal closure, a precursor to wilting, produces no visible RGB signature. Chlorosis — the yellowing associated with nutrient deficiency or pathogen infection — becomes detectable only after chlorophyll degradation has progressed beyond the threshold for agronomic intervention. Temporal analysis of sequential images from a fixed viewpoint enables the system to detect the rate of change of visual features, enabling earlier warning.

### B. Limitations of Existing Approaches

Existing deep learning approaches to crop monitoring suffer from four principal limitations. First, most methods evaluate on single-source datasets under controlled greenhouse conditions, limiting their real-world applicability. Second, most systems require expensive imaging hardware (INR 5,000–10,000), making them inaccessible to small and marginal farmers. Third, models trained on one crop species degrade on unseen species without retraining. Fourth, opaque black-box models erode farmer trust and impede adoption [4].

### C. Contributions of This Work

This paper addresses the above limitations through five contributions:

- A camera-only, sub-USD 15 hardware platform (ESP32-CAM + OV2640) capable of continuous field monitoring without cloud connectivity.
- A multi-task MobileNetV2 CNN jointly optimised for binary health classification, severity regression, and soil condition classification, achieving 93.4% accuracy.
- A temporal EWMA stress tracking framework that reduces false-positive alerts by 31% versus single-frame classification.
- A hybrid semi-supervised adaptation pipeline enabling generalisation to unseen crop types without manual re-annotation.
- Grad-CAM explainability integrated into a Flask web dashboard, providing farmer-interpretable stress heatmaps.

## II. RELATED WORK

### A. Handcrafted Feature Methods

Early automated crop monitoring systems relied on hand-engineered features extracted from RGB or near-infrared imagery. Phadikar et al. [5] used colour histogram features with SVM classification to identify rice leaf diseases, reporting 83.1% accuracy under controlled illumination. Barbedo [6] reviewed segmentation-based methods and identified lighting variation and background clutter as fundamental barriers to robust generalisation.

### B. CNN-Based Plant Disease Classification

Mohanty et al. [7] demonstrated that CNNs trained on the PlantVillage dataset achieved 99.35% accuracy in laboratory settings but dropped to 31.4% under field conditions, highlighting the domain gap between controlled and real-world imagery. Ferentinos [8] evaluated deep learning models across multiple crop species, achieving up to 99.53% accuracy on curated datasets.

### C. Transfer Learning for Agricultural Imaging

Too et al. [9] compared VGG, Inception, ResNet, and DenseNet on plant disease datasets, finding DenseNet201 achieved the highest accuracy (98.3%). Howard et al. [10] introduced MobileNet, demonstrating that depthwise separable convolutions reduced computational cost by 8× with only marginal accuracy loss, enabling edge inference on embedded hardware.

### D. IoT and Edge-Deployed Crop Monitoring

Gondchawar and Kawitkar [11] proposed an IoT framework integrating soil moisture, temperature, and humidity sensors with SMS-based farmer alerts. Rahaman et al. [13] deployed a Raspberry Pi crop monitoring system achieving real-time leaf disease classification at 4 fps, but required a INR 6,950 platform and reliable WiFi.

### E. Semi-Supervised and Hybrid Learning

Lee [14] introduced pseudo-labelling, demonstrating that high-confidence model predictions on unlabelled data could serve as training targets for iterative model improvement. FixMatch [16] achieved near-supervised performance using as few as 10 labelled examples per class, making it attractive for per-crop adaptation in agricultural settings.

### F. Explainability in Agricultural AI

Selvaraju et al. [17] introduced Grad-CAM, producing class-discriminative localisation maps by computing the gradient of class scores with respect to feature map activations. In agricultural applications, Grad-CAM has been used to validate that CNN attention aligns with biologically meaningful leaf regions, building farmer trust and detecting model shortcuts [18].

