Smart Agricultural Technology 13 (2026) 101677

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Optimizing artificial lighting for convolutional neural network-based crop monitoring with low-cost RGB imaging in indoor cultivation

Matteo Landolfo, Fabio Perotti, Alessandro Pistillo, Giuseppina Pennisi *, Giorgio Gianquinto, Francesco Orsini

Department of Agricultural and Food Sciences, Alma Mater Studiorum, Bologna University, Viale Giuseppe Fanin, 40-50, Bologna 40127, Italy

ARTICLE INFO

Keywords:
Computer Vision
Convolutional Neural Network (CNN)
Matthews Correlation Coefficient (MCC)
Artificial Lighting
Indoor Farming
Image Sensor

ABSTRACT

This study investigated the effect of different red:blue (R:B) spectral light ratios on the performance of a multi-task convolutional neural network (CNN) model developed for the automatic classification of four horticultural species and their corresponding phenological stages under controlled artificial lighting conditions. The model was trained and tested using RGB images acquired under five distinct spectral treatments (R:B 1, 3, 5, 7, and 9), and its performance was evaluated using accuracy, precision, recall, F1-score, and Matthews correlation coefficient (MCC). For species classification, the best results were obtained with an R:B 1, achieving an accuracy of 86 %, precision of 87 %, recall of 85 %, F1-score of 85 %, and MCC of 0.81. In terms of phenological stage classification, the highest performance was observed at R:B 3 and R:B 5, both yielding 93 % accuracy and F1-score, precision and recall above 92 %, and an MCC of 0.86. These findings demonstrate that the multi-task CNN model is capable of learning robust and generalizable representations, maintaining high classification performance even under non-optimal spectral conditions. The integration of optimized artificial lighting with intelligent classifiers proves to be a strategic approach for automated monitoring systems in indoor and precision agriculture. Future research should explore the impact of additional spectral components (e.g., green or far-red wavelengths) and the adoption of more advanced neural architectures to further enhance the system's robustness and scalability.

1. Introduction

In recent decades, agriculture has faced unprecedented challenges, primarily driven by the growth of the global population and the effects of climate change. By 2050, the global population is projected to reach 9.7 billion, resulting in a 70 % increase in food demand compared to current levels (United Nations, 2017). The rising demand for food places pressure on agricultural systems already compromised by urbanization, soil degradation, and intensive farming practices [1]. Climate change exacerbates these challenges by increasing crop vulnerability to extreme weather events, posing serious risks to global food security [2,3]. This situation demands a significant shift in agricultural practices, adopting more sustainable and efficient technologies to proactively address these challenges. One of the most promising recent proposals is precision agriculture, which leverages advanced technologies to monitor and optimize resources in real-time. Research has shown that these techniques can reduce water usage by up to 50 % and fertilizer use by up to 30 % in open-field agriculture [4]. However, the application of such solutions is not limited solely to extensive open-field agriculture but is also particularly relevant in controlled environment systems (indoor farming), where the integration of advanced technologies enables precise control of environmental parameters and optimized resource

All authors agree that: This research presents an accurate account of the work performed, all data presented are accurate and methodologies detailed enough to permit others to replicate the work. This manuscript represents entirely original works and or if work and/or words of others have been used, that this has been appropriately cited or quoted and permission has been obtained where necessary. This material has not been published in whole or in part elsewhere. The manuscript is not currently being considered for publication in another journal. That generative AI and AI-assisted technologies have not been utilized in the writing process or if used, disclosed in the manuscript the use of AI and AI-assisted technologies and a statement will appear in the published work. That generative AI and AI-assisted technologies have not been used to create or alter images unless specifically used as part of the research design where such use must be described in a reproducible manner in the methods section. All authors have been personally and actively involved in substantive work leading to the manuscript and will hold themselves jointly and individually responsible for its content.

* Corresponding author.
E-mail address: giuseppina.pennisi@unibo.it (G. Pennisi).

https://doi.org/10.1016/j.atech.2025.101677
Received 6 August 2025; Received in revised form 25 November 2025; Accepted 27 November 2025
Available online 28 November 2025

utilization [5,6]. Recent studies indicate that new cultivation systems, supported by automation and artificial intelligence technologies, are essential for effectively addressing global challenges [7].

Vertical farms (VFs), or Plant Factories with Artificial Lighting (PFAL), represent a promising response to global agricultural challenges, offering a sustainable solution through cultivation in indoor environments [8]. This agricultural production model overcomes the limitations of traditional farming by eliminating the influence of variable climatic factors and allowing for continuous production throughout the year [6]. Vertical farms efficiently utilize limited urban spaces, employing advanced technologies to manage environmental parameters, lighting, irrigation, and nutrient delivery, thereby improving productivity per area unit [9]. Resource efficiency is one of the main advantages of vertical farms. These systems employ soil-less cultivation techniques, such as hydroponics, aeroponics, and aquaponics, to drastically reduce water consumption and fertilizers compared to traditional agriculture [6]. According to [10], hydroponic systems can reduce water usage by up to 90 % while maintaining high productivity levels. Additionally, the aeroponic system, which directly nebulizes nutrients to the plant roots, has proven even more efficient in water and nutrient usage, further increasing sustainability [11]. VFs are particularly suited for urban farming, helping to reduce the distance between production sites and consumers, thus lowering transportation costs and decreasing food distribution-related emissions [9]. This approach supports the growth of local food economies, enhancing access to fresh and nutritious food. Integrating these indoor cultivation systems with Internet of Things (IoT) technologies allows for real-time monitoring of critical variables, such as plant growth, environmental conditions, and nutrient levels, optimizing decision-making processes and minimizing waste [12]. However, the effectiveness of these techniques strongly depends on the accurate management of resources, as any errors in nutrient delivery can severely compromise crop health [13].

Although vertical farms offer innovative solutions for indoor cultivation, they still face significant challenges that limit their large-scale adoption. One of the main issues is the high energy cost, primarily related to the intensive use of artificial lighting required to support photosynthesis and ensure optimal plant growth. Despite advances in light-emitting diode (LED) technology, lighting remains one of the most energy-consuming components, significantly impacting operational costs [14]. Managing the controlled environment also requires constant monitoring and adjustment of temperature, humidity, and CO_2 levels, resulting in additional energy consumption to maintain optimal conditions for crop growth [15]. Another critical aspect is the integrated management of resources such as water and nutrients. While hydroponic technologies reduce resource use compared to traditional agriculture, dynamic and efficient management of these resources requires advanced control technologies capable of adapting in real-time to the plants' needs. However, the success of these systems depends on optimizing the cultivation parameters to meet the physiological needs of the plants. One of the most relevant challenges is the adaptation of these parameters based on the cultivated species and its specific phenological stage, as each growth phase requires different environmental conditions to ensure optimal development (Kozai, 2019). Regulating cultivation conditions in vertical farming systems is mainly based on experimental data and manual or semi-automatic control systems [16]. Although these approaches have proven somewhat effective, they present significant limitations regarding efficiency and scalability. The need to continuously adapt the settings based on the species and growth phase of the plant consumes time and resources, making it difficult to implement fully automated systems on a large scale [17].

