

CNN intelligent early warning for apple skin lesion image acquired by infrared video sensors^①

Tan Wenxue (谭文学)^{* **}, Zhao Chunjiang^{② ***}, Wu Huarui^{***}

(* College of Computer Science, Beijing University of Technology, Beijing 100022, P. R. China)

(** College of Computer Science, Hunan University of Arts and Science, Changde 415000, P. R. China)

(*** National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, P. R. China)

Abstract

Video sensors and agricultural IoT (internet of things) have been widely used in the informatized orchards. In order to realize intelligent-unattended early warning for disease-pest, this paper presents convolutional neural network (CNN) early warning for apple skin lesion image, which is real-time acquired by infrared video sensor. More specifically, as to skin lesion image, a suite of processing methods is devised to simulate the disturbance of variable orientation and light condition which occurs in orchards. It designs a method to recognize apple pathologic images based on CNN, and formulates a self-adaptive momentum rule to update CNN parameters. For example, a series of experiments are carried out on the recognition of fruit lesion image of apple trees for early warning. The results demonstrate that compared with the shallow learning algorithms and other involved, well-known deep learning methods, the recognition accuracy of the proposal is up to 96.08%, with a fairly quick convergence, and it also presents satisfying smoothness and stableness after convergence. In addition, statistics on different benchmark datasets prove that it is fairly effective to other image patterns concerned.

Key words: lesion image, self-adaptive momentum(SM) convolutional neural network(CNN), deep learning, early warning, agri-sensor

0 Introduction

China, a big country in production of fruit, has a large planting area of fruit trees. Fruit industry is of great significance in Chinese commercial agriculture. With the development of equipment agriculture construction, expert system based on knowledge data-base (KDB) has been widely applied in the process of agricultural production^[1,2]. However, as to the agricultural production environment where almost all smart devices are networked by IoT, limitations of the KDB system become increasingly obvious. For knowledge rule and extraction of disease-pest feature has to be mainly operated and accomplished by human experts manually^[3]. Accuracy and standardization degree of the expression is influenced to a great extent by human factors, which directly affects the efficiency of system decision. Agricultural informatization system and sensing network deployment give facilities for the real-

time acquisition, processing and transmission of fruit growth information. While KDB system cannot often respond to the real-time information in a timely manner, because its description format (video, images) is not compatible with the knowledge representation paradigm used in the KDB system (standardized and wording knowledge expression in professional terminology). In addition, another case happens very often is that the time lag induced by data format transferring causes losses to growers.

In response to these problems, the automatic feature extraction of lesion patterns and real-time disease-pest recognition have become a hotspot of the research of agricultural information technology. It attracts the attention of people to research on the machine learning methods which can be a potential application in lesion pattern feature extraction and recognition. Liu, et al researched on discrimination analysis of apple during vacuum freeze-drying using near infrared diffuse reflectance spectroscopy and principal component analy-

① Supported by the National Natural Science Foundation of China (No. 61271257), Beijing National Science Foundation (No. 4151001) and Hunan Education Department Project (No. 16A131).

② To whom correspondence should be addressed. E-mail: zhaocjncita@163.com

Received on Jan. 8, 2015

sis^[4]. During a growth cycle of apple, growers always pay relatively more attention to prevention and treatment of the fruit skin diseases. Yongsheng, et al presented an apple recognition method based on color difference R-G^[5], using which the image of apples under different light conditions was segmented effectively, the center and radius were accurately extracted for apple shape recognition, with the accuracy up to 97%. The aforementioned provides a helpful reference for the feasibility of the feature extraction for disease-pest recognition based on the image of diseased apples.

In the past decades, machine learning techniques based on neural network played an increasingly important role in pattern recognition and dimensional reduction^[6,7]. As a key step for pattern recognition, feature extraction directly influences accuracy of recognition. Because of the richness and diversity of natural data (images, pictographic symbols and other patterns), the ambiguity and indefiniteness of natural language, it is almost impossible for professional experts to set up a versatile method for feature extraction just by means of handcraft. Image operation, convolution and sampling can maintain the invariance of the extracted sift, surf, and other meta image feature points to image transformation including rotation, translation, scaling and distortion to some extent, which is very suitable for extracting features of the image pattern acquired in a complicated, unstable scene. Based on that, convolution and sampling operation are introduced into neural network, and research on how to exploit them to extract automatically a group of pathological features consistent with disease-pest recognition task from the image of diseased apple, and how to perform a real-time disease-pest early warning based on the orchard sensor network for image acquisition.

