Object detection and classification using few-shot learning in smart agriculture: A scoping mini review

Nitiyaa Ragu[1] and Jason Teo[2,3]*

[1]Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia,
[2]Advanced Machine Intelligence Research Group, Faculty of Computing and Informatics, Universiti
Malaysia Sabah, Kota Kinabalu, Malaysia, [3]Evolutionary Computing Laboratory, Faculty of
Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia

Smart agriculture is the application of modern information and communication
technologies (ICT) to agriculture, leading to what we might call a third green
revolution. These include object detection and classification such as plants,
leaves, weeds, fruits as well as animals and pests in the agricultural domain.
Object detection, one of the most fundamental and difficult issues in computer
vision has attracted a lot of attention lately. Its evolution over the previous
two decades can be seen as the pinnacle of computer vision advancement.
The detection of objects can be done *via* digital image processing. Machine
learning has achieved significant advances in the field of digital image
processing in current years, significantly outperforming previous techniques.
One of the techniques that is popular is Few-Shot Learning (FSL). FSL is a type
of meta-learning in which a learner is given practice on several related tasks
during the meta-training phase to be able to generalize successfully to new but
related activities with a limited number of instances during the meta-testing
phase. Here, the application of FSL in smart agriculture, with particular in the
detection and classification is reported. The aim is to review the state of the
art of currently available FSL models, networks, classifications, and offer some
insights into possible future avenues of research. It is found that FSL shows
a higher accuracy of 99.48% in vegetable disease recognition on a limited
dataset. It is also shown that FSL is reliable to use with very few instances and
less training time.

KEYWORDS

smart agriculture, object detection, computer vision, digital image processing,
machine learning, few-shot learning

Few-Short Learning in agriculture

In the realm of computer vision, detecting small objects such as plants, leaves, weeds,
fruits, and pests is a critical study topic. It is a technology that employs computer vision
equipment to acquire photos to determine whether the obtained images include weeds,
pests or even plant diseases (Lee et al., 2017). Computer vision-based object detection
technology is currently being used in agriculture and has partially replaced traditional
naked-eye identification (Liu and Wang, 2021). During the growth process, pests, and
weeds frequently harm crops which leads to various diseases. They will become more

dangerous if control is not done in a timely or effective manner. This will have an impact on crop yields and result in significant losses in agricultural income. People rely on their experiences while using traditional detection methods (Zhang S. et al., 2021).

Numerous agricultural plants, weeds, fruits, pests, and diseases have similar morphology, making it difficult to recognize them. Only judging by human eye observation is time-consuming, debilitating, biased, imprecise, and costly (Ashtiani et al., 2021). Nonprofessional agricultural employees will be harmed by their personal experiences since they will be unable to appropriately identify the species. As a result, pest control is an important aspect of agricultural productivity, and detecting agricultural pests is critical for agricultural development (Zhang S. et al., 2021).

As a result, an accurate detection technique is very desirable, and FSL is an approach that is precise in object detection. FSL learning is a sub-area in machine learning. FSL is a sort of machine learning problem (defined by E, T, and P), in which E, Experience, comprises a small number of supervised instances for T, Task, measured by P, Performance (Wang et al., 2020). FSL is used to quickly train the model on a little amount of tagged data (Fei-Fei et al., 2006). The fundamental concept behind FSL is comparable to how humans learn: it makes use of previously learned knowledge to pick up new skills while minimizing the quantity of training data needed. Research findings (Duan et al., 2021) demonstrate that FSL can address several machine learning's drawbacks, such as the need for huge datasets to be collected and labeled, by dramatically lowering the quantity of data necessary to train machine learning models. Due to the small number of datasets, the data are divided into two splits [train and test (without validation) or train and validation (without test)] in many types of research in the field of artificial intelligence, which can be a weakness. The data set is divided into two subsets because, in small data sets, an additional split could result in a smaller training set that is more susceptible to overfitting (Ashtiani et al., 2021).

A framework put forth by Fink (2004) allows object classifiers to be learned from a single example, which drastically reduces the amount of data gathering required for image classification tasks. Additionally, FSL enables the training of models that are appropriate for a few uncommon circumstances (Duan et al., 2021). A good example is the significant advancements in drug development that have been made as a result of the most recent FSL research advancements (Jiménez-Luna et al., 2020). Clinical biological data for a given drug are extremely scarce when predicting whether it is harmful, which frequently makes use of classic machine learning algorithms difficult.

FSL is widely used for scene classification (Alajaji et al., 2020), text classification (Muthukumar, 2021), image classification (Li X. et al., 2020) and image retrieval (Zhong Q. et al., 2020). Even though software and computational technologies have long been utilized for detection in agriculture,

FSL's capacity to identify them with limited training data makes them suitable for the task. This is ground-breaking because, they can be detected under various conditions of lighting, orientation, and background (Nuthalapati and Tunga, 2021).

FSL has been proven in agriculture to be an effective method for plant disease identification (Argüeso et al., 2020). Recent FSL approaches in agriculture are mainly used for plant disease detection (Wang and Wang, 2019; Li and Chao, 2021), fruit detection and classification (Janarthan et al., 2020; Ng et al., 2022), leaf identification and classification (Afifi et al., 2020; Jadon, 2020; Tassis and Krohling, 2022) and pest detection (Li and Yang, 2020, 2021; Nuthalapati and Tunga, 2021) using the meta-train set which is subsequently used to build embeddings for the samples in the meta-test set.