Furthermore, the variability of lighting conditions can compromise the quality of images acquired by computer vision systems, negatively affecting the performance of recognition algorithms [18]. This aspect represents a particularly relevant challenge, as the reliability of analyses strongly depends on the system's ability to compensate for lighting variations and extract accurate information from images [19]. To address these issues, developing an advanced monitoring system capable of automatically identifying the cultivated species and its corresponding phenological stage is essential, providing real-time information to optimize growth conditions.

Integrating computer vision into indoor farming offers a unique opportunity to overcome the limitations of traditional methods, enhancing efficiency and automation in crop management [20]. Implementing a computer vision system for species and phenological stage classification in indoor cultivation represents an essential step toward the automation of precision agriculture. The main objective of this study is to develop a classification system based on a computer vision algorithm that is robust to variations introduced by different light treatments used in indoor farming. Specifically, the aims are:

- To evaluate how different light spectra affect the quality of images captured under indoor conditions.
- To determine which light spectrum provides the most reliable data for classifying horticultural species and their phenological stages, while minimizing visual distortions.

The approach combines RGB imaging techniques with algorithmic analysis to significantly contribute to the existing literature on the interaction between lighting and computer vision in indoor farming, as lighting conditions can alter the appearance of objects or the overall content of images, negatively impacting the performance of recognition algorithms. By analyzing how different light spectra influence image quality and developing classification algorithms that are robust to lighting variations, the study aims to provide practical guidelines for optimizing lighting systems and optical sensors. The expected results will enhance the efficiency and reliability of automated monitoring systems in indoor farming, promoting more sustainable and productive precision agriculture.

2. Materials and methods

2.1. Experimental setup and design

The experiment was conducted at AlmaVFarm, the experimental vertical farm (Fig. 1A) at the Department of Agricultural and Food Sciences at the University of Bologna (Bologna, Italy). The experiment was divided into germination (2 weeks) and growth (3 weeks), with environmental parameters optimized for each phase. During the germination phase, the facility maintained a relative humidity of 85 ± 5 % and a temperature of 22 ± 1 °C. The lighting conditions included a 16-hour photoperiod with a light intensity of 180 ± 10 μmol m^{-2} s^{-1}, ensuring uniform and optimal seedling emergence and early root development conditions. In the subsequent growth phase, the relative humidity was reduced to 65/70 ± 10 % day/night to support transpiration and photosynthesis, while the temperature was regulated to 24/21 ± 1 °C (day/night). The light intensity was increased to 215 ± 10 μmol m^{-2} s^{-1}, maintaining the same 16-hour photoperiod. Additionally, supplemental CO_2 was provided inside the growth chamber to maintain a stable concentration of 850 ppm. An ebb-and-flow hydroponic system was used, divided into five sectors with a 3-level tray layout (each tray covering an area of 0.53 m^2). The system operated with a closed water cycle: the nutrient solution from the main reservoir circulated daily for 10 min, flooding each tray and returning to the main reservoir when not absorbed by the plants or substrate. The nutrient solution's electrical conductivity (EC) and pH were manually checked three times a week using a tester (HI98130, Hanna Instruments Italia Srl, Ronchi di Villafranca Padovana, Italy) in the main reservoir. Adjustments were carried out through a fertigation system (Nido One Pro, Nido Srl, Rimini, Italy), ensuring consistent maintenance of the desired values without the need for continuous manual intervention. The nutrient solution used throughout the growth cycle had the following composition: N—NO_3: 15.44 mM; N—NH_4: 1.93 mM; P: 1.93 mM; K: 9.65 mM; S-SO_4: 2.32 mM;

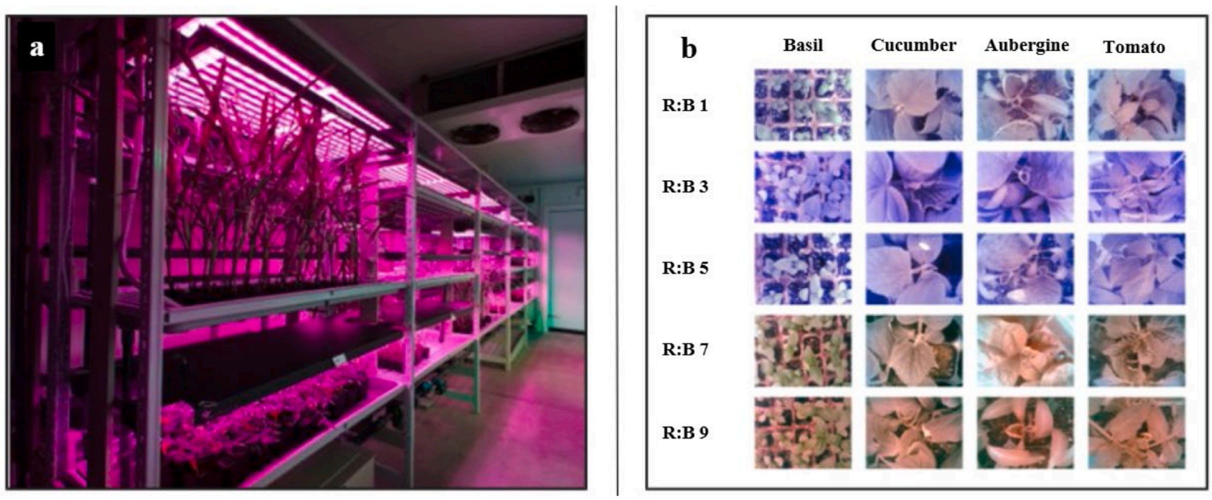

Fig. 1. (a) Experimental vertical farm at the department of agricultural and food sciences, University of Bologna. (b) Example photos captured for each species (from left to right: basil, cucumber, aubergine, and tomato) and for each light treatment (R:B 1, R:B 3, R:B 5, R:B 7, R:B 9).

Ca: 4.34 mM; Mg: 0.97 mM; Fe: 45.7 μM; Cu: 15.1 μM; Zn: 14.8 μM; B: 30.0 μM; Mn: 46.3 μM; Mo: 0.7 μM.

The study focused on four horticultural species at the seedling stage: basil (*Ocimum basilicum* var. Genovese), cucumber (*Cucumis sativus* var. Marketmore 76), aubergine (*Solanum melongena* var. Violetta lunga di Romagna), and tomato (*Solanum lycopersicum* var. ciliegino). The plants were cultivated for the entire duration of the study (5 weeks) in 104-cell trays (0.02 L cell^{-1}), using a substrate mix composed of peat and vermiculite in a 70:30 (v/v) ratio to provide an optimal environment for root development during germination and vegetative growth.

