Journal of Information Hiding and Multimedia Signal Processing ©2009 ISSN 2073-4212
Ubiquitous International Volume **15**, Number **4**, December **2024**

Semantic Connection-Based Learning for Dragon fruit disease classification

Duong Manh Ha
Office of Ministry of Agriculture and Rural Development, Ha Noi 100000, Viet Nam
ha.duong.manh@gmail.com

Tran Hung
Center of researching and technology transferring on irrigation
for the Northern mountainous region in Vietnam (CTIM), Viet Nam
hungtlpb@gmail.com

Nguyen Xuan Kieu
Center of researching and technology transferring on irrigation
for the Northern mountainous region in Vietnam (CTIM), Viet Nam
nguyenxuankieu@gmail.com

Nguyen Gia Vuong
Center of researching and technology transferring on irrigation
for the Northern mountainous region in Vietnam (CTIM), Viet Nam
nguyengiavuonghd@gmail.com

Quynh Dao Thi Thuy*
Faculty of Information Technology
Posts and Telecommunications Institute of Technology (PTIT)
Intelligent Computing for Sustainable Development Laboratory (IC4SD, PTIT)
Ha Noi 100000, Viet Nam
quynhdtt@ptit.edu.vn

*Corresponding author: Quynh Dao Thi Thuy
Received October 12, 2024, revised October 30, 2024, accepted November 1, 2024.

Abstract. *Identifying diseased leaves is one of the most critical tasks in agriculture. In Vietnam, dragon fruit is particularly susceptible to various leaf diseases during its development. Recently, artificial intelligence techniques have advanced significantly and achieved notable success in image recognition. Therefore, applying AI techniques to identify diseased leaves, especially those of dragon fruit, is essential. One issue is that dragon fruit leaves can suffer from numerous diseases. For instance, while there are 11 known diseases affecting dragon fruit leaves, we only have data for 5 of them. Consequently, training a deep learning model will only be effective for these 5 diseases, and adding data for other diseased leaves can affect accuracy. To update the model without conflicts, we have developed a semantic connection-based learning model aimed at creating separate classification spaces for each label. Accordingly, we designed a semantic connection model to integrate new features from both old and new classes to enhance accuracy. Our experiments demonstrate superior performance, achieving 92% accuracy when adding new classes.*
Keywords: identifying diseased leaves; semantic connection-based learning; deep learning; leaf dragon disease classification

---

1. **Introduction.** In recent years, the applications of deep learning in the agricultural sector have achieved significant milestones [1][2][3]. Specifically, in the area of pest and disease identification on plant leaves, many researchers worldwide have been implementing new technologies such as AI and deep learning [4][5]. Identifying diseases on dragon fruit stems is one of the challenging yet fascinating problems. Although it presents many challenges, developing a deep learning model that can assist in the early detection of diseases on dragon fruit stems would be a small yet beneficial success for farmers and the agricultural industry in general.

The advent of artificial intelligence technologies has brought remarkable efficiency to practical applications in image recognition problems [6][7]. With the increasing modernity of technology, data is also growing at an astonishing rate, necessitating the development of systems capable of accommodating this influx of new data. Accordingly, the greatest obstacle in deep learning models is that learning new data classes can affect previously learned models, leading to the forgetting of previously learned features [8][9][10]. Therefore, some studies have inspired [11][12] the construction of deep learning models capable of learning new data without forgetting the old features, demonstrating superior performance and meeting the demands of today's practical applications [13][14].

Convolutional neural network (CNN) models [15][16] generally generalize the features and learn to produce fixed weights to perform classification tasks. Since training is mostly based on fixed classes, the prediction process is relatively good and reliable. However, learning new features can affect old features, causing conflicts between the new and old features, resulting in suboptimal prediction models [17][18][19].

Some models [20][21] have been developed to learn new features while retaining old ones. Most of these models involve a trade-off between stability in accuracy and flexibility in adding new data features. In an era of increasing data, we need to enhance flexibility, such as expandable network models [22] that can address flexibility when adding new features and new labels. Most of these models are optimized to ensure that adding new labels does not affect old labels through various mechanisms such as fixing old features [23] or combining features into tasks [24]; additionally, there are many other mechanisms.

To create deep learning models capable of recognizing various diseases on dragon fruit plants with diverse classes and learning new classes in the future, we have optimized the expandable network based on semantic connectivity learning to avoid feature loss between tasks. Here, the key issue is to build a semantic connection space to represent tasks without relying too heavily on samples, leading to reduced computational costs.