1 Lesion image acquisition and sensor network

Equip an orchard in Daxing district with a latest agricultural IoT network including video sensor device, and video sensing makes use of an infrared video sensor shown in Fig.1(a). It is a spherical device embedded with a camera which can uniformly rotate horizontally or vertically at an adjustable angular speed and the camera supports 600 lines HD (high definition) and 27 times zoom. In addition, the device provides an array of infrared light source with an effective irradiation distance of 120m. Fruit trees in the orchard are arranged in a matrix, and in general fruits are uniformly distributed at random in the tree crown, most of which are hemispherical and slightly tilted horizontally, shown in Fig.1(b). Moreover, in view of the best dis-

tance and perspective for effect of the image acquired, the primary sensor for the crown is positioned horizontally at the centroid of the polygon composed by the earth snap-in points of the trees in corresponding monitoring area, and its height is the average of that of the tree crowns, as is exhibited by Fig.1(c). Thus, orchard sensors consist of an array, and through the transmission wire and remote control device shown in Fig.1(d), all information of data acquisition nodes converges at the integrated manipulation center, which is remotely accessed by intelligent terminals using Internet for online access-processing. In addition, the planting site is also equipped with additional agricultural information sensing equipments, which are assigned with real-time monitoring of environmental parameters such as air temperature, air humidity, light and soil temperature. All the sensing information converges through the data bus, which is potential to realize real-time early warning for disease-pest based on multi-source information.



Fig. 1 Sphere video sensor of uniform rotation

In a growth cycle, fruit crops often develop external local biological lesion, which is always the result of the effect of fertilizers, pest or the influence of meteorological condition (temperature, sunlight, humidity etc.)^[8], soil environment (moisture, heavy metal content of soil), biological characteristics (water intake of root and leaf, etc.), or the impact of agricultural planting measures^[9]. Usual regions are infected easily with disease covering trunk and leaf, fruit and root. The diseased organs can be photographed by the image sensors of the IoT, or the mobile handheld ima-

ging equipments. For example, some fruit skin images of the apples infected with common diseases are exhibited in Fig. 2.

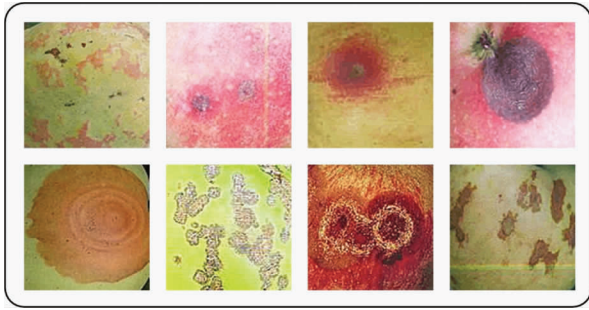


Fig. 2 Lesion images of apple sphere

2 Dimensional reduction and image processing

2.1 Simulation of direction disturbance

In an orchard, distribution of the image sensing equipments follows a certain layout. And the relative position of equipment to fruits in the monitoring area is closely related to current distribution of growing fruits^[10]. Image is a case of 2-dimensional system. Regard direction of gravity as reference, as to a certain lesion object under observation, different image patterns are acquired by sensors from different perspectives. Thus, reference direction and a sensed image form an angle. In order to establish an adaptability of recognition system to the lesion images, of which angles are variable, an expanded image dataset is constructed from the originative standardized image using rotation transformation through 4 different angles respectively. As is shown in Fig. 3, Fig. 3(a) is the originative, and Fig. 3(b) is the corresponding gray image of 256-scale after a clockwise rotation through 90°.

