Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 246 (2024) 606–615

Procedia Computer Science

www.elsevier.com/locate/procedia

28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2024)

Incremental and Zero-Shot Machine Learning for Vietnamese Medicinal Plant Image Classification

Trien Phat Tran[a,*], Fareed Ud Din[a], Ljiljana Brankovic[a], Cesar Sanin[b], Susan M Hester[a,c], Minh Duc Hoang Le[d]

[a]*The University of New England, Armidale NSW 2351, Australia*
[b]*The University of Newcastle, Callaghan NSW 2308, Australia*
[c]*The University of Melbourne, Parkville VIC 3052, Australia*
[d]*Thanh Dong University, Hai Duong, Vietnam*

Abstract

This paper presents a study on the use of incremental and zero-shot learning for classifying Vietnamese medicinal plants using image analysis. Traditional machine learning methods often struggle with the constant emergence of new plant species and variability in appearances. Our methodology combines incremental learning, which continuously updates the model with new data while retaining prior knowledge, and zero-shot learning, which classifies unseen plant species by leveraging semantic similarities. Evaluated on a unique dataset from Vietnam, our approach shows improved adaptability, robustness, and reduced dependency on extensive labeled data, making it suitable for dynamic environments like medicinal plant identification.

Keywords: incremental machine learning; zero-shot learning; image classification, medicinal plants; openai clip

1. Introduction

1.1. Context and Importance

The field of image classification has witnessed remarkable advancements with the advent of deep learning techniques, such as convolutional neural networks (CNNs) [1]. State-of-the-art architectures like VGGNet [2], ResNet [3], and EfficientNet [4] have achieved impressive performance on various benchmarks. However, these models often require extensive labeled training data and struggle to adapt to new or evolving classes without retraining from scratch.

* Corresponding author. Tel.: +61-404-645-878.
E-mail address: ttran72@myune.edu.au

1877-0509
10.1016/j.procs.2024.09.469

This limitation presents significant challenges in dynamic environments like medicinal plant identification, where new species or variations may emerge continually.

Incremental learning, also known as continual learning, offers a promising solution by enabling models to acquire new information from a changing data stream while retaining previously learned knowledge [5]. This capability is crucial in agricultural settings where new plant species or diseases can emerge over time. Furthermore, zero-shot learning (ZSL) techniques allow models to classify instances into classes that were not seen during the training phase [6, 7], further enhancing their adaptability and reducing the need for extensive data labeling.

1.2. Research Gap

There is a notable gap in the application of incremental and zero-shot learning for the classification of medicinal and herbal plants, as current literature primarily focuses on general plant species. This presents an opportunity to enhance the classification of rare or newly discovered medicinal plants by integrating these learning methods. Such a combination could leverage diverse imagery from the public to improve model robustness and generalization, filling a critical gap in the research and potentially transforming the field.

1.3. Research question

How can incremental machine learning and zero-shot learning enhance the classification of medicinal plants in resource-constrained environments?

This research aims to investigate the potential to harness incremental and zero-shot learning respectively for the task of classifying Vietnamese medicinal plants, particularly in settings where resources and labeled data are scarce. The motivation behind this question lies in addressing the challenges associated with the constant evolution of data and the introduction of new plant species that may not have previously been included in training datasets.

2. Related studies

2.1. Incremental learning

In our study, a detailed search for recent literature combining ”incremental learning” or ”incremental online learning” with ”medicinal” or ”herbal plants” revealed no results. This indicates a research gap at the nexus of these fields, with existing studies primarily focusing on plant disease detection and classification.

Methodological Insights

The reviewed papers indicate that incremental learning techniques have gained significant attention in the field of plant image classification. These methods enable models to continually learn from new data without forgetting previously acquired knowledge, a phenomenon known as catastrophic forgetting. Notable approaches include class-incremental learning [10], coupled with knowledge distillation techniques like Dynamic Y-KD [10] and Gaussian Mixture Models [17]. Other methodologies incorporate transfer learning [9, 8, 15], allowing the model to leverage knowledge from related tasks or domains. Convolutional Neural Networks (CNNs) and Long Short-Term Memory Recurrent Neural Networks (LSTM-RNNs) have been widely employed in conjunction with incremental learning frameworks [16, 15].

Advantages and Generalization of Incremental Learning

Incremental learning offers several advantages in the context of medicinal plant image classification. It enables models to adapt to changing environments and continuously improve their performance as new data becomes available [17, 15]. This is particularly beneficial in dynamic agricultural settings where new plant species or diseases may emerge over time. Additionally, incremental learning can reduce the computational complexity and memory requirements associated with training on large datasets [16]. By leveraging transfer learning and zero-shot learning techniques, models can generalize to unseen classes [9], enhancing their adaptability and reducing the need for extensive data labelling.