1.1. **Related works.** *Deep Learning Models for Incremental Learning of Classes:* Some models are based on old label features, and when new label features are added, they do not affect existing classes [25][26]. Methods that can maintain features by storing them in a database [27] have the drawback of high storage costs and the need to retrain from scratch. Research groups working on plant disease recognition [28] frequently use this method in training their models. Data mapping methods through models that reflect old features during training [29] incur high computational costs. Methods for adjusting model parameters [30] to increase unbiased evaluations of more feature-rich labels have also been explored. Recently, some models [31][32] based on teacher-student frameworks help incorporate new labels, ensuring balanced prediction processes without bias towards old labels in plant disease recognition problems. When adding new labels, deep learning models often create large feature map frameworks to learn similar classifiers, thereby adjusting features across all classes (both old and new labels). Two main factors affect model expansion: the memory cost for training multiple features and the computational time required when adding new labels.

Pre-trained Deep Learning Models: We can leverage pre-trained models from research groups worldwide to address the issue of new and old label features [33][34]. In recent years, many recommendation models [35][36] have been extensively trained and have proven effective for image recognition tasks, making these backbones advantageous to use. Some feature extraction models for storage [37][38] have also significantly improved the ability to learn important features in images, especially for plant disease images. Many groups have enhanced loss functions [39][40] to support the recognition of leaf diseases, and these functions can be reused to evaluate models for dragon fruit disease recognition. Additionally, incorporating multi-modal [41][42] approaches to handle various pests and labels from different models into a comprehensive model is a growing trend. Using pre-trained models to transfer necessary parameters into the dragon fruit disease recognition model is an effective and accurate approach.

1.2. **Motivation and contribution.** In this paper, we propose a model based on Semantic Connection Learning (SCL) to address the aforementioned challenges. To enhance efficiency between new and old tasks, we construct a feature label connection space based on the semantics of each feature, aiding in the recognition of individual tasks for each label. This approach allows for effective learning of new classes without affecting previously learned classes. These working spaces are learned by adjusting pre-trained models, and to address training costs and capacity issues, we utilize parameter copying from pre-trained models to improve the process of continuing to learn new labels. This way, we can leverage the strength of old labels with new ones while efficiently and optimally aggregating informational features from multiple labels.

The main innovations and contributions of this work include:

1. Construct training feature spaces to train the proposed Semantic Connection-based Learning (SCL) model. We refine pre-trained models to save training costs and resources while achieving desired accuracy when adding new labels.
2. Build a dataset for diseases on dragon fruit, named D-Dragon. D-Dragon is a dataset we collected from various dragon fruit fields in Vietnam. The images we collected include dragon fruit, dragon fruit stems, and various pests and diseases on dragon fruit plants.
3. Analyze the effectiveness of the training model in adding new labels, and objectively evaluate the accuracy on the D-Dragon dataset against current state-of-the-art models.

2. Proposed method.

2.1. **Problem Definition.** The semantic connection learning model is a continuous learning model for classification that creates a unified classifier [10] Suppose there is a sequence of different classes $\boldsymbol{L}$ in the training set $D_{training}$, denoted as $\boldsymbol{L} = L_1, L_2, ..., L_N$, where $L_i = \boldsymbol{X}, y_i, i \in N$, is the $i-th$ training set with label y_i. $\boldsymbol{X}$ will include data features belonging to class y_i, and y_i belongs to $\boldsymbol{Y}$, where $\boldsymbol{Y}$ is the label space of the dragon fruit disease recognition problem. Our objective is to ensure that when adding the sequence of classes $\boldsymbol{L}$ into the model training, the model's performance remains unaffected and it can still accurately classify dragon fruit diseases.

2.2. **Model Architecture.** As described in Figure 1, we propose a new model for Semantic Connection-based Learning (SCL) to strengthen the model with both old and new labels. Initially, the model is trained using available labeled datasets, and then during the model's usage phase, when additional data is introduced, a new model is created with the new labels integrated into the working space. The SCL model is trained to create

