Deep Learning and Applications in Object Classification

Student1: Xiaotian Wang(2020180015), Student2: Weidong Zhao(2020124075),
Student3: Yibao Miao(2020124126) and Student4 Xin Gao(2020124148)

May 24, 2021

1 Introduction

With the increasing enthusiasm for deep learning, various deep learning frameworks and network models have been proposed. Deep learning has a good performance in many fields, especially in computer vision. It has achieved excellent performance in image classification, target detection, semantic segmentation and other tasks.

However, the training quality of deep neural networks is strongly affected by the number of labeled samples in the training set. In practical problems, it is extremely expensive to obtain tens of thousands of labeled training samples, but it is very easy to obtain unlabeled samples. People use a variety of ways to study the use of unlabeled samples. Among them, semi supervised learning based on a small number of labeled samples and a large number of unlabeled samples is more suitable for real-world applications, and has recently become a hot new direction in the field of deep learning.

Semi supervised learning is a powerful method to train a large number of data without a large number of tags. Semi supervised learning reduces the need for labeled data by providing a way to utilize unlabeled data. Since unlabeled data can usually be obtained with minimal manpower, any performance improvement brought by semi supervised learning is usually low-cost. People's enthusiasm for semi supervised learning is also growing, and a large number of semi supervised learning methods based on deep learning framework are proposed.

In semi supervised learning, there is a basic assumption that it is beneficial to use unlabeled data samples in the training process. It can make less marked training more robust, and even outperform supervised learning in some cases.

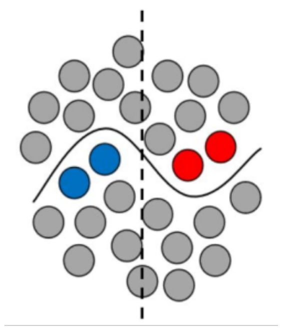

Figure 1: Schematic diagram of semi-supervision function

Figure 1 briefly illustrates the benefits of using unlabeled data in semi supervised learning. Red and blue circles represent different categories of labeled data points. Gray circles represent unmarked data points. If only a small number of labeled samples are available, only the potential true distribution (solid line) can be assumed (dotted line). This true distribution can only be determined when unlabeled data points are considered and decision boundaries are clarified.

It is generally believed that semi supervised learning was proposed by Shah Shahani and landgrebe in 1994. Semi supervised learning has developed rapidly in the late 20th century and the early 21st century with the emergence of great demand for the use of unmarked data in practical applications. In a short time, the representative work of semi supervised learning, such as bifurcation based method, semi supervised support vector machine and graph semi supervised learning, has made great achievements. However, in this paper, we only focus on the semi supervised learning method based on deep neural network.

Hinton et al. Put forward the concept of ”deep learning” in 2006. Deep learning maps low-level features to high-level features through deep neural network (DNN), and abstracts features through high-level, so as to discover the distributed feature representation of data. After more than ten years of continuous development, deep learning has made great breakthroughs in various fields. Deep network also has more and more powerful function expression ability and feature extraction ability.

Semi supervised learning based on deep learning framework is a new research idea in recent years, which can be understood as a deep learning algorithm used in the training data mixed with a small amount of labeled data and a large number of unlabeled data. Through several years of continuous research, semi supervised learning algorithms based on deep learning framework have been developed into three categories. It includes pre training the network with unlabeled data and fine-tuning the network with marked data; The labeled data is used to train the network, and the depth feature obtained from the network is used to do semi supervised algorithm; End to end semi supervised depth model.

The first two methods use labeled samples and unlabeled samples, but for the deep neural network itself, its training still works in the way of supervised learning. Only the third method is semi supervised learning. This paper will also focus on this topic.

The rest of the paper is structured as follows. The second chapter introduces the background theory of semi supervised algorithm. The third chapter is the application of semi supervised learning. The last chapter draws a conclusion.

2 Background theory

Characteristics and harmfulness of agricultural diseases and insect pests Crop pest mainly refers to crop pests and crop diseases in the growth of crops. According to their performance, there are mainly the following characteristics: first, chewing mouth pests. Chewing mouth insect mainly refers to the moth eating the leaves and stem of crops, which leads to the hollows and bite marks of different degrees of crops. For example, if the chewing mouth insect is encountered in corn planting, the production will be reduced to different degrees; Second, the stinger mouth insect. The main purpose of this paper is to introduce the phenomenon that the insect can absorb the juice of crops through the bursal prickles, which leads to the phenomenon of withering leaves and dead branches of crops.