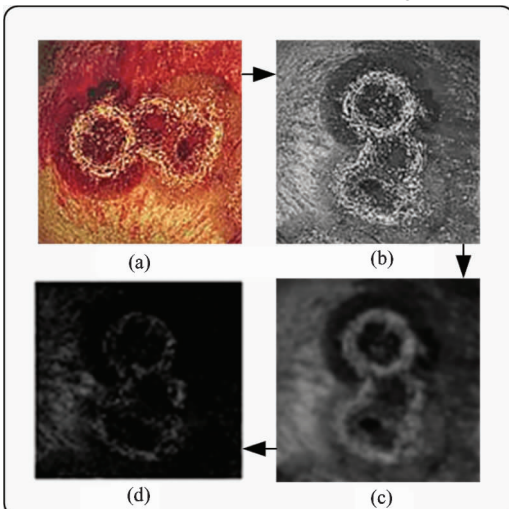


Fig. 3 Disturbance and dimensional reduction

2.2 Brightness disturbance

Because of the interference of many factors, the orchard light condition often becomes complex. For example, variable sunlight orientation from the periodic motion of sun, the random appearance of cloud, the unpredictable weather of somber, sunshining, rain, fog, and a backlighting sensor, etc. All of these factors are potential to give an impact to the brightness and balance of lesion image. Similarly, imaging system acquires some different image pattern of a certain object. For the adaptability of recognition system to lesion images acquired in a variety of light background, it is a must to train it using some learning samples derived from the normalized ones. To an originative image, adjust its intensity value which is saturated in a certain brightness interval by computer program and the selected intervals are $[0.2\ 0.4]$, $[0.4\ 0.6]$, $[0.6\ 0.8]$ and the default. This is a simulation of brightness disturbance. For interval $[0.6\ 0.8]$, Fig. 3(c) represents the corresponding disturbed image of the aforementioned, rotated sample. Brightness adjustment will cause some noise, which goes against recognition of image. And it is a must to consider the critical point of consequent noise as much as possible when to determine brightness intervals.

Thus, combining direction and brightness, each sample derives 16 disturbed patterns, which are approximate depictions of one and the same lesion in different scenes.

2.3 Dimensional reduction and sparsity

Dimensional reduction is always used to map a pattern vector in a higher dimensional space to one represented in a lower dimensional space, and the precondition is that the relative space relationship of pattern features is unchanged^[11]. As to recognition of pattern in a lower-dimensional space, data dimensional and computational complexity decreases more greatly than the originative pattern, which effectively improves recognition accuracy and downsizes the error caused by redundant information of pattern. Lesion image is a typical pattern of high dimensional relative to the machine learning ability, and its corresponding sparse reduced version can be acquired by dimensional reduction. Methods commonly used for dimensional reduction cover Laplacian-Eigenmap^[12], which is nonlinear, and PCA^[13], which is linear. For example, Fig. 3(d) exhibits the dimensionally reduced lesion image aforementioned in the last subsection. This image graphs the apple black rot, which is a common disease to apples. What should be noted is that in the matrix of gray-scale value, the blackest point denotes its intensi-

ty to approach 0, while the whitest point denotes it to approach 255. After the $[0, 1]$ normalization, it turns into a sparse matrix. The processed image represents the contour and texture of originative lesion by and large. In parallel with Fig. 3(b), number of non-zero value sharply decrease while zero value and value approximate to zero form the main of matrix. The dimension of pattern is reduced effectively.

3 Convolutional network designs

3.1 Deep learning network

Multilayer perceptron is a machine learning network, idea of which is inspired from bionics. Composed of some layers of single function, the network can only passively adjust weights according to training error, and cannot automatically extract features of pattern and gradually deeply process base features. The number of network weights w_n equals to the sum of the weight amount of each layer, as Eq. (1).

$$w_n = \sum_{i=2}^L n_{i-1} \times n_i \quad (1)$$

where n denotes the number of node in a layer. Learning from unstructured pattern needs hundreds of inputs, which incur tens of thousands weights. Such a large number of free parameters inflate volume of the potential object functions expressed by the learning network, thus a training set of bigger volume is required to cover all possible feasible solutions. In addition, the expanded demand of memory causes that the task cannot run in the embedded systems which are restricted in memory size, and that the network is incompetent to deal with more complicated learning tasks for the expanding of layers and nodes endlessly is unfeasible.

