

## Article

# An Intelligent Sorting Method of Film in Cotton Combining Hyperspectral Imaging and the AlexNet-PCA Algorithm

Quang Li , Ling Zhao \*, Xin Yu, Zongbin Liu and Yiqing Zhang 

College of Mechanical and Automotive Engineering, Liaocheng University, Liaocheng 252000, China; li18754773982@163.com (Q.L.); yx15553285576@163.com (X.Y.); liu\_zongbin@163.com (Z.L.); yiqingzhang@lcu.edu.cn (Y.Z.)

\* Correspondence: zhaoling@lcu.edu.cn

**Abstract:** Long-staple cotton from Xinjiang is renowned for its exceptional quality. However, it is susceptible to contamination with plastic film during mechanical picking. To address the issue of tricky removal of film in seed cotton, a technique based on hyperspectral images and AlexNet-PCA is proposed to identify the colorless and transparent film of the seed cotton. The method consists of black and white correction of hyperspectral images, dimensionality reduction of hyperspectral data, and training and testing of convolutional neural network (CNN) models. The key technique is to find the optimal way to reduce the dimensionality of the hyperspectral data, thus reducing the computational cost. The biggest innovation of the paper is the combination of CNNs and dimensionality reduction methods to achieve high-precision intelligent recognition of transparent plastic films. Experiments with three dimensionality reduction methods and three CNN architectures are conducted to seek the optimal model for plastic film recognition. The results demonstrate that AlexNet-PCA-12 achieves the highest recognition accuracy and cost performance in dimensionality reduction. In the practical application sorting tests, the method proposed in this paper achieved a 97.02% removal rate of plastic film, which provides a modern theoretical model and effective method for high-precision identification of heteropolymers in seed cotton.

**Keywords:** seed cotton; film; hyperspectral image; dimension reduction; convolutional neural network



**Citation:** Li, Q.; Zhao, L.; Yu, X.; Liu, Z.; Zhang, Y. An Intelligent Sorting Method of Film in Cotton Combining Hyperspectral Imaging and the AlexNet-PCA Algorithm. *Sensors* **2023**, *23*, 7041. <https://doi.org/10.3390/s23167041>

Academic Editors: Bin Fan and Wenqi Ren

Received: 6 July 2023

Revised: 2 August 2023

Accepted: 7 August 2023

Published: 9 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cotton plays an irreplaceable part in the livelihood of the general population. Xinjiang is China's largest major producer of long-staple cotton. However, due to the low rainfall and strong light, drip irrigation under the film is often adopted to boost yield, which is prone to mixing with impurities such as plastic film during mechanical picking. In the spinning and weaving processes, the residual film combined with seed cotton can result in a significant number of flaws, which can impact the strength and coloring effect of the yarn and lead to financial losses for the textile sector [1].

The existing mainstream cotton film removal processes include Mechanical separation, Electrostatic separation, and Optical color separation. Whitelock D. P. et al. investigated major impurity removal equipment in the US cotton industry. A rotating spiked cylinder was utilized to eliminate significant impurities from the seed cotton. These impurities were subsequently gathered in a separate box by means of a grid strip or screen. [2]. Zhang et al. used computational fluid dynamics (CFD) to model the electrostatic separation of mechanical cotton harvesting and residual plastic film by flying the experimental sample into an electric field at different speeds and applying different electric field forces [3]. With the increasing prevalence of machine vision, optical color separation has become a popular method for the intelligent classification of agricultural products. In a study conducted by Li et al., a machine vision system was utilized to gather information on the color, shape, and texture of foreign fibers in cotton. The resulting data were adopted to achieve a classification accuracy of 92.34% through multi-class support vector machine (MSVM) [4].

However, mechanical classification is challenging in the aspect of assuring accuracy and small-size film classification. Electrostatic separation becomes unstable for long-term work because of environmental conditions. Optical color selection relies on color and form characteristics, making it challenging to effectively classify film which is colorless, transparent, or irregularly shaped. Hence, it is imperative to investigate a dependable technique for identifying transparent films in seed cotton.

Hyperspectral imaging combines advanced knowledge from multiple disciplines to achieve a perfect fusion of traditional two-dimensional imaging techniques and spectroscopy. Guo and Ma described the linear relationship between spectra and data by the partial least squares (PLS) method, which could realize the analysis of adulterated rice and the prediction of pork meat fatty acids [5,6]. Zhang, Jiang, et al. employed the support vector machine (SVM) in combination with shortwave infrared hyperspectral techniques for cotton foreign matter classification, which significantly improved the detection rate of plastic films in cotton compared to conventional methods [7–9].

The above literature has yielded promising results. However, extracting features from hyperspectral images requires manual intervention and has limitations in feature mining. In addition, manual feature extraction of hyperspectral images requires considerable expertise and has subjectivity in feature mining and selection. Therefore, it is highly significant for hyperspectral image features to probe an automatic feature extraction method.

Deep learning is an advanced technology applied to image processing. It has the capability to automatically detect and analyze complex information, which helps to extract deeper features. The use of hyperspectral data greatly enhances the accuracy and efficiency of image recognition [10,11]. However, it is important to note that hyperspectral data can be affected by elevated latitudes and severe information redundancy issues. To efficiently extract feature information to support the training of deep learning models, data dimensionality reduction is commonly applied to improve the data processing speed [12]. Jia et al. employed a method for dimensionality reduction of hyperspectral images by flexible Gabor-based superpixel-level unsupervised linear discriminant analysis (LDA), which reduced a large amount of flexible Gabor (FG) features and increased the peculiarity of image features [13]. Kang et al. proposed a method based on PCA-EPFs for hyperspectral image (HSI) classification, which used principal component analysis (PCA) to reduce the dimension of the superimposing Edge-preserving features (EPFs). The literature not only represented the EPFs in the mean square sense but also highlighted the divisibility of pixels in EPFs [14]. To reduce the dimension of hyperspectral remote sensing images, Daniela Lupu et al. established an independent component analysis (ICA) method based on a stochastic higher-order Taylor approximation-based algorithm, which could identify local maxima and facilitate minibatching [15]. The previous researchers have utilized LDA, PCA, and ICA techniques for reducing the dimensionality of hyperspectral data. The experimental results have demonstrated excellent outcomes, effectively enhancing the efficiency of image processing.

Convolutional neural network (CNN) is the most frequently employed deep learning model that performs excellent classification effect in feature extraction of hyperspectral data; it can be used to solve the problem of plastic film in seed cotton [16,17]. LeNet, AlexNet, and VGGNet are frequently employed neural network models in CNN which achieve high classification and recognition accuracy with great fusion with hyperspectral images. Hüseyin Fırat et al. proposed a method to effectively classify hyperspectral remote sensing images (HRSIs) based on PCA dimension reduction and LeNet-5 of the 3D-CNN model. The results showed that a 100% recognition and classification effect was obtained in all experimental data [18]. Jiang et al. obtained hyperspectral images of different types of pesticide residues and used the fusion of the AlexNet-CNN deep learning network to detect post-harvest pesticide residues in apples. The test results showed that when the number of training epochs was 10, the detection accuracy was 99.09% [19]. Zhao et al. recorded the waterlogging of cotton after seeding with hyperspectral images. Based on the comparison experiment of GoogLeNet Inception-v3 (GLNI-v3) and VGG-16 conducted

by CNN, the classification accuracy of VGG-16 was 97.00% higher than that of GLNI-v3, and the method could provide theoretical support for the evaluation of cotton loss after waterlogging [20]. The aforementioned literature demonstrates that the CNN-based models (LeNet, AlexNet, VGGNet) mentioned above exhibit strong generalization and adaptive capabilities in processing hyperspectral images, resulting in effective application outcomes.

The combination of hyperspectral imaging and CNN techniques is commonly applied to the classification of remote-sensing images. However, there have been few reports of methods to identify the residual film in seed cotton. The academic paper presents a novel approach for removing film in seed cotton, which combines hyperspectral images and deep learning algorithm. The innovations are listed as follows:

(1) The study establishes an optimal method for dimensionality reduction of hyperspectral data, which can reduce redundant hyperspectral characteristic information and reduce the time and cost of subsequent neural network training.

(2) The study integrates hyperspectral imaging technology and deep learning algorithm to obtain the optimal AlexNet-PCA-12 model which can effectively remove the colorless and transparent film in seed cotton in the practical application.

The remainder of the paper is structured as follows. Section 2 describes the hyperspectral sorting system, the theory of dimensionality reduction and CNN. Section 3 illustrates the discussion of the results of dimension reduction and CNN experiments. Conclusions and viewpoints are provided in Section 4.

## 2. Materials and Methods

### 2.1. Hyperspectral Sorting System

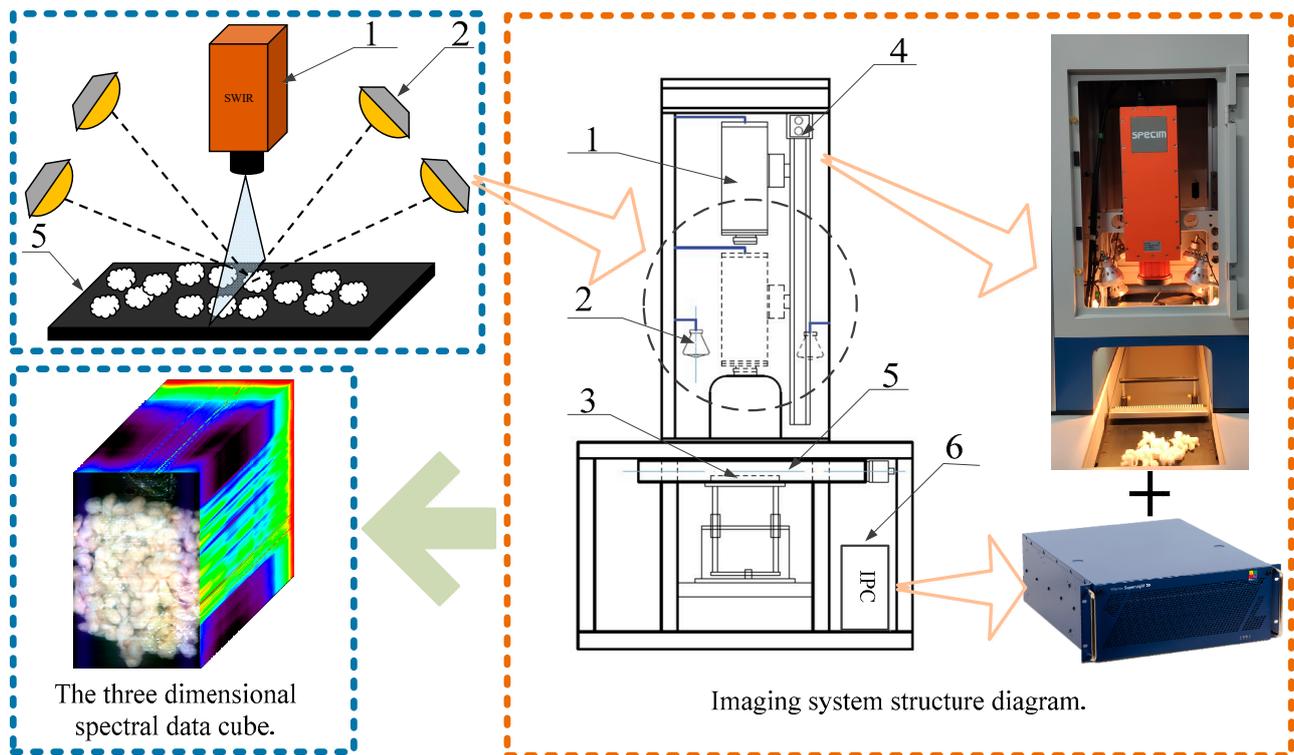
#### 2.1.1. Experimental Materials

Gaia Sorter-Dual, a full-band hyperspectral Sorter, is used in conjunction with the hyperspectral camera “Image- $\lambda$ -N25E-SWIR”. A total of 10 kg of machine-picked long-staple cotton from southern Xinjiang and 50 pieces of film of different sizes are picked out by skilled workers.