Meanwhile, the insect pests may carry pathogens, bring infectious diseases to crops, reduce crop yield and increase the difficulty of pest control; Third, the plant diseases and insect pests of strain. The bacterial crop diseases and insect pests mainly refer to the diseases and insect pests caused by soil, water source, plant seedlings and other factors, such as the lack of water and light, etc., which are common in rice planting and potato planting. Agricultural diseases and insect pests have a certain harm to agricultural production and crop quality, the specific performance is as follows: first, reduce the yield of crops. The most direct impact of diseases and insect pests is the growth of crops. In serious cases, dead branches and dead branches will appear, which will lead to no yield and no harvest; Second, it hinders the steady growth of agricultural economy. The output of crops has a direct impact on the steady growth of agricultural economy. The reduction of output will affect the income of farmers, but also cause local crop demand tension, pull up prices and hinder the steady improvement of agricultural economy; Third, improper prevention and control will affect the sowing of crops in the next year. Improper prevention and control of agricultural diseases and insect pests will seriously affect the second year sowing of crops. For example, improper use of some pesticides will directly cause land pollution, resulting in the land can not be sown for several consecutive years.

The image classification methods based on semi supervised deep learning can be divided into four categories: multi view training, consistent regularization, diverse hybrid training, and semi supervised learning combined with Gan network. Next, we will introduce each category, and give typical network models for analysis and comparison.

2.1 Multi view training

Suppose that each data can be classified from different views, and then use these classifiers trained from different perspectives to classify the unlabeled samples, and then select the credible unlabeled samples and their pseudo tags to join the training set. In this case, the goal of multi view training is to learn a unique prediction function to model the data in each view, and jointly optimize all the functions used to improve the generalization performance. Ideally, the results of different views complement each other, and different models can cooperate with each other to improve each other's performance.

2.2 Uniform regularization

Because in semi supervised learning, unlabeled data does not have its own category tags. Therefore, a consistent regularization method without marking information is developed. The main idea of consistent regularization is that for an input, even if it is slightly disturbed, its prediction should be consistent. In the specific semi supervised algorithm, the most basic idea is to use the standard cross entropy loss function to calculate the error for the labeled samples, while for the unlabeled samples, we need to use the consistency regularization, such as the minimum mean square error or KL divergence. The core idea of each algorithm is unchanged, That is to minimize the distance between unlabeled data and its disturbed output, but there are many changes in the form of calculated output. Finally, we need to add the weight coefficient, combined with the supervised and unsupervised two parts of the loss, so as to build the overall semi supervised loss function.

2.3 Multiple hybrid methods

The hybrid method tries to integrate the ideas of semi supervised learning in a framework, so as to get better performance in task recognition.

2.4 Esemi supervised Gan network

Generative adversarial networks (GAN) is a depth generation model proposed by goodflow et al. [30] in 2014. It can be widely used in the field of image vision, and has outstanding effects in the field of image, such as generating high-resolution realistic images, image restoration, wind migration, etc. When there are only a few labeled samples, it is easy to combine the idea of semi supervised learning with Gan network, and achieve good experimental performance in image classification.

3 Applications

3.1 Disease and its image features

When a lesion occurs in a certain part, the appearance of the organ is generally more significant. Apple's common diseases and their representative diseases are shown in the picture. Some lesions, a single lesion is relatively small, but there will be a large area of lesion distribution, such as bitter pox and scab; Some lesions, generally only a single lesion, but the lesion area is large, its identification needs to collect complete lesions, such as ringworm disease; For example, anthracnose appears in the form of star shaped spots in the early stage, ulcers in the middle stage, and ulcers in the late stage.

Plant protection experts often classify it as mild, moderate and severe, such as Phytophthora rot; Some of the lesions are regular in shape and round in shape, such as round lines and Phytophthora. The round surface with prominent center is like anthrax, and the round hole is like black star and bitter pox; Some lesions were irregular cloud like in shape and area, such as rust fruit disease; Lesions often show different colors. Scab shows dark green on green fruits and dark red on ripe fruits.

3.2 Data extraction

Crop disease image is the abnormal visual and optical performance of crop infected parts. In plant diseases.In physics, the abnormal performance of the host plant itself is called the symptom, and the characteristics of the pathogen are called the symptom.

Specifically, the common diseases are moldy, powdery, punctate, purulent, particulate matter.It is characterized by degeneration, erosion, spots, deformity and wilting. In the image information collection, recording, transmission conditions do not have and general.