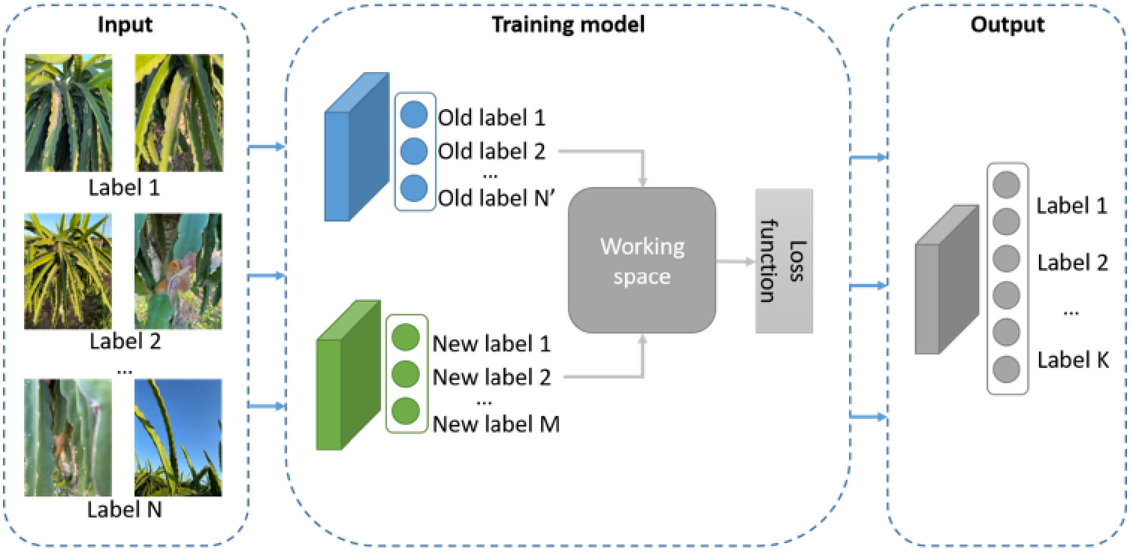

FIGURE 1. Overview of the proposed Semantic Connection-based Learning (SCL) algorithm

working spaces that are unbiased towards either old or new labels, allowing the model to learn both simultaneously. Additionally, the SCL model aggregates unassigned labels for transfer learning and synthesis, ensuring that the model can continue to assimilate the aggregated knowledge within the working space.

2.3. **Pre-training Objectives.** In today's modern world, reusing pre-trained models [43] is quite common. Instead of fine-tuning and training from scratch, we can leverage pre-trained models to maintain previously trained data. The basic idea is to inherit some important pre-trained weights or copy critical layers and then modify the final layers to suit the specific problem at hand. We will construct the models as follows:

$$model\_old = copy\_layer(model\_pretrain(L_i)), i \in N \tag{1}$$

After obtaining the pre-trained model to continue learning with new classes, to avoid conflicts with old classes, we will expand the model by freezing certain layers and extending connections between models before training. Here, the working space will connect the models:

$$model\_new = working\_space(keep(model\_old1, model\_old2, ..., model\_oldN)), model\_old(L_i) \tag{2}$$

Once the new model is created, the working space updates this model to encode information about the tasks into future predictions, optimizing to avoid conflicts between old and new labels. Moreover, the SCL model also promises to prevent overwriting new label information onto old labels. This benefits the exploration of new data in the future, enhancing operational capabilities and optimizing training costs and usage.

3. Model Description.

3.1. **Semantic Connection-based Learning.** From the CNN backbones [44], we obtain comprehensive feature vector embeddings. We construct classifiers on the [45] prototype to support the prediction model. Specifically, each workspace will contain embeddings information to perform specific tasks corresponding to each label. Essentially, after converting from images to feature vector embeddings, the data becomes significantly lighter and requires fewer parameters compared to the original image parameters. The cost of storing these embeddings is also quite low.

Once the overall embedding set is obtained, we will establish connections between the embedding space features and the corresponding label sets. Here, prototype extraction will be performed from the i-th layer in the connection set workspace:

$$P_i = \frac{1}{N}\sum_{j=1}^{N} I(y_j = i) \tag{3}$$

where, N is the number of data samples in class i. P_i is the connection set of prototypes in class i. I represents the prototype extraction method.

When a new class is introduced, the connection set and the workspace will carry out their tasks. Here, the recalculations will be performed in the new workspace with new connection sets. For example, if we have workspace W_1 with connection sets $\boldsymbol{P} = P_1, P_2, ..., P_N$, adding a new label P_{new} will update W_1 and $\boldsymbol{P}$ by recalculating the weights and aggregating the prototypes of the old classes in the new workspace.

To avoid conflicts between the old and new workspaces, the connection sets will carry out their tasks to find the new connection sets with the closest similarity. When adding a new connection set P_i to the old workspace, to find the best weight set for the new workspace, there needs to be a similarity measure between the connection set P_i and one of the old connection sets P_j, calculated across the classes using the workspace prototypes:

$$sim_{i,j} = \frac{P_i}{||P_i||_2}\frac{P_j^T}{||P_j||_2} \tag{4}$$

The similarity reflects the local relationship between the connection sets of the old classes and the new classes. Additionally, similarity is a way to share the workspace between the old and new workspaces. After calculating the similarity, we will establish relationships between the classes and thereby construct the new workspace.

3.2. **Working Space.** The essence of the workspace is to act as a transitional area from one connection set to another across classes, with the input being the feature vector embeddings of images. The workspace aggregates multiple embedded prototypes from different connection sets. These connection sets extract specific features according to the tasks of each class. Thus, the prototypes are suitable for performing classification tasks and supporting inference in the final prediction layer.