As shown in Figure 1, the hyperspectral imaging system can obtain the hyperspectral images of seed cotton mixed with the film: the resolution is 384 pixels  $\times$  600 pixels, the spectral range is 1000~2500 nm, 288 bands. The hyperspectral camera is positioned directly above the platform, with four halogen lamps symmetrically placed around it. The angle of irradiation of the halogen lamp can be adjusted arbitrarily. All halogen sources are adjusted to the position directly below the camera. The distance regulating mechanism is responsible for controlling the vertical motion of the hyperspectral camera in order to adjust the camera’s image surface. Additionally, the transfer platform can continuously move horizontally to capture continuous one-dimensional images. Since experimental subjects come in different sizes, the electronic control platform allows for vertical movement to create storage space for the subjects. The collected hyperspectral images can be regarded as 230,400 pieces of sample data, including 92,456 seed cotton samples, 63,478 film on cotton samples, 62,897 background samples and 11,569 film on background samples. The specific operation steps are as follows:

(1) The samples are placed on the transfer platform and irradiated uniform light from a halogen lamp. The reflected light is then captured by a hyperspectral camera, which provides one-dimensional spectral information.

(2) The transfer platform moves horizontally to obtain continuous one-dimensional spectral information, which is then transmitted to an industrial computer to generate hyperspectral images containing all the spectral information.



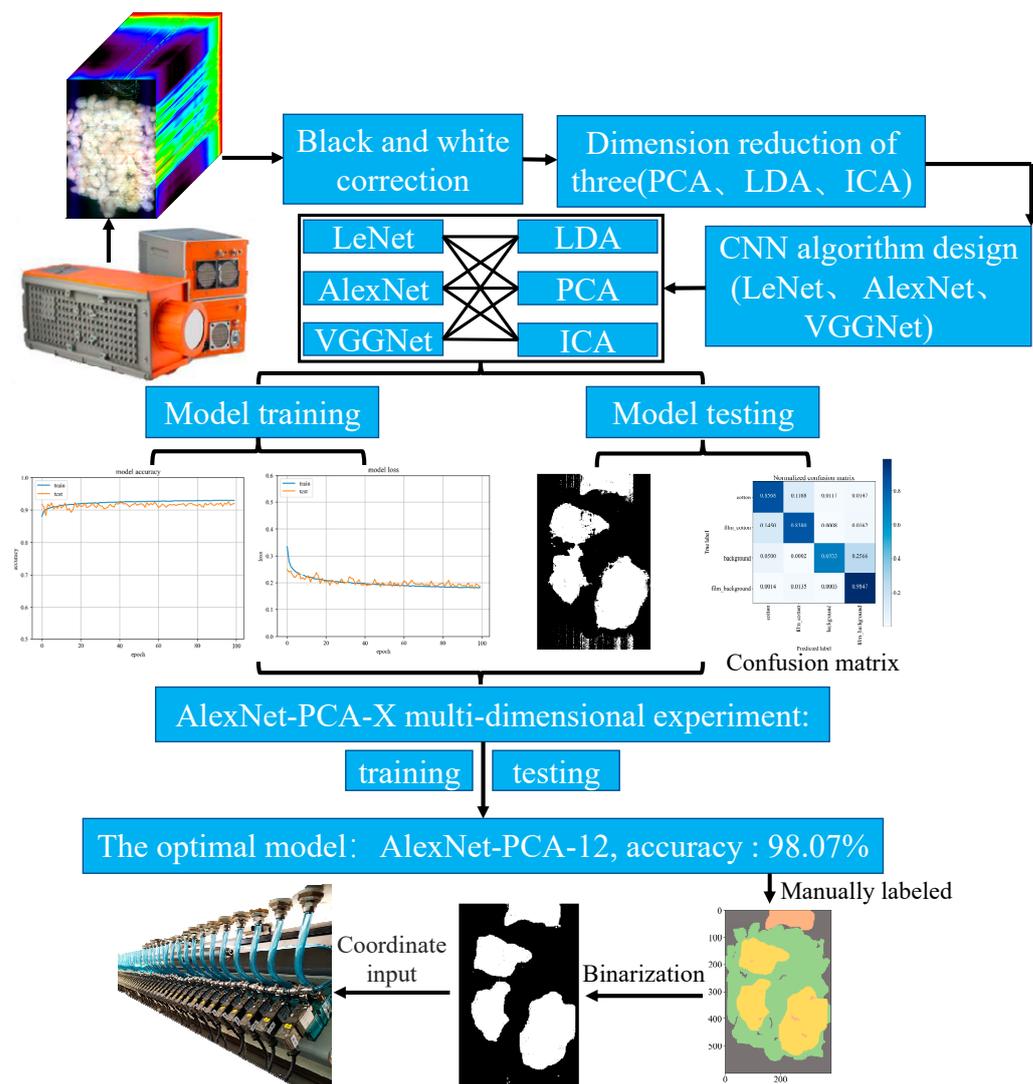
**Figure 1.** Hyperspectral imaging system. (1) Hyperspectral camera, (2) Halogen lamp, (3) Electronic control platform, (4) Distance regulating mechanism, (5) Transfer platform, (6) Industrial computer.

### 2.1.2. Algorithm Environment

Hardware environment for Intel® Core (TM)i7-6700 CPU by Intel Semiconductor Co., Ltd. in Dalian, China, 16 GB RAM, and NVIDIA GeForce RTX 2080Ti by Taiwan Integrated Circuit Manufacturing Co., Ltd. in Taiwan, China, 11 GB was obtained. Software environment for tensorflow-gpu 2.0.0, spectral 0.22.1, sklearn 0.23.2, matplotlib 3.2.2, kears 2.3.1, cuda 10.2.89, cudnn 7.6.5 was used employing the Python 3.6 programming language.

### 2.1.3. Technical Route

The technical route of the film sorting system is illustrated in Figure 2, where seed cotton mixed with the film is given to the study subject. Firstly, a hyperspectral camera is used to collect 1000–2500 nm hyperspectral images. Secondly, the experimental validation involves nine models, including black and white correction, dimension reduction, and CNN training and testing. The purpose is to determine the best models for hyperspectral data dimension reduction and CNN. The binarization is established to display the recognition outcome of the optimal AlexNet-PCA-12 model, which has an eminent recognition accuracy of 98.07%. Finally, the coordinates of the film are fed into a high-speed spray valve to complete the film removal in the practical application sorting tests.



**Figure 2.** The technical route of the film sorting system.

### 2.2. Black and White Correction of Hyperspectral Images

The stability of data can be impacted by environmental factors including light intensity and angular variations. In addition, there is a dark current in the camera and noise interference in the acquisition. Therefore, the hyperspectral images need to be corrected, which can remove ambient light interference and most of the noise in the image and effectively improve the classification and recognition accuracy of the subsequent model. The original hyperspectral images can be corrected by [20]

$$I_{ref} = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}}, \tag{1}$$

where  $I_{ref}$  represents the corrected image,  $I_{raw}$  is the original image,  $I_{white}$  denotes standard correction image,  $I_{dark}$  indicates background correction image.

### 2.3. Dimension Reduction of Hyperspectral Data

Dimension reduction can effectively eliminate noise and irrelevant information while also preventing data redundancy and dimension explosion caused by high-dimensional data during algorithmic processing [21]. At present, the dimension reduction methods of hyperspectral data mainly conclude linear discriminant analysis (LDA), principal component analysis (PCA), independent component analysis (ICA), etc. By extracting and

mapping the main feature bands of the original data, these methods can effectively reduce the operating cost of the algorithm while ensuring the recognition accuracy of the algorithm.

### 2.3.1. Linear Discriminant Analysis

LDA is a linear learning method that employs pattern recognition, machine learning, and other techniques to extract similar features of two objects or events from multiple datasets. These features are then combined to more accurately identify the differences between them [22].

Hyperspectral data contains an LDA multi-classification task, which needs to project the vector  $x$  of the  $D$  dimension to  $y$  of the  $d$  ( $d < D$ ) dimension, and the projection equation can be provided by

$$y = W^T x, \quad (2)$$

where  $W$  is the projection matrix and the projection direction of each column vector is provided by  $w_i$ .

Multi-classification task data sets  $X$  can be written as

$$X = \{x_1^{(1)}, x_2^{(2)}, \dots, x_{M_1}^{(1)}, x_1^{(2)}, \dots, x_{M_N}^{(N)}\}, \quad (3)$$

where  $N$  represents the number of sample types,  $i$  indicates the kind of sample,  $x_j^{(i)}$  denotes the  $j$  sample of class  $i$ ,  $M_i$  is the number of class  $i$  training samples ( $i = 1, 2, \dots, N$ ).

The in-class divergence matrix  $S_w$  is obtained as [22]

$$S_w = \sum_{i=1}^N \sum_{j=1}^{M_i} p(i, j) (x_j^{(i)} - \mu_i) (x_j^{(i)} - \mu_i)^T, \quad (4)$$

where  $\mu_i$  presents the mean of training samples of class  $i$ ,  $p(i, j)$  is the probability of  $x_j^{(i)}$ .

The overall divergence matrix  $S_t$  is given by [22]

$$S_t = \sum_{i=1}^N \sum_{j=1}^{M_i} p(i, j) (x_j^{(i)} - \mu) (x_j^{(i)} - \mu)^T, \quad (5)$$

where  $\mu$  represents the mean value of all training samples.

Interclass divergence matrix  $S_b$  [22] is

$$\begin{aligned} S_b &= S_t - S_w \\ &= \sum_{i=1}^N p(i) (\mu_i - \mu) (\mu_i - \mu)^T, \end{aligned} \quad (6)$$

where  $p(i)$  denotes the probability of class  $i$ .

Then, we obtain the objective function  $J$  [22]:

$$J = \frac{W^T S_b W}{W^T S_w W} \quad (7)$$

$$S_W^{-1} S_b W = \lambda W$$

The projection matrix  $W$  of the  $d$  dimension can be obtained by calculating the largest  $d$  eigenvalues of  $S_W^{-1} S_b$  and the corresponding  $d$  eigenvectors;  $d$  ( $d < N$ ) is the dimension after dimensionality reduction of hyperspectral data.

### 2.3.2. Principal Component Analysis

PCA is a dimension reduction algorithm based on the discrete Karhunen–Loeve transform for extracting the main feature components of multivariate data information [23].

Although the majority of the noise in the image can be removed using PCA, it has greater advantages in terms of time complexity.