Challenges and Limitations

Despite its advantages, incremental learning faces several challenges and limitations. Catastrophic forgetting remains a significant issue, where models struggle to retain previously learned knowledge while incorporating new

information [10]. This is exacerbated by the trade-off between plasticity and stability, as increased adaptability to new data can lead to instability and forgetting of old knowledge [10]. Furthermore, the performance of incremental learning methods on unseen classes is often limited by the representativeness and variability of the training data, as well as the effectiveness of the semantic spaces constructed to bridge the gap between seen and unseen classes [9]. Additionally, some incremental learning approaches may incur higher memory and computational costs during inference [10], which can be a limitation in real-time applications.

2.2. Zero-shot learning

Methodological Advancements The study of zero-shot learning (ZSL) has seen significant methodological advancements in recent years. Belissent et al. [14] explored techniques such as Zero-Shot Learning, Transfer Learning, Kernel Extreme Learning Machine (KELM), and Crossover Optimization. Singh et al. [11] proposed a combination of Zero-Shot Transfer Learning, Robust Discriminative Losses, and Convolutional Neural Networks (CNNs). Zabihzadeh et al. [12] investigated a diverse range of methods, including Zero-Shot Learning (ZSL), Deep Metric Learning (DML), Few-Shot Learning (FSL), Data Augmentation-based Methods, Meta Learning-based Methods, Metric Learning-based Methods, Siamese Network, Soft-Triple Loss, Inception V3, GDFL (General Discriminative Feature Learning), and Proxy-based DML Methods. Additionally, Zhong et al. [13] explored Conditional Adversarial Autoencoders (CAAE), Adversarial Training, and Generative Models for ZSL.

Advantages and Applications Zero-shot learning offers several advantages and applications in the domain of medicinal plant image classification. It addresses the challenge of identifying classes where no labeled examples are available, enabling the model to leverage knowledge from a labeled source domain to classify classes in the target domain without direct training [11]. This approach is valuable for expanding classification capabilities to previously unseen objects or classes [11]. ZSL promotes generalization with minimal supervision, making it useful when labeled data is limited or unavailable in the target domain [11]. Furthermore, it enhances the model's adaptability to handle the dynamic nature of agricultural disease environments where new pathogens can emerge [13]. [12] highlighted the generalization capabilities, discrimination power, and resource efficiency of their ZS-DML method for plant disease detection.

Challenges and Limitations Despite its advantages, zero-shot learning faces several challenges and limitations. [11] discussed the dependency on semantic attributes, limited generalization to unseen classes, sensitivity to domain shift, limitations in fine-grained tasks, and the lack of interpretability. [13] highlighted the dependency on attribute quality, the domain shift problem, the balance between seen and unseen classes, generative model challenges, complexity of implementation, and scalability to larger and more diverse sets as potential limitations. Additionally, [13] mentioned the challenges of ensuring that synthetic data generated by generative approaches like CAAE is diverse and representative enough to effectively train the model.

3. Proposed Methodology

3.1. About the dataset

Our dataset, known as Med Herb Lens [18], comprised images of medicinal plants captured manually in Vietnam using a mobile device. It was organised into seven preliminary categories, each represented by a unique numerical identifier. Each category included approximately 100 images, totalling around 707 images. The dataset was dynamic, with plans to expand by adding more categories and images as additional data becomes available through crowdsourcing in future work.

Class descriptions mapping:

- Averrhoa carambola: A medium-sized tree with distinctive five-angled, yellow to green star-shaped fruits, often eaten fresh or used in cooking.,

- Piper sarmentosum: A low-growing plant with heart-shaped, glossy leaves that are widely used as a wrap in Southeast Asian cuisine.,

- Piper betle: A vine with glossy, heart-shaped leaves, commonly chewed with areca nut as a traditional custom in many Asian cultures.,

- Stachytarpheta jamaicensis: A shrub with elongated spikes of small, deep blue to purple flowers, commonly known as blue porterweed.,
- Polyscias fruticosa: A shrub with dense, finely divided leaves, often used as an ornamental plant in tropical gardens.,
- Paederia tomentosa: A vine known for its foul-smelling leaves when crushed and used traditionally in some Asian herbal medicines.,
- Cordyline fruticosa: An evergreen plant with striking red or purple leaves, often used in tropical landscapes or as indoor ornamental plants.