This article reviews the limitations of deep learning, advantages of FSL, types of FSL networks, their applications in the agricultural sector and their corresponding findings.

Limitations of deep learning

In agriculture, deep learning-based automatic classification of several categories is a hot topic of research (Too et al., 2019; Li Y. et al., 2020). Yet, in general, people still lean on non-automated classification by experts for the classification of distinct species. This is due in part to the fact that deep learning networks based on Convolutional Neural Networks (CNNs) demand thousands of labeled examples per target category for training, and labeling samples on such a large scale necessitates domain experts (Krizhevsky et al., 2017).

Furthermore, following training, the number of categories that a trained CNN-based model can recognize remains constant. To increase the number of categories that the network can recognize, it must be fine-tuned by adding fresh examples from other classes (Chen et al., 2019). Fine-tuning is a transfer learning approach which is used to benefit from the pre-trained network by adjusting the parameters to the data set. Due to the pre-established weights in a pre-trained network, fine-tuning is quicker than training from start (Ashtiani et al., 2021). Also, adequate data is required during training to avoid the network from overfitting (Gidaris and Komodakis, 2018) whereas humans can acquire new skills with little to no guidance using one or a few examples only (Lake et al., 2011).

Advantages of Few-Shot Learning

Humans can generalize new information based on a small number of examples, whereas artificial intelligence typically needs thousands of examples to produce comparable results (Duan et al., 2021). Researchers sought to create a machine learning model that, after learning a significant amount of data for some categories, could quickly learn a new category with only a few sample data, drawing inspiration from humans' rapid

learning abilities. This is the issue that FSL seeks to address. This task only offers a small number of observable examples, much like the production of character samples (Lake et al., 2015).

FSL has lately received a lot of attention as a way for networks to learn from a few samples. FSL tries to solve the challenge of classification with a small number of training examples, and it is gaining traction in a variety of domains (Lake et al., 2011). In FSL, there are two sets of labeled image data which are meta-train and meta-test, with image classes that are mutually exclusive in both sets (Argüeso et al., 2020). The goal is to develop a classifier on the visual classes in the meta-test set that can classify a particular query sample even with very few labeled examples using the data in the meta-train set and learn transferrable information (Altae-Tran et al., 2017).

Additionally, cross-domain FSL is defined as when the domain of visual classes in the meta-train set differs from the domain of visual classes in the meta-test set (Argüeso et al., 2020). One of the primary issues in cross-domain FSL is feature distribution differences between domains. Likewise, both the meta-train and meta-test sets in mixed-domain FSL contain classes from many domains, whereas single-domain FSL (Li and Yang, 2021) only contains instances from a single domain.

According to research (Duan et al., 2021), FSL can learn about uncommon circumstances where it is challenging to find information about supervision by combining one or a few examples with previously learned knowledge. This can solve the issue of the machine learning algorithm's need for thousands of supervised examples to guarantee the model's generalizability. Additionally, FSL can help expedite the training cycle and lower the high cost of data preparation (Parnami and Lee, 2022).

Types of networks in Few-Shot Learning

According to the traditional paradigm, an algorithm learns whether task performance gets better with practice when given a specific task. The Meta-Learning paradigm consists of several tasks. An algorithm that is learning to determine whether its efficiency at each task increases with practice and the number of tasks is called a meta-learning algorithm (Chen et al., 2021). Numerous Meta-Learning techniques have been released by researchers in recent years to address FSL classification issues. Metric-Learning (Yang and Jin, 2006) and Gradient-Based Meta-Learning (Khodak et al., 2019) methods can be used to categorize them all into two main classes.

Metric-learning is the term used to describe the process of learning a distance function over objects (Chen et al., 2021). Algorithms using metrics learn to compare data samples in general. They categorize query samples for an FSL classification problem according to how similar they are to the support samples. A base-learner and a meta-learner are created for the gradient-based technique. A base-learner is a model that is initialized and taught within each episode by the meta-learner, whereas a meta-learner is a model that learns across episodes (Cao et al., 2019). Table 1 shows the meta-learning algorithms with its category respectively.

One of the most well-liked meta-learning algorithms, Model-Agnostic Meta-Learning (MAML) has made significant advancements in the field of meta-learning research. Furthermore, meta-learning is a two-part learning process where the model first prepares to learn new tasks and then actually learns the new task from a few samples in the second stage (Yavari, 2020). The general goal of MAML is to identify suitable starting weight values so that the model can quickly pick up new tasks.

The first Metric-Learning method created to address FSL issues was Matching Networks (Chen et al., 2021). When using the Matching Networks approach to resolve an FSL job, a large base dataset is required (Li X. et al., 2020). This dataset is divided into episodes. Every image from the query and support sets is given to a CNN, which produces embeddings for each one. The SoftMax of the cosine distance between each query image's embeddings and the support-set embeddings is used to classify each image. It also uses Long Short-Term Memory (LSTM) for testing and Bidirectional Long Short-Term Memory (biLSTM) for training (Ye et al., 2020). The resulting classification's Cross-Entropy Loss is backpropagated through the CNN. Matching Networks can learn to construct picture embeddings in this way. Matching Networks can categorize photographs using this method without having any special prior knowledge of classes (Vinyals et al., 2016). Simply comparing several instances of the classes is used for everything.