**Data conversion.** While reading is performed in the hyperspectral image data, each band data is converted into a one-dimensional vector. The hyperspectral image data are assumed to have a total of  $N$  bands with a  $w \times h$  resolution, which can be represented as a matrix of  $(w \times h) \times N$ . Here, the band  $i$  can be expressed as

$$x^i = [x_1^i, x_2^i, \dots, x_{w \times h}^i], (i = 1, 2 \dots N). \quad (8)$$

**For the eigenspace.** The mean vector of all bands is calculated as [24]

$$\bar{x} = \frac{1}{N} \sum_i^N x^i. \quad (9)$$

The distance vector between each band and the average band can be obtained as

$$d_i = x^i - \bar{x}. \quad (10)$$

We set the matrix  $B$  as

$$B = [d_1, d_2, \dots, d_N]. \quad (11)$$

Then, the covariance matrix can be obtained as follows [24]:

$$\frac{1}{N} BB^T = \frac{1}{N} \sum_i^N d_i d_i^T. \quad (12)$$

The transpose matrix in Formula (12) can be written as

$$(BB^T)^T = B^T B. \quad (13)$$

Since Formula (12) is a high-dimensional vector of  $(w \times h) \times (w \times h)$ , the calculation of eigenvectors of the first  $Z (Z \leq N)$  large eigenvalues of the covariance matrix is too large, while Formula (13) is a low-dimensional vector of  $N \times N$ , and therefore its eigenvalue can be calculated first [25]:

$$v_j = B u_j \lambda_j^{-\frac{1}{2}}, \quad (j = 1, 2, \dots, Z), \quad (14)$$

where  $\lambda_j$  presents the eigenvalue of Formula (12) and  $u_j$  is the eigenvector of Formula (13). The eigenspace  $v_j$  can be formed by the eigenvalues of Formula (13):

$$W = \{v_1, v_2, \dots, v_Z\}. \quad (15)$$

**Projection and similarity detection.** The difference vector between each band and the average band is projected into the eigenspace, and the eigenvector  $i$  is expressed as

$$P_i = W^T d_i, \quad (i = 1, 2, \dots, N). \quad (16)$$

The Euclidean distance is written as [25]

$$\varepsilon_i = \|P_i - P_k\|^2, \quad (i, k = 1, 2, \dots, N). \quad (17)$$

When using PCA dimension reduction, similarity between images is determined by the Euclidean distance. A smaller Euclidean distance indicates a greater similarity and better results. After this operation,  $n$  eigenvector  $P$  with minimum Euclidean distance is tested to form a fresh hyperspectral data set, where  $n < N$  is the dimension of hyperspectral data after dimensionality reduction.

### 2.3.3. Independent Component Analysis

ICA is a method to find data intrinsic components from multi-dimensional statistical data which focuses on data analysis from independent sources, decomposing multivariate signals into different non-Gaussian signals [26]. Hyperspectral image data  $X$  can be regarded as a two-dimensional matrix with  $N$  rows and  $L$  columns ( $L = w \times h$ ). Hyperspectral data with band  $n$  ( $n < N$ ) can be obtained through ICA to achieve the purpose of dimension reduction.

ICA of  $X$  can be expressed as [15]

$$X = AS = \sum_{d=1}^N a_d s_d, \quad (18)$$

where  $N$  is the number of bands.  $d$  denotes band index number ( $d = 1, 2, \dots, N$ ),  $A = (a_1, a_2, \dots, a_d, \dots, a_N)$  presents a mixing matrix,  $a_d = (a_{1d}, a_{2d}, \dots, a_{Nd})^T$  denotes the column vector of  $A$ ,  $S = (s_1, s_2, \dots, s_d, \dots, s_N)^T$  indicates an independent component matrix,  $s_d = (s_{d1}, s_{d2}, \dots, s_{dN})^T$  is the row vector of  $S$ .

We set  $W = A^{-1}$  according to Formula (18) [15]:

$$S = A^{-1}X = WX, \quad (19)$$

where  $(w_1, w_2, \dots, w_d, \dots, w_N)$  is defined as the transformation matrix  $W$ ,  $(w_{1d}, w_{2d}, \dots, w_{Nd})^T$  presents the column vector of  $W$ . The independent component  $S$  is obtained by finding the appropriate transformation matrix  $W$  for the independent statistical and non-Gaussian properties of each component according to the principle of the central limit theorem.

Depending on the choice of the objective function, ICA includes FastICA, Projection pursuit, and Infomax, which mainly extract independent components by increasing non-Gaussian properties, reducing mutual information, and performing maximum likelihood estimation [15]. The FastICA approach which adopts batch processing to incorporate a huge quantity of sample data into the iterative process is utilized to optimize independent components. It also establishes negative entropy as a non-Gaussian measure of random variables. The steps of solving independent components by the FastICA algorithm can be described as follows:

**(1) Bleaching data.** We set the average value of hyperspectral image data  $X$  as  $\bar{X}$  and perform decentralized processing on the data to obtain

$$P = X - \bar{X}. \quad (20)$$

The covariance matrix for  $P$  can be written as [27]

$$C = \text{cov}(P, P^T). \quad (21)$$

The eigenvalue  $\lambda$  and eigenvalue diagonal matrix  $D$  are calculated through  $|\lambda I - C| = 0$  where  $I$  denotes the unit vector; the eigenvector  $E$  can be found by  $(\lambda I - C)E = 0$ .

There is a bleaching transformation matrix  $U = D^{-1/2}E^T$ , and the data after bleaching are obtained as [27]

$$Z = U \times P. \quad (22)$$

**(2) Finding the matrix  $W$ .** We let  $k$  be the number of iterations, and the iterative computation of  $w^{(k)}$  can be expressed as [27]

$$w^{(k)} = E \left\{ ZG \left( w_d^{(k-1)T} Z \right) \right\} - E \left\{ Zg \left( w_d^{(k-1)T} Z \right) \right\} \times w_d^{(k-1)}, \quad (23)$$

where  $G(t) = \tanh(t) = (e^t - e^{-t}) / (e^t + e^{-t})$  denotes a hyperbolic tangent function,  $g(t)$  is the first derivative of  $G(t)$ ,  $E(\bullet)$  indicates mean function.

We orthogonalize and standardize the matrix  $W$  [27]:

$$\begin{aligned} \sum_{j=1}^{d-1} \left( w_d^{(k)T} w_j \right) w_j &\rightarrow w_d^{(k)} \\ \frac{w_d^{(k)}}{\|w_d^{(k)}\|} &\rightarrow w_d^{(k)} \end{aligned} \quad (24)$$

For any real number  $\varepsilon$  greater than 0, if  $\left| w_d^{(k)T} w_d^{(k-1)} - 1 \right| < \varepsilon$ , the  $w_d$  converges; otherwise,  $k = k + 1$  takes Formula (23) to continue the iteration. The column vector  $w_d$  of  $W$  can be obtained from Formula (24), and  $d$  ( $d = 1, 2, \dots, N$ ) is the index number of each band.

When  $d = N$ , the matrix  $W$  is calculated as follows:

$$W = (w_1, w_2, \dots, w_d, \dots, w_N). \quad (25)$$

By taking the matrix  $W$  into Formula (19), the independent component  $S$  can be solved.

**(3) Selecting band.** The matrix  $W$  is defined as  $(w_1, w_2, \dots, w_j, \dots, w_N)$ , the column vector  $d$  of  $W$  is defined as  $(w_{1j}, w_{2j}, \dots, w_{ij}, \dots, w_{Nj})^T$  ( $i, j = 1, 2, \dots, N$ ), where  $w_{ij}$  indicates the capacity of the  $j$  band containing  $i$  independent component information. By calculating the average absolute weight factor, it can assess how much of each band contains independent component information:

$$\bar{w}_j = \frac{1}{N} \sum_{i=1}^N |w_{ij}|. \quad (26)$$

The gained  $n$  bands with the largest average weight coefficient  $\bar{w}_j$  are formed into a new low-dimensional image to achieve the dimensionality reduction of the hyperspectral image;  $n$  ( $n < N$ ) is the number of bands after the dimensionality reduction of the hyperspectral data.

#### 2.4. Construction of the Convolutional Neural Network

The convolutional neural network mainly concludes with an input layer, convolution layer, pooling layer, fully connected layer, and output layer, which can effectively solve the over-fitting problem [28]. The research illustrates a 2D-CNN-based method for hyperspectral image classification which can reduce the training cost while ensuring high classification and recognition accuracy.

**Convolution layer.** The convolution layer applies a convolution kernel to transform the input matrix into a unit matrix for the next layer. During forward propagation, the convolution kernel computes the nodes in the right unit matrix by using the nodes in the left input matrix [29]. Multiple convolution kernels are used to convolve with input image data, and a series of feature graphs are obtained through an activation function after biasing [30]. In the paper, the ReLU activation function is utilized to map the input of neurons to the output; its nonlinear characteristics are introduced into the neural network, enabling its application to various nonlinear models. The convolution formula is expressed as follows [29]:

$$X_j^l = f \left( \sum_{i \in M_j} X_i^{l-1} \bullet w_{ij}^l + b_j^l \right), \quad (27)$$

where  $X_j^l$  denotes the  $j$  element of the  $l$  layer,  $M_j$  stands for  $j$  convolution area of the  $l - 1$  layer feature map,  $X_i^{l-1}$  presents the elements,  $w_{ij}^l$  is the weight of the corresponding

convolutional kernel matrix,  $b_j^l$  is the offset item,  $f(\cdot)$  indicates the activation function,  $\sum_{i \in M_j} X_i^{l-1} \bullet w_{ij}^l$  is the convolution formula.

**Pooling layer.** If all the features obtained through convolution are inputted into the classifier, a significant amount of computation is required to handle it. In this case, the Pooling function is required to process the feature maps obtained by convolution, and the Max pooling method is utilized in this paper. The pooled element matrix can reduce the dimension of the feature information obtained from the convolution layer and reduce the size of the matrix in the direction of height and width while ensuring the invariance of the feature scale. Meanwhile, the number of parameters of the whole neural network can be reduced, thus improving the generalization ability of the model [31].

**Fully connected layer.** With multi-layer convolution and pooling processing, images are gradually extracted with higher-level and more abstract feature information, which is classified by fully connected layers [32]. After unrolling the input feature vector into one dimension, the fully connected layer outputs the result via weighted summation and activation functions. The output formula is [29]

$$y^k = f(w^k x^{k-1} + b^k), \quad (28)$$

where  $k$  is the serial number of the network layer,  $y^k$  is the output,  $x^{k-1}$  represents the expanded one-dimensional eigenvector,  $w^k$  stands for the weight coefficient,  $b^k$  is the offset item.  $f(\cdot)$  is a model for probabilistic computation and an activation function suitable for classification tasks, which can be formulated as follows:

$$y = \text{softmax}(w_{ij}x_j + b). \quad (29)$$

The softmax loss function is structured in the fully connected layer to measure the solving accuracy of the problem, and loss function is adopted to describe the degree of dissatisfaction with the classification result. The effect of the neural network model is defined by the loss function. The tinier the loss value, the tinier the deviation between the result obtained by the model and the real value [33].

The purpose of neural network optimization is to accurately and timely update the parameters. Two optimization methods are employed for neural networks in the paper: the first step is the Gradient descent algorithm, and the second is the Back propagation algorithm. The optimization method of Gradient descent is to randomly select a function on the training data during the iteration process, which ensures the rapid update of parameters in each iteration. The back propagation algorithm based on the gradient descent algorithm can not only calculate vector gradients, but also calculate multidimensional tensors [34].

To avoid training overfitting, the Dropout function is used in the fully connected layers to make the output of neurons in the hidden layer drop to zero with a certain percentage probability. Dropout disables some hidden layer nodes that do not participate in the forward propagation process of the CNN. Due to the stochastic nature of the Dropout, each sample input to the network corresponds to a different network structure, but all these structures share weight. Since a neuron cannot depend on additional specific neurons, it reduces the complexity of inter-neuron adaptation and enables them to learn deeper features [35].