3.2. Incremental Machine Learning model

Our approach could be described as incremental batch learning. In this model [19], we added new data in batches for further training which is an effective strategy for updating models with new data without retraining from scratch. The incremental learning approach involved the following steps:

1. **Data Preparation:** The dataset was organized into two distinct sets: an initial set with five classes and an incremental set. The incremental one contained images of two new/unseen classes and new/unseen images of the initial set's two other classes. Each set was further divided into training and testing subsets.
2. **Model Initialization:** A pre-trained ResNet-18 model [21] from the PyTorch torchvision.models library was employed as the base model. The final fully connected layer was replaced with a new linear layer to match the number of classes in the initial dataset.
3. **Initial Training:** The initialized model was trained on the initial training set using the cross-entropy loss function and stochastic gradient descent (SGD) optimizer with a learning rate of 0.001 and momentum of 0.9. The model was trained for 3 epochs, and its performance was evaluated on the initial test set.
   The formula for cross-entropy loss in a classification context is:

$$L = -\sum_{c=1}^{M} y_{o,c} \log(p_{o,c}) \tag{1}$$

   Here, M is the number of classes, y is a binary indicator (0 or 1) if class label c is the correct classification for observation o, and p is the predicted probability that observation o is of class c.
   The update rule for SGD with momentum is given by:

$$\begin{aligned} v_t &= \gamma v_{t-1} + \eta \nabla_\theta J(\theta) \\ \theta &= \theta - v_t \end{aligned} \tag{2}$$

   where θ represents the parameters of the model, $\nabla_\theta J(\theta)$ is the gradient of the objective function with respect to the parameters, η is the learning rate, v_t is the update vector at time t, and γ is the momentum factor.
4. **Incremental Training:** After the initial training phase, the model was further trained on the incremental training set. This step aimed to expand the model's knowledge to classify the new classes introduced in the incremental dataset. The training process followed the same hyperparameters and settings as the initial training phase, running for an additional 3 epochs.
5. **Evaluation:** The incrementally trained model's performance was evaluated on the incremental test set. Various metrics, including accuracy, precision, recall, F1-score, and a confusion matrix, were computed and reported.
6. **Model Saving:** Finally, the trained model's state dictionary was saved to a file (.pth) for future use or further analysis.

3.3. Zero-Shot Learning dataset classification using CLIP

The zero-shot learning approach [20] aimed to leverage OpenAI's pre-trained CLIP (Contrastive Language-Image Pre-Training) models [22] to classify medicinal plant images without any additional fine-tuning on the target dataset. The process involved the following steps:

1. **Model Loading:** Several pre-trained CLIP models were loaded from the clip library, including RN50, RN101, RN50x4, RN50x16, RN50x64, ViT-B/32, ViT-B/16, ViT-L/14, and ViT-L/14@336px [22]. These models were pre-trained on a large dataset of image-text pairs, enabling them to capture visual-semantic representations.
2. **Data Preparation:** The dataset was organized in a folder structure, with each class represented by a subfolder containing the corresponding images. Each subfolder was re-named after the plant's scientific name.
3. **Class Descriptions:** A dictionary of class descriptions was defined, mapping each class name to a textual description of the corresponding medicinal plant. These descriptions were used as the textual input to the CLIP model for zero-shot classification.
4. **Image Preprocessing:** Images were loaded and preprocessed using the preprocess function provided by the CLIP library. This step resized and normalized the images to match the input requirements of the CLIP model.
5. **Zero-Shot Classification:** For each class, the images were loaded and preprocessed. Then, the CLIP model encoded both the image and the corresponding class descriptions into feature representations. The cosine similarity between the image and text features was calculated, and the class with the highest similarity score was assigned as the predicted label. The formula for cosine similarity is:

$$\text{Cosine Similarity}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \tag{3}$$

   where $\mathbf{u}$ and $\mathbf{v}$ are the vector representations of the image and text features, respectively.
6. **Evaluation:** The performance of the zero-shot classification was evaluated using various metrics, including accuracy, precision, recall, F1-score, and a confusion matrix. These metrics were calculated by comparing the predicted labels with the true labels (class names).
7. **Visualization:** The code plotted the cosine similarities between image and text features, as well as the average accuracy for each class. The confusion matrix was visualized using Seaborn's heatmap.

4. Results and Analysis

4.1. Incremental Machine Learning model

4.1.1. Initial Training Phase

During the initial training phase, the model exhibited a high level of accuracy, achieving a test accuracy of 97.75%. The corresponding test loss was relatively low at 0.2416, indicating effective learning from the training dataset. Precision and recall were closely matched at 0.9786 and 0.9775, respectively, resulting in an F1 Score of 0.9769. This suggests a balanced performance across the prediction of different classes, which is critical in maintaining the reliability of the model in varied scenarios.