The training process involves finding a workspace that can appropriately coordinate for each task to encode information for specific tasks. Extracting embedded information from input images is the process of connecting and aggregating the prototypes of previous classes. Ultimately, the workspace will implement complete and balanced classifiers across classes to make predictions. Therefore, after the training process is completed, the workspace will contain only the connection sets, not the feature vector embeddings.

3.3. **Loss function.** Our model's primary objective is the classification of pest and disease images on dragon fruit plants. We calculate dual loss functions between the label sets of the unified model y and the normalized connection sets y' generated by the trained model. The ESL loss function is constructed as follows:

$$L_{ESL}(y, y') = \frac{1}{N}\sum_{j=1}^{N}(y^j - y'^j)^2 \tag{5}$$

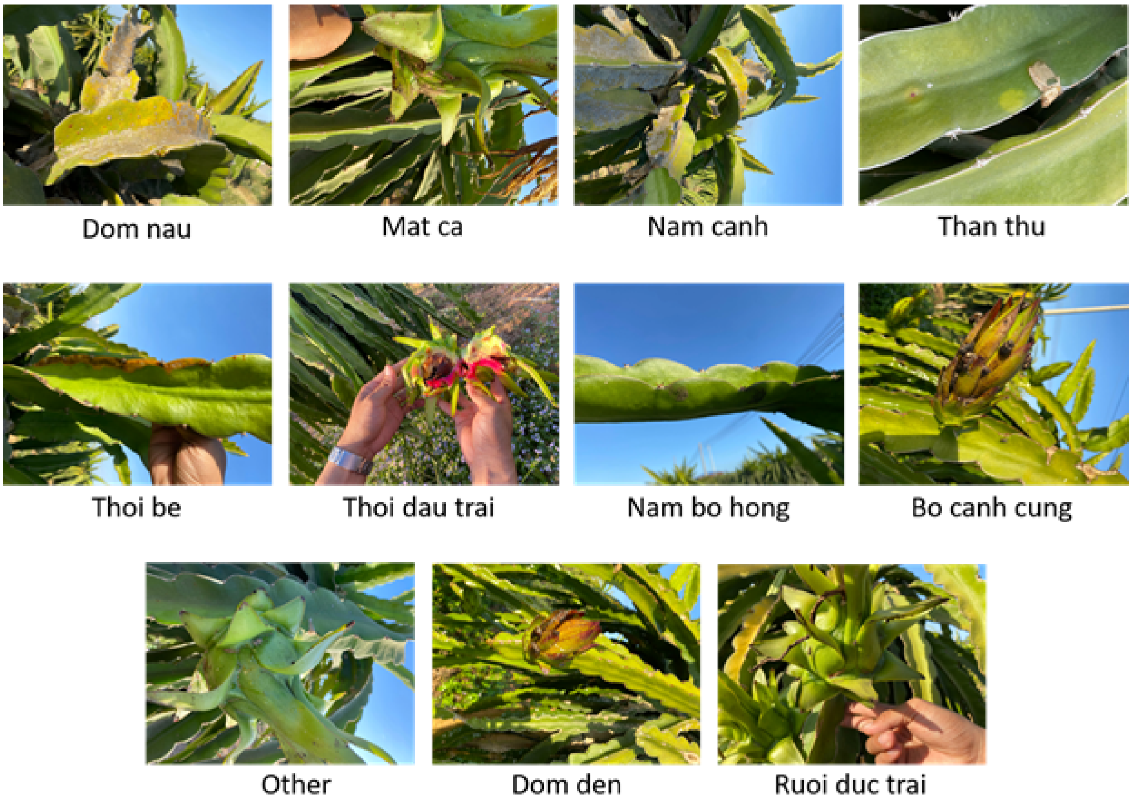

FIGURE 2. Dataset D-Dragon

To ensure effective updates of the workspaces, we construct a loss function to evaluate each updated connection set. If an old class with an old connection set has high accuracy, and when a new class is added, we will use the old model's objective function as the regression target. Otherwise, the output of the model with the updated connection set will be chosen as the replacement. This way, we will assess the comprehensive convergence of the connection sets and the overall ESL model. The loss function of ESL for the connection sets is as follows:

$$L_{connection-based}(c, c') = \sum_{k} smooth_{Loss}(c_k - c'_k) \tag{6}$$

4. Experimental results and analyses.

4.1. **Dataset Description.** We collected 11 labels of unhealthy conditions on dragon fruit tree stems and 1 label of healthy dragon fruit trees, with approximately 500 images collected for each label. We named this dataset D-Dragon, to analyze and evaluate the performance of the proposed method in classifying diseases on dragon fruit tree stems. Based on the features extracted from the D-Dragon dataset, we calculate for pest-infected and non-pest-infected images, and if a disease is detected, we determine the specific type of disease from one of the dragon fruit tree pest types. Figure 2 shows some sample images corresponding to each type of pest.