### 2.5. Design of Intelligent Recognition Algorithm for Film in Seed Cotton

In this section, three CNN models based on LeNet, AlexNet, and VGGNet are constructed for hyperspectral image recognition. The CNN schematic is shown in Figure 3. The schematic involves two steps. Firstly, the hyperspectral data are used to train the model and extract useful image features. Secondly, the trained features are applied to the testing set for verification, and the resulting recognition accuracy is outputted. Additionally,

the network parameters are regulated through gradient descent and back propagation algorithms, which ascertain network parameters in time.

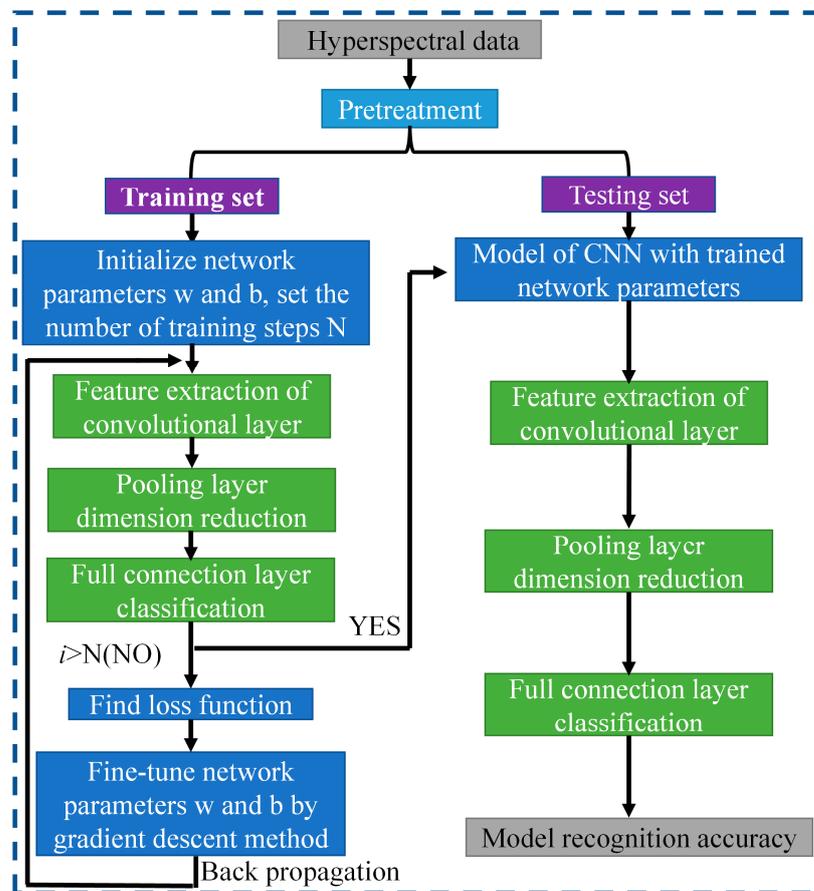


Figure 3. CNN schematic chart.

To achieve optimal recognition results for hyperspectral image recognition, LeNet, AlexNet, and VGGNet models are altered accordingly. The specific parameters for each model are outlined in Tables 1–3. For facilitating CNN to input hyperspectral data and output recognition accuracy, CNN is set corresponding to the input layer and output layer, specifically as follows: In the input layer,  $5 \times 5$  indicates the data size of the input convolutional network by manual division and D denotes the data dimension obtained after adopting different dimensionality reduction algorithms. In the output layer, the Softmax loss function outputs the probabilities of four units, which include “cotton”, “film on cotton”, “background”, and “film on background”.

Table 1. LeNet structural parameter.

Type	Variables	Kernel Parameter	Data Output
Input layer		$5 \times 5 \times D$ hyperspectral data set	
1-Conv	3 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
2-Pool	Max pooling	Size: $3 \times 3$ . All zero-filling. Step: 2	
3-Conv	9 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
4-Pool	Max pooling	Size: $3 \times 3$ . All zero-filling. Step: 2	Dropout drops 25% weight
Flatten layer		Convert multi-dimensional input into one dimension	
FC		Input neuron number: 108. Output neuron number: 18	
Output layer		Softmax loss function outputs the probabilities of four units.	

LeNet structure mainly consists of 2 Convs, 2 Pools, and 1 FC. “Conv” is convolution layer, “Pool” denotes pooling layer, “FC” indicates fully connected layer.

**Table 2.** AlexNet structural parameter.

Type	Variables	Kernel Parameter	Data Output
Input layer		$5 \times 5 \times D$ hyperspectral data set	
1-Conv	3 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
1-Pool	Max pooling	Size: $3 \times 3$ . All zero-filling. Step: 2	
2-Conv	9 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
2-Pool	Max pooling	Size: $3 \times 3$ . All zero-filling. Step: 2	
3-Conv	12 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
4-Conv	12 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
5-Conv	9 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
5-Pool	Max pooling	Size: $3 \times 3$ . All zero-filling. Step: 2	
Flatten layer		Convert multi-dimensional input into one dimension	
FC		Input neuron number: 27. Output neuron number: 60	
Output layer		Dropout drops 50% weight, Softmax loss function outputs the probabilities of four units.	

AlexNet structure mainly consists of 3 convolution groups (including 1 Conv and 1 Pool), 2 Convs, and 1 FC. "Conv" is convolution layer, "Pool" denotes pooling layer, "FC" indicates fully connected layer.

**Table 3.** VGGNet structural parameter.

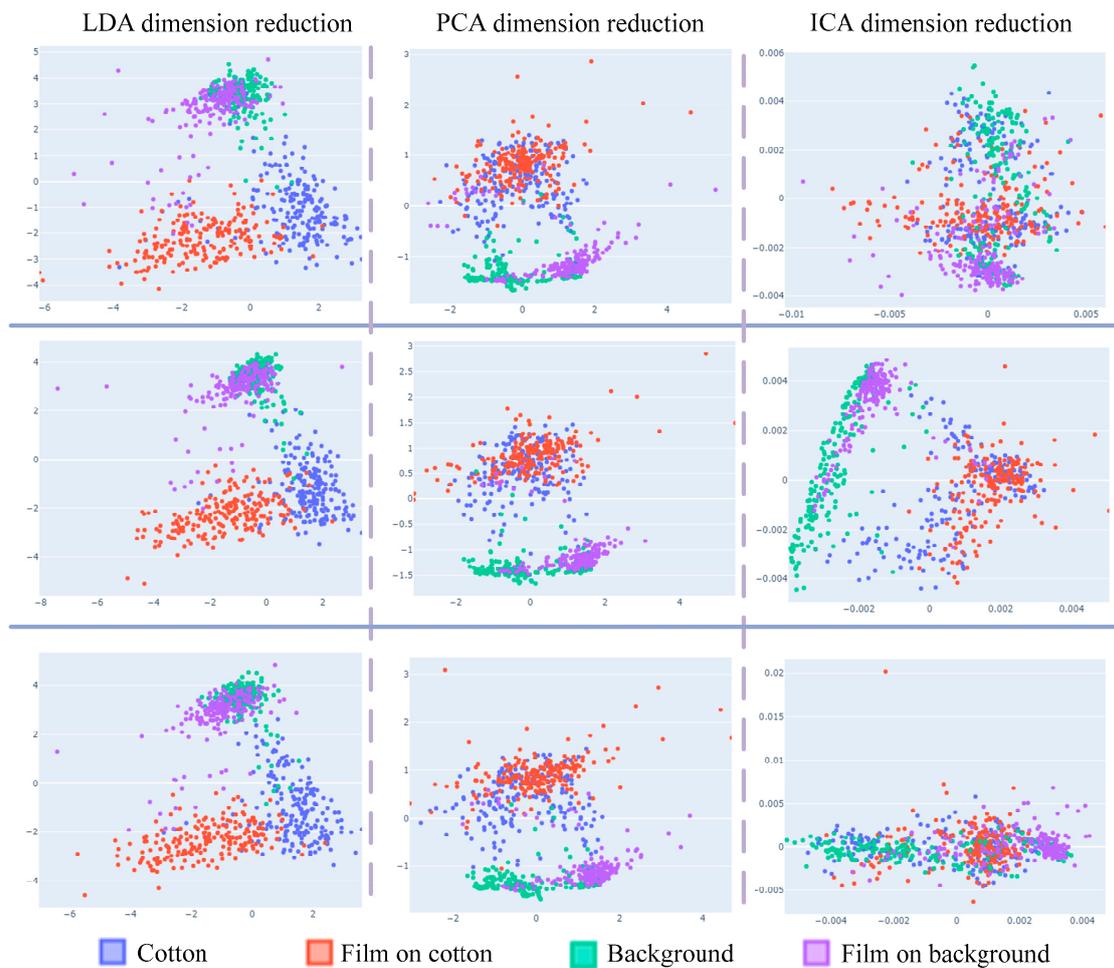
Type	Variables	Kernel Parameter	Data Output
Input layer		$5 \times 5 \times D$ hyperspectral data set	
1-Conv	3 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
1-Conv	3 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
2-Pool	Max pooling	Size: $2 \times 2$ . All zero-filling. Step: 2	Dropout drops 20% weight
3-Conv	6 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
3-Conv	6 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
4-Pool	Max pooling	Size: $2 \times 2$ . All zero-filling. Step: 2	Dropout drops 20% weight
5-Conv	12 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
5-Conv	12 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
5-Conv	12 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
6-Pool	Max pooling	Size: $2 \times 2$ . All zero-filling. Step: 2	Dropout drops 20% weight
7-Conv	24 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
7-Conv	24 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
7-Conv	24 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
8-Pool	Max pooling	Size: $2 \times 2$ . All zero-filling. Step: 2	Dropout drops 20% weight
9-Conv	24 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
9-Conv	24 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
9-Conv	24 convolution kernels	Size: $3 \times 3$ . All zero-filling. Step: 1	ReLU activation function
10-Pool	Max pooling	Size: $2 \times 2$ . All zero-filling. Step: 2	Dropout drops 20% weight
Flatten layer		Convert multi-dimensional input into one dimension	
FC		Input neuron number: 72. Output neuron number: 24	
Output layer		Dropout drops 20% weight, Softmax loss function outputs the probabilities of four units.	

VGGNet structure mainly consists of 5 convolution groups (including 2 or 3 Convs), 5 Pools, and 1 FC. "Conv" is convolution layer, "Pool" denotes pooling layer, "FC" indicates fully connected layer.

### 3. Results and Discussion

#### 3.1. Design of Intelligent Recognition Algorithm for Film in Seed Cotton

To verify the generalization of the dimensional reduction, scatter plots are presented in Figure 4. The plots depict the application of different dimensional reduction methods on the same hyperspectral data from the three-dimensional reduction experiments. The scatter plots of dimensionality reduction for different batches of the same data under the same experimental conditions can be concluded as follows:



**Figure 4.** The scatter plot of three-dimensional reduction experiments.

(1) Considering only the first two samples, LDA data have obvious clustering and separability, but LDA data cannot classify sample “background” accurately.

(2) ICA data classified the four types of samples differently in different batches, so it is not general to data from different batches of dimensionality reduction, and the trained model cannot achieve ideal results on the test.

(3) Considering only the first two samples, PCA dimension reduction has distinct aggregation and separability on “background” and “film on background”, while the data coincidence of the two samples “cotton” and “film on cotton” has no classification.

(4) The result shows that LDA has outstanding classification results with a dimensionality reduction of two for hyperspectral data. However, LDA can only reduce the data to three dimensions. Therefore, when the computer performance is satisfied, PCA obtains higher recognition accuracy when it is used to retain more dimensions.