Prototypical Networks learn a metric space in which each class may be classified by computing distances between prototype representations (Yavari, 2020). They represent a more straightforward inductive bias, which is advantageous in this low-data environment. The notion behind Prototypical Networks is that each class has an embedding in which points cluster around a single prototype representation (Pan et al., 2019). To accomplish so, a neural network is used to learn a non-linear mapping of the input into an embedding space, and a class's prototype is taken to be the mean of its embedding space support set. The next step is to classify an embedded query point by simply identifying the closest class prototype (Ji et al., 2020).

Another straightforward and efficient technique for both one-shot and FSL that adheres to the metric-learning paradigm is the Relation Network (Yavari, 2020). Relation Network uses an embedding function f to extract the features from the inputs, just like the Prototypical and Siamese Networks (Sung et al., 2018). As an embedding function, CNN is utilized because the inputs are images. The Relation Network includes a relation function in addition to the embedding function. This relation function evaluates how closely the query sample is connected to the various classes in the support set (Sun et al., 2018).

TABLE 1 Types of meta learning algorithms.

Networks	Category	Embedding function for testing	Embedding function for training
Model-agnostic meta learning (Wang and Wang, 2019)	Gradient-based meta learning	Deep neural network	Deep neural network
Matching network (Wang and Wang, 2019)	Metric learning	CNN, LSTM	CNN, biLSTM
Prototypical network (Yavari, 2020)	Metric learning	CNN	CNN
Relation network (Yavari, 2020)	Metric learning	CNN	CNN
Siamese network (Yavari, 2020)	Metric learning	CNN	CNN

A Siamese Network is made up of two networks that receive different inputs but are connected at the top by an energy function (Yavari, 2020). This function calculates a metric between each side's highest-level feature representation. The dimensions of the twin networks are inextricably linked (Dong and Shen, 2018). The network's structure is duplicated in both the top and bottom parts, resulting in twin networks with shared weight matrices at each layer. The Siamese Network's goal is to evaluate two photos and determine whether they are similar (Melekhov et al., 2016). The process for comparing a Siamese network is quite straightforward.

Other methods for small data training

There are also methods other than FSL for small data training. Knowledge may be successfully exploited in sophisticated applications by being appropriately organized and represented in knowledge graphs (Chen et al., 2020). Knowledge graphs can efficiently manage, mine, and organize knowledge from massive amounts of data, enhancing the quality of information services and delivering more intelligent services to users. It is used to store connected descriptions and details about things like events, objects, actual circumstances, and theoretical or abstract ideas (Pujara et al., 2013). This approach simultaneously encodes semantics underlying the data collection in addition to serving as data storage. The arrangement and structuring of significant data points from the data set to integrate information gathered from multiple sources is the main purpose of the knowledge graphs model. Labels are added to a knowledge graph to associate certain meanings. Nodes and edges are the two basic parts of a graph (Wang et al., 2017). Nodes are made up of two or more objects, and edges show how they are related to one another.

By using enough prior knowledge from related fields to complete new tasks in the target domain, transfer learning simulates the human visual system despite the size of the data set (Shao et al., 2014). In transfer learning, the source domain and the target domain are two different sorts of domains that can both benefit from the training and testing data (Weiss et al.,

2016). The source domain contains training examples, which are under a different distribution from the target domain data, and the target domain contains the testing instances, which are the classification system's task. For a transfer learning task, there is typically just one target domain, although there may be one or more source domains (Torrey and Shavlik, 2010).

Another popular method is self-supervised learning (SSL). SSL has becoming more common since it can save on the expense of annotating huge datasets (Jaiswal et al., 2020). It has the ability to employ the learnt representations for numerous downstream tasks and use self-defined pseudo labels as supervision. Particularly, contrastive learning has lately taken center stage in self-supervised learning for computer vision, natural language processing (NLP), and other fields. It attempts to push embeddings from various samples away while trying to embed augmented copies of the same sample closely together (Zhai et al., 2019).

Methodology

A literature search has been carried out by querying Google Scholar, IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library and Scopus. According to studies, Google Scholar's reach and coverage are greater than those of similar academic websites. Then, using exclusion criteria, a small number of databases is chosen to search for pertinent papers. The search was conducted using the terms "few shot learning", "smart agriculture", "object detection", "few shot classification", "one shot learning", "low shot learning", "Meta-Learning", "Prototypical Network", "Siamese Network", "MAML", "Relation Network", and "Matching Network".

The search was also limited to recent research papers in the years 2010 and above. Repeated things between search engine results were searched by employing a straightforward script based on the edit distance between article titles. Around 150 articles were returned by all the search engines with 58 articles in IEEE Xplore, 50 articles in ScienceDirect, 21 articles in Google Scholar, 15 articles in SpringerLink, and six articles in Scopus. The resulting papers were very diverse, and most of them fell outside the purview of the review. As a result, studies that had been published and that dealt with FSL in agriculture

and reported noteworthy findings considering recent literature have been chosen. The search was restricted to keywords, titles, and abstracts.

Exclusion criteria were used to screen out works that had nothing to do with the literature review. The following standards are disregarded in this analysis: (a) The ensuing articles have nothing to do with FSL in agriculture, (b) Articles that are not written in English, (c) Articles that are duplicated in the databases, (d) Articles that do not have significant results and problems. This filtered to 30 articles and during the process, references of the selected papers were also looked to find other works that fit the inclusion criteria. Finally, the resulting articles were organized by considering the models and types of objects in the agriculture field.