For data partitioning during the training process, we follow the standard evaluation protocol [46]. We will divide the data by class. Five classes will be used for initial training, and the remaining classes will be split into two groups to be included in subsequent training phases. We adhere to the method of random class order shuffling when partitioning data for continued training to ensure fairness across all classes.

TABLE 1. Comparison to traditional exemplar-based Incremental Learning of Classes methods

Methods	Exemplars	Datasets	
		ImageNet-mini	D-Dragon (ours)
iCaRL [50]	5 / class	70.56	82.64
DER [51]	5 / class	78.25	89.73
MEMO [52]	5 / class	73.84	85.68
EASE [53]	5 / class	79.67	90.24
SCL (ours)	5 / class	**80.34**	**91.89**

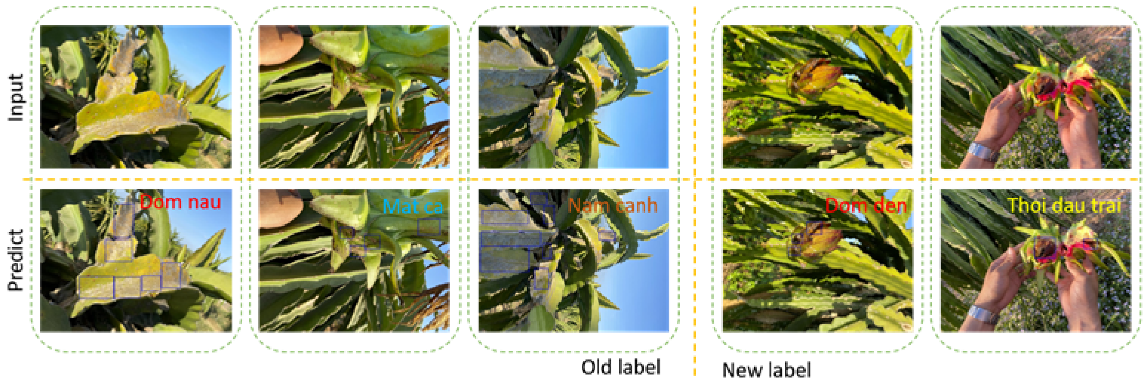

FIGURE 3. Prediction of SCL model

4.2. **Experimental Setting.** We conducted experiments on an Nvidia 4090 25GB GPU, using the Pytorch framework [47] Additionally, we leveraged the ViT [48] representative model to fine-tune the pre-trained model. We also evaluated the performance on two datasets: ImageNet-mini [49] and our dataset, P-D-Dragon. In our SCL model, we trained the model using the SGD optimizer and found that after 38 epochs, the training process nearly converged. The learning rate was set to 0.001, and the balance parameter was set to 0.01.

The research in this paper focuses on two main questions:

- Research Question 1 (RQ1): How much does SCL improve performance on the two datasets, ImageNet-mini and D-Dragon, compared to other baseline methods?
- Research Question 2 (RQ2): How close are the predictions of the SCL model to the ground truth values?

4.3. **Performance Compare of SCL and other baselines (RQ1).** In this section, we compare SCL with other advanced methods based on the concept we mentioned on two datasets for evaluation. Table 1 reports the comparison results of different methods based on the Incremental Learning of Classes approach. We can see that our SCL method achieves the best accuracy results among the five comparable methods, significantly outperforming current SOTA methods such as EASE. Our SCL method, as shown in the last row of Table 1, demonstrates 1-2% higher accuracy compared to the SOTA method EASE on both proposed evaluation datasets. Additionally, we also mention the number of classes, where we initially train with 5 labels, and subsequent labels are gradually introduced to obtain the average accuracy during the training and evaluation process. As shown in Table 1, SCL performs quite well compared to competitors, and the experiments allocate labels reasonably.

4.4. **Qualitative Study (RQ2).** We conducted the study by evaluating specific labels in the test set and obtained results from the SCL model. The input consists of images of diseased dragon fruit, and the output is the prediction results of the SCL model as shown in Figure 3. The SCL model takes an input image of diseased dragon fruit with a size of 1024x480. Below are illustrative output images from the test set of the D-Dragon dataset. Corresponding to each label above, we evaluated five labels with names in Vietnamese: "Dom nau," "Mat ca," "Nam canh," "Dom den," and "Thoi dau trai," representing various diseases on the leaves and fruits of dragon fruit plants. These are typical diseases that damage farmers' dragon fruit crops. The prediction results of the SCL model are displayed below, next to the actual images.

We initially trained the model with 5 labels: "Dom nau," "Mat ca," and "Nam canh." We then added 2 more labels, "Dom den" and "Thoi dau trai," enabling the model to predict 7 labels. From these five examples, it can be seen that the SCL model classifies the diseases of dragon fruit very well. For the most part, the model's performance is not adversely affected by the addition of new labels and achieves satisfactory results.