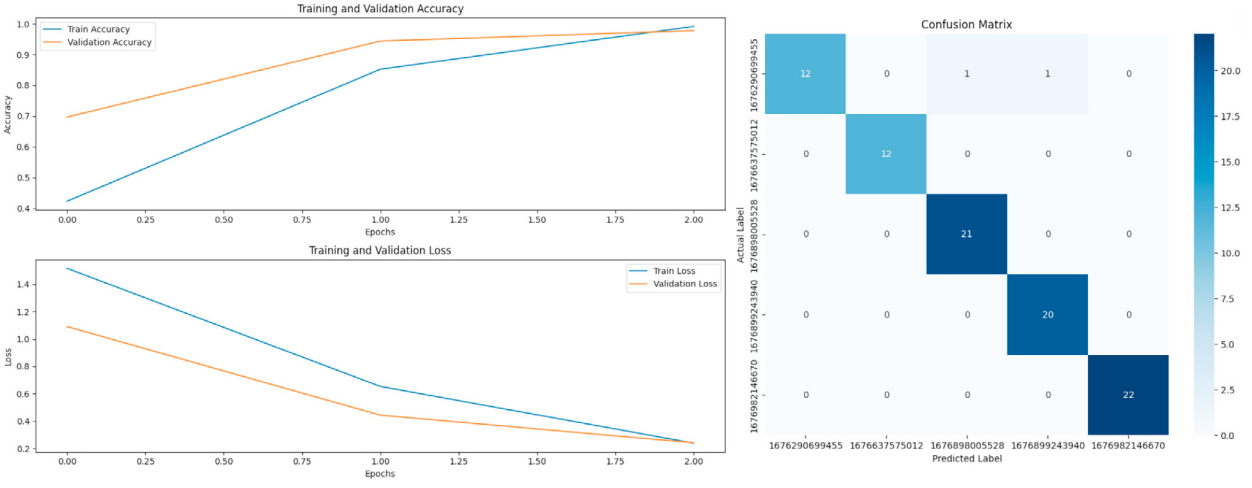

Fig. 1. (a) training and validation accuracy/loss; (b) confusion matrix.

4.1.2. Incremental Training Phase

The classification report from this phase reveals that the model performed exceptionally well across most classes, with perfect scores in precision for three out of five classes and recall scores reaching 1.00 for two classes. However, the class labelled 1676290699455 displayed a recall of 0.86, the lowest among the group, indicating some instances of this class were not identified correctly.

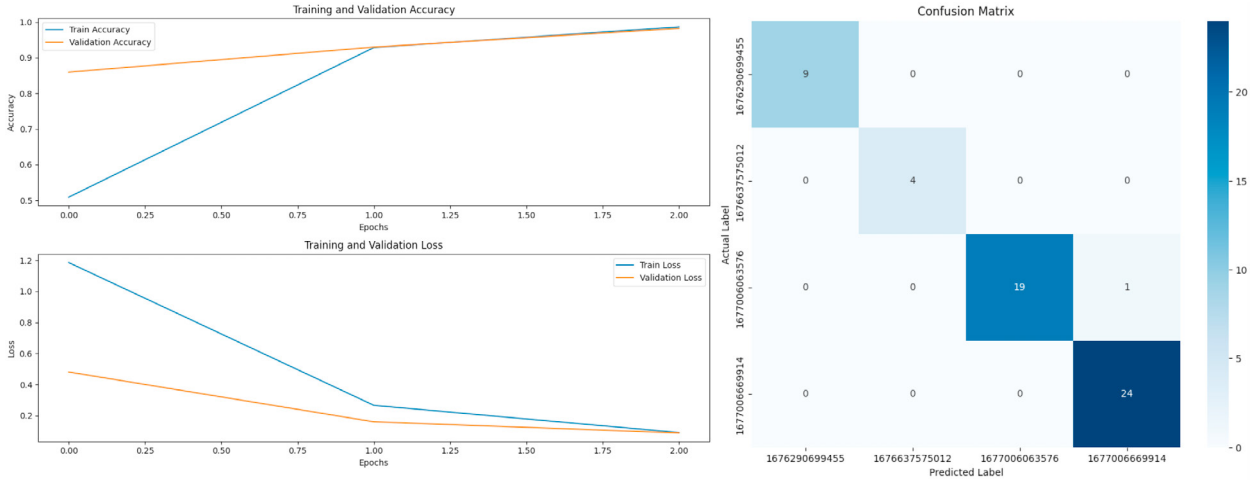

Fig. 2. (a) training and validation accuracy/loss; (b) confusion matrix.