Results and discussion

FSL was first used in the field of computer vision, where it produced promising results in picture classification. It does not require large samples during training, which reduces the time and cost of obtaining samples. As a result, it is now being welcomed in many fields such as smart agriculture (Yang et al., 2022). FSL has shown compelling results in smart agriculture such as plant disease classification, plant leaf classification and plant counting. Early diagnosis of plant diseases is crucial for preventing plagues and mitigating their consequences on crops.

Deep learning is used to create the most accurate automatic algorithms for identifying plant diseases using plant field pictures. Traditional machine learning and deep learning techniques, on the other hand, have certain drawbacks, one of which is the requirement for large-scale datasets for models to generalize well (Wang et al., 2020). These methods necessitate the collecting and annotation of huge image datasets, which is usually impossible to achieve due to technological or financial constraints. As a result, new strategies were presented with the goal of developing models that generalized well with fewer data. As a result, meta-learning approaches (Hospedales et al., 2021) and FSL methods, were proposed. The Table 2 shows the results of the findings.

Research on identification has begun over two decades. Deep learning was used in these studies, and they were a huge success. However, there are a few challenges that deserve greater attention and research, including the demand for a big training set, training time, CNN architecture, pre-processing, feature extractions, and performance evaluation measures. Some of the experiments will not work in the wild due to camouflaging of the backgrounds (Liu et al., 2019; Kuzuhara et al., 2020; Yao et al., 2020).

To address these problems, authors have utilized machine learning in the identification of pests as this study makes the machine to think like humans. Machine learning has shown satisfactory results in most studies (Brunelli et al., 2019; Tachai et al., 2021; Zhang H. et al., 2021). However, machine learning also uses CNN feature extractions for training which shows improvement. Classifiers such as Support Vector Machine (SVM) is the best classifier that enhances the accuracy (Kasinathan et al., 2021).

To fill the gap between large training samples and time consumption, FSL was implemented. FSL has become popular in the computer vision field in both classification and detection. FSL was introduced very recently, and the results generated by this meta-learning approach are favorable. In the smart agriculture sector, FSL is widely used in plant disease identification (Wang and Wang, 2019; Li and Chao, 2021), plant counting (Karami et al., 2020), leaf classification (Afifi et al., 2020; Jadon, 2020; Tassis and Krohling, 2022) and fruit ripeness classification (Janarthan et al., 2020; Ng et al., 2022). The results for these experiments have exceeded 90%.

For plant leaf classification using the Siamese network, the accuracy of 95.32%, 91.37% and 91.75% were achieved for Flavia, Swedish and Leafsnap respectively (Wang and Wang, 2019). For maize plant counting, the overall precision was above 95% (Karami et al., 2020). For plant disease classification using triplet loss, the accuracy was 91.4% and using the Siamese network the accuracy was 90.6% (Argüeso et al., 2020). In vegetable disease recognition 99.48% accuracy was reached (Wang et al., 2021). For pest detection, have utilized a prototypical network and triplet loss with a terminal realization for cotton pests. They used two datasets for comparison and achieved 95.4 and 96.2% respectively (Li and Yang, 2020; Nuthalapati and Tunga, 2021).

However, some FSL techniques show lower accuracies due to their limitations. For fruit ripeness classification, they achieved 75% accuracy for only five samples (Ng et al., 2022). Another plant disease recognition has achieved an accuracy of around 55% for zero-shot (Zhong F. et al., 2020). For pests, one study achieved 60.3% (Nuthalapati and Tunga, 2021). For animal detection, the authors achieved 55.61 and 71.03% for 1 and 5 shots respectively due to no semantic annotations in testing (Zhang H. et al., 2021). Another bird detection study achieved 60.19 and 75.75% for 1 and 5 shots respectively (Liu et al., 2019).

Even though some research has lower accuracies, they are still satisfactory because FSL only uses very less instances in the training of 1, 2, 3, 5, and 10 images for detection and classification (Zhong F. et al., 2020). This proves that FSL is giving a good result with very less training samples compared to CNN which requires more training samples and training time.

Main findings and ways forward

Deep learning and machine learning is the process of utilizing algorithms to direct a computer in building a suitable model based on existing data and then using that model to evaluate novel situations. The models' applicability are significantly constrained by the massive amounts of labeled data that are often needed to train it because it has multiple parameters. Large amounts of tagged data are frequently

TABLE 2 Studies of few-shot learning in smart agriculture.