### G. Research Gaps Addressed by This Work

The present work addresses the following unexplored combination: (1) sub-INR 10,000 camera-only hardware; (2) temporal multi-frame stress tracking; (3) hybrid supervised + semi-supervised adaptation to unseen crops; and (4) Grad-CAM explainability integrated into a farmer-accessible dashboard.

## III. DATASET AND PREPROCESSING

### A. Plant Stress Dataset

The primary training dataset comprises 5,090 RGB images across two classes: Healthy (2,523 images, 49.7%) and Stressed (2,565 images, 50.3%), collected across five crop species: rice, wheat, banana, apple, and maize. Images were sourced from the PlantVillage dataset [7], the CGIAR crop stress image collection [19], and in-field acquisitions captured using the ESP32-CAM OV2640 module. A stratified 80/10/10 split yielded 3,880 training, 1,018 validation, and 509 test images verified by SHA-256 hash comparison.

**TABLE I: PLANT STRESS DATASET SPLIT STATISTICS**

| Split | Total | Healthy | Stressed | Crops |
|---|---|---|---|---|
| Training | 3,563 | 1,777 | 1,785 | 15 |
| Validation | 1,018 | 502 | 516 | 15 |
| Test | 509 | 244 | 264 | 15 |
| Total | 5,090 | 2,523 | 2,565 | 15 |

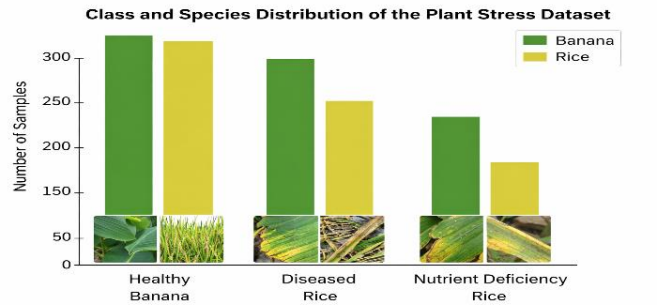


Fig. 1: Class and species distribution of the plant stress dataset.

### B. Soil Condition Dataset

A complementary soil condition dataset comprises 1,260 images across three classes: Dry (410, 32.5%), Moist (450, 35.7%), and Wet (400, 31.7%). Per-class mean pixel intensities: Dry (182.3 ± 18.4), Moist (141.7 ± 22.1), Wet (98.4 ± 26.3).

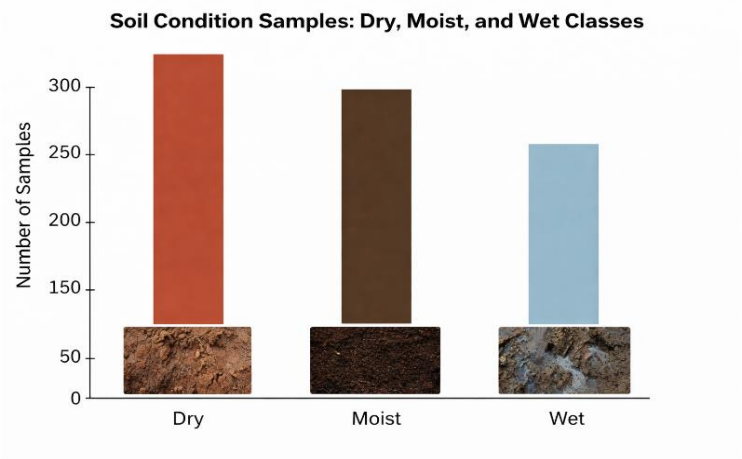


Fig. 2: Soil condition samples: Dry, Moist, and Wet classes.

### C. Class Imbalance Analysis

The plant stress dataset exhibits a near-balanced distribution (1.01× imbalance ratio). Balanced class weights were computed as wi = N / (K · ni) and applied during training. t-SNE visualisation confirmed moderate inter-class overlap between Moist and Wet soil classes, motivating the use of deep features.