In the subsequent incremental training phase, the model was updated with new data, potentially including different distributions or new classes, and it showed an improved overall performance. The test accuracy slightly increased to 98.25% with a notably lower test loss of 0.0887. Precision and recall values increased to 0.9832 and 0.9825, respectively, and the F1 Score also saw a marginal improvement to 0.9824. These metrics suggest that the model not only retained its previous knowledge but also adapted well to the new data.

The classification report for this phase indicates that all classes achieved high precision and recall rates, with several classes obtaining perfect scores. Notably, class 1677006063576 had a slightly lower recall of 0.95, suggesting minor challenges in classifying all instances of this class correctly.

4.1.3. Comparative Analysis

Comparing the two phases, the incremental training phase demonstrated enhancements in accuracy, precision, recall, and F1 Score, suggesting that the model effectively integrated new information without significant loss of previously learned information. The reduction in test loss underscores improved model efficiency post-incremental training.

Table 1. Incremental learning training results.

Metric	Initial training	Incremental training
Loss	0.2386	0.0909
Accuracy	0.9913	0.9866
Test Loss	0.2416	0.0887
Test Accuracy	0.9775	0.9825
Precision	0.9786	0.9832
Recall	0.9775	0.9825
F1 Score	0.9769	0.9824

Moreover, the analysis reveals that the model's ability to generalize across different sets of data improved with incremental training. This enhancement could be attributed to the additional data that may have provided more representative examples of the underlying patterns across classes.

In summary, the training results illustrate the model's robustness and adaptability to new data, underscoring the effectiveness of the chosen training methodology in fostering a resilient and accurate predictive model.

4.2. Zero-shot learning classification using CLIP

4.2.1. Cosine similarities

Cosine similarities reflect the alignment between image features and text descriptions across numerous instances. Here are some observations:

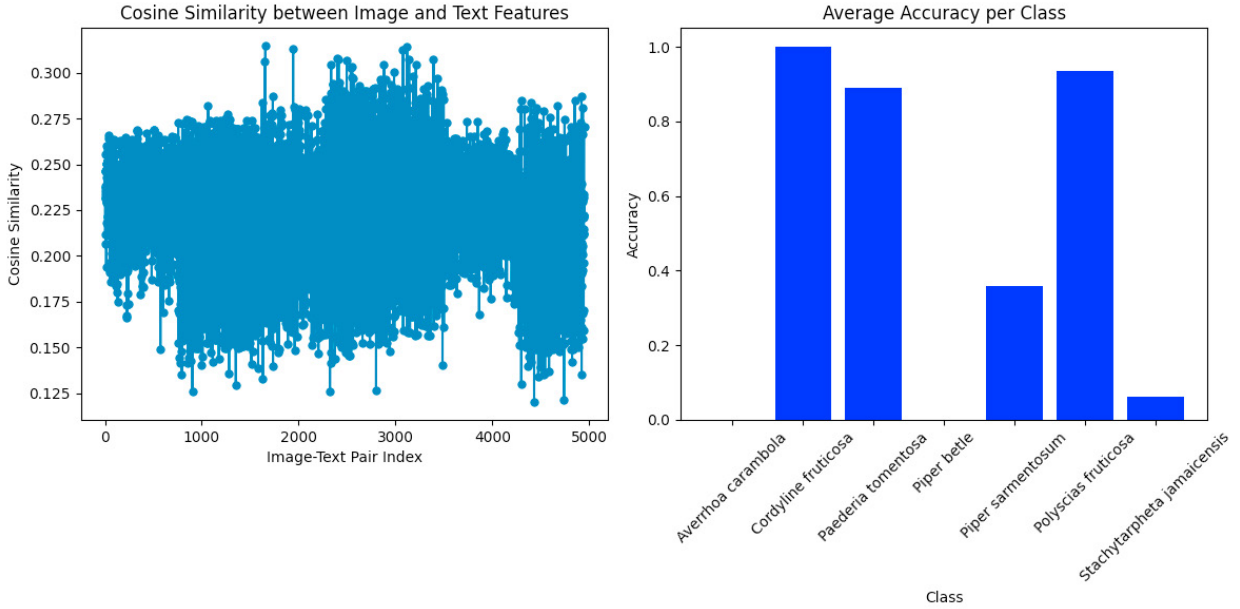

Fig. 3. (a) cosine similarities between text description and image; (b) average accuracy per class.