According to early results of the research on retinal and based on convolution theory, some scholars have brought forward an innovative learning method, which performs an automatic extraction of feature of receptive field. By the response from the convolution operator to the local receptive field, neurons can extract many basic visual features, which are constantly and gradually-deeply processed in the subsequent layer, and then get some higher-order features.

This process has not been realized by the classical learning method, and it is deep learning from pattern features. These neurons compose a Convolution-Subsampling hierarchy of multilayer, which gradually deepens extraction and process of pattern feature, and outputs a novelty learning network, namely convolutional neural network (CNN).

3.2 Image Convolution Operation

2D convolution of discrete field is shown as Eq. (2), where \mathbf{K} is a convolution kernel matrix, which convolves with matrix \mathbf{A} as Eq. (2). The computation process is denoted by $\mathbf{A} \times \mathbf{K}$ and the result is called the convolution of \mathbf{A} by \mathbf{K} . Convolution kernel is also known as a filter, and different filters are employed to extract different corresponding features.

$$\sum_{v=-\infty}^{+\infty} \sum_{u=-\infty}^{+\infty} \mathbf{A}[u, v] \cdot \mathbf{K}[m-u, n-v] = \mathbf{A} \times \mathbf{K} \quad (2)$$

If \mathbf{K} is factorable, as Eq. (3), then, the convolution impacted upon \mathbf{A} can be decomposed into 2 times of convolution. Inversely, an advanced, higher-order filter can be produced by connection of simple ones in series as Eq. (4).

$$\mathbf{K} = \mathbf{a}^T \mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \cdot [b_1 \quad b_2 \quad b_3] \quad (3)$$

$$\mathbf{A} \times \mathbf{K} = \mathbf{A} \times (\mathbf{a}^T \mathbf{b}) = (\mathbf{A} \times \mathbf{a}^T) \times \mathbf{b} \quad (4)$$

3.3 Architecture design of deep network

Here, an architecture design of lesion image recognition network is presented. In general, each non-terminal node has two ends, input and output. "Layer" is defined as a set of the nodes in a column, and "connection" denotes a weight or a convolution operator, which is used to connect the output end of a node (the prior of connection) and the input end of another node (the posterior of connection). Information is one-way backward transferred from the prior to the posterior. Association between layers is embodied by a connection, while a connection is used to bridge between two adjacent layers. Structure of the layer refers to its neuron activation function. A void layer has no nonlinear transference function, and by which information is linearly transferred to the posterior. Many layers are stacked up into a learning network, and the data structure for saving parameters of connection is defined in its subsequent layer's declaration of machine language code.

Fig. 4 exhibits the design of 5-layer convolutional network. Considering facilitating comparing experimental progress of lesion image with other benchmark data sets, the lesion image is formatted as 28×28 pixels. Type annotation "S" ("C") says that the layer is a sampling (convolutional) layer. "Full connection" denotes a Gaussian connection, which executes the last process before the aftermath of network. Layer and connection object is represented by an array variable in

a “cell” structure of Matlab to program data storage.

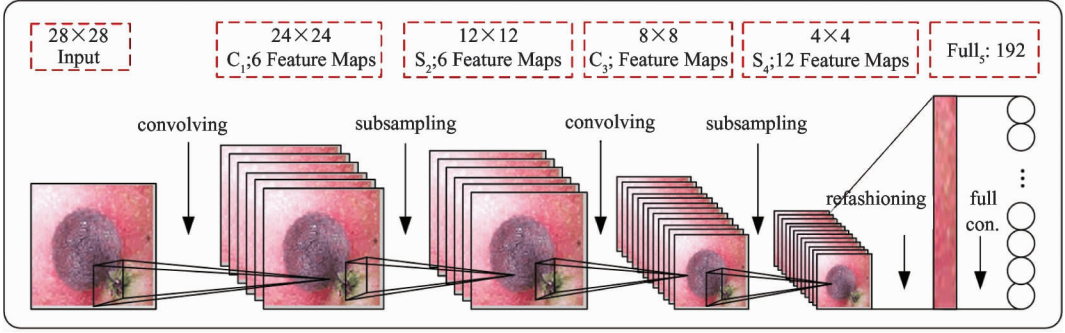


Fig. 4 Convolution neural network for apple lesion image recognition

Matlab sentence “Struct (‘type’, ‘C’, ‘output-maps’, 6, ‘kernel-size’, 5)” defines a convolutional layer, i. e. whose prior connection is a convolution operation. It outputs 6 feature maps, with a convolution kernel which is a square matrix of 5×5 . The resolution of the outputted feature map by this convolution is decreased horizontally and vertically by “kernel-size-1” pixels.