2.2. Light treatments: intensity, quality, and photoperiod

Five different light spectra (Table 1) were provided by dimmable LED lamps (Flygrow®, Flytech Srl, Belluno, Italy) from the start of the cycle (0 Days After Sowing, DAS) until the final harvest (36 DAS). Each light treatment was replicated on three trays. Specifically, the light spectra were characterized by the exclusive application of light within the red (peak at 663 nm) and blue (peak at 470 nm) regions, resulting in the following incremental R:B ratios: R:B 1, R:B 3, R:B 5, R:B 7, and R:B 9 (Fig. 1B). Before the experiment began, a handheld spectrophotometer was used to determine the spectral properties (CL-500A, Konica Minolta, Chiyoda, Tokyo, Japan) of each treatment. Both photoperiod (expressed as light duration per day, h d^{-1}) and light intensity (expressed as Photosynthetic Photon Flux Density, PPFD, in μmol m^{-2} s^{-1}) remained constant for all light treatments, set to 16 h d^{-1} and 215 ± 10 μmol m^{-2} s^{-1}, respectively. The PPFD was accurately measured using a PAR meter (Apogee Instruments, Logan, UT, USA).

2.3. Image acquisition

Images of the plants (Fig. 1B) were taken daily (excluding weekends) from the moment of sowing until the end of the study. Two growth stages were considered: the germination phase (going from 1 to 14 DAS) and the vegetative phase (from 15 to 36 DAS). For each species and growth stage, images were acquired from different angles to ensure a complete representation of the morphological characteristics of the plants, which is crucial for the subsequent classification step using the CNN algorithm. The RGB images were captured using the ESP32-CAM module (ESP32-CAM, Espressif Systems, Shanghai, China), a versatile device offering ease of use, low cost, and operational flexibility. The ESP32-CAM was equipped with a CMOS OV2640 image sensor (OmniVision Technologies, Santa Clara, USA), supporting a maximum resolution of 1600 × 1200 pixels; however, a resolution of 640 × 480 pixels was selected for the experiment. This resolution was sufficient for capturing the morphological details of the plants while optimizing acquisition time. The captured images were stored directly on an SD card integrated into the ESP32-CAM module, ensuring secure and local data storage. After each acquisition session, the photos were transferred to a central database, facilitating subsequent analyses and management of the entire dataset. The ESP32-CAM allowed manual adjustment of parameters such as lighting and focus, adapting to different environmental conditions. This enabled the creation of a complete and diverse dataset, essential for training and validating a robust CNN model for plant species classification. The choice of the Arduino ESP32-CAM module reduced costs. It allowed testing of the classification algorithm's effectiveness with low-resolution images, demonstrating that the model is sufficiently robust for low-cost applications while maintaining good accuracy.

Table 1
Experimental lighting treatments.

Light Treatments	Red (%)	Blue (%)	Light Intensity PPFD (μmol m^{-2} s^{-1})	Photoperiod (h d^{-1})
R:B 1	50	50	215 ± 10	16
R:B 3	75	25	215 ± 10	16
R:B 5	83.3	16.7	215 ± 10	16
R:B 7	87.5	12.5	215 ± 10	16
R:B 9	90	10	215 ± 10	16

2.4. Dataset annotation

The entire image dataset was manually structured through an annotation system, categorizing each file based on the plant species (basil, cucumber, aubergine, or tomato) and the growth stage (germination or vegetative). As a result, two main folders were created for each plant species:

- Plant images in the germination phase, from day 1 to day 14 after sowing (DAS);
- Plant images in the vegetative phase, from day 15 to day 36 DAS.

The final dataset consisted of a total of 2400 images for the germination phase and 4800 images for the vegetative phase. The vegetative phase was thus overrepresented compared to the germination phase, leading to an imbalanced dataset. To address this imbalance and ensure better model generalization, a data augmentation strategy was applied to the minority class (germination phase). Data augmentation [21] was carried out through geometric transformations (rotations, translations, zooms, horizontal and vertical flips) to generate 2400 new variants of the existing images without the need to collect new data. This approach increased the number of images in the germination class, balancing the dataset and improving the model's performance in recognizing plants during this initial phase. The images were labeled following a hierarchical directory structure, facilitating the integration of the dataset into the convolutional neural network architecture.

2.5. CNN algorithm architecture

The architecture of the convolutional neural network developed in this study was designed to address a multi-task learning problem, with the goal of simultaneously classifying both the plant species and its phenological stage (Table 2). This approach allows for the leveraging of shared information between the two tasks, improving model efficiency and prediction quality [22].

The model follows a hierarchical architecture where the initial convolutional layers extract general representations from the input images. The input consists of RGB images with dimensions of 256 × 192 × 3, which are processed through a series of transformations. Batch normalization, applied after each convolutional layer, ensures a uniform distribution of pixel values, promoting faster and more stable convergence during training [23]. The initial convolutional layers capture simple features such as edges and textures, while deeper layers learn more complex morphological patterns, including leaf shape and plant structure [24]. Dimensionality reduction is achieved through max pooling and global average pooling operations, which compress the feature maps while retaining the most salient information for classification. This approach enhances computational efficiency by reducing the number of parameters and mitigating overfitting without sacrificing critical information [25]. An additional regularization mechanism is provided by spatial dropout applied at the third convolutional block, which helps prevent overfitting the training data. Following feature extraction, the two-dimensional feature maps are reduced to a one-dimensional vector via global average pooling, eliminating the need for a traditional flattening layer. The network subsequently branches into two heads, each dedicated to a specific multi-task classification objective. The first head predicts the plant species, concluding with an output layer comprising four units corresponding to the species under study, employing a softmax activation function to convert raw outputs into class probabilities. The second head is responsible for classifying the phenological stage, with a final output layer consisting of two units, and softmax activation is utilized to distinguish between the two stages considered.

Table 2
List of layers in the CNN classification algorithm architecture.