At the same time, plant protection experts inherited and exchanged the achievements of civilization in this field through the industrial words. What is ”mildew” and what is ”color change” are generally recognized by industry experts,They can achieve the transformation between the literal symbols and the expressed phenomena at will, and according to the visual performance of crop diseases,correct diagnosis of disease types, and take reasonable plant protection measures.

If we can use the computer to automatically extract the lesion features from the crop image, and then let the plant through the learning process,If the machine learns the domain knowledge of plant protection experts, then the computer can complete crop diseases like human experts.

The nursing early warning response measures were given. And its working speed, uninterrupted working time, service life. The advantages of low cost are far better than that of human beings.

3.3 Principal component analysis

Whether the principal component method can be used for feature extraction of crop disease image with the object of disease pattern recognition is a problem worthy of study.

The discrimination of Apple mold by near infrared diffuse reflectance spectroscopy was discussed.

In order to evaluate the feasibility of heart disease, principal component analysis (PCA) was used to extract spectral features of heart disease Fisher discriminant function, the correct rate is 87 percentage.

Li Guifeng explored the method of rapid and non-destructive detection of Apple texture quality by near-infrared spectroscopy. There 240 apple samples were collected by near-infrared spectroscopy, and the damage characteristics in Apple texture map were successfully extracted by principal component method. Relevant research shows that it is feasible to apply PCA method of linear dimension reduction to feature extraction of crop disease image data in practice.

3.3.1 Mean and covariance

Mean value is also called mathematical expectation, which describes the average level of a sample space on a feature. It can be seen that mathematical expectation reflects the characteristics of a certain data set, and for a certain distribution, it is the statistical characteristics of distribution, belonging to the characteristics of set level.

$$E[X]=\sum_{i=1}^{\infty} x_i p_i$$

$$E[X]=\int_{-\infty}^{\infty} x f(x)\mathrm{d}x$$

Figure 2: Formula 1 corresponds to discrete random variables, and formula 2 corresponds to continuous random variables

Standard deviation, also known as deviation, is the arithmetic square root of variance. In a certain feature of the sample, it describes the average level of mean deviation, reflects the stability of the sample, and is also defined on the data set to describe the statistical characteristics of the sample set or distribution.

$$\sigma=\sqrt{E[(X-\mu)^2]}=\sqrt{E[X^2]-(E[X])^2}$$

Figure 3: The calculation formula of standard deviation is shown in the figure

3.3.2 Eigenvalues and eigenvectors

The eigenvalue of the matrix reflects the effect of the matrix. As shown in, when the effect of the matrix multiplied by a vector on the left and the number of vectors are equivalent to a scalar, the scalar is called the eigenvalue of the matrix, and the vector is the corresponding eigenvector of the eigenvalue. It is widely used in matrix analysis and transformation.

The eigenvalues are arranged from the highest to the lowest, which gives the weight order and importance order of the components, just like the decimal integer, from the highest to the individual bit. In this way, under the premise of a certain accuracy, we can abandon some low-order and low-order components in exchange for simple expression and description. It can be seen on the sorted eigenvector table. If we abandon some lower order eigenvectors, we will abandon some lower order eigenvectors, and the remaining vectors are called eigenvectors. In the covariance matrix of samples, the eigenvalues show the correlation strength between sample attributes, which provides useful information for data correlation analysis.

$$\mathbf{A}_{recon}' = \mathbf{A}_{minusmean} \times \mathbf{A}_{PCA} \times \mathbf{A}_{PCA}^{T}$$

$$\mathbf{A}_{recon} = \mathbf{A}_{recon}'^{T}(i,:) + mean_{A}(i,:)$$

Figure 4:

Dimensionality reduction is based on matrix calculation, and it does not need to iterate repeatedly to pursue convergence. Therefore, its complexity is mainly memory complexity. The covariance matrix and the eigenvalue vector of the matrix are obtained by matrix calculation. The speed of this process is affected by the dimension and number of samples. Under the general computing environment, it generally has a fast response speed, so it is often able to get the reduced dimension data set quickly. The classification model of support vector machine is essentially an optimization problem, and its complexity only depends on the number of constraints and training samples.

3.4 Dimension reduction and reconstruction of PCA

Arrange the eigenvalues from high to low to get the weight order (importance order) of the components.

From top to bottom. In this way, on the premise of sacrificing a certain precision, we can discard some low-order and low-order components in exchange for simple sample expression and description. After discarding, the remaining vectors are called eigenvectors. How to get a better balance between the precision of restoration and the complexity of computation by discarding the components.