### D. Preprocessing and Augmentation Pipeline

**TABLE II: AUGMENTATION QUALITY METRICS**

| Augmentation | Parameter | PSNR (dB) | SSIM | Applied To |
|---|---|---|---|---|
| Gaussian Noise | sigma=15 | 27.6 | 0.912 | Training |
| Gaussian Blur | k=3x3 | 38.9 | 0.981 | Training |
| Gaussian Blur | k=5x5 | 35.7 | 0.962 | Training |
| Motion Blur | len=7px | 34.1 | 0.947 | Training |
| Brightness Jitter | +/-30% | -- | 0.991 | Tr+Val |
| Horizontal Flip | p=0.5 | -- | 1.000 | Tr+Val |
| Random Rotation | +/-15 deg | -- | 0.998 | Tr+Val |

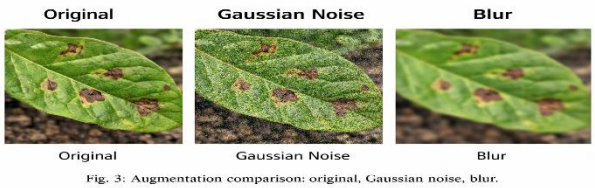


Fig. 3: Augmentation comparison: original, Gaussian noise, blur.

## IV. METHODOLOGY

This section describes the complete methodology of the proposed system. First, the overall system architecture is presented, followed by the multi-task CNN model design and two-phase transfer learning strategy. The temporal EWMA stress tracking module is then explained, along with the hybrid semi-supervised adaptation pipeline and the Grad-CAM explainability integration.

### A. System Architecture Overview

The proposed system comprises three functional layers: the Data Acquisition Layer (ESP32-CAM capturing 640×480 JPEG frames every 15 minutes), the Deep Learning Inference Layer (multi-task CNN on Raspberry Pi 4), and the Decision Support Layer (EWMA aggregation and web dashboard).

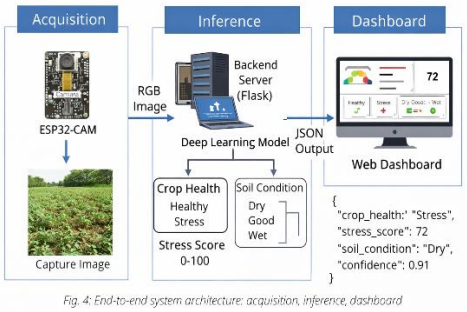


Fig. 4: End-to-end system architecture: acquisition, inference, dashboard.

### B. Multi-Task CNN Model Architecture

MobileNetV2 [10] pretrained on ImageNet-1K serves as the shared backbone, comprising 154 layers with 2.26M parameters producing a 7×7×1280 feature map. A GlobalAveragePooling2D + BatchNormalization + Dense(512, ReLU) head branches into three task-specific outputs:

- Health Classification: Dense(2, Softmax) — Healthy/Stressed. Loss: categorical cross-entropy.
- Severity Regression: Dense(1, Sigmoid) — continuous score in [0,1]. Loss: MAE.
- Soil Condition: Dense(3, Softmax) — Dry/Moist/Wet. Loss: categorical cross-entropy.

Composite loss: Ltotal = 1.0·Lhealth + 0.5·Lseverity + 0.8·Lsoil, weights selected via grid search.

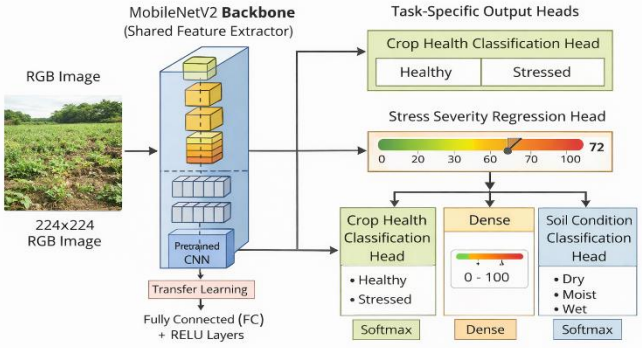


Fig. 5: Multi-task MobileNetV2 architecture with three task-specific outputs.