References	Data source	Techniques	Accuracy	Purpose	Limitation
Wang and Wang (2019)	Flavia Dataset; Swedish Dataset; Leafsnap Dataset	Siamese Network; Simultaneous Two-way Convolutional Neural Network; Spatial Structure Optimizer (SSO); K-Nearest Neighbor (KNN)	Flavia = 95.32%; Swedish = 91.37%; Leafsnap = 91.75%	To handle a leaf classification problem with a small sample size using Siamese Network	Has to improve the generalization ability of the model
Karami et al. (2020)	Maize Plant Dataset	Modified CenterNet Architecture	>95%	To identify and count maize plant by using a smaller number of images in a field using FSL	Adapting generative adversarial network (GAN) in future
Ng et al. (2022)	Fruit Dataset	Meta-learning paradigm; Few-shot classification	>75%	To classify new varieties of fruits with only a few training examples	Not stated
Argüeso et al. (2020)	PlantVillage Dataset	CNN; SVM Classifier; Siamese Network; Few-shot and baseline transfer learning; Root Mean Square optimizer (RMSprop)	Triplet Loss; Accuracy = 91.4%; Precision = 93.8%; Recall = 92.6%; F-Score = 92.3%; Siamese; Accuracy = 90.6%; Precision = 93.8%; Recall = 93.1%; F-Score = 91.2%	To classify plant diseases to overcome unfeasibility of deep learning	Convergence problems
Zhong F. et al. (2020)	CUB Dataset; AWA2 Dataset; APY Dataset	Conditional Adversarial Autoencoders (CAAE); Kullback-Leibler (KL) Divergence Penalty	Zero Shot; Unseen = 55.4%; Seen = 51.6%; Harmonic = 53.4%	Recognition of Citrus aurantium L. diseases using few and zero shot learning	Only uses a discriminator to replace KL-divergence in variational autoencoders
Li and Chao (2021)	PlantVillage Dataset	Semi-supervised few-shot learning	5-shot = 90%; 10-shot = 92.6%	To resolve plant leaf disease recognition problem using semi-supervised FSL	Did not consider wrong labels
Wang et al. (2021)	Self-Collected Dataset	Image text collaborative representation learning (ITC-Net)	Accuracy = 99.48%; Precision = 98.90%; Sensitivity = 98.78%; Specificity = 99.66%	Recognition of vegetable illness in complicated backgrounds	The feature spaces of image and text are still independent
Jadon (2020)	Mini-Leaves Diseases Dataset; Sugarcane Dataset	Stacked Siamese Matching (SSM) Net	Mini Leaves = 92.7%; Sugarcane = 94.3%; Comparison with VGG16; Mini Leaves = 0.91; Sugarcane = 0.90	Mini-leaves diseases and sugarcane illness identification using combined methods of Siamese and Matching networks	Not stated
Hu et al. (2019)	Tea Leaf Disease Image Dataset	Conditional deep convolutional generative adversarial networks (C-DCGAN); Support vector machine (SVM)	90%	Tea leaf disease identification with low shot learning	Low generalization performance
Janarthan et al. (2020)	Citrus Fruits and Leaves Dataset	Deep Siamese Network	95.04%	To create a classification strategy for diagnosing citrus diseases	Not Stated
Afifi et al. (2020)	PlantVillage Dataset; Coffee Leaf Dataset	Triplet network; Deep adversarial metric learning (DAML)	99%	Automatic identification of plant diseases with limited data	Focusing on the full image instead of focusing on the affected part

(Continued)

TABLE 2 (Continued)

References	Data source	Techniques	Accuracy	Purpose	Limitation
Tassis and Krohling (2022)	Leaf Dataset; Symptoms Dataset	Prototypical networks; Triplet networks; feature extraction	1st Experiment = 96.03%; 2nd Experiment = 96.72%; 3rd Experiment = 93.25%	FSL in classification and severity assessment tasks with biotic stressors in coffee leaves	During training, there is a lack of performance evaluation of the approaches in unobserved classrooms
Li and Yang (2020)	National Bureau of Agricultural Insect Resources (NBAIR) Dataset; Li's dataset of natural scenes	Prototypical Network; CNN; Triplet loss; Terminal realization	95.4% (National Bureau of Agricultural Insect Resources dataset); 96.2% (Li's dataset of natural scenes)	Few shot cotton pest recognition and terminal realization	Not stated
Nuthalapati and Tunga (2021)	Plant and Pests (PP) dataset; PlantVillage dataset; Plants in Wild (PiW) dataset; Cross dataset	Few-shot learning using transformer	Plant and Pests = 60.3%; PlantVillage = 88.5%; Plants in Wild = 93.3%; Cross = 86.1%	To classify diverse pests, plants, and illness	Not stated
Li and Yang (2021)	PlantVillage Dataset	Few-shot classification; CNN	Single domain,; Ntrain = Ntest = 3 = 90.4%,; Ntrain = Ntest = 5 = 82.5%.; Cross domain 1,; Ntrain = Ntest = 3 = 86.5%,; Ntrain = Ntest = 5 = 81%.; Cross domain 2,; Ntrain = Ntest = 3 is 54.8%,; Ntrain = Ntest = 5 is 43.9%.	Task-driven meta-learning few shot classification for pests and plants and to investigate single domain and cross-domain classification	Not stated
Zhang H. et al. (2021)	MiniImagenet Dataset; CUB-200-2011 Dataset; Flower102 Dataset	Absolute-relative Learning	1-shot = 55.61%; 5-shot = 71.03%	Absolute-relative learning paradigm in supervised and unsupervised FSL	Only used semantic annotations as labels in training, and do not use them during testing
Liu et al. (2019)	ImageNet Dataset	Few-Shot Unsupervised Image-to-Image Translation (FUNIT); Generative Adversarial Networks (GAN)	Animal Face; 1-shot = 82.36%; 5-shot = 96.05%; North American Birds; 1-shot = 60.19%; 5-shot = 75.75%	Few-shot unsupervised image-to-image translation of animal faces, birds, flowers, and foods	FUNIT fails to achieve translation when the look of novel object classes differs considerably from that of the source classes

exceedingly expensive, challenging, or even impossible to obtain. FSL is a crucial component of the development of machine learning since it is a crucial question if a good model can be created by training with only a minimal amount of labeled data. It may be applied to learn uncommon situations using a limited collection of characteristics, which is a method that can successfully address the issue of dataset imbalance in intrusion detection.