Similarly, “Struct (‘type’, ‘S’, ‘scale’, 2)” defines a sampling layer, whose prior connection is a sampling operation. Compression ratio of the sampling is “scale = 2”, namely, the feature map is returned by sampling 1 row(column) from each 2 consecutive rows (columns). And its resolution is decreased horizontally and vertically to a half of the original in pixels.

In general, symbols in the top rectangle of Fig. 4 respectively denote type, sequence number of layer, number and size of the feature map produced by the prior connection of layer. Obviously, the learning network forms a “double pyramid” after introducing convolution and sampling. That is to say that feature map (F. Maps) resolution decreases backward, while its number increases backward. Geometrical invariance of transformation from the learning network is developed by gradually reducing the feature space resolution, and the consequent information loss is compensated from advancing abundance of feature images.

4 Learning based self-adaptive momentum

Optimization of network parameters is constructed by propagation of the full connection weigh error and kernel error of each convolution. From layer $l = 2$ to n , select kernel of convolution connection and exploit kernel error to update it, and the corresponding adjusted convolution kernel is produced as

$$\mathbf{K}_{ij}^{(n)} = \mathbf{K}_{ij}^{(n-1)} + \Delta \mathbf{K}_{ij} \quad (5)$$

If layer l is the full connection layer, then task of

learning is to adjust weight w_{ji} . Let $E_d(\cdot)$ be a loss function, which uses the sum of error square, where “ \cdot ” represents the free parameters to be determined by training. net_j , which is the input of the activation part j . δ_j the fastest declination rate of $E_d(\cdot)$ in respect to net_j as Eq. (6) denotes the updating coefficient of part j .

$$\delta_j = - \frac{\partial E_d}{\partial net_j} \quad (6)$$

The adjustment minim of i th input-weight for part j Δw_{ji} is calculated by Eq. (7) from δ_j , the input x_{ji} , which is transferred along the weight w_{ji} , and learning rate η . It is called native error ∂w_{ji} . On this base, the n th updated weight $w_{ji}^{(n)}$ is returned by Δw_{ji} and the $(n - 1)$ th weight $w_{ji}^{(n-1)}$ according to Eq. (8).

$$\partial w_{ji} = \eta \delta_j x_{ji} \quad (7)$$

$$w_{ji}^{(n)} = \partial w_{ji} + w_{ji}^{(n-1)}; \Delta w_{ji}^{(n)} = \partial w_{ji} \quad (8)$$

4.1 Constant momentum learning

For expediting the convergence of learning network, Dahanayake, et al put forward a momentum machine learning method^[14]. The method enables learning track to escape away from the region where there may be a local minimum of loss function at a fairly brisk pace, thus to bypass the trap of local optimal. According to the momentum method, α denotes a constant of interval $[0, 1]$, and the method regresses to Eq. (7) where $\alpha = 0$. In general, α is a nonzero, the designated adjustment error is produced in the way that the native error is added to a momentum component $\alpha \Delta w_{ji}^{(n-1)}$, which is a multiply of α (Inertia coefficient) and last adjustment $\Delta w_{ji}^{(n-1)}$ as Eq. (9), and is termed constant momentum (CM) method.

$$\Delta w_{ji}^{(n)} = \partial w_{ji} + \alpha \Delta w_{ji}^{(n-1)} \quad (9)$$

4.2 Self-adaptive momentum rule

With respect to the convolution network, any of

the parameters to be updated in the learning process is denoted by τ_{ji} . In considering the sign of momentum component $\alpha\Delta w_{ji}^{(n-1)}$, a case is that it is just determined by $\Delta w_{ji}^{(n-1)}$, where inertia coefficient α is set at some one positive constant. Relative to native errors, it is sometimes downwind and sometimes upwind, which slows down the convergence of the learning process.