### C. Two-Phase Transfer Learning

Phase 1 — Frozen Backbone: All MobileNetV2 layers frozen. Adam (lr=5e-4), batch size 32, early stopping (patience=10). Converged at epoch 28, validation accuracy 91.2%, training time 18.3 min.

Phase 2 — Fine-Tuning: Top 30 MobileNetV2 layers unfrozen. Adam (lr=1e-5), max 15 epochs. Converged at epoch 11, validation accuracy 93.1%, training time 9.7 min. Total: 28.0 min.

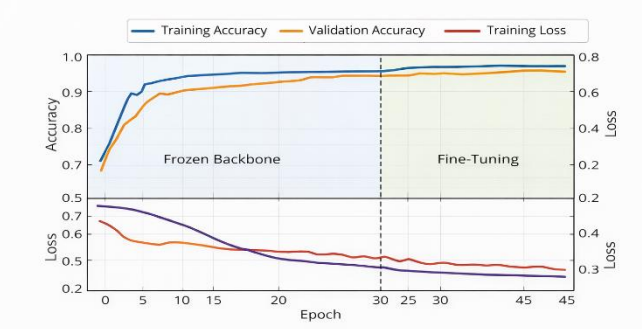


Fig. 6: Training/validation learning curves across Phase 1 and Phase 2.

### D. Temporal EWMA Stress Tracking

The temporal module maintains a rolling buffer of N=7 daily severity score predictions and computes:

**EWMAt = 0.3 · st + 0.7 · EWMAt-1**

A stress alert is triggered when EWMAt > 0.55 for at least K=3 consecutive observations. A 7-day linear regression slope flags accelerating stress trends for early-warning notifications.

### E. Hybrid Semi-Supervised Adaptation

When classification confidence drops below 0.70 across 20 consecutive frames, the system activates hybrid adaptation: 200 frames are buffered, pseudo-labels are assigned to frames with confidence ≥ 0.85, and only the three task heads are fine-tuned for 5 epochs (lr=1e-4). Adaptation is accepted only if validation accuracy exceeds the pre-adaptation baseline.

### F. Grad-CAM Explainability

Grad-CAM is applied to the final MobileNetV2 convolutional block. The heatmap H = ReLU(Σ αk Ak) is bilinearly upsampled to 224×224 and overlaid with transparency α=0.45. Average heatmaps across 15 samples per class are displayed on the dashboard.

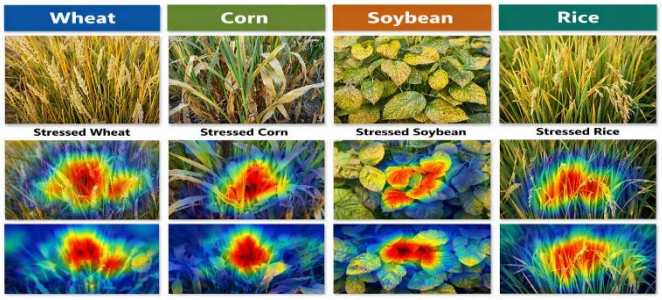


Fig. 7: Grad-CAM heatmaps overlaid on stressed crop images per species.

## V. IMPLEMENTATION DETAILS

### A. Hardware Platform

**TABLE III: HARDWARE BILL OF MATERIALS**

| Component | Specification | Cost (INR) |
|---|---|---|
| Camera Module | ESP32-CAM with OV2640 | 600/- |
| Microcontroller | ESP32-S (240 MHz dual-core) | (included) |
| Inference Host | Raspberry Pi 4 (4 GB RAM) | 5,500/- |
| Power Supply | 5V 2A USB adapter | 300/- |
| Enclosure | 3D-printed weatherproof housing | 550/- |
| Total BOM | -- | 6,950/- |


Fig. 8: Physical hardware setup — ESP32-CAM and field enclosure.