5. **Conclusion.** In this work, we aim to improve the method of incremental learning of new labels based on semantic learning by connecting labels and features to enhance learning capabilities in the real world. This paper proposes the Semantic Connection-based Learning (SCL) method for expanding classes with a pre-trained model to support the recognition of diseases on dragon fruit plants. Specifically, we have attempted to build a diverse workspace through connection sets. The aggregation of semantic features allows the model to understand the prototype features of classes and create integral connection sets that support adding new labels without affecting the prediction quality of existing labels. Additionally, we have built a dataset of diseases on dragon fruit, a significant agricultural product in Vietnam. We conducted experiments to demonstrate the effectiveness of our SCL model. Although the SCL model fundamentally achieves good accuracy, it requires considerable storage space for model retention, which is memory-intensive. Therefore, our future work will focus on designing a more streamlined model and optimizing the connection sets to minimize memory usage.

Acknowledgement. This research was funded by the scientific project under code KC-4.0-11/19-25 of the Ministry of Science and Technology of Vietnam.

REFERENCES

[1] A. S. Chains and E. Bilali, “Sharma2020 a systematic literature review on machine learning applications for sustainable agriculture supply chain performance _ elsevier enhanced reader,” pp. 1–42, 2019.

[2] Y. Liu, X. Ma, L. Shu, G. P. Hancke, and A. M. Abu-Mahfouz, “From industry 4.0 to agriculture 4.0: Current status, enabling technologies, and research challenges,” *IEEE Trans. Ind. Informatics*, vol. 17, no. 6, pp. 4322–4334, 2021, doi: 10.1109/TII.2020.3003910.

[3] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, “Machine learning applications for precision agriculture: A comprehensive review,” *IEEE Access*, vol. 9, pp. 4843–4873, 2021, doi: 10.1109/ACCESS.2020.3048415.

[4] M. Chithambarathanu and M. K. Jeyakumar, “Survey on crop pest detection using deep learning and machine learning approaches,” *Springer US*, vol. 82, no. 27, pp. 4843–4873, 2021, doi: 10.1007/s11042-023-15221-3.

[5] A. Ahmad, D. Saraswat, and A. El Gamal, “A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools,” *Smart Agric. Technol*, vol. 3, p. 100083, 2023, doi: 10.1016/j.atech.2022.100083.

[6] E. J. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nat. Med*, vol. 25, no. 1, pp. 44–56, 2019, doi: 10.1038/s41591-018-0300-7.

[7] S. K. Jagatheesaperumal, M. Rahouti, K. Ahmad, A. Al-Fuqaha, and M. Guizani, “The duo of artificial intelligence and big data for industry 4.0: Applications, techniques, challenges, and future research directions,” *IEEE Internet Things J*, vol. 9, no. 15, pp. 12 861–12 885, 2022, doi: 10.1109/JIOT.2021.3139827.

[8] O. Serradilla, E. Zugasti, J. Rodriguez, and U. Zurutuza, “Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects,” *Appl. Intell*, vol. 52, no. 10, pp. 10 934–10 964, 2022, doi: 10.1007/s10489-021-03004-y.

[9] M. De Lange et al, “A continual learning survey: Defying forgetting in classification tasks,” *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 44, no. 7, pp. 3366–3385, 2022, doi: 10.1109/TPAMI.2021.3057446.

[10] A. Douillard, Y. Chen, A. Dapogny and M. Cord, “Plop: Learning without forgetting for continual semantic segmentation,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit*, pp. 4039–4049, 2021, doi: 10.1109/CVPR46437.2021.00403.

[11] Y. Cong, M. Zhao, J. Li, S. Wang and L. Carin, “Gan memory with no forgetting,” *Adv. Neural Inf. Process. Syst*, 2020.

[12] S. F. Ahmed et al, “Deep learning modelling techniques: current progress, applications, advantages, and challenges,” *Springer Netherlands*, vol. 56, no. 11, 2023, doi: 10.1007/s10462-023-10466-8.

[13] I. H. Sarker, “Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions,” *SN Comput. Sci*, vol. 2, no. 6, pp. 1–20, 2021, doi: 10.1007/s42979-021-00815-1.

[14] A. N. Muhammad, A. M. Aseere, H. Chiroma, H. Shah, A. Y. Gital and I. A. T. Hashem, “Deep learning application in smart cities: recent development, taxonomy, challenges and research prospects,” vol. 33, no. 7, 2021, doi: 10.1007/s00521-020-05151-8.

[15] Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng and Jun Zhou, “A survey of convolutional neural networks: analysis, applications, and prospects,” *IEEE Trans. neural networks Learn. Syst*, vol. 33, no. 12, pp. 6999–7019, 2021.