On this basis, a straight strategy is proposed to adopt a self-adaptive variable to revise inertia coefficient. Let *sign* be the function of returning the sign of the multiply of native error and last adjustment $\Delta\tau_{ji}^{(n-1)}$, and its use secures a dynamic accordance of sign between the momentum component and the native error. That is to say, no matter what sign of $\Delta\tau_{ji}^{(n-1)}$ and $\partial\tau_{ji}$, the revised momentum component $\alpha\Delta\tau_{ji}^{(n-1)}$ shares the sign with $\partial\tau_{ji}$ at all time, which is shown in Table 1.

Table 1 Parameter signs of self-adaptive momentum

$\partial\tau_{ji}$	$\Delta\tau_{ji}^{(n-1)}$	α	$\alpha\Delta\tau_{ji}^{(n-1)}$
+	+	+	+
+	-	-	+
-	+	-	-
-	-	+	-

As to the magnitude of inertia coefficient, it is assigned the ratio of the square of last adjustment minim to the addition of the aforementioned square and square of the native error as Eq. (10). The ratio can change the size of inertial coefficient flexibly. Once the native error $\partial\tau_{ji}$ increases, inertial α decrease accordingly; while $\partial\tau_{ji}$ decreases, the appropriate increase of α is returned. Thus the revised momentum component $\alpha\Delta\tau_{ji}^{(n-1)}$ always dynamically maintains a steady tempo, and here it is termed the self-adaptive momentum (SM). Eq. (11) represents the consequent adjustment $\Delta\tau_{ji}^{(n)}$, which is helpful to the learning process not only to crash out of trap of local minimum in a fairly wide step but also to avoid missing convergence point.

$$\alpha = \text{sign}(\partial\tau_{ji} \cdot \Delta\tau_{ji}^{(n-1)}) \frac{(\Delta\tau_{ji}^{(n-1)})^2}{(\Delta\tau_{ji}^{(n-1)})^2 + \partial\tau_{ji}^2} \quad (10)$$

$$\Delta\tau_{ji}^{(n)} = \partial\tau_{ji} + \alpha\Delta\tau_{ji}^{(n-1)} \quad (11)$$

5 Experimental results and discussion

Following instruction of the botanists on fruit trees, and by means of video sensors, researchers have acquired 250 image samples from the diseased apples as follows. Scab skin, black rot, scar skin and ring

spot^[15], all of which often occur to apple skin. After normalization and processing of disturbance, a base dataset of apple lesion image comes into existence. All algorithms concerned are programmed using Matlab. In order to explore the improvement of the recognition performance, convergence, the generalization ability and effectiveness, a serial of testing experiments are carried out.

5.1 Recognition performance

As to recognition performance, the experimental result of the proposed method (SM CNN) is compared with 4 learning method including LeNet of 5 layer (LeNet-5), Boosted-LeNet-4 (LeNet-4)^[16], multilayer neural network with 3 layers (MNN) and *k*-Nearest Neighbour (*k*NN). Based on the aforementioned dataset, along the ladder of increasing iterative epoch, a sequence of experiments is carried out on respective algorithm, and the corresponding accuracy is graphed in Fig. 5. As a result, despite not being the best in 5 algorithms, the accuracy of SM CNN is very close to it. And in parallel with the conventional neural network of multi-layer and *k*NN, it presents overwhelming odds.