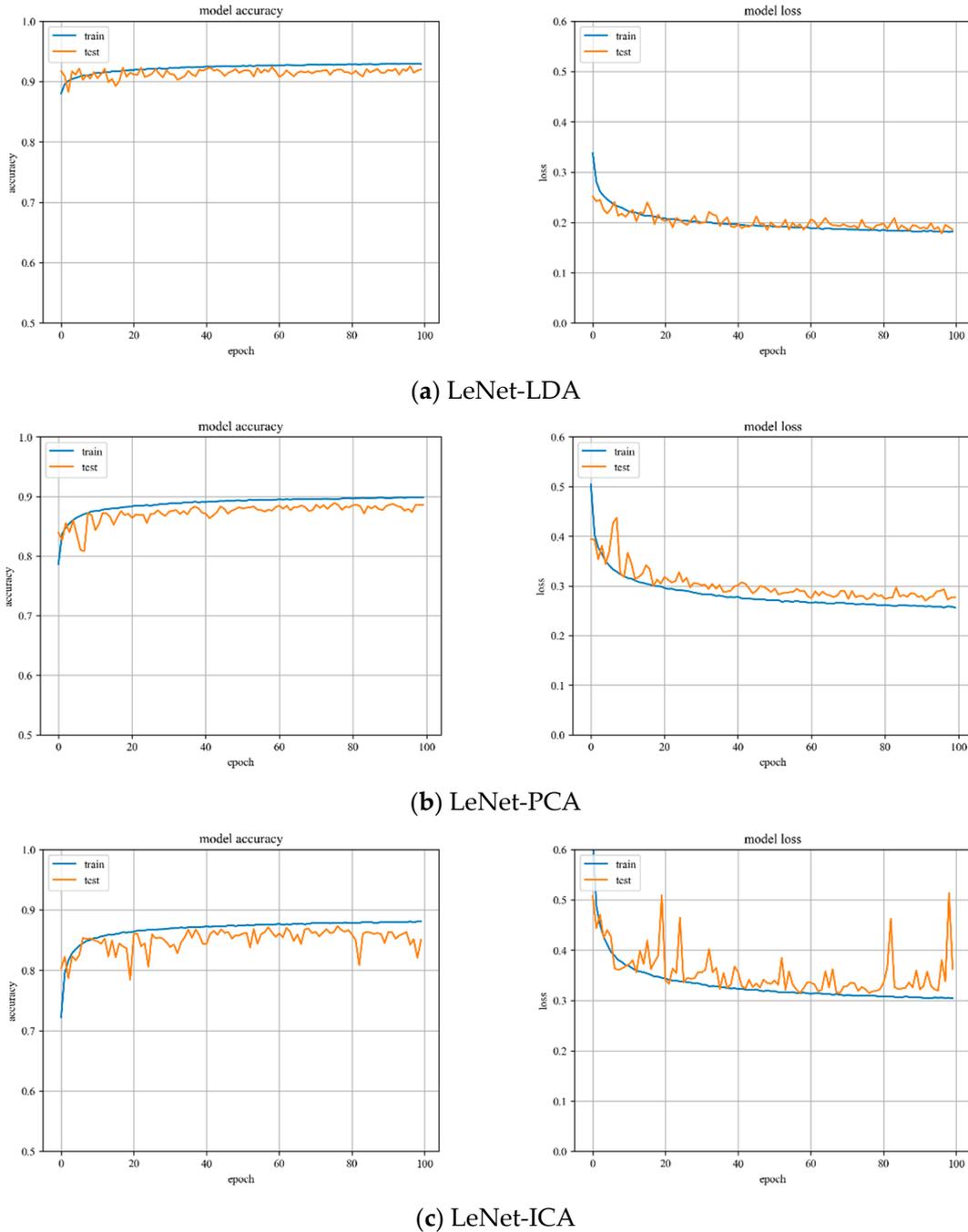
### 3.2. CNN Model Training

Comparing three different dimension reduction methods (LDA, PCA, and ICA) after the hyperspectral data to a three-dimensional effect, three different structures are adopted CNN (LeNet, AlexNet, and VGGNet) for training and testing accuracy.

**LeNet model training.** Variations of training and testing accuracy of LeNet with the number of training epochs, training and testing loss curves are shown in Figure 5.

In Figure 5a, LDA recognition accuracy on the test set is about 92%, and the loss value is 0.15~0.2. In Figure 5b, PCA recognition accuracy on the test set is about 89%, and the loss value is 0.25~0.3. In Figure 5c, ICA recognition accuracy on the test set is about 85%, and the loss value is 0.3~0.35.

Despite the fact that LDA and PCA are relatively stable to changes throughout the training phase, PCA slightly underperforms the LeNet model with LDA hyperspectral data reduction. However, the LeNet model with ICA hyperspectral data reduction has the worst stability of the three.

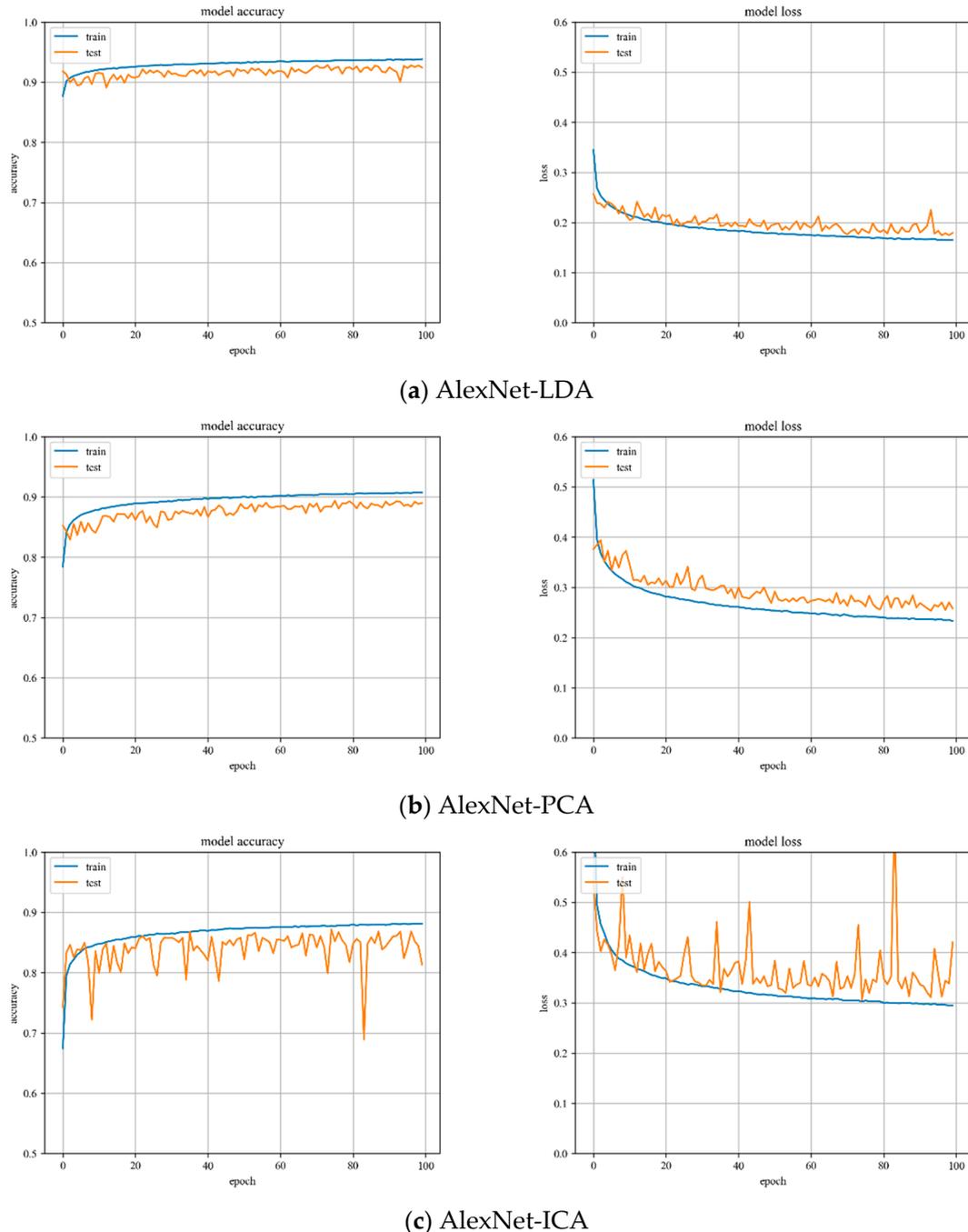


**Figure 5.** Accuracy curve and loss curve of LeNet model.

**AlexNet model training.** Variations of training and testing accuracy of AlexNet with the number of training epochs, training and testing loss curves are shown in Figure 6.

In Figure 6a, LDA recognition accuracy on the test set is about 93%, and the loss value is 0.15~0.2. In Figure 6b, PCA recognition accuracy on the test set is about 90%, and the loss value is 0.25~0.3. In Figure 6c, ICA recognition accuracy on the test set is about 88%, and the loss value is 0.3~0.35.

Despite the fact that LDA and PCA are relatively stable to changes throughout the training phase, PCA slightly underperforms the AlexNet model with LDA hyperspectral data reduction. However, the AlexNet model with ICA hyperspectral data reduction has the worst stability of the three.



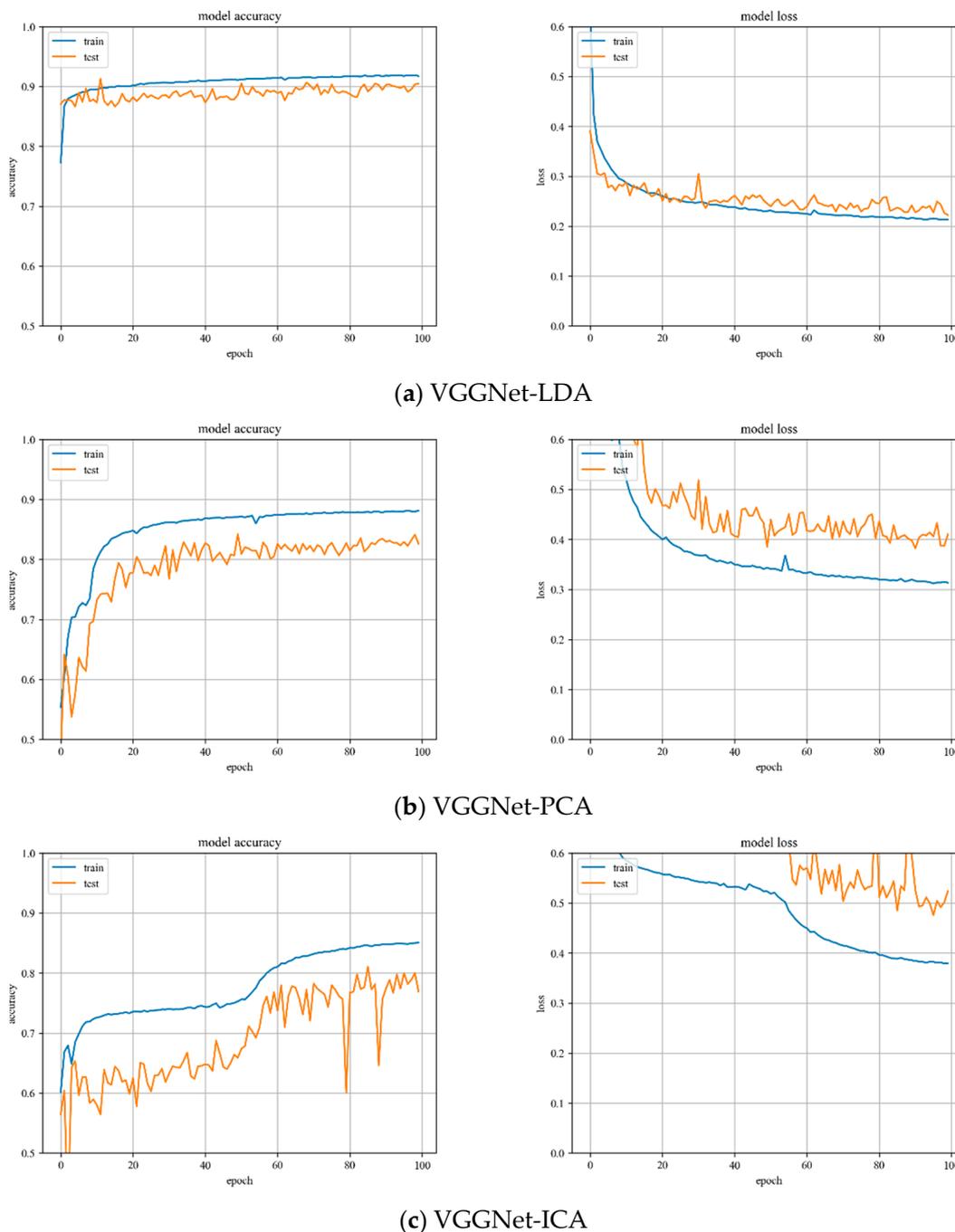
**Figure 6.** Accuracy curve and loss curve of AlexNet model.