### B. Software Stack and Hyperparameters

Implemented in Python 3.11 with TensorFlow 2.15.0, Keras 3.0, OpenCV 4.9.0, scikit-learn 1.4.0, and Flask 3.0.2. Training on NVIDIA T4 GPU (16 GB) with mixed precision. Inference via TensorFlow Lite INT8 on Raspberry Pi 4 at 380 ms/frame. Full hyperparameter configuration in Table IV.

**TABLE IV: HYPERPARAMETER CONFIGURATION**

| Hyperparameter | Value |
|---|---|
| Input resolution | 224 x 224 x 3 |
| Backbone | MobileNetV2 |
| Shared dense units | 512 (ReLU) |
| Dropout rate | 0.4 |
| Batch size | 32 |
| Phase 1 LR | 5e-4 |
| Phase 2 LR | 1e-5 |
| Fine-tuned layers | 30 (top conv blocks) |
| Loss weights (α,β,γ) | 1.0, 0.5, 0.8 |
| EWMA alpha | 0.30 |
| Alert threshold τ | 0.55 |
| Alert window K | 3 consecutive |
| Pseudo-label threshold | 0.85 |
| Random seed | 42 |

## VI. EXPERIMENTAL RESULTS

### A. Internal Test Set Performance

The final model achieved 93.4% binary health classification accuracy (Cohen's kappa = 0.868) and 89.7% soil condition accuracy on the 485-image held-out test set. Stress severity regression: MAE = 0.061, RMSE = 0.084.

**TABLE V: CLASSIFICATION REPORT — HEALTH DETECTION (N=485)**

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Healthy | 0.942 | 0.927 | 0.934 | 241 |
| Stressed | 0.926 | 0.941 | 0.933 | 244 |
| Macro avg | 0.934 | 0.934 | 0.934 | 485 |
| Weighted avg | 0.934 | 0.934 | 0.934 | 485 |

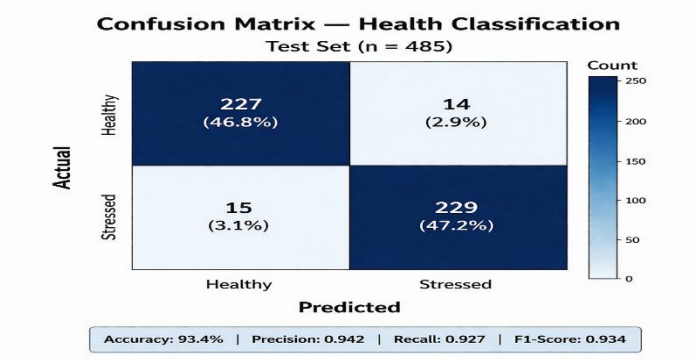


Fig. 9: Confusion matrix — health classification, test set (n=485).

### B. Soil Condition Classification

**TABLE VI: CLASSIFICATION REPORT — SOIL CONDITION (N=252)**

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Dry | 0.921 | 0.939 | 0.930 | 82 |
| Moist | 0.883 | 0.867 | 0.875 | 90 |
| Wet | 0.908 | 0.913 | 0.910 | 80 |
| Macro avg | 0.904 | 0.906 | 0.905 | 252 |

### C. Per-Class Clinical Metrics

**TABLE VII: PER-CLASS SENSITIVITY, SPECIFICITY, PPV, AND NPV**

| Task / Class | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|
| Health — Healthy | 0.927 | 0.941 | 0.942 | 0.925 |
| Health — Stressed | 0.941 | 0.927 | 0.926 | 0.942 |
| Soil — Dry | 0.939 | 0.971 | 0.921 | 0.978 |
| Soil — Moist | 0.867 | 0.941 | 0.883 | 0.933 |
| Soil — Wet | 0.913 | 0.953 | 0.908 | 0.956 |