In recent years detection and classification in the agricultural domain has received a lot of attention. Beginning with deep learning and machine learning, they are showing excellent results in identifying the objects. This approach is very useful for limited datasets and time for detection as well as classification of various things including plant diseases and pests. Most of the authors used the current state of the art in FSL such as the Siamese network, Prototypical network, and Model-Agnostic Meta Learning network.

Most of the FSL approaches show favorable results of more than 90%. One research has reached 99.48% accuracy in vegetable disease detection using a self-collected dataset. Even though the accuracy is very excellent in this research, the authors stated that the feature spaces of image and text are still independent, and it needs to be improved in the future. Triplet network and deep adversarial metric learning have also given accuracies of 99% with full coffee leaf images. Besides FSL, one shot learning has also shown a very satisfying result of 95.32% using the Siamese network. The low generalization ability of the model is an issue in this research and the gradient descent in the training must be stopped at the appropriate point as well as adding a regularization term to the error function to smooth out the mappings.

Most of the results of plant datasets have higher accuracies than animal datasets. However, pest detection using a prototypical network has achieved 96.2% accuracy. Another pest detection only achieved 60.3% due to camouflaging images

in the background which can be improved by focusing on the insects only. The least accuracies that have been obtained are with animal detection, 55.61% for one shot. However, when five shots were used, the accuracy increased to 71.03%. This proves that when more instances are used, the accuracy increases. Despite that, Few-Shot Unsupervised Image-to-Image Translation (FUNIT) has produced an accuracy of 82.36% for one shot and 96.05% for five shots using animal faces. For birds, an accuracy of 60.19% for one shot and 75.75% for five shots was obtained.

According to studies reviewed here, FSL can significantly boost the detection performance on small datasets. The popularity of FSL for smart farming applications is increasing rapidly due to the straightforward implementation yet highly promising results obtained in agricultural applications.

Author contributions

NR: writing—original draft and investigation. JT: writing—review and editing, supervision, and conceptualization.

All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Afifi, A., Alhumam, A., and Abdelwahab, A. (2020). Convolutional neural network for automatic identification of plant diseases with limited data. *Plants.* 10, 28. doi: 10.3390/plants10010028

Alajaji, D., Alhichri, H. S., Ammour, N., and Alajlan, N. (2020). "Few-shot learning for remote sensing scene classification," in *2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS).* IEEE. p. 81–84. doi: 10.1109/M2GARSS47143.2020.9105154

Altae-Tran, H., Ramsundar, B., Pappu, A. S., and Pande, V. (2017). Low data drug discovery with one-shot learning. *ACS Central Sci.* 3, 283–293. doi: 10.1021/acscentsci.6b00367

Argüeso, D., Picon, A., Irusta, U., Medela, A., San-Emeterio, M. G., Bereciartua, A., et al. (2020). Few-Shot Learning approach for plant disease classification using images taken in the field. *Comput. Elect. Agri.* 175, 105542. doi: 10.1016/j.compag.2020.105542

Ashtiani, S. H. M., Javanmardi, S., Jahanbanifard, M., Martynenko, A., and Verbeek, F. J. (2021). Detection of mulberry ripeness stages using deep learning models. *IEEE Access.* 9, 100380–100394. doi: 10.1109/ACCESS.2021.30 96550

Brunelli, D., Albanese, A., d'Acunto, D., and Nardello, M. (2019). Energy neutral machine learning based iot device for pest detection in precision agriculture. *IEEE Int. Things Magaz.* 2, 10–13. doi: 10.1109/IOTM.0001.1900037

Cao, Z., Zhang, T., Diao, W., Zhang, Y., Lyu, X., Fu, K., et al. (2019). Meta-seg: A generalized meta-learning framework for multi-class few-shot semantic segmentation. *IEEE Access.* 7, 166109–166121. doi: 10.1109/ACCESS.2019.29 53465

Chen, W. Y., Liu, Y. C., Kira, Z., Wang, Y. C. F., and Huang, J. B. (2019). A closer look at few-shot classification. *arXiv.*

Chen, X., Jia, S., and Xiang, Y. (2020). A review: knowledge reasoning over knowledge graph. *Expert Syst. Appl.* 141, 112948. doi: 10.1016/j.eswa.2019.112948

Chen, Y., Liu, Z., Xu, H., Darrell, T., and Wang, X. (2021). "Meta-baseline: Exploring simple meta-learning for few-shot learning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (Montreal, QC: IEEE), 9062–9071. doi: 10.1109/ICCV48922.2021.00893

Dong, X., and Shen, J. (2018). "Triplet loss in iamese network for object tracking," in *Proceedings of the European conference on computer vision (ECCV)* (p. 459–474).

Duan, R., Li, D., Tong, Q., Yang, T., Liu, X., and Liu, X. (2021). A survey of few-shot learning: an effective method for intrusion detection. *Sec. Communi. Net.* 2021. doi: 10.1155/2021/4259629

Fei-Fei, L., Fergus, R., and Perona, P. (2006). One-shot learning of object categories. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 594–611. doi: 10.1109/TPAMI.2006.79

Fink, M. (2004). Object classification from a single example utilizing class relevance metrics. *Adv. Neural Inf. Process. Syst.* 17.

Gidaris, S., and Komodakis, N. (2018). "Dynamic few-shot visual learning without forgetting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT: IEEE), 4367–4375.