No.	Layer	Description
1	Input Layer	RGB images 256 × 192 × 3
2	Conv Layer 1	32 filters 3 × 3, ReLU, "same" padding
3	Batch Normalization 1	Batch normalization
4	Max Pooling 1	2 × 2 pooling
5	Conv Layer 2	64 filters 3 × 3, ReLU, "same" padding
6	Batch Normalization 2	Batch normalization
7	Max Pooling 2	2 × 2 pooling
8	Conv Layer 3	128 filters 3 × 3, ReLU, "same" padding
9	Batch Normalization 3	Batch normalization
10	Spatial Dropout	Spatial dropout 0.3
11	Max Pooling 3	2 × 2 pooling
12	Conv Layer 4	256 filters 3 × 3, ReLU, "same" padding
13	Batch Normalization 4	Batch normalization
14	Global Average Pooling	Dimensionality reduction
15	Dense Layer 1	256 neurons, ReLU
16	Dropout	Dropout 0.5
17	Dense Layer 2	128 neurons, ReLU
18	Output Layer 1 (Species)	4 units, softmax
19	Output Layer 2 (Phenology)	2 units, softmax

Integrating multi-task learning within this architecture allows the shared convolutional layers to learn more generalizable and transferable representations, thereby improving classification robustness and reducing the demand for extensive training datasets [26]. Moreover, jointly optimizing both tasks reduces the risk of overfitting, particularly in scenarios with limited data, as the model is encouraged to identify features relevant to both objectives [27]. From a computational standpoint, this solution enables both classifications to be performed in a single inference pass, decreasing processing time compared to separate models and enhancing overall system efficiency [22].

2.6. Training and evaluation

The multi-task learning CNN model was trained on an RGB image dataset of 9.600 samples, including original images and images generated through data augmentation. This process improved the diversity of the training data, helping the model generalize better on new images and reducing the risk of overfitting. The dataset was split into two sets to ensure the model's good generalization ability: 80 % for training and 20 % for evaluation of unseen data. This splitting strategy is widely adopted in literature to improve the generalization ability of neural networks and prevent overfitting [24]. As the model was designed to perform two classification tasks simultaneously—identifying the plant species and determining its phenological stage—two separate loss functions were adopted, one for each output. Categorical cross-entropy was used for both classifications. It is ideal for multi-class classification problems as it minimizes the divergence between the predicted probability distribution and the actual class distribution [28]. A cross-validation strategy was implemented to obtain more robust and reliable results. Since the dataset was balanced through data augmentation, K-Fold Cross-Validation was adopted, which divides the dataset into K subsets (folds) and iteratively uses each fold as a test set. In contrast, the remaining K-1 folds are used for training. This method ensures a more reliable evaluation of the model's performance, reducing dependency on a single data split and improving generalization ability [29]. In this study, $K = 5$ was chosen as it represents a good compromise between evaluation stability and computational cost. The model optimization was carried out using the Adam algorithm, known for its ability to combine the advantages of Adaptive Gradient Descent (AdaGrad) and Root Mean Square Propagation (RMSProp), making it particularly effective for managing deep networks and ensuring rapid convergence without unstable oscillations [30]. The learning rate was set to 0.001, a value commonly used for deep CNNs, which balances the learning speed with optimization stability [31]. The training was performed with a batch size of 32 images, a choice that allows for a good trade-off between model accuracy and computational efficiency. The learning process was conducted for a maximum of 200 epochs, with early stopping implemented. This was configured to halt training if validation accuracy showed no significant improvements for 50 consecutive epochs, thus avoiding the risk of overfitting to the training data.

2.7. Confusion matrix

A confusion matrix was used to evaluate the multi-task CNN model's performance more in-depth, an essential tool for analyzing the classifier's behavior concerning different classes [32]. The confusion matrix was built separately for each classification task (plant species and phenological stage) and for each lighting condition, allowing for identifying any systematic error trends and the degree of confusion between classes. For each iteration of K-Fold Cross-Validation, a confusion matrix was generated for the test fold, and subsequently, an aggregated average matrix was calculated across all folds. This methodology provided a more stable and reliable evaluation of the model's performance, reducing variability due to the dataset splitting into training and test sets. In the confusion matrix, each row represents the actual instances belonging to a specific class, while each column indicates the number of times the model assigned those instances to each of the predicted classes. The following elements were calculated:

- True Positives (TP): instances correctly classified in their respective class.
- False Positives (FP): instances incorrectly assigned to a different class.
- False Negatives (FN): instances belonging to a class but incorrectly assigned to another.
- True Negatives (TN): instances correctly excluded from a given class.

For a more intuitive representation of the model's performance, normalized confusion matrices were produced, where each value was divided by the total number of samples belonging to the respective true class. This approach allowed for a more easily interpretable distribution of the predictions. This analysis provided a more detailed understanding of the errors made by the model, highlighting possible classes more prone to misclassification, the degree of separability between categories, and the influence of different lighting treatments on the recognition quality [33].

2.8. Evaluation metrics

The performance of the CNN model was evaluated using specific metrics for multi-class classification and potentially imbalanced datasets. Since the model performs two distinct classifications (plant species and phenological stage), the metrics were calculated separately for each task. After each cross-validation iteration, the evaluation metrics were derived based on the confusion matrix averaged across different folds. The primary metric used is accuracy (Eq. (1)), which represents the fraction of correctly classified instances out of the total [34]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision (Eq. (3)) is a metric that quantifies the fraction of elements classified as positive that are actually positive. This value is particularly relevant in problems where false alarms must be minimized, such as disease classification or recognizing rare species [35]. The formula is:

$$\mathrm{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Recall (Eq. (4)), which measures the model's ability to correctly identify positive instances [36], is calculated as:

$$\mathrm{Recall} = \frac{TP}{TP + FN} \tag{3}$$

Another key metric is the F1-score (Eq. (2)), which provides a trade-off between precision and recall [35] and is calculated as:

$$F1Score = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{4}$$

For a more comprehensive evaluation, the Matthews Correlation Coefficient (MCC) (Eq. (5)) was also calculated, which provides an indication of classification quality considering all elements of the confusion matrix:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5}$$

This metric is considered particularly robust in classification as it considers all elements of the confusion matrix [36]. The metrics were calculated separately for each lighting condition (red:blue light ratio) to understand the influence of different environmental conditions on the model's ability to distinguish between species and growth stages.

2.9. Development environment

The multi-task classification algorithm was developed using Python 3.8, chosen for its compatibility with the advanced libraries required to implement and train convolutional neural networks. Python is a de facto standard in the field of artificial intelligence and machine learning, thanks to its versatility and the wide range of tools available for data manipulation and deep learning. The TensorFlow and Keras libraries were used to construct the multi-task CNN model, providing a flexible, high-performance framework for deep learning. TensorFlow efficiently manages the computational flow required for training the network, while Keras simplifies the definition of the architecture and facilitates optimization with its high-level interface [37]. The model training was accelerated through support for NVIDIA GPUs, allowing a significant reduction in processing time compared to execution on CPUs. Numerical data management and manipulation were handled by NumPy, which is essential for normalizing input data and efficiently representing multi-dimensional arrays. OpenCV was employed for image processing operations, including resizing, format conversion, and normalization, and it is a widely used library for image processing in computer vision applications. The algorithm implementation and model training were conducted on a Windows 11 workstation with an Intel Core i7 processor, 16 GB of RAM, and an NVIDIA RTX 3060 GPU with 6 GB of RAM.