**VGGNet model training.** Variation of training and testing accuracy of VGGNet with the number of training epochs, training and testing loss curves are shown in Figure 7.

In Figure 7a, LDA recognition accuracy on the test set is about 90%, and the loss value is 0.2~0.25. The LDA model is relatively stable to changes throughout the training process and has excellent model stability.

In Figure 7b, PCA recognition accuracy on the test set is about 84%, and the loss value is about 0.4. In Figure 7c, ICA recognition accuracy on the test set is about 80%, and the loss

value fluctuates widely. Both models have minor stability during the training phase. The ICA results are inferior than the VGGNet model with PCA hyperspectral data reduction.



**Figure 7.** Accuracy curve and loss curve of VGGNet model.

### 3.3. CNN Model Testing

The confusion matrix for the test samples for the different algorithmic models are illustrated in Tables 4–6. It can be seen that 1 is the cotton, 2 represents the film on cotton, 3 indicates the background, and 4 denotes the film on background, the diagonal expresses the probability of correct classification. The experimental data analysis is as follows:

(1) Using LDA and PCA dimensionality reduction hyperspectral data, it can be determined that the three kinds of CNN models have higher recognition accuracy for test samples. However, there are some errors in the classification of film samples on cotton and film samples on background, which is consistent with the conclusion of the scatter plots above.

(2) Since the hyperspectral data for ICA dimension reduction is not the same batch as the data during training, the extracted dimension information is unstable and the recognition effect is confused. Therefore, it cannot be applied to hyperspectral image recognition, which is consistent with the conclusion of the scatter plots above.

**Table 4.** LDA sample confusion matrix (%).

Model			Predictive			
	Actual		1	2	3	4
LeNet		1	91.89	5.72	2.22	0.17
		2	4.78	93.26	0.12	1.83
		3	3.99	0.05	90.49	5.47
		4	0.09	0.86	2.07	96.97
AlexNet		1	93.10	4.60	2.08	0.22
		2	5.17	92.87	0.18	1.77
		3	3.60	0.03	90.03	6.34
		4	0.00	0.81	1.24	97.95
VGGNet		1	90.44	6.84	2.48	0.23
		2	4.06	95.13	0.04	0.77
		3	4.40	0.07	84.55	10.98
		4	0.03	1.90	0.95	97.12

**Table 5.** PCA sample confusion matrix (%).

Model			Predictive			
	Actual		1	2	3	4
LeNet		1	89.13	8.86	1.57	0.44
		2	15.31	83.88	0.08	0.72
		3	3.83	0.08	91.15	4.94
		4	0.23	1.33	1.82	96.63
AlexNet		1	91.80	6.72	1.18	0.30
		2	14.98	84.41	0.08	0.53
		3	3.97	0.10	87.81	8.11
		4	0.09	1.35	0.17	98.39
VGGNet		1	87.95	8.75	2.11	1.19
		2	16.43	82.05	0.08	1.44
		3	4.11	0.02	72.40	23.47
		4	0.40	0.81	0.00	98.79

**Table 6.** ICA sample confusion matrix (%).

Model			Predictive			
	Actual		1	2	3	4
LeNet		1	75.24	0.07	23.07	1.62
		2	69.72	2.80	14.28	13.20
		3	90.51	0.75	8.62	0.12
		4	74.90	23.22	0.13	1.75
AlexNet		1	70.05	5.87	16.41	7.67
		2	32.78	32.72	15.84	18.65
		3	78.95	0.40	19.75	0.90
		4	65.08	24.49	5.19	5.25
VGGNet		1	48.65	18.68	21.23	11.44
		2	29.46	38.76	8.64	23.14
		3	28.10	10.37	51.14	10.39
		4	18.10	58.86	9.01	14.03

The Overall Accuracy (OA) of the test samples is illustrated in Table 7, representing the percentage of all samples that are accurately predicted. The results can be summarized as follows:

(1) When the hyperspectral data are reduced to three dimensions, the average OA of LDA is 91.68%, while PCA has an average OA of 87.08%; on the other hand, ICA has a lower average OA of 40.35%. Based on these results, it can be concluded that LDA demonstrates superior performance in terms of dimensionality reduction.

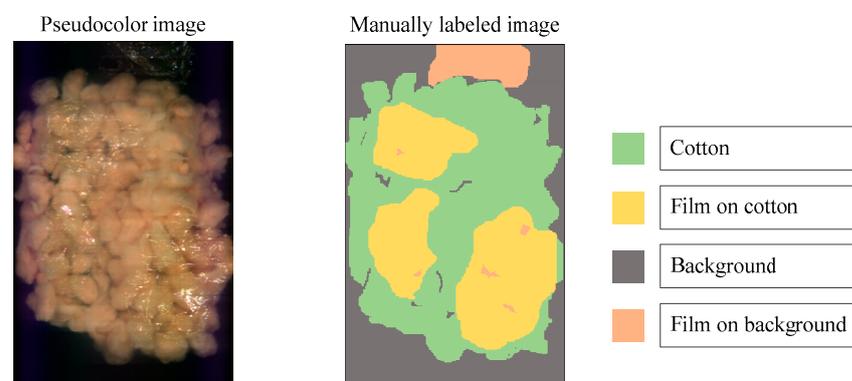
(2) The data in the table show that the CNN-based AlexNet model can achieve excellent recognition effects when the data are dimensionally reduced.

(3) When the dimension reduction of ICA is 3, it exhibits poor performance in terms of average OA compared to the other two dimensionality reduction methods. However, PCA can retain more dimension information to improve the recognition accuracy, which has more potential in practical applications.

**Table 7.** Overall accuracy of the test samples (%).

	LDA	PCA	ICA
LeNet	92.14	88.61	33.41
AlexNet	92.45	89.00	42.78
VGGNet	90.46	83.63	44.86
Average	91.68	87.08	40.35

To distinguish the classification effects more intuitively, three bands are selected to display the hyperspectral data as pseudocolor images. Additionally, the spectral toolkit is used in the model tests to plot the predictions in the form of a two-dimensional image. The pseudocolor and manually labeled images are shown in Figure 8.



**Figure 8.** Pseudocolor and manually labeled images.

Considering the actual sorting system only needed to locate the spatial coordinate position of the film, the classification results are combined from four categories into two categories: “film on cotton” and “film on background” are classified as film, and “cotton” and “background” are classified as non-film. The binarized images are shown in Figures 9–11. It can be seen that the combination of the AlexNet neural network structure and the LDA algorithm indicate the best recognition results, while the VGGNet neural network structure and the ICA algorithm denote the worst recognition results.

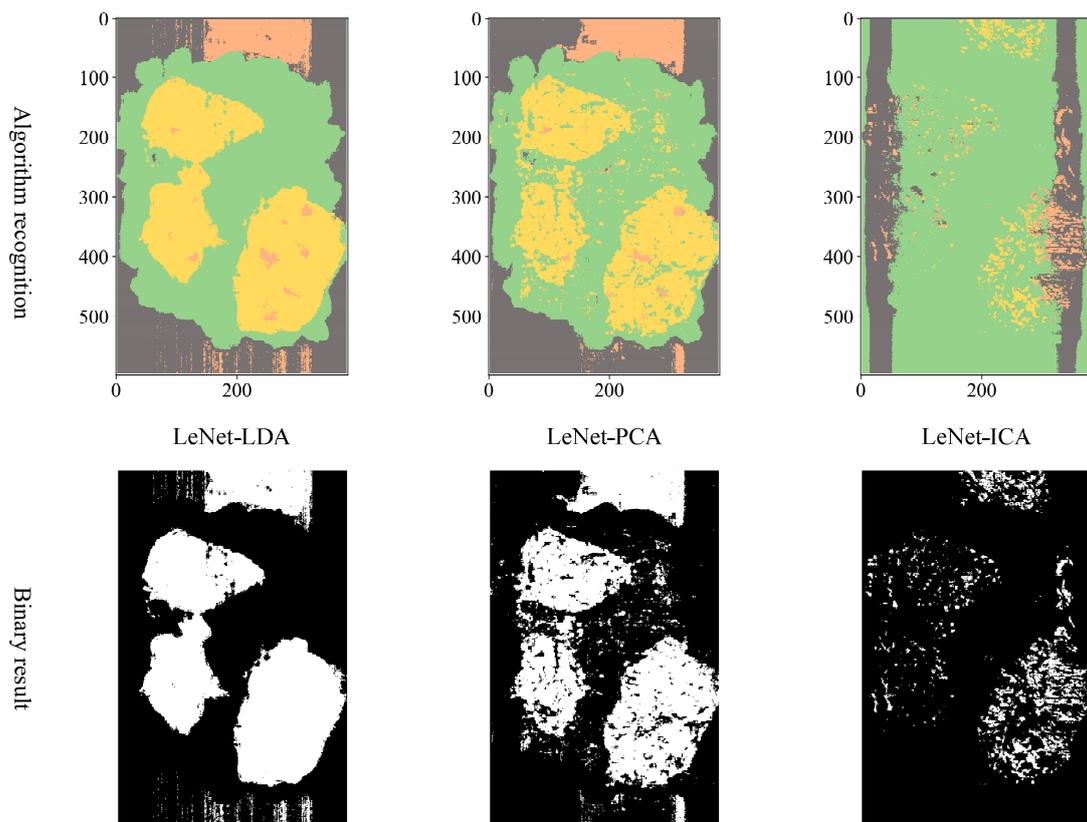


Figure 9. LeNet binarization images.