In general, the sum of correlations selected from large to small is not less than one percent of the total correlation.

In this chapter, a feature extraction method based on eigenvalue overflow discarding is proposed, and an error distance analysis method is designed for the reconstruction evaluation of dimensionality reduction performance. Dimensionality reduction experiments are carried out on Apple lesion image and ORL benchmark data set. The results show that PCA feature extraction with 95 percentage, threshold has better reconstruction performance; In the second stage of the experiment, the disease pattern feature set extracted from the first stage experiment is used as the recognition network to carry out the cross validation recognition experiment on the characteristic disease pattern. The results show that the minimum accuracy can reach 97 percentage under the conditions of 2 percentage and 5 percentage discount and appropriate network parameters, and the algorithm shows quite optimistic usability; Under the premise of reasonable selection of training parameters, the PCA dimension reduction pattern is effective.

$$(\sum_{i=1}^{x} \lambda_i)/(\sum_{i=1}^{\infty} \lambda_i) \ge \tau$$

Figure 5:

3.5 Semi supervision method

The principle of Boltzmann machine is based on statistical physics, which is a kind of modeling from energy function. Energy function expresses the statistical relationship between attributes through Boltzmann distribution.

3.5.1 Maxwell Boltzmann distribution

Boltzmann distribution, also known as Maxwell Boltzmann distribution, is a probability distribution, which is widely used in physics. The velocity of any single particle is constantly changing due to the collision with other particles. The collision time is happening, and the particle energy is also changing. In a certain equilibrium state, the percentage of particles whose velocity is in a certain range remains relatively stable. In other words, the probability of particles in a certain energy state obeys a certain distribution. If the system is in or close to equilibrium, such as constant temperature process, the Maxwell Boltzmann distribution specifies this proportion.

$$P(E_i) = \frac{N_i}{N} = \frac{g_i e^{(-E_i/kT)}}{\sum_j g_j e^{(-E_j/kT)}}$$

Figure 6:

The kinetic energy of a particle corresponds to the velocity, and it corresponds to the energy one by one. Therefore, the probability density of the energy states can correspond to the rate distribution curve. The energy function is a continuous function, and the probability distribution function of the rate is also a continuous function. For example, at room temperature, the velocity distribution functions of several inert gases are as follows

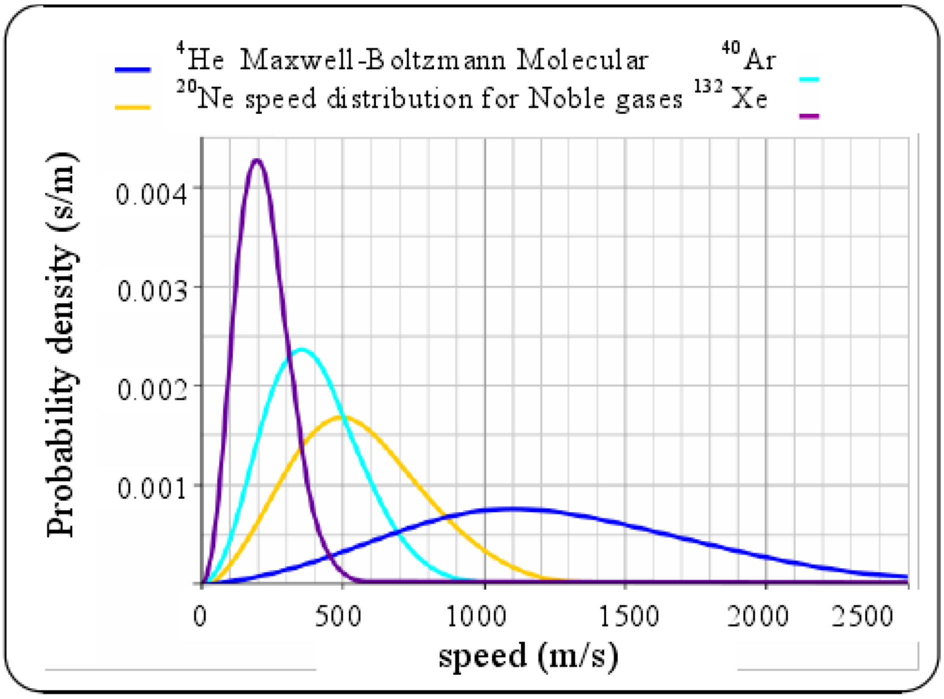

Figure 7:

3.5.2 Boltzmann machine learning model

A single Boltzmann machine structure is a bipartite graph. Bipartite graph, also known as bipartite graph, is a special model in graph theory. Let G = (V, E) be an undirected graph if the vertex v can be divided into two disjoint graphs.