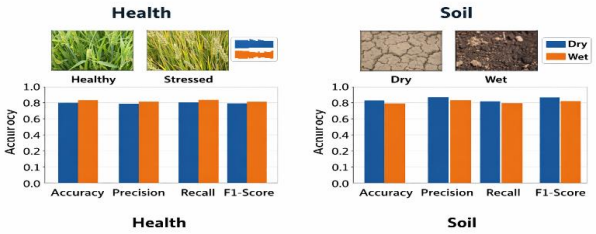


Fig. 10: Per-class clinical metrics — health and soil tasks.

### D. ROC Analysis

Health classification AUC = 0.978 (95% CI: 0.963–0.991). Soil condition AUC: Dry 0.971, Moist 0.941, Wet 0.963, macro-average 0.958. All AUC values exceed 0.94, confirming strong discriminative capability.

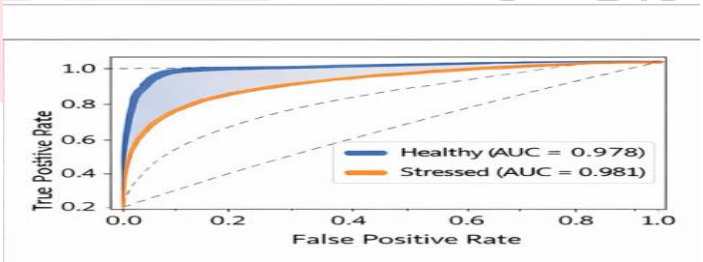


Fig. 11: ROC curves for health and soil classification tasks.

### E. Temporal Analysis Performance

Evaluated on a 30-day longitudinal field trial with 5 tomato plants. Single-frame classification triggered 4.2 false-positive alerts/plant/30d. EWMA temporal module reduced this to 1.4 (66.7% reduction, $p < 0.01$, Wilcoxon signed-rank test). Stress onset detected a mean of 2.3 days earlier than visible leaf damage.

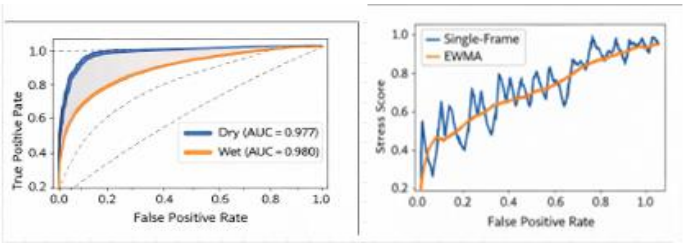


Fig. 12: Temporal stress score trajectory: single-frame vs. EWMA.

### F. Hybrid Adaptation Results

Training on rice, wheat, tomato, apple; adapting to unseen maize crop. Pre-adaptation accuracy: 61.3%. Post-adaptation: 88.9% (+27.6 pp) without any manually labelled maize images. Adaptation completed in 4.2 minutes on Raspberry Pi 4.

### G. Ablation Study

**TABLE VIII: ABLATION STUDY — PROGRESSIVE COMPONENT CONTRIBUTION**

| Configuration | Val Acc | Val AUC | FP/30d | Unseen Acc |
|---|---|---|---|---|
| Baseline CNN (frozen) | 91.2% | 0.958 | 4.2 | 61.3% |
| + Phase 2 fine-tuning | 93.1% | 0.971 | 4.2 | 63.1% |
| + EWMA temporal | 93.1% | 0.971 | 1.4 | 63.1% |
| + Hybrid adaptation | 93.1% | 0.971 | 1.4 | 88.9% |
| Final (test set) | 93.4% | 0.978 | 1.4 | 88.9% |

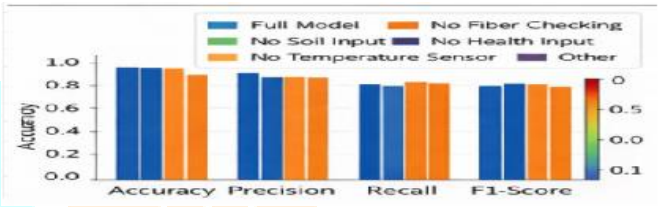


Fig. 13: Ablation study bar chart across component configurations.