- Range of Values: The cosine similarities range from around 0.16 to about 0.33, suggesting a moderate to substantial level of alignment or match between the images and corresponding text across the dataset. Lower values indicate less similarity or alignment, while higher values suggest greater congruence between the textual descriptions and image features.
- Distribution: The majority of cosine similarity scores are concentrated in the mid-range (around 0.25 to 0.30). This suggests that for most pairs, there is a reasonable degree of matching between the images and text, but there may be room for improvement in terms of achieving higher similarity scores, which would indicate a better match.
- Extremes and Variability: The minimum and maximum values (near 0.16 and 0.33, respectively) show that there is some variability in how well some images match their descriptions compared to others. This could be due to various factors such as the quality of images, the descriptiveness and accuracy of text, or inherent difficulties in capturing certain features visually or describing them textually.
- Improvement Opportunities: To enhance the model's performance, it might be helpful to analyze cases with particularly low similarity scores to understand potential mismatches or shortcomings in the dataset or model processing. Similarly, reviewing high-scoring pairs could provide insights into what constitutes a 'good' match and how to replicate this success across other data points.

4.2.2. Model Architectures and Size Variations

- ResNet Models (RN50, RN101, RN50x4, RN50x16, RN50x64): These models show a broad range of accuracy, precision, recall, and F1 scores. Notably, RN50x64 achieved the highest accuracy (0.42) and F1 score (0.32)

among ResNet configurations. The larger-scale model, RN50x64, appears to perform better in terms of precision and recall compared to its smaller counterparts, indicating that scaling up can enhance performance.

- Vision Transformer Models (ViT-B/32, ViT-B/16, ViT-L/14, ViT-L/14@336px): ViT models generally provided improved results over most ResNet configurations, with the ViT-L/14@336px model achieving the highest accuracy (0.48), precision (0.48), recall (0.46), and F1 score (0.38) across all models tested. This suggests that ViT architectures, especially at larger scales or higher resolutions, are more adept at capturing the nuances of complex image-text relationships.

Table 2. OpenAI's CLIP models classification results.

Model	Accuracy	Precision (macro)	Recall (macro)	F1 Score (macro)
RN50	0.31	0.38	0.28	0.23
RN101	0.22	0.15	0.21	0.12
RN50x4	0.4	0.29	0.39	0.3
RN50x16	0.23	0.41	0.22	0.13
RN50x64	0.42	0.4	0.41	0.32
ViT-B/32	0.41	0.26	0.38	0.27
ViT-B/16	0.42	0.29	0.39	0.3
ViT-L/14	0.42	0.35	0.4	0.33
ViT-L/14@336px	0.48	0.48	0.46	0.38

- Accuracy: Overall, the low Accuracy indicates a need for further tuning or a different approach. ViT models generally outperform ResNet models, with ViT-L/14@336px leading. This reflects the transformer's capacity to better generalize from the training data under the constraints of this fine-grained dataset.
- Precision and Recall: The low Precision and low Recall indicate that a large number of false positives are present and many actual positives are missed. There is a noticeable variance in precision and recall across models. ViT models, particularly at larger scales, show a better balance between precision and recall, indicating a more robust model for both identifying relevant classes and minimizing misclassification.
- F1 Score: This is a weighted average of precision and recall. A low F1 score reflects poor performance, particularly in balancing precision and recall. The F1 scores are significantly higher in ViT models, especially ViT-L/14@336px, which achieves the best balance between precision and recall, essential for datasets with uneven class distributions.

4.2.3. Confusion Matrix Insights

Detailed examination of confusion matrices reveals specific classes that are consistently misclassified across different models, suggesting possible similarities in visual features or insufficient training data for those classes. For instance, significant confusion occurs in models with lower overall accuracy, where classes like 'Averrhoa carambola' and 'Polyscias fruticosa' are frequently misidentified.