Hospedales, T., Antoniou, A., Micaelli, P., and Storkey, A. (2021). Meta-learning in neural networks: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 5149–5169. doi: 10.1109/TPAMI.2021.3079209

Hu, G., Wu, H., Zhang, Y., and Wan, M. (2019). A low shot learning method for tea leaf's disease identification. *Comput. Electron. Agric.* 163, 104852. doi: 10.1016/j.compag.2019.104852

Jadon, S. (2020). "SSM-net for plants disease identification in low data regime," in *2020 IEEE/ITU International Conference on Artificial Intelligence for Good* (Geneva: IEEE), 158–163.

Jaiswal, A., Babu, A. R., Zadeh, M. Z., Banerjee, D., and Makedon, F. (2020). A survey on contrastive self-supervised learning. *Technologies.* 9, 2. doi: 10.48550/arXiv.2011.00362

Janarthan, S., Thuseethan, S., Rajasegarar, S., Lyu, Q., Zheng, Y., and Yearwood, J. (2020). Deep metric learning based citrus disease classification with sparse data. *IEEE Access.* 8, 162588–162600. doi: 10.1109/ACCESS.2020.3021487

Ji, Z., Chai, X., Yu, Y., Pang, Y., and Zhang, Z. (2020). Improved prototypical networks for few-shot learning. *Pattern Recognit. Lett.* 140, 81–87. doi: 10.1016/j.patrec.2020.07.015

Jiménez-Luna, J., Grisoni, F., and Schneider, G. (2020). Drug discovery with explainable artificial intelligence. *Nat. Machine Intelli.* 2, 573–584. doi: 10.1038/s42256-020-00236-4

Karami, A., Crawford, M., and Delp, E. J. (2020). "Automatic plant counting and location based on a few-shot learning technique," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* p. 5872–5886.

Kasinathan, T., Singaraju, D., and Uyyala, S. R. (2021). "Insect classification and detection in field crops using modern machine learning techniques," in *Information Processing in Agriculture* p. 446–457.

Khodak, M., Balcan, M. F. F., and Talwalkar, A. S. (2019). Adaptive gradient-based meta-learning methods. *Adv. Neural Inf. Process. Syst.* 32. doi: 10.48550/arXiv.1906.02717

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Commun. ACM.* 60, 84–90. doi: 10.1145/3065386

Kuzuhara, H., Takimoto, H., Sato, Y., and Kanagawa, A. (2020). "Insect Pest Detection and Identification Method Based on Deep Learning for Realizing a Pest Control System," in *2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. p. 709-714.

Lake, B., Salakhutdinov, R., Gross, J., and Tenenbaum, J. (2011). "One shot learning of simple visual concepts," in *Proceedings of the annual meeting of the cognitive science society*. p. 33.

Lake, B. M., Salakhutdinov, R., and Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*. 350, 1332–1338. doi: 10.1126/science.aab3050

Lee, S. H., Chan, C. S., Mayo, S. J., and Remagnino, P. (2017). How deep learning extracts and learns leaf features for plant classification. *Pattern Recognit.* 71, 1–13. doi: 10.1016/j.patcog.2017.05.015

Li, X., Pu, F., Yang, R., Gui, R., and Xu, X. (2020). AMN: Attention metric network for one-shot remote sensing image scene classification. *Remote Sens.* 12, 4046. doi: 10.3390/rs12244046

Li, Y., and Chao, X. (2021). Semi-supervised few-shot learning approach for plant diseases recognition. *Plant Methods* 17, 1–10. doi: 10.1186/s13007-021-00770-1

Li, Y., and Yang, J. (2020). Few-shot cotton pest recognition and terminal realization. *Comput. Electron. Agric.* 169, 105240. doi: 10.1016/j.compag.2020.105240

Li, Y., and Yang, J. (2021). Meta-learning baselines and database for few-shot classification in agriculture. *Comput. Electron. Agric.* 182, 106055. doi: 10.1016/j.compag.2021.106055

Li, Y., Zhang, L., Wei, W., and Zhang, Y. (2020). "Deep self-supervised learning for few-shot hyperspectral image classification," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE.* p. 501–504.

Liu, J., and Wang, X. (2021). Plant diseases and pests detection based on deep learning: a review. *Plant Methods.* 17, 1–18. doi: 10.1186/s13007-021-00722-9

Liu, L., Wang, R., Xie, C., Yang, P., Wang, F., Sudirman, S., et al. (2019). PestNet: An end-to-end deep learning approach for large-scale multi-class pest detection and classification. *IEEE Access.* 7, 45301–45312. doi: 10.1109/ACCESS.2019.2909522

Melekhov, I., Kannala, J., and Rahtu, E. (2016). "Siamese network features for image matching," in *2016 23rd international conference on pattern recognition (ICPR)*. IEEE. p. 378–383

Muthukumar, N. (2021). "Few-shot learning text classification in federated environments," in *2021 Smart Technologies, Communication and Robotics (STCR)*. p. 1-3.

Ng, H. F., Lo, J. J., Lin, C. Y., Tan, H. K., Chuah, J. H., and Leung, K. H. (2022). "Fruit ripeness classification with few-shot learning," in *Proceedings of the 11th International Conference on Robotics, Vision, Signal Processing and Power Applications*. Springer, Singapore. p. 715–720.