[16] T. Kattenborn, J. Leitloff, F. Schiefer and S. Hinz, “Review on convolutional neural networks (cnn) in vegetation remote sensing,” *ISPRS J. Photogramm. Remote Sens*, vol. 173, pp. 24–29, 2021, doi: 10.1016/j.isprsjprs.2020.12.010.

[17] L. Fregoso-Aparicio, J. Noguez, L. Montesinos and J. A. García-García, “Machine learning and deep learning predictive models for type 2 diabetes: a systematic review,” *Diabetol. Metab. Syndr.*, vol. 13, no. 1, 2021, doi: 10.1186/s13098-021-00767-9.

[18] K. Hendrickx, L. Perini, D. Van der Plas, W. Meert and J. Davis, “Machine learning with a reject option: a survey,” *Mach. Learn*, vol. 113, no. 5, pp. 3073–3110, 2024, doi: 10.1007/s10994-024-06534-x.

[19] F. A. M. Al-Yarimi, N. M. A. Munassar, M. H. M. Bamashmos and M. Y. S. Ali, “Feature optimization by discrete weights for heart disease prediction using supervised learning,” *Soft Comput*, vol. 25, no. 3, pp. 1821–1831, 2021, doi: 10.1007/s00500-020-05253-4.

[20] G. Lin, H. Chu and H. Lai, "Towards better plasticity-stability trade-off in incremental learning: A simple linear connector," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit*, pp. 89–98, 2022, doi: 10.1109/CVPR52688.2022.00019.

[21] R. Krishnan and P. Balaprakash, "Formalizing the generalization-forgetting trade-off in continual learning," *Adv. Neural Inf. Process. Syst*, vol. 21, no. 1, pp. 17 284–17 297, 2021.

[22] C. Xie, Z. Zhang, Y. Wu, F. Zhu, R. Zhao and S. Liang, "Described object detection: Liberating object detection with flexible expressions," *Adv. Neural Inf. Process. Syst*, vol. 36, pp. 1–13, 2023.

[23] G. Zheng, A. H. Awadallah and S. Dumais, "Meta label correction for noisy label learning," *35th AAAI Conf. Artif. Intell. AAAI 2021*, vol. 12B, pp. 11 053–11 061, 2021, doi: 10.1609/aaai.v35i12.17319.

[24] Y. Bai et al, "Understanding and improving early stopping for learning with noisy labels," *Adv. Neural Inf. Process. Syst*, vol. 29, pp. 24 392–24 403, 2021.

[25] K. Zhu and J. Wu, "Residual attention: A simple but effective method for multi-label recognition," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 184–193, 2021, doi: 10.1109/ICCV48922.2021.00025.

[26] Y. Wang et al, "Multi-label classification with label graph superimposing," *AAAI 2020 - 34th AAAI Conf. Artif. Intell*, pp. 12 265–12 272, 2020, doi: 10.1609/aaai.v34i07.6909.

[27] W. Chen et al, "Deep learning for instance retrieval: A survey," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 45, no. 6, pp. 7270–7292, 2023, doi: 10.1109/TPAMI.2022.3218591.

[28] A. Tabbakh and S. S. Barpanda, "A deep features extraction model based on the transfer learning model and vision transformer 'tlmvit' for plant disease classification," *IEEE Access*, vol. 11, no. April, pp. 45 377–45 392, 2023, doi: 10.1109/ACCESS.2023.3273317.

[29] Y. Y. Tsai, P. Y. Chen and T. Y. Ho, "Transfer learning without knowing: Reprogramming black-box machine learning models with scarce data and limited resources," *37th Int. Conf. Mach. Learn. ICML 2020*, vol. PartF168147-13, pp. 9556–9566, 2020.

[30] K. T. Tran, X. S. Vu, K. Nguyen and H. D. Nguyen, "Neupronet: neural profiling networks for sound classification," *Neural Comput. Appl*, vol. 36, no. 11, pp. 5873–5887, 2024, doi: 10.1007/s00521-023-09361-8.

[31] B. Liu, S. Wei, F. Zhang, N. Guo, H. Fan and W. Yao, "Tomato leaf disease recognition based on multi-task distillation learning," *Front. Plant Sci*, vol. 14, no. January, pp. 1–19, 2023, doi: 10.3389/fpls.2023.1330527.

[32] W. Liu, H. Wang, X. Shen and I. W. Tsang, "The emerging trends of multi-label learning," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 44, no. 11, pp. 7955–7974, 2022, doi: 10.1109/TPAMI.2021.3119334.

[33] J. Lanchantin, T. Wang, V. Ordonez and Y. Qi, "General multi-label image classification with transformers," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit*, pp. 16 473–16 483, 2021, doi: 10.1109/CVPR46437.2021.01621.