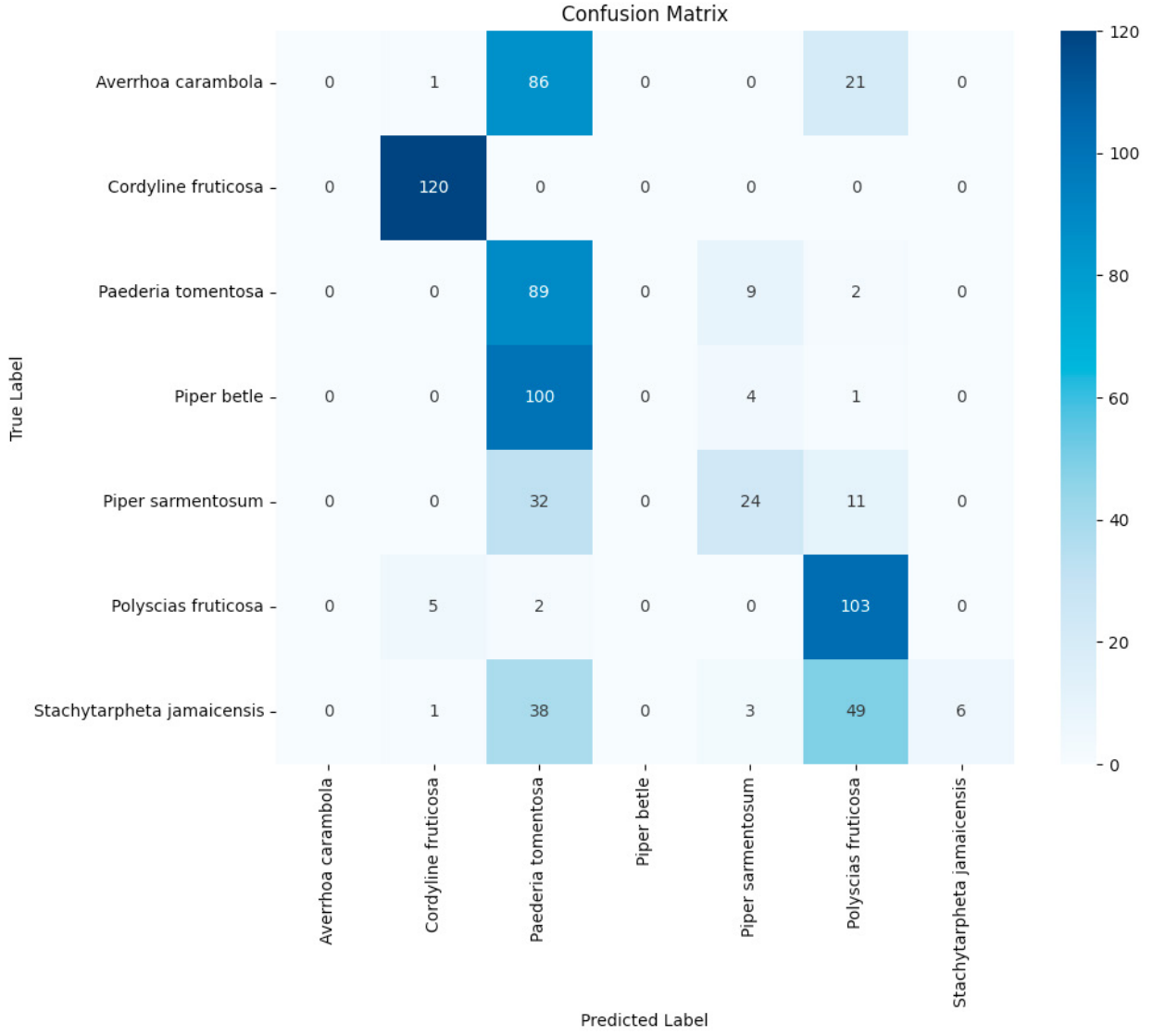

Fig. 4. Enter Caption

5. Conclusion and future work

While the incremental learning model performed well, OpenAI's CLIP struggled with fine-grained classification of closely related plant species. This limitation likely stems from CLIP's generalist training data, which may lack detailed examples and descriptions for distinguishing subcategories. Potential solutions include fine-tuning CLIP on specialized medicinal plant datasets, augmenting training data with detailed descriptions, engineering domain-specific features, and implementing attention mechanisms.

Looking ahead, key future directions involve:

- Expanding the dataset through crowdsourcing for enhanced representativeness
- Optimizing incremental learning strategies with advanced techniques to mitigate catastrophic forgetting
- Fine-tuning zero-shot models like CLIP on specialized medicinal plant data to improve performance
- Exploring hybrid approaches combining incremental and zero-shot learning strengths

- Adapting methodologies to new domains like plant disease detection or biodiversity monitoring

By addressing these areas, we can enhance machine learning capabilities for accurate medicinal plant classification and promote sustainable practices. Overcoming the limitations of current models on fine-grained tasks will be crucial for realizing the full potential of these techniques in specialized domains with rapidly evolving data.

Acknowledgements

The literature search and study selection process was conducted manually, with the assistance of Excel spreadsheet organization and tracking. While Adobe Acrobat Reader AI Assistant facilitated the data extraction process, data synthesis critical analysis and interpretation were manually performed by the researchers. The final paper was prepared using Overleaf.