In the bipartite graph of Boltzmann machine, a and B, the set of a point constitutes the visible layer variable space, which is usually represented by X. it has multiple components and is usually described by line vector. A single variable is represented by line vector and the dimension of X vector. In the training process, is known, also known as the sample.

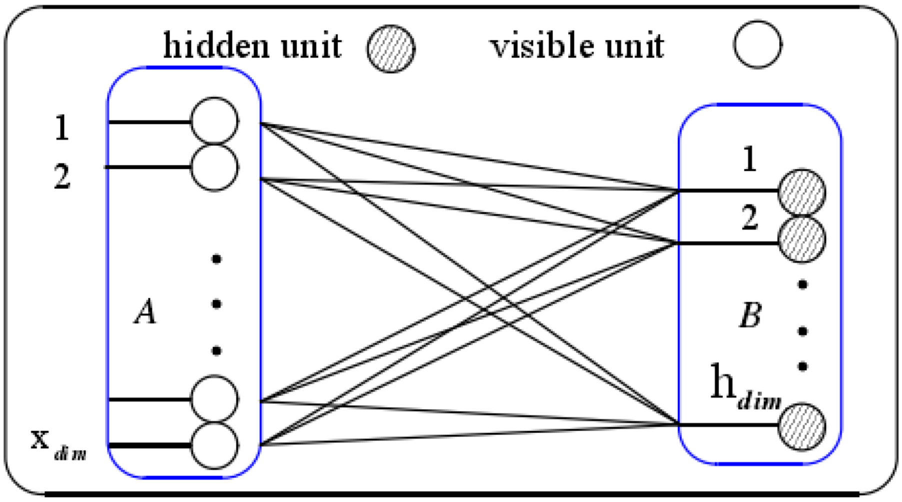

Figure 8:

The other is the hidden layer, which is usually represented by H and described by row vector. A single variable is represented by row vector and described by vector dimension. W: The weight matrix connecting the visible layer and the hidden layer. When the output (state value) hi of a single node in the hidden layer is given by its adjacent visible layer vector, and the value of hi presents a certain probability distribution, the sampling result under the probability distribution is called forward sampling; On the contrary, the state value Xi of a single node in the visible layer is the result of probability distribution sampling under the condition that the transmission vector of its adjacent hidden layer is given, which is called backward sampling. If x (DIM) > H (DIM) The essence of forward sampling is dimension reduction, while the essence of backward sampling is reconstruction.

$$\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{x_{dim}})$$

$$\boldsymbol{h} = (\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_{h_{dim}})$$

Figure 9:

3.5.3 Energy function of Boltzmann machine model

Energy function, for a given x, or h, no matter what transformation Boltzmann machine model makes, it can transmit some information to the outside world and make some output. Take x as the particle mass, h as the rate obtained in the model, and think that (x, h) corresponds to a certain energy state. At this time, the energy function enters the graph, and the model parameters B, C, W are collectively referred to.

$$E_\theta(x,h) = -xb^T - ch^T - x\mathbf{W}h^T$$

Figure 10:

The interaction between the hidden layer and the visible layer is expressed as an energy function.One is the two-layer state vector. Through the interaction of the weight matrix, xwh is the weight connecting the visible layer I and the hidden layer node j.

$$Z_\theta = \sum_{h \in \hat{H}, x \in \hat{X}} e^{-E_\theta(x,h)}$$

$$\sum_{x \in \hat{X}} \sum_{h \in \hat{H}} P_\theta(x,h) = 1$$

Figure 11:

The physical meaning of "energy" in the model is completely different from that in the Boltzmann physical system. In statistical mechanics, Boltzmann distribution is a one-dimensional distribution (pure gas, each particle has the same mass [104], but the velocity presents a distribution in a certain range). It only considers the kinetic energy of the speed of linear motion, and does not discuss the change and distribution of particle mass.