3. Results

The analysis of the performance of the multi-task learning-based CNN model allowed the evaluation of the accuracy in classifying horticultural species and their phenological stages under different lighting conditions. The results were obtained by evaluating the accuracy, precision, recall, F1-score and Matthews Correlation Coefficient (MCC) metrics for each considered light treatment.

3.1. Confusion matrices

The average confusion matrix was computed for each lighting treatment by aggregating the results obtained across the five cross-validation folds. This procedure enables a more stable and reliable evaluation of the distribution of correct and incorrect predictions made by the model. The confusion matrices related to horticultural species classification (Fig. 2) show a marked concentration of values along the main diagonal, indicating the model's strong ability to distinguish between classes. Misclassifications are limited and primarily involve confusion between species with similar morphological characteristics. Lighting treatments with R:B ratios of 1 and 5 exhibit the highest density of correct predictions, consistent with the elevated accuracy and F1-score metrics observed. Regarding phenological stage classification (Fig. 3), the matrices also display an intense diagonal concentration of predictions, with marginal errors occurring mostly between adjacent stages, where visual distinctions can be subtler. Treatments with R:B ratios of 3 and 5 show the best distribution, characterized by clear class separation and reduced misclassifications.

3.2. Performance of R:B spectral ratios

The model's performance in classifying horticultural species is summarized in Table 3, which reports the average values of accuracy, precision, recall, F1-score, and MCC for each light treatment. Contrary to initial hypotheses based on preliminary data, the metrics demonstrate a remarkable robustness of the model, even in the presence of significant variations in the light spectrum. Classification accuracy for species ranged between approximately 70 % and 86 %, with the highest values observed under the R:B 1 treatment. This stability is also reflected in the other metrics. Precision and recall remained high and consistent with the overall trend, generally exceeding 71 % and reaching approximately 87 % in the best-performing treatment (R:B 1). The F1-score, which

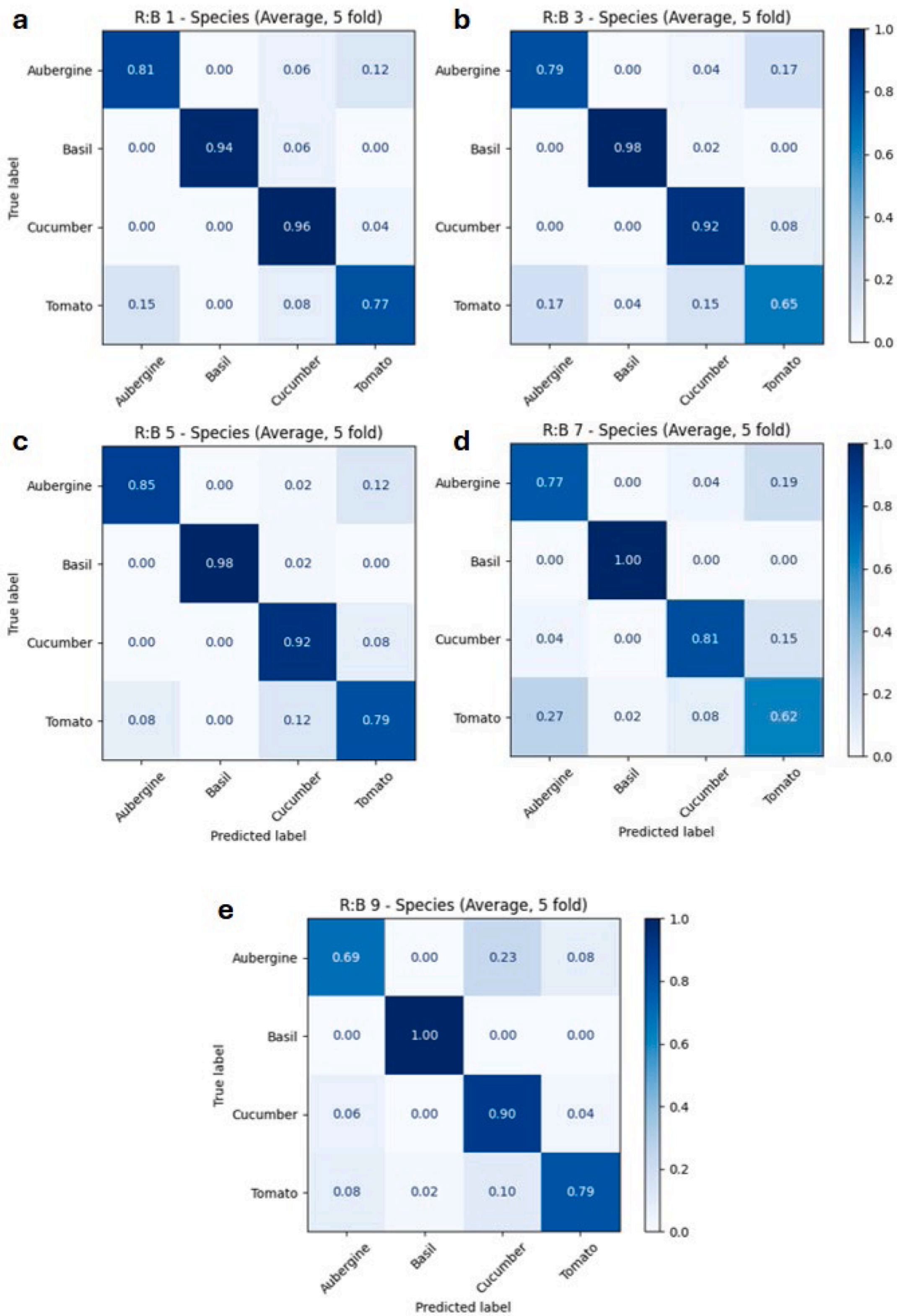

Fig. 2. Species classification - normalized confusion matrices for each light treatment: R:B 1 (a), R:B 3 (b), R:B 5 (c), R:B 7 (d), R:B 9 (e).

balances precision and recall, followed a similar pattern, with the highest value (85 %) again recorded under R:B 1. The Matthews Correlation Coefficient further confirmed this robustness, with values above 0.71 across all treatments and peaks reaching 0.81 under R:B 1.

3.3. *Performance of growth stages*

The results related to phenological stage classification, presented in Table 4, show a trend consistent with that observed for species classification, with high and relatively stable performance across the different light treatments.