Nuthalapati, S. V., and Tunga, A. (2021). "Multi-domain few-shot learning and dataset for agricultural applications," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. p. 1399–1408.

Pan, Y., Yao, T., Li, Y., Wang, Y., Ngo, C. W., and Mei, T. (2019). "Transferrable prototypical networks for unsupervised domain adaptation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. p. 2239–2247.

Parnami, A., and Lee, M. (2022). Learning from few examples: a summary of approaches to few-shot learning. *arXiv*. doi: 10.48550/arXiv.2203.04291

Pujara, J., Miao, H., Getoor, L., and Cohen, W. (2013). Knowledge graph identification. In International semantic web conference. Berlin, Heidelberg: Springer. p. 542–557

Shao, L., Zhu, F., and Li, X. (2014). Transfer learning for visual categorization: a survey. *IEEE*. 26, 1019–1034. doi: 10.1109/TNNLS.2014.2330900

Sun, C., Shrivastava, A., Vondrick, C., Murphy, K., Sukthankar, R., and Schmid, C. (2018). "Actor-centric relation network," in *Proceedings of the European Conference on Computer Vision (ECCV)* p. 318–334.

Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P. H., and Hospedales, T. M. (2018). "Learning to compare: Relation network for few-shot learning," in *Proceedings of the IEEE conference on computer vision and pattern recognition (pp.* 1199–1208).

Tachai, N., Yato, P., Muangpan, T., Srijiranon, K., and Eiamkanitchat, N. (2021). "KaleCare: Smart Farm for Kale with Pests Detection System using Machine Learning," in *2021 16th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)* p. 1–6.

Tassis, L. M., and Krohling, R. A. (2022). Few-shot learning for biotic stress classification of coffee leaves. *Artif. Intell.* 6, 57–67. doi: 10.1016/j.aiia.2022.04.001

Too, E. C., Li, Y., Kwao, P., Njuki, S., Mosomi, M. E., and Kibet, J. (2019). Deep pruned nets for efficient image-based plants disease classification. *J. Intell. Fuzzy Syst.* 37, 4003–4019. doi: 10.3233/JIFS-190184

Torrey, L., and Shavlik, J. (2010). "Transfer learning," in *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques* (IGI Global), 242–264.

Vinyals, O., Blundell, C., Lillicrap, T., and Wierstra, D. (2016). Matching networks for one shot learning. *Adv. Neural Inf. Process. Syst.* 29.

Wang, B., and Wang, D. (2019). Plant leaves classification: a few-shot learning method based on siamese network. *IEEE Access.* 7, 151754–151763. doi: 10.1109/ACCESS.2019.2947510

Wang, C., Zhou, J., Zhao, C., Li, J., Teng, G., and Wu, H. (2021). Few-shot vegetable disease recognition model based on image text collaborative representation learning. *Comp. Electronics Agri.* 184, 106098. doi: 10.1016/j.compag.2021.106098

Wang, Q., Mao, Z., Wang, B., and Guo, L. (2017). Knowledge graph embedding: a survey of approaches and applications. *IEEE Trans. Knowl. Data Eng.* 29, 2724–2743.

Wang, Y., Yao, Q., Kwok, J. T., and Ni, L. M. (2020). Generalizing from a few examples: a survey on few-shot learning. *ACM Comp. Surv.* 53, 1–34. doi: 10.1145/3386252

Weiss, K., Khoshgoftaar, T. M., and Wang, D. (2016). A survey of transfer learning. *J. Big data* 3, 1–40. doi: 10.1186/s40537-016-0043-6

Yang, J., Guo, X., Li, Y., Marinello, F., Ercisli, S., and Zhang, Z. (2022). A survey of few-shot learning in smart agriculture: developments, applications, and challenges. *Plant Method.* 18, 1–12. doi: 10.1186/s13007-022-00866-2

Yang, L., and Jin, R. (2006). Distance metric learning: a comprehensive survey. *Michigan State Uni.* 2, 4.

Yao, Y., Zhang, Y., and Nie, W. (2020). "Pest Detection in Crop Images Based on OTSU Algorithm and Deep Convolutional Neural Network," in *2020 International Conference on Virtual Reality and Intelligent Systems (ICVRIS)*. IEEE. p. 442–445

Yavari, N. (2020). *Few-Shot Learning with Deep Neural Networks for Visual Quality Control: Evaluations on a Production Line.*

Ye, H. J., Hu, H., Zhan, D. C., and Sha, F. (2020). "Few-shot learning via embedding adaptation with set-to-set functions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Seattle, WA: IEEE), 8808–8817.

Zhai, X., Oliver, A., Kolesnikov, A., and Beyer, L. (2019). "S4l: Self-supervised semi-supervised learning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. p. 1476–1485.

Zhang, H., Koniusz, P., Jian, S., Li, H., and Torr, P. H. (2021). "Rethinking class relations: Absolute-relative supervised and unsupervised few-shot learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Nashville, TN: IEEE), 9432–9441.

Zhang, S., Zhu, J., and Li, N. (2021). "Agricultural Pest Detection System Based on Machine Learning," in *2021 IEEE 4th International Conference on Electronics Technology (ICET)*. IEEE. p. 1192-1196.

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. *Comput Electron Agric.* 179, 105828. doi: 10.1016/j.compag.2020.105828

Zhong, Q., Chen, L., and Qian, Y. (2020). "Few-shot learning for remote sensing image retrieval with maml," in *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE. p. 2446–2450.