References

[1] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998) "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* **86** (11): 2278–2324.
[2] Simonyan, Karen and Zisserman, Andrew. (2015) "Very Deep Convolutional Networks for Large-Scale Image Recognition." In *International Conference on Learning Representations*.
[3] He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, and Sun, Jian. (2016) "Deep Residual Learning for Image Recognition." In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 770–778.
[4] Tan, Mingxing and Le, Quoc V. (2019) "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." *CoRR*, abs/1905.11946.
[5] van de Ven, Gido M., Tuytelaars, Tinne, and Tolias, Andreas S. (2022) "Three types of incremental learning." *Nature Machine Intelligence* **4** (12): 1185–1197.
[6] Wang, W., Zheng, V., Yu, H., and Miao, C. (2019) "A Survey of Zero-Shot Learning: Settings, Methods, and Applications." *ACM Transactions on Intelligent Systems and Technology* **10** (2): 1–37.
[7] Saad, E., Paprzycki, M., Ganzha, M., Bădică, A., Costin Bădică, Fidanova, S., Lirkov, I., and Ivanović, M. (2022) "Generalized Zero-Shot Learning for Image Classification—Comparing Performance of Popular Approaches." *Information* **13** (12): 561.
[8] Minervini, M., Abdelsamea, M. M., and Tsaftaris, S. A. (2014) "Image-based plant phenotyping with incremental learning and active contours." *Ecological Informatics* **23**: 35–48.
[9] Chore, A., and Thankachan, D. (2023) "Nutrient Defect Detection In Plant Leaf Imaging Analysis Using Incremental Learning Approach With Multifrequency Visible Light Approach." *Journal of Electrical Engineering & Technology* **18** (2): 1369–1387.
[10] Sahu, S. K., and Pandey, M. (2023) "Hybrid Xception transfer learning with crossover optimized kernel extreme learning machine for accurate plant leaf disease detection." *Soft Computing (Berlin, Germany)* **27** (19): 13797–13811.
[11] Singh, R. Satya Rajendra, and Sanodiya, R. K. (2023) "Zero-shot Transfer Learning Framework for Plant Leaf Disease Classification." *IEEE Access* **11**: 1–1.
[12] Zabihzadeh, D., and Masoudifar, M. (2023) "ZS-DML: Zero-Shot Deep Metric Learning approach for plant leaf disease classification." *Multimedia Tools and Applications*.
[13] Zhong, F., Chen, Z., Zhang, Y., and Xia, F. (2020) "Zero- and few-shot learning for diseases recognition of Citrus aurantium L. using conditional adversarial autoencoders." *Computers and Electronics in Agriculture* **179**: 105828.
[14] Belissent, N., Peña, J. M., Mesías-Ruiz, G. A., Shawe-Taylor, J., and Pérez-Ortiz, M. (2024) "Transfer and zero-shot learning for scalable weed detection and classification in UAV images." *Knowledge-Based Systems* **292**: 111586.
[15] Page-Fortin, M. (2023) "Class-Incremental Learning of Plant and Disease Detection: Growing Branches with Knowledge Distillation." In *2023 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*: 593–603.
[16] Hung, P. D., and Su, N. T. (n.d.) "Incremental Learning for Classifying Vietnamese Herbal Plant." In *Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications*, Springer Singapore, pp. 434–442.
[17] Ouadfel, S., Mousser, W., Ghoul, I., & Taleb-Ahmed, A. (2022) "Incremental deep learning model for plant leaf diseases detection." In *Artificial Neural Networks for Renewable Energy Systems and Real-World Applications*, 207–222. Elsevier.
[18] Tran, T., Le, M. (2024). "Oriental Medicinal Herb Images [Data set]." Kaggle. Available online: https://www.kaggle.com/datasets/trientran/oriental-medicinal-herb-images
[19] Tran, T., Din, F., Brankovic, L., Sanin, C., Hester, S. (2024). "Incremental Machine Learning For Medicinal plants [Notebook]." Google Colab. Available online: https://colab.research.google.com/drive/1LVGqEWpq8sZ-LMYSJRsmZVCG2h2BH1-z
[20] Tran, T., Din, F., Brankovic, L., Sanin, C., Hester, S. (2024). "Zero-shot learning for medicinal plant image classification using CLIP [Notebook]." Google Colab. Available online: https://colab.research.google.com/drive/1xlDFdwDcyvkMZDx0QZxE2rrhOeyPfrLe
[21] PyTorch. (2024). "ResNet18 Model Documentation." Available online: https://pytorch.org/vision/main/models/generated/torchvision.models.resnet18.html
[22] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... and Sutskever, I. (2021, July). "Learning transferable visual models from natural language supervision." In *International Conference on Machine Learning* (pp. 8748-8763). PMLR.