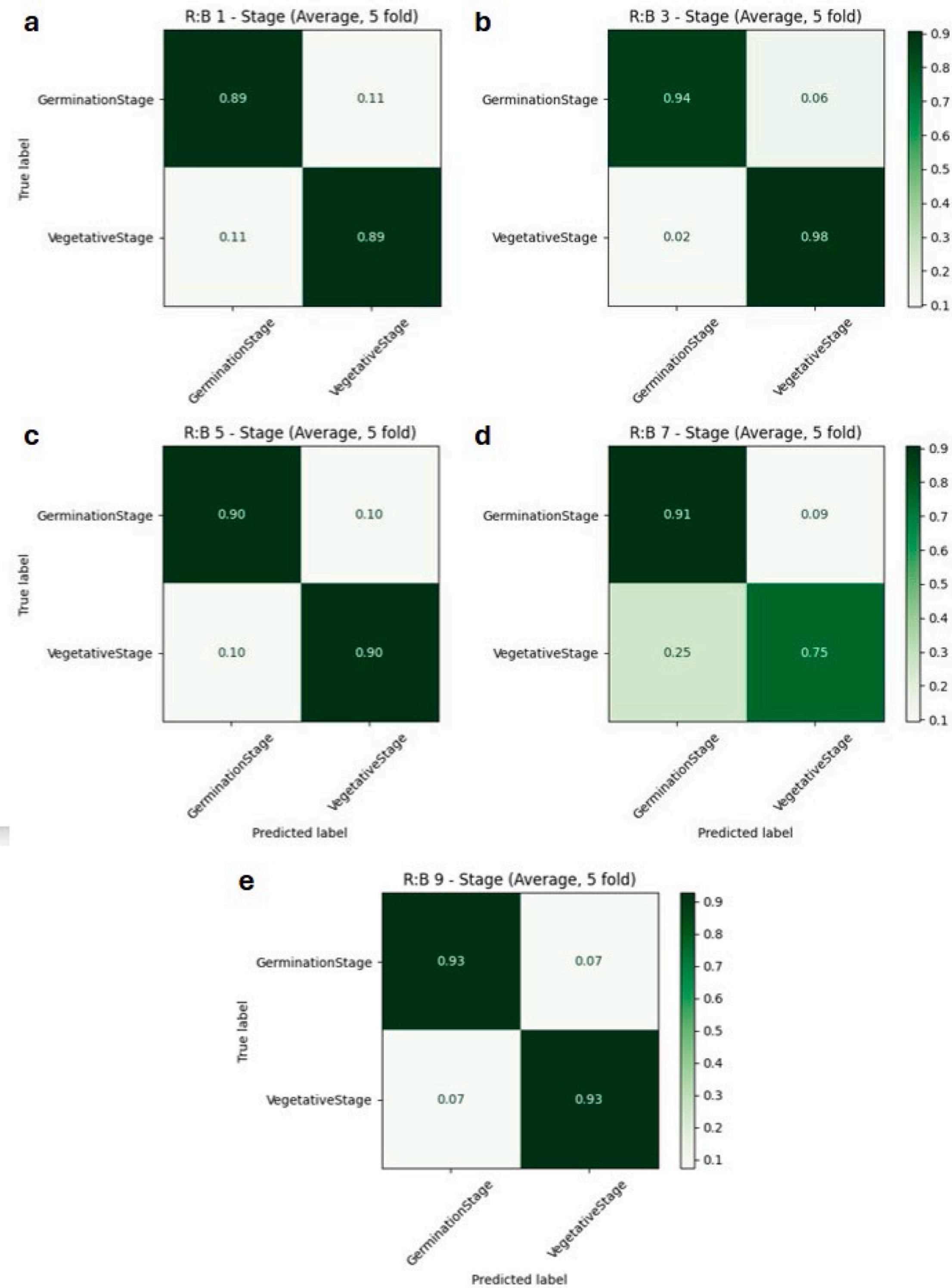

Fig. 3. Phenology classification - normalized confusion matrices for each light treatment: R:B 1 (a), R:B 3 (b), R:B 5 (c), R:B 7 (d), R:B 9 (e).

Classification accuracy ranged between approximately 82 % and 93 %, with the highest values recorded under the R:B 3 and R:B 5 treatments, indicating a superior ability of the model to recognize phenological stages under intermediate or balanced lighting conditions. Precision and recall also exhibited high values, ranging from approximately 82 % to 93 %, with the best results again observed under the R:B 3 and R:B 5 treatments. The F1-score, which balances precision and recall, remained high across all treatments, reaching peak values of 93 % and 92 % for R:B 3 and R:B 5, respectively. Finally, the Matthews Correlation Coefficient (MCC) ranged from 0.77 to 0.86, with the highest values once again obtained under the R:B 3 and R:B 5 conditions.

4. Discussion

Analysis of the performance metrics demonstrates that the multi-task convolutional neural network classification model maintained high and

Table 3
Model performance for species classification under each light treatment.

Light treatments	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MCC (dimensionless)
R:B 1	86	87	85	85	0.81
R:B 3	70	79	71	69	0.63
R:B 5	81	84	80	79	0.75
R:B 7	78	78	78	77	0.71
R:B 9	81	85	83	80	0.77

Table 4
Model performance for phenological stage classification under each light treatment.

Light treatments	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	MCC (dimensionless)
R:B 1	90	90	89	89	0.79
R:B 3	93	93	93	93	0.86
R:B 5	93	92	93	92	0.86
R:B 7	82	82	83	81	0.65
R:B 9	88	88	89	88	0.77

relatively stable performance across all light treatments, for both horticultural species and phenological stage classification. The observed variability among different red:blue spectral ratios was limited, indicating the model's general robustness to variations in artificial lighting conditions. This behavior can be attributed to two main factors: first, the visual characteristics of the species and phenological stages remain sufficiently distinctive despite changes in lighting conditions; second, the employed multi-task CNN architecture effectively learned robust and generalizable features, minimizing the impact of illumination-induced spectral variability.

Although lighting affects the chromatic properties of RGB images, interspecific and phenological stage differences remain pronounced enough to be reliably detected by the model. Consequently, the captured images contain rich visual information adequate to support effective classification, as reflected by the consistently high performance metrics across all treatments.

From a computer vision and artificial intelligence perspective, CNNs equipped with architectural enhancements such as batch normalization, dropout, and global average pooling are designed to learn features invariant to minor visual domain transformations (Krizhevsky et al., 2012; [23]. This confers resilience against variations in illumination, viewpoint, and other small image perturbations. Specifically, the multi-task approach promotes intermediate feature sharing between species and phenological stage classification tasks, enhancing learning efficiency and generalization capability. Shared, more abstract feature representations are less sensitive to non-structural spectral variations, such as those caused by different R:B ratios. This aligns with previous findings showing that multi-task models can acquire more robust and generalizable representations than single-task counterparts [26,38].

These findings suggest that CNN-based computer vision systems can operate reliably in controlled environments with variable artificial lighting, provided that essential visual information is preserved. This has important implications for deploying automated monitoring systems in indoor agriculture and vertical farming, where lighting is often optimized for plant growth rather than image acquisition.