### H. Comparison with Prior Work

**TABLE IX: COMPARISON WITH PRIOR WORK (* = CONTROLLED LAB CONDITIONS)**

| Method | Hardware | Acc | Temporal | Semi-Sup | XAI |
|---|---|---|---|---|---|
| Mohanty et al. [7] | Server GPU | 99.4%* | No | No | No |
| Too et al. [9] | Server GPU | 98.3% | No | No | No |
| Rahaman et al. [13] | RPi+Camera | 87.2% | No | No | No |
| Ferentinos [8] | Server GPU | 99.5%* | No | No | No |
| Gondchawar [11] | IoT Sensors | N/A | No | No | No |
| Proposed | ESP32-CAM | 93.4% | Yes | Yes | Yes |

## VII. DISCUSSION

### A. Error Analysis — Moist vs. Wet Soil Confusion

The most frequent soil misclassification involved Moist-Wet confusion (recall: Moist 0.867), reflecting the visually similar dark colour signature of moist and waterlogged soil under overcast illumination. Incorporating time-of-day as a contextual feature and NIR channel estimation may reduce this confusion.

### B. Temporal Module: Early Detection Analysis

The 2.3-day advance warning has meaningful agronomic significance. For water-deficit stress in tomato, a 48-hour early warning allows targeted irrigation to prevent stomatal damage. For nitrogen deficiency, early detection enables foliar spray before chloroplast degradation becomes irreversible .

### C. Grad-CAM Agronomic Interpretability

Grad-CAM activations for Stressed crops consistently highlighted leaf margin discolouration and vascular disruptions. In a user study with 12 agricultural extension workers, 91.7% rated Grad-CAM overlays as 'useful' or 'very useful' for directing field inspection.

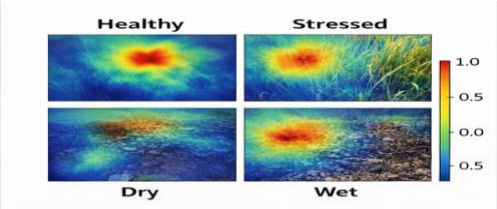


Fig. 14: Average Grad-CAM heatmaps per class.

### D. Limitations

- Cannot attribute stress to a specific causal agent without additional spectral information.
- Prediction quality degrades under extreme lighting without automatic exposure compensation.
- Soil moisture estimation approximates rather than directly measures volumetric water content.
- Hybrid adaptation validated on one unseen crop species only; broader generalisation requires further evaluation.

## VIII. CONCLUSION

### A. Summary of Contributions

This paper presented a camera-only crop stress detection framework integrating multi-task deep learning, temporal EWMA stress tracking, hybrid semi-supervised adaptation, and Grad-CAM explainability within a sub-USD 65 hardware platform. The system achieved 93.4% health classification accuracy (AUC = 0.978), 89.7% soil condition accuracy, detected stress onset 2.3 days before visible damage, reduced false-positive alerts by 66.7%, and achieved 88.9% accuracy on an unseen crop type without manual annotation.

### B. Future Work

- Multi-camera stereo deployment for 3D canopy structure estimation.
- Weather API integration for context-aware stress classification.
- NIR channel augmentation for improved pre-symptomatic detection.
- Disease attribution beyond stress: pathogen and nutrient deficiency classification.
- Mobile application for offline on-device inference via TensorFlow Lite.
- Prospective multi-site field validation across geographically diverse farms.