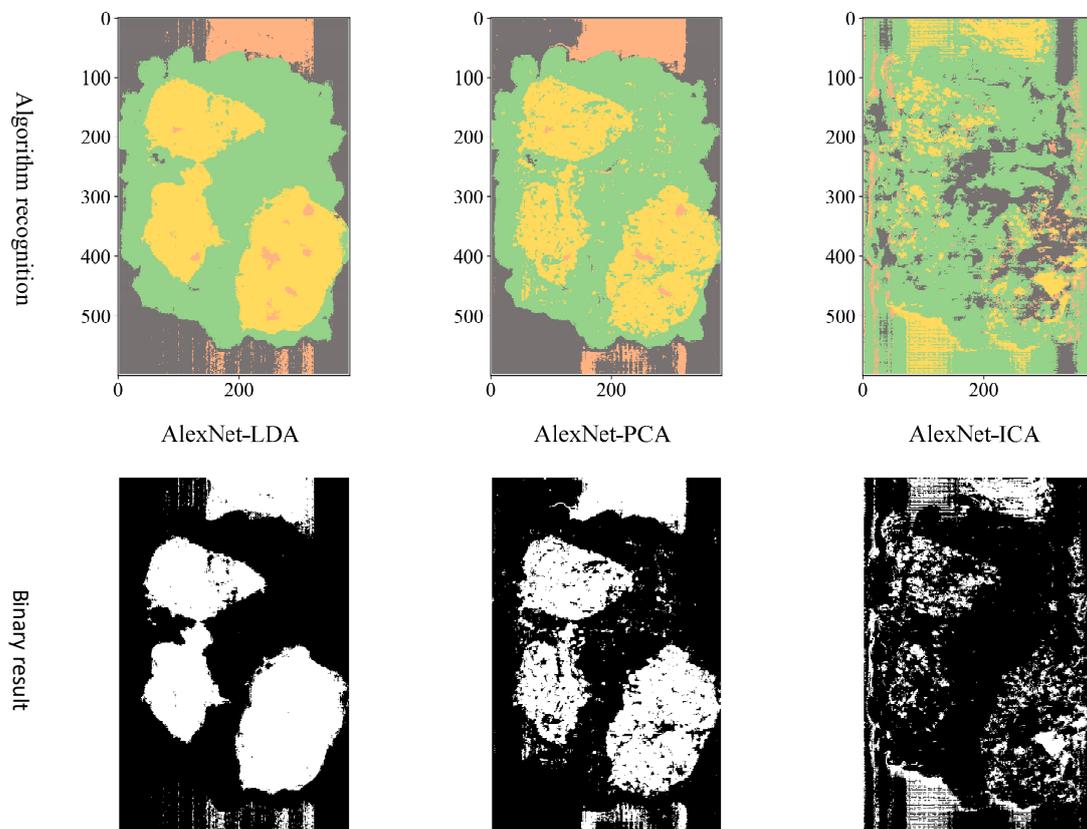


Figure 10. AlexNet binarization images.

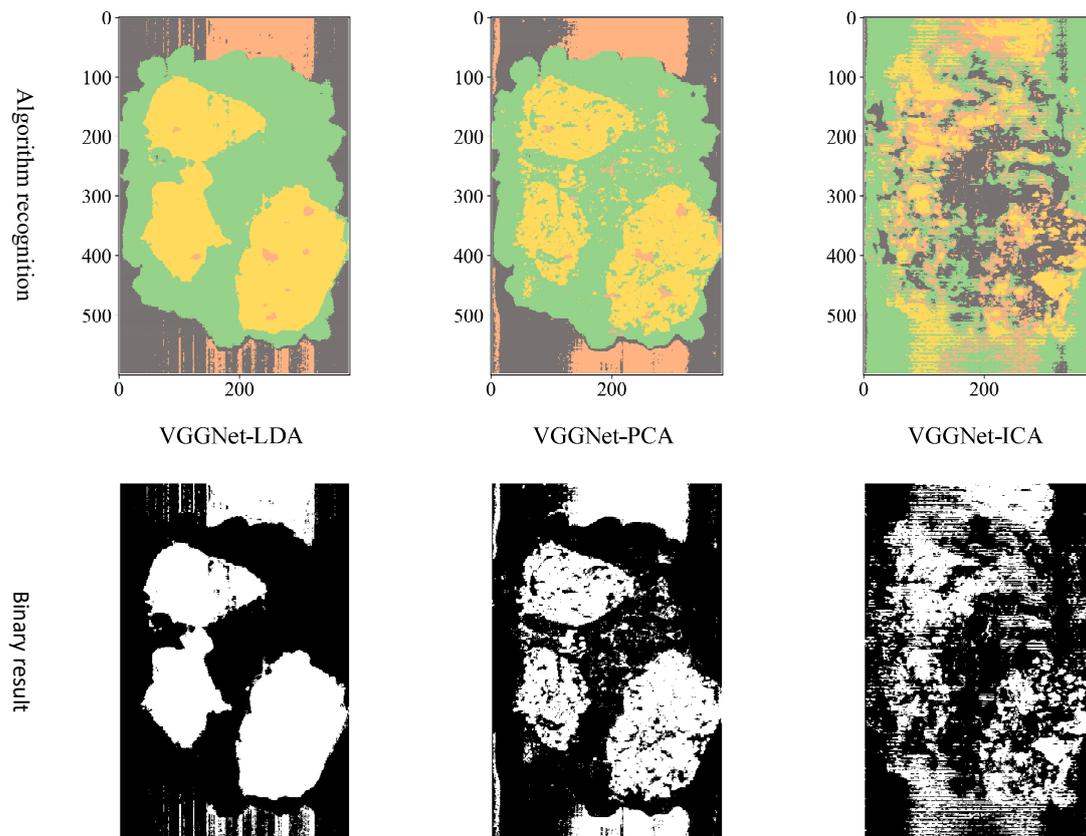


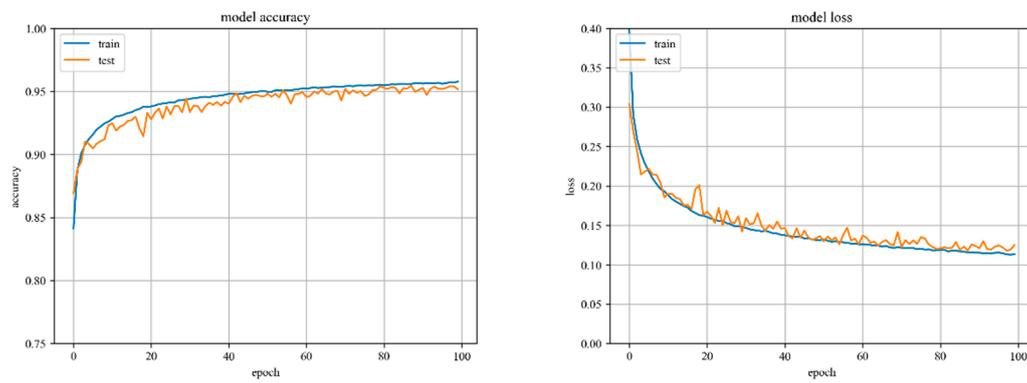
Figure 11. VGGNet binarization images.

Regarding the reduction to three dimensions, the above experiments validate the classification effect of different dimensionality reduction methods on hyperspectral data. The results show that LDA achieves the highest performance in terms of aggregation and separability of features preserved by dimensionality reduction of hyperspectral data. With limited device conditions for hyperspectral images, it is advisable to opt for LDA dimension reduction. However, due to the limitations of the LDA algorithm, the data can only be reduced to three dimensions. Therefore, when the computer performance meets the requirements, PCA achieves higher recognition accuracy when more dimensions are retained. In summary, the AlexNet-PCA multi-dimensional algorithm is experimented with to obtain the highest recognition accuracy for seed cotton mixed with the film.

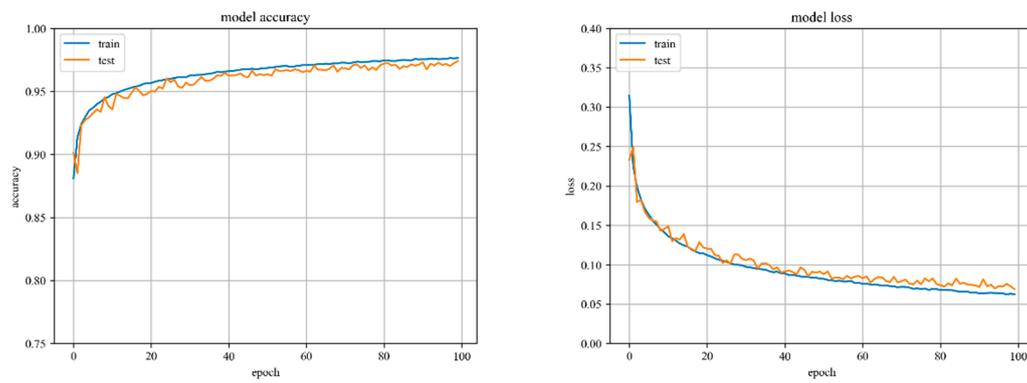
### 3.4. AlexNet-PCA Multi-Dimensional Algorithm Experiment

#### 3.4.1. AlexNet-PCA Model Training

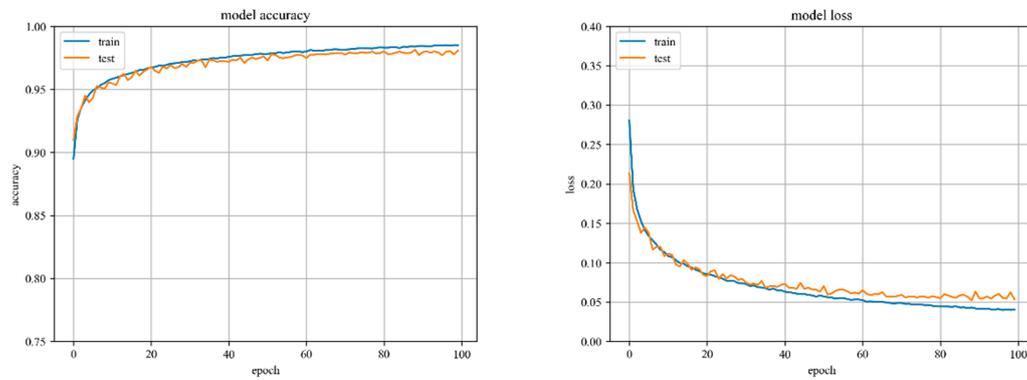
In Figure 12, the accuracy and loss value curves of the AlexNet model are shown when PCA is used to reduce the dimensionality by 6, 9, 12, and 15. The accuracy curve of the test set starts to converge at the training process of up to 40 iterations and mostly peaks at the training process of up to 60 iterations. The variation is stable throughout the training process and the model has great stability.



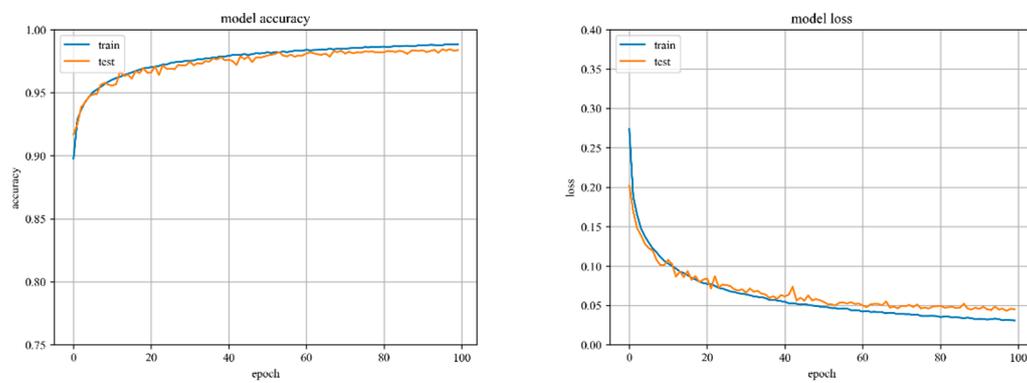
(a) AlexNet-PCA-6



(b) AlexNet-PCA-9



(c) AlexNet-PCA-12



(d) AlexNet-PCA-15

**Figure 12.** Accuracy curve and loss curve of AlexNet-PCA-X model.

### 3.4.2. AlexNet-PCA Model Testing

As shown in Table 8, the experimental data analysis can be summarized as follows:

(1) The AlexNet-PCA algorithm for “cotton” and “background” has a minor number of errors in sample recognition classification. It can be attributed to the edge junction containing the reflection spectrum of both the cotton and the background.