5. Conclusions

This study evaluated the influence of different red:blue spectral ratios on the performance of a multi-task CNN model for the automatic classification of horticultural species and their phenological stages, using RGB images acquired under controlled lighting conditions. The results demonstrated that the model maintains high and stable performance across all tested light treatments, with limited variability attributable to spectral composition. Morphological differences among horticultural species are sufficiently pronounced to allow accurate visual classification even under varying illumination conditions. It is plausible to hypothesize that spectral variations could have a greater impact in the case of classification at the variety level within a single species, making automatic discrimination more challenging due to the reduced morphological distance between classes. Furthermore, the multi-task CNN architecture proved capable of learning robust and generalizable feature representations, effectively compensating for illumination-induced variability. This study showed that CNN-based computer vision systems can operate reliably under suboptimal lighting conditions. Such robustness is particularly relevant in controlled agricultural environments (e.g., vertical farming, automated greenhouses), where lighting is often designed primarily for plant physiology rather than optical image acquisition. The implications of these findings are twofold. On one hand, they confirm the effectiveness of combining low-cost RGB sensors with deep learning models for automatic crop classification. On the other hand, they highlight the importance of an integrated approach to lighting design that considers both plant physiological needs and the requirements of computer vision systems.

Future research could investigate the effect of additional spectral components, such as green light or the far-red band, assessing their influence on image quality and consequently on the performance of automatic classification models. These components may be especially relevant when addressing classification between species and among varieties within the same species, where morphological differences are subtler and more susceptible to illumination effects. Concurrently, integrating advanced deep learning approaches, such as self-supervised learning, reinforcement learning, or generative models for data augmentation, could enhance model generalization under variable or suboptimal environmental conditions. Overall, this study lays the foundation for developing resilient and scalable systems for precision agriculture, emphasizing the critical importance of deliberate lighting design in synergy with artificial intelligence architectures to maximize the reliability of automatic crop classification.

Ethics in publishing statement

I testify on behalf of all co-authors that our article submitted followed ethical principles in publishing.

CRediT authorship contribution statement

Matteo Landolfo: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fabio Perotti:** Investigation, Data curation. **Alessandro Pistillo:** Investigation. **Giuseppina Pennisi:** Writing – review & editing, Visualization, Validation, Methodology, Conceptualization. **Giorgio Gianquinto:** Writing – review & editing, Supervision. **Francesco Orsini:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationshipsthat could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by the Italian Ministry of Education and Research (MUR), within the call for Research Projects of National Interest (PRIN), within the project "VFARM - Sustainable Vertical Farming" (Project code: 2020ELWM82, CUP: J33C20002350001) and the Agritech National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 103217/06/2022, CN00000022).

Data availability

Data will be made available on request.