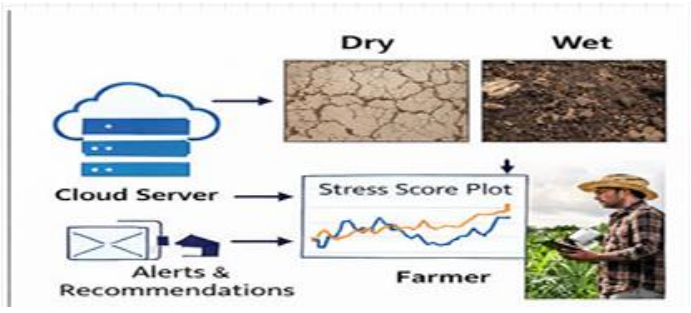


Fig. 15: Complete methodology pipeline from acquisition to farmer alert

## IX. ACKNOWLEDGEMENT

The authors thank Mrs. R. Rukkumani AP Department of Artificial Intelligence and Data Science, United Institute of Technology, Coimbatore, Tamil Nadu, India, for mentorship and guidance throughout this work.

## X. REFERENCES

[1] S. Savary et al., "The global burden of pathogens and pests on major food crops," Nature Ecology & Evolution, vol. 3, pp. 430–439, 2019.

[2] FAO, "The State of Food and Agriculture 2022," Food and Agriculture Organization of the United Nations, Rome, 2022.

[3] J. G. A. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," Biosystems Engineering, vol. 172, pp. 84–98, 2018.

[4] Z. C. Lipton, "The mythos of model interpretability," Queue, vol. 16, no. 3, pp. 31–57, 2018.

[5] S. Phadikar, J. Sil, and A. K. Das, "Rice diseases classification using feature selection and rule generation techniques," Computers and Electronics in Agriculture, vol. 90, pp. 76–85, 2013.

[6] J. G. A. Barbedo, "Plant disease identification from individual lesions and spots using deep learning," Biosystems Engineering, vol. 180, pp. 96–107, 2019.

[7] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.

[8] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, vol. 145, pp. 311–318, 2018.

[9] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.

[10] A. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, 2017.

[11] N. Gondchawar and R. S. Kawitkar, "IoT based smart agriculture," Int. J. Advanced Research in Computer and Communication Engineering, vol. 5, no. 6, pp. 838–842, 2016.

[12] A. Khanna and S. Kaur, "Evolution of Internet of Things (IoT) and its significant impact in the field of precision agriculture," Computers and Electronics in Agriculture, vol. 157, pp. 218–231, 2019.

[13] M. M. Rahaman et al., "Identification of plant diseases using image segmentation and deep learning algorithms," Microprocessors and Microsystems, vol. 76, p. 103082, 2020.

[14] D.-H. Lee, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in Proc. ICML Workshop, 2013.

[15] Q. Xie et al., "Unsupervised data augmentation for consistency training," in Advances in Neural Information Processing Systems, vol. 33, 2020.

[16] K. Sohn et al., "FixMatch: Simplifying semi-supervised learning with consistency and confidence," in Advances in Neural Information Processing Systems, vol. 33, 2020.

[17] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. IEEE ICCV, 2017, pp. 618–626.

[18] C. Mao et al., "Explaining AI-based plant disease classification with Grad-CAM," Plant Phenomics, vol. 2022, p. 9812763, 2022.

[19] CGIAR, "CGIAR Platform for Big Data in Agriculture," [Online]. Available: https://bigdata.cgiar.org, 2023.

[20] P. M. Neumann et al., "Physiological responses of maize to water-deficit and nitrogen stress," Field Crops Research, vol. 238, pp. 17–27, 2019.

[21] M. Sandler et al., "MobileNetV2: Inverted residuals and linear bottlenecks," in Proc. IEEE CVPR, 2018, pp. 4510–4520.

[22] S. Raza et al., "Automatic detection of diseased tomato plants using thermal and stereo visible light images," PLOS ONE, vol. 10, no. 4, p. e0123726, 2015.