(2) Misclassification is observed when using the AlexNet-PCA algorithm to identify the samples of “cotton” and “film on cotton”, “background” and “film on background”. It can be attributed to the weak reflection nature of the film, which leads to an indistinct discrimination of features.

**Table 8.** AlexNet-PCA-X sample confusion matrix (%).

Model	Actual \ Predictive	Predictive			
		1	2	3	4
AlexNet-PCA-6	1	94.53	4.12	1.27	0.08
	2	3.18	96.21	0.11	0.49
	3	2.30	0.01	94.44	3.25
	4	0.03	0.95	0.29	98.73
AlexNet-PCA-9	1	96.88	1.67	1.37	0.07
	2	1.95	97.71	0.05	0.29
	3	1.06	0.03	97.68	1.23
	4	0.03	0.81	0.43	98.73
AlexNet-PCA-12	1	98.04	1.20	0.73	0.04
	2	1.00	98.75	0.02	0.23
	3	1.55	0.04	97.34	1.07
	4	0.03	0.92	0.43	98.62
AlexNet-PCA-15	1	98.38	0.50	1.11	0.01
	2	1.35	98.45	0.04	0.16
	3	0.96	0.01	98.38	0.66
	4	0.00	0.63	1.27	98.10

Especially for the PCA dimension selection, a set of linearly increasing dimensions 3, 6, 9 and 12 is chosen for the AlexNet-PCA multi-dimensional algorithm experiment. The linearly increasing dimensions are conducive to the smooth change in the image curve between overall accuracy and dimension; hence, the experimental results are more intuitive.

The Overall Accuracy (OA) of the test sample is shown in Table 9, representing the percentage of all samples that are accurately predicted. As the number of dimensions retained by PCA increases, the OA of the samples keep increasing. Figure 13 illustrates the OA of the samples as a function of the dimensionality reduction of PCA. The data in Table 9 and Figure 13 show the following:

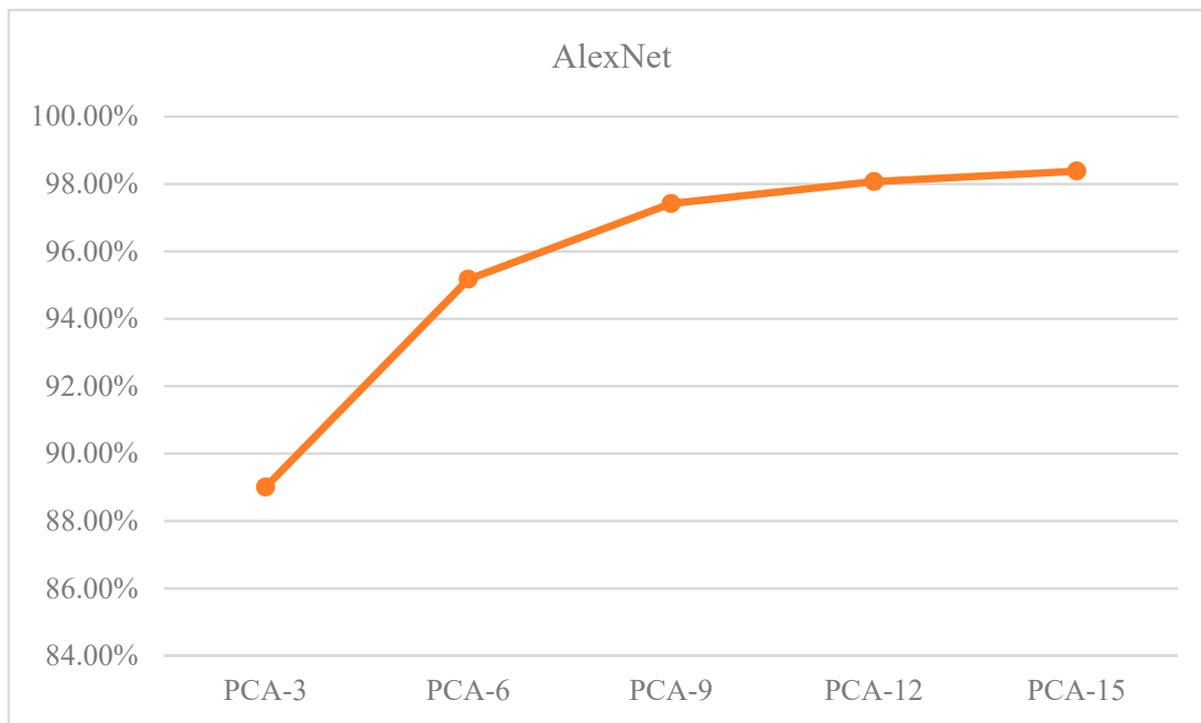
(1) The increase in PCA dimensionality has an inverse relationship with the increase in accuracy.

(2) When the PCA dimension is set to 12, the proposed algorithm achieves a recognition accuracy of over 98%. Additionally, the overall classification accuracy of the samples begins to converge.

With the increase in PCA dimensionality reduction, the complexity of the neural network model also increases. However, the complexity of the model can lead to overfitting, which in turn can decrease the generalization ability of the model. In the study, we primarily utilize the Dropout method to avoid the overfitting problem. Dropout effectively weakens the connections between neuronal nodes, which reduces the network’s reliance on individual neurons and thereby enhances model generalization ability.

**Table 9.** Overall accuracy of AlexNet-PCA-X test samples (%).

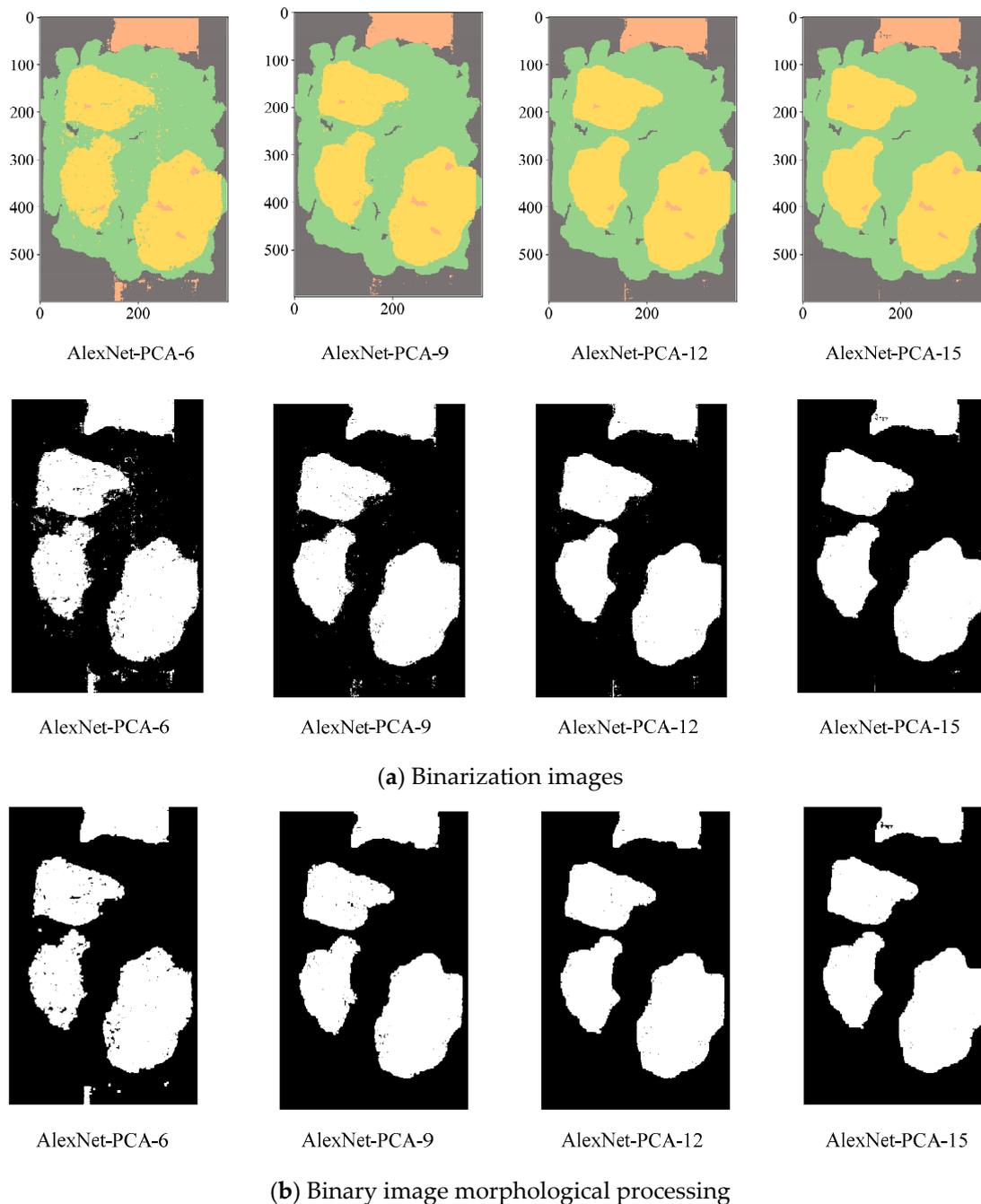
Dimension	PCA-3	PCA-6	PCA-9	PCA-12	PCA-15
AlexNet	89.00	95.17	97.42	98.07	98.38

**Figure 13.** Relationship between OA and PCA.

The binarized images are shown in Figure 14a. As demonstrated in Figure 14b, the morphological method is utilized to perform an open operation on the binary image, which effectively minimizes the noise caused by light, dust, and artificial marks. As a result, the binary image contains the eliminated artifacts of identified small areas and image edges. The Figure 14 results show that:

(1) Despite reducing the dimensionality to six utilizing PCA, the post-processing results still exhibit significant imperfections. However, when the dimension reduction is increased to 12, the post-processed image results successfully meet the requirement of providing coordinates. With a dimension reduction of 15, there is no significant difference between the post-processed image results compared to those obtained with a dimension reduction of 12.

(2) Considering the relationship between speed of accuracy improvement, computer performance, image processing results, dimension reduction, and training cost, PCA with a dimension reduction of 12 is the optimal solution for computer performance.



**Figure 14.** AlexNet-PCA-X binarization images.

### 3.4.3. Practical Application Testing of Model AlexNet-PCA-12

As can be seen from the above, the AlexNet-PCA-12 model with the optimal recognition accuracy is obtained experimentally. To verify the feasibility of the research, an application sorting test of the algorithm is conducted in a cotton factory in Aksu, Xinjiang. As depicted in Figure 15, the computer platform running the algorithm obtains the actual coordinates of the film and inserts them into the industrial control center, which controls the response time of the high-speed spray valve to complete the film removal.

Table 10 shows the data of several sorting experiments: the overall removal rate of film is 97.02%, and the cotton sorting amount can reach 3.0 t/h, which meets the requirements of practical application.