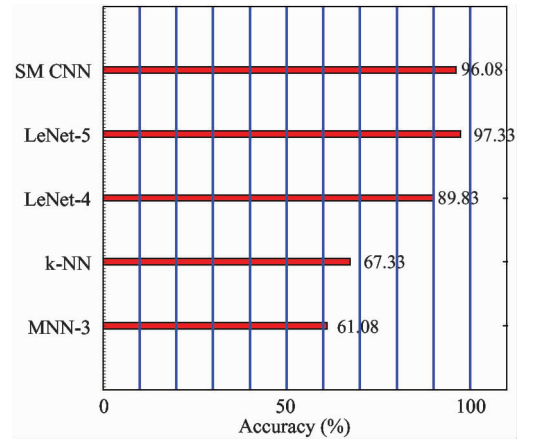


Fig. 5 Accuracy comparison among algorithms

5.2 Convergence

In this experiment, 3 best algorithms are selected in respect of recognition performance, and research on the variation of error rate with training epochs. The results are illustrated in Fig. 6, where the points labelled by a dashed sign “+” circled by an elliptic, denote the starting point of convergence of algorithm. It can be seen that compared with LeNet-5, which presents the best recognition performance and LeNet-4, the convergence point of SM CNN occurs earliest, at 16×10^2 rounds about. SM CNN advances convergence by 4×10^2 rounds and 7×10^2 rounds respectively than the rest ones. As far as the smoothness and fluctuation after

convergence is concerned, performance of SM CNN method is more competitive.

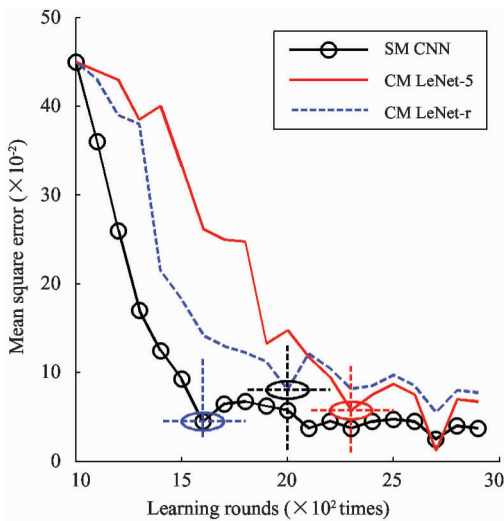


Fig. 6 MSE and convergence

5.3 Comparison of effectiveness

Specific to recognition of apple lesion image, performance of algorithm has been tested from previous experiment. However, as different benchmark data sets are concerned, how will SM CNN work? Adopt 2 additional bench data sets to experiment, and the results are illustrated in Fig. 7. As supplements, the 2 datasets are respectively MNIST Zip Digit^[17], which is the handwritten digit dataset of America zip code, and ORL-Face, which is the face image dataset provided by Olivetti research laboratory. It can be seen that for zip-digit and apple lesion image, the proposal presents a fairly good precision and convergence; while specific to ORL-Face set, its stable precision decreases obviously.

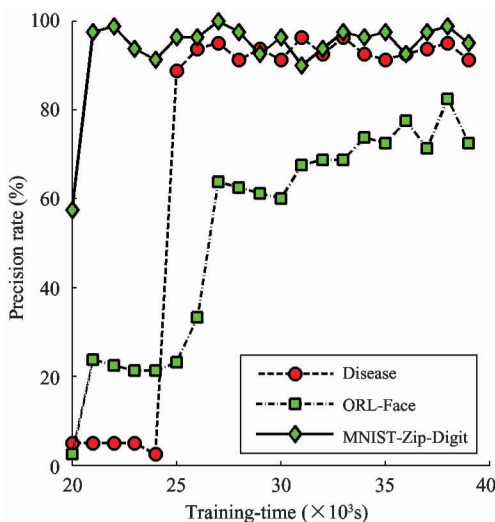


Fig. 7 Precision upon benchmark image sets

However, precision of 75% demonstrates it is effective in general, but it leaves a lot to be desired. An acceptable explanation is that: by comparison, digit is a simple pattern, which only presents some linear features; and face is a relatively complex pattern, which is influenced by expression, covering (cap, glasses). The complexity gives a non-negligible impact on precision of recognition.

6 Conclusions

Based on the deep machine learning and orchard infrared video sensor, the paper presents an intelligent early warning method, SM CNN, which is grounded upon the convolutional network for recognition of fruit skin lesion image. It also designs the architecture of network and self-adaptive momentum rule for parameter learning. Take example by apple and systematically launch the research and experiment on the recognition of fruit skin lesion image sensed by a sensor. The results demonstrate that compared with the shallow learning algorithms and other involved, generally accepted deep learning method, the recognition accuracy of SM CNN is up to 96.08%, with a fairly quick convergence, and it presents satisfying smoothness and fluctuation after convergence. In addition, statistics on the different benchmark datasets prove it is fairly effective to recognize image pattern.