[34] Z. Liu et al, "Improved fine-tuning by better leveraging pre-training data," *Adv. Neural Inf. Process. Syst*, vol. 35, no. NeurIPS, pp. 1–14, 2022.

[35] L. Yuan, Q. Hou, Z. Jiang, J. Feng and S. Yan, "Volo: Vision outlooker for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 45, no. 5, pp. 6575–6586, 2023, doi: 10.1109/TPAMI.2022.3206108.

[36] Y. Mei et al, "Pyramid attention network for image restoration," *Int. J. Comput. Vis*, vol. 131, no. 12, pp. 3207–3225, 2023, doi: 10.1007/s11263-023-01843-5.

[37] W. Li, T. Zheng, Z. Yang, M. Li, C. Sun and X. Yang, "Classification and detection of insects from field images using deep learning for smart pest management: A systematic review," *Int. J. Comput. Vis*, vol. 66, 2021, doi: 10.1016/j.ecoinf.2021.101460.

[38] X. Cao, Z. Wei, Y. Gao and Y. Huo, "Recognition of common insect in field based on deep learning," *J. Phys. Conf. Ser.*, vol. 1634, no. 1, 2020, doi: 10.1088/1742-6596/1634/1/012034.

[39] N. A. Al-Mekhlafi et al, "Metabolomic and antioxidant properties of different varieties and origins of dragon fruit," *Microchem. J.*, vol. 160, no. 1, 2021, doi: 10.1016/j.microc.2020.105687.

[40] Z. Unal, "Smart farming becomes even smarter with deep learning - a bibliographical analysis," *IEEE Access*, vol. 8, pp. 105 587–105 609, 2020, doi: 10.1109/ACCESS.2020.3000175.

[41] P. Date, "Smart agriculture: Advancements on plant disease detection with big data and deep learning," 2022.

[42] M. Ahmad et al, "A comprehensive survey for hyperspectral image classification: The evolution from conventional to transformers," 2024, available: http://arxiv.org/abs/2404.14955.

[43] T. Y. Wu et al, "Class-incremental learning with strong pre-trained models," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit*, vol. 2022-June, pp. 9591–9600, 2022, doi: 10.1109/CVPR52688.2022.00938.

[44] A. P. Fard and M. H. Mahoor, "Ad-corre: Adaptive correlation-based loss for facial expression recognition in the wild," *IEEE Access*, vol. 10, pp. 26 756–26 768, 2022, doi: 10.1109/ACCESS.2022.3156598.

[45] X. Wang, J. Liang, Y. Xiao and W. Wang, "Prototypical concept representation," *IEEE Trans. Knowl. Data Eng*, vol. 35, no. 7, pp. 7357–7370, 2023, doi: 10.1109/TKDE.2022.3180886.

[46] M. Ristin, M. Guillaumin, J. Gall and L. Van Gool, "Incremental learning of random forests for large-scale image classification," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 38, no. 3, pp. 490–503, 2016, doi: 10.1109/TPAMI.2015.2459678.

[47] L. Zimmer, M. Lindauer and F. Hutter, "Auto-pytorch: Multi-fidelity metalearning for efficient and robust autodl," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 43, no. 9, pp. 3079–3090, 2021, doi: 10.1109/TPAMI.2021.3067763.

[48] A. Parvaiz, M. A. Khalid, R. Zafar, H. Ameer, M. Ali and M. M. Fraz, "Vision transformers in medical computer vision—a contemplative retrospection," *Eng. Appl. Artif. Intell*, vol. 122, pp. 0–3, 2023, doi: 10.1016/j.engappai.2023.106126.

[49] H. Yu, X. Cheng and W. Peng, "Toplight: Lightweight neural networks with task-oriented pretraining for visible-infrared recognition," pp. 3541–3550, 2023, doi: 10.1109/cvpr52729.2023.00345.

[50] X. Liu et al, "Generative feature replay for class-incremental learning," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work*, vol. 2020–June, pp. 915–924, 2020, doi: 10.1109/CVPRW50498.2020.00121.

[51] S. Yan, J. Xie and X. He, "Der: Dynamically expandable representation for class incremental learning," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3013–3022, 2021, doi: 10.1109/CVPR46437.2021.00303.

[52] M. Zhang, S. Levine and C. Finn, "Memo: Test time robustness via adaptation and augmentation," *Adv. Neural Inf. Process. Syst.*, vol. 35, no. NeurIPS, pp. 1–14, 2022.

[53] D.-W. Zhou, H.-L. Sun, H.-J. Ye and D.-C. Zhan, "Expandable subspace ensemble for pre-trained model-based class-incremental learning," pp. 23 554–23 564, 2024, available: http://arxiv.org/abs/2403.12030.