References

[1] J.A. Foley, N. Ramankutty, K.A. Brauman, E.S. Cassidy, J.S. Gerber, M. Johnston, D.P. Zaks, Solutions for a cultivated planet, Nature 478 (7369) (2011) 337–342, https://doi.org/10.1038/nature10452.
[2] H.C.J. Godfray, J.R. Beddington, I.R. Crute, L. Haddad, D. Lawrence, J.F. Muir, C. Toulmin, Food security: the challenge of feeding 9 billion people, Science 327 (5967) (2010) 812–818, https://doi.org/10.1126/science.1185383.
[3] A. Raza, A. Razzaq, S.S. Mehmood, X. Zou, X. Zhang, Y. Lv, J. Xu, Impact of climate change on crops adaptation and strategies to tackle its outcome: a review, Plants 8 (2) (2019) 34, https://doi.org/10.3390/plants8020034.
[4] A.T. Balafoutis, B. Beck, S. Fountas, Z. Tsiropoulos, J. Vangeyte, T. van der Wal, S. M. Pedersen, Smart farming technologies–description, taxonomy and economic impact, Technol. Econ. Perspect. (2017) 21–77, https://doi.org/10.1007/978-3-319-68715-5_2.
[5] J. Jägermeyr, Agriculture's historic twin-challenge toward sustainable water use and food supply for all, Front. Sustain. Food Syst. 4 (2020) 35, https://doi.org/10.3389/fsufs.2020.00035.
[6] Kozai T., Niu G. & Takagaki M. (2019). Plant factory. In: Kozai, T., Takagaki, M. & Niu, G. (Eds.). An Indoor Vertical Farming System for Efficient Quality Food Production. Academic Press (pp. 93–112). 10.1016/B978-0-12-816691-8.00005-4.
[7] T. Searchinger, R. Waite, C. Hanson, J. Ranganathan, P. Dumas, E. Matthews, C. Klirs, Creating a sustainable food future: a menu of solutions to feed nearly 10 billion people by 2050, Final rep. (2019). research.wri.org/sites/default/files/2019-07/WRR_Food_Full_Report_0.pdf.
[8] T. Kozai, Plant factory in Japan-current situation and perspectives, Chron. Hortic. 53 (2013) 8–11, https://doi.org/10.2183/pjab.89.447.
[9] K. Al-Kodmany, The vertical farm: a review of developments and implications for the vertical city, Buildings 8 (2) (2018) 24, https://doi.org/10.3390/buildings8020024.
[10] J.E. Son, H.J. Kim, T.I. Ahn, Hydroponic systems. Plant Factory, Academic Press, 2020, pp. 273–283, https://doi.org/10.1016/B978-0-12-816691-8.00020-0.
[11] I.A. Lakhiar, G. Jianmin, T.N. Syed, F.A. Chandio, N.A. Buttar, W.A. Qureshi, Monitoring and control systems in agriculture using intelligent sensor techniques: a review of the aeroponic system, J. Sens. 2018 (1) (2018) 8672769, https://doi.org/10.1155/2018/8672769.
[12] G. Guerrero-Ulloa, A. Méndez-García, V. Torres-Lindao, V. Zamora-Mecías, C. Rodríguez-Domínguez, M.J. Hornos, Internet of Things (IoT)-based indoor plant care system, J Ambient. Intell. Smart Env. 15 (1) (2023) 47–62, https://doi.org/10.3233/AIS-220483.
[13] R.N. Roy, A. Finck, G.J. Blair, H.L.S. Tandon, Plant nutrition for food security. A guide for integrated nutrient management, FAO Fertil. Plant Nutr. Bull. 16 (368) (2006) 201–214, openknowledge.fao.org/handle/20.500.14283/a0443e.
[14] S.H. Van Delden, M. SharathKumar, M. Butturini, L.J.A. Graamans, E. Heuvelink, M. Kacira, L.F.M. Marcelis, Current status and future challenges in implementing and upscaling vertical farming systems, Nat. Food. 2 (12) (2021) 944–956, https://doi.org/10.1038/s43016-021-00402-w.
[15] A. Arcasi, A.W. Mauro, G. Napoli, F. Tariello, G.P. Vanoli, Energy and cost analysis for a crop production in a vertical farm, Appl. Therm. Eng. 239 (2024) 122129, https://doi.org/10.1016/j.applthermaleng.2023.122129.
[16] M.N. Halgamuge, A. Bojovschi, P.M. Fisher, T.C. Le, S. Adeloju, S. Murphy, Internet of Things and Autonomous Control For Vertical Cultivation Walls Towards Smart Food Growing: A review, 61, Urban Forestry & Urban Greening, 2021 127094, https://doi.org/10.1016/j.ufug.2021.127094.
[17] J. Monteiro, J. Barata, M. Veloso, L. Veloso, J. Nunes, A scalable digital twin for vertical farming, J. Ambient. Intell. Humaniz. Comput. 14 (10) (2023) 13981–13996, https://doi.org/10.1007/s12652-022-04106-2.
[18] C. Liu, W.T. Freeman, R. Szeliski, S.B. Kang, Noise estimation from a single image, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) 1, IEEE, 2006, pp. 901–908, https://doi.org/10.1109/CVPR.2006.207.
[19] M.S. Drew, J. Wei, Z.N. Li, Illumination-invariant color object recognition via compressed chromaticity histograms of color-channel-normalized images, in: Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), IEEE, 1998, pp. 533–540, https://doi.org/10.1109/ICCV.1998.710768.
[20] A. Kamilaris, F.X. Prenafeta-Boldú, A review of the use of convolutional neural networks in agriculture, J. Agric. Sci. 156 (3) (2018) 312–322, https://doi.org/10.1017/S0021859618000436.
[21] Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621, 10.48550/arXiv.1712.04621.
[22] R. Caruana, Multi-Task Learning, 28, Machine learning, 1997, pp. 41–75, https://doi.org/10.1023/A:1007379606734.
[23] Ioffe, S., & Szegedy, C. (2015,June). Batch normalization: accelerating deep network training by reducing internal covariate shift. In International conference on machine learning (pp. 448–456). Pmlr. Proceedings.mlr.press/v37/ioffe15.html.
[24] S. Loussaief, A. Abdelkrim, Machine learning framework for image classification, in: 20167th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), IEEE, 2016, pp. 58–61, https://doi.org/10.25046/aj030101.
[25] Kohler, M., & Langer, S. (2020). Statistical theory for image classification using deep convolutional neural networks with cross-entropy loss under the hierarchical max-pooling model. arXiv preprint arXiv:2011.13602, 10.48550/arXiv.2011.13602.
[26] Ruder, S. (2017). An overview of multi-task learning in deep neural networks. arXiv preprint arXiv:1706.05098, 10.48550/arXiv.1706.05098.
[27] Kendall, A., Gal, Y., & Cipolla, R. (2018). Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7482–7491), 10.1109/CVPR.2018.00781.
[28] Panchapagesan, S., Sun, M., Khare, A., Matsoukas, S., Mandal, A., Hoffmeister, B., & Vitaladevuni, S. (2016). Multi-task learning and weighted cross-entropy for DNN-based keyword spotting, 10.21437/Interspeech.2016-1485.
[29] C.M. Bishop, Machine Learning, 128, Mach. Learn, 2006, 1-2819119.
[30] Kingma, D.P., & Ba, J. (2014). Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980, 10.48550/arXiv.1412.6980.
[31] Y. Bengio, Practical recommendations for gradient-based training of deep architectures, in: G. Montavon, G.B. Orr, Müller (Eds.), Neural Networks: Tricks of the Trade, 2nd ed., Springer, 2012, pp. 437–478, https://doi.org/10.1007/978-3-642-35289-8_26.
[32] S. Visa, B. Ramsay, A.L. Ralescu, E. Van Der Knaap, Confusion matrix-based feature selection, Maics 710 (1) (2011) 120–127.
[33] A.D. Forbes, Classification-algorithm evaluation: five performance measures based onconfusion matrices, J. Clin. Monit. 11 (1995) 189–206, https://doi.org/10.1007/BF01617722.
[34] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006. ISBN: 9780387310732.
[35] Yacouby, R., & Axman, D. (2020,November). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. In Proceedings of the first workshop on evaluation and comparison of NLP systems (pp. 79–91), 10.18653/v1/2020.eval4nlp-1.9.
[36] Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for multi-class classification: an overview. arXiv preprint arXiv:2008.05756, 10.48550/arXiv.2008.05756.
[37] F.J.J. Joseph, S. Nonsiri, A. Monsakul, Keras and TensorFlow: a hands-on experience, Adv. Deep Learn. Eng. Sci. (2021) 85–111, https://doi.org/10.1007/978-3-030-66519-7_4.
[38] Misra, I., Shrivastava, A., Gupta, A., & Hebert, M. (2016). Cross-stitch networks for multi-task learning. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3994–4003).

Further reading

[39] T.G. Benton, R. Bailey, The paradox of productivity: agricultural productivity promotes food system inefficiency, Glob. Sustain. 2 (2019) e6, doi:10.1017/sus.2019.3.
[40] J. Gupta, S. Pathak, G. Kumar, Deep learning (CNN) and transfer learning: a review, Journal of Physics: Conference Series, 2273, IOP Publishing, 2022. 012029, doi:10.1088/1742-6596/2273/1/012029.
[41] Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R.R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580. doi:10.48550/arXiv.1207.0580.
[42] Z. Li, R. Guo, M. Li, Y. Chen, G. Li, A review of computer vision technologies for plant phenotyping, Comput. Electron. Agric. 176 (2020) 105672, https://doi.org/10.1016/j.compag.2020.105672.
[43] G. Pennisi, S. Blasioli, A. Cellini, L. Maia, A. Crepaldi, I. Braschi, G. Gianquinto, Unraveling the role of red: blue LED lights on resource use efficiency and nutritional properties of indoor grown sweet basil, Front. Plant Sci. 10 (2019) 305, https://doi.org/10.3389/fpls.2019.00305.
[44] B. Sasmal, G. Das, P. Mallick, S. Dey, S. Ghorai, S. Jana, C. Jana, Advancements and challenges in agriculture: a comprehensive review of machine learning and IoT applications in vertical farming and controlled environment agriculture, Big Data Comput. Vis. 4 (2) (2024) 67–94, https://doi.org/10.22105/bdcv.2024.474315.1183.
[45] A. Silwal, T. Parhar, F. Yandun, H. Baweja, G. Kantor. A robust illumination-invariant camera system for agricultural applications, IEEE, 2021, pp. 3292–3298, https://doi.org/10.48550/arXiv.2101.02190.
[46] H. Tian, T. Wang, Y. Liu, X. Qiao, Y. Li, Computer vision technology in agricultural automation—a review, Inf. Process. Agric. 7 (1) (2020) 1–19, https://doi.org/10.1016/j.inpa.2019.09.006.