Specific to the application to production, in combination with the agri-sensor IoT installed in the orchards for agricultural information monitoring, SM CNN is potential to realize an auxiliary early warning to grower for disease-pest based on lesion image automatic recognition, and to alert them for a timely response of production management. Moreover, compared with the KDB expert system, it can automatically extract features from lesion image pattern, thus realize an intelligent, unattended alert for disease-pest, which releases users from human intervention and advances a real-time of response to warning for a better agriculture production income.

References

- [1] Clancey W J. The epistemology of a rule-based expert system—a framework for explanation. *Artificial intelligence*, 1983, 20(3) : 215-251
- [2] Romeo J, Pajares G, Montalvo M, et al. A new expert system for greenness identification in agricultural images. *Expert Systems with Applications*, 2013, 40 (6) : 2275-2286
- [3] Wang J, Gao B B, Wang Z Q, et al. Intelligent geospatial information retrieval for agricultural experts—a knowledge-based method. *Sensor Letters*, 2010, 8(1) : 178-183

- [4] Liu X B, Yuming G, Lihong F. On-line monitoring of moisture ratio for apple during vacuum freeze-drying based on image texture analysis. *Transactions of the Chinese Society of Agricultural Engineering*, 2012, 28(21) : 229-235
- [5] Yongsheng S, Jun Q, Gang L, et al. Recognition and shape features extraction of apples based on machine vision. *Transactions of the Chinese Society for Agricultural Machinery*, 2009, 08 : 161-165
- [6] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11) : 2278-2324
- [7] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313(5786) : 504-507
- [8] Lemmetty A, Soukainen M, Tuovinen T. First report of 'candidatus phytoplasma mali,' the causal agent of apple proliferation disease, in apple trees in Finland. *Plant Disease*, 2013, 97(10) : 1376-1376
- [9] Dutot M, Nelson L M, Tyson R C. Predicting the spread of postharvest disease in stored fruit, with application to apples. *Postharvest Biology and Technology*, 2013, 85 : 45-56
- [10] Liu Y, Yang Y, Lv X P, et al. A self-learning sensor fault detection framework for industry monitoring IoT. *Mathematical Problems in Engineering*, 2013
- [11] Taylor G W, Hinton G E, Roweis S T. Two distributed-state models for generating high-dimensional time series. *Journal of Machine Learning Research*, 2011, 12 : 1025-1068
- [12] Pan R, Zhang X. A note on laplacian eigenmaps. *Journal of Shanghai Jiaotong University (Science)*, 2009, 05 : 632-634
- [13] Kadappa V, Negi A. Computational and space complexity analysis of SubXPCA. *Pattern Recognition*, 2013, 46(8) : 2169-2174
- [14] Dahanayake B W, Upton A R M. Derivation of momentum LMS learning algorithms by minimizing objective functions. In: *Proceedings of the Conference on Neural Networks*, San Francisco, USA, 1993, 2 : 831-835
- [15] Hartevelde D O C, Akinsanmi O A, Chandra K, et al. Timing of infection and development of alternaria diseases in the canopy of apple trees. *Plant Disease*, 2014, 98(3) : 401-408
- [16] Al-Jawfi R. Handwriting arabic character recognition LeNet using neural network. *International Arab Journal of Information Technology*, 2009, 6(3) : 304-309
- [17] Xie Y, Zhang W S, Qu Y Y, et al. Discriminative subspace learning with sparse representation view-based model for robust visual tracking. *Pattern Recognition*, 2014, 47(3) : 1383-1394

Tan Wenxue, born in 1973. He is a Ph. D candidate of College of Computer Science, Beijing University of Technology. He graduated with Master's of Science in Information technology and Earth Exploring from East China Institute of technology, Jiang-xi, Mainland of P. R. China, 2003. His current research interests include agriculture information technology and artificial intelligence, cloud information security.