train($\overline{S}$,RBM(b, W, c)){
1.initialization : n_{epoch}, η
2.initialization : $c = 0, b = 0, \mathbf{W} = 0$
for $i_{epoch} = 1$ to n_{epoch}\{ $\overline{S} = \bigcup_{i_{batch}=1}^{n_{batch}} \overline{B}_{i_{batch}}$;
for $i_{batch} = 1$ to n_{batch}\{
feedback _ CD($\overline{B}_{i_{batch}}$,RBM(b,W,c));
$c = c + \frac{1}{s_{batch}} \cdot \eta \cdot \Delta c;\ \mathbf{W} = \mathbf{W} + \frac{1}{s_{batch}} \cdot \eta \cdot \Delta \mathbf{W};\ b = b + \frac{1}{s_{batch}} \cdot \eta \cdot \Delta b;$
}
}
}

Figure 12:

3.6 Comparative summary

In the cross validation experiment of disease recognition, the data show that: the SVN recognition method of disease pattern based on RBM dimension reduction is effective, and the accuracy is slightly inferior to PCA method in terms of quantity. After PCA dimension reduction, the disease pattern retains 67 main attributes, while RBM dimension reduction retains 30 attributes, which only depends on the structure of RBM network, The residual attributes affect the recognition performance to some extent.

This method introduces the idea of ”machine learning” into the feature extraction process, so that the machine can accept the guidance and improve the extraction performance. Compared with the ”invariable” feature extraction idea based on numerical analysis and calculation, this method has a new and innovative significance. The positive performance and improvement of convergence, robustness and effectiveness of the method also indicates that it has positive practical significance for the construction and development of crop disease intelligent early warning and response system.

4 Conclusion

Concousion: Agricultural sensor networks have acquired massive amounts of valuable image data. How to learn and process these data to discover novel agricultural knowledge models has become an advanced research topic in agricultural information technology. Using deep machine learning as a technical means, the following conclusions are drawn.

1.Penalty Correction Support Vector Machine Method Improves Learning Effect When the support vector classification method processes unbalanced samples, different targets, especially sparse targets, often lead to significant differences in learning error rates. The support vector machine method based on the penalty correction of the Lagrangian coefficient analysis method, for sparse samples, the method can improve the learning effect under the premise of stabilizing the overall performance; the stress test shows that the error rate and the generalization ability of the learning machine increase with the correction coefficient And the sample size converges quickly, which can significantly improve the classification performance of unbalanced sample sets

2.Feature extraction method based on eigenvalue overflow and discard The collection environment of agricultural field video sensing equipment is complicated, and there are many kinds of interference. Fixed-point and mobile methods are difficult to ensure that the representative samples that can be sampled under each specific environmental state show a uniform and complete distribution in the image collection, and there is no guarantee Network training uniformity. Focusing on numerical calculation and dimensionality reduction, based on principal component analysis, a feature extraction method based on eigenvalue overflow and discarding is proposed. This method shows good reconstruction performance; in SVN for disease image dimensionality reduction recognition, in the case of 2 ”fold” and 5 ”fold”, it shows satisfactory accuracy.

Future research directions: 1. The integration of agricultural multi-source information and knowledge Multi-source agricultural information learning and outputting multi-source agricultural knowledge, how to organically integrate them, and collaboratively carry out disease identification, diagnosis, and early warning is an urgent research topic for the application and development of agricultural

Internet of Things.

2. Incremental deep learning of agricultural IoT perception data In the agricultural Internet of Things, it is a very promising topic to study how to implement the incremental deep learning of agricultural perception data in an efficient, asymptotic, and real-time manner. []

References

[1] A. Krizhevsky, G. Hinton *et al.*, "Learning multiple layers of features from tiny images," 2009.

[2] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *Medical image analysis*, vol. 58, p. 101552, 2019.

[3] H.-J. Yoo, "Deep convolution neural networks in computer vision: a review," *IEIE Transactions on Smart Processing and Computing*, vol. 4, no. 1, pp. 35–43, 2015.

[4] Y. Grandvalet, Y. Bengio *et al.*, "Semi-supervised learning by entropy minimization." in *CAP*, 2005, pp. 281–296.

[5] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, "Randaugment: Practical automated data augmentation with a reduced search space," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 702–703.

[6] K. Sohn, D. Berthelot, C.-L. Li, Z. Zhang, N. Carlini, E. D. Cubuk, A. Kurakin, H. Zhang, and C. Raffel, "Fixmatch: Simplifying semi-supervised learning with consistency and confidence," *arXiv preprint arXiv:2001.07685*, 2020.

[7] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," *arXiv preprint arXiv:1606.03498*, 2016.

[8] A. Coates, A. Ng, and H. Lee, "An analysis of single-layer networks in unsupervised feature learning," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2011, pp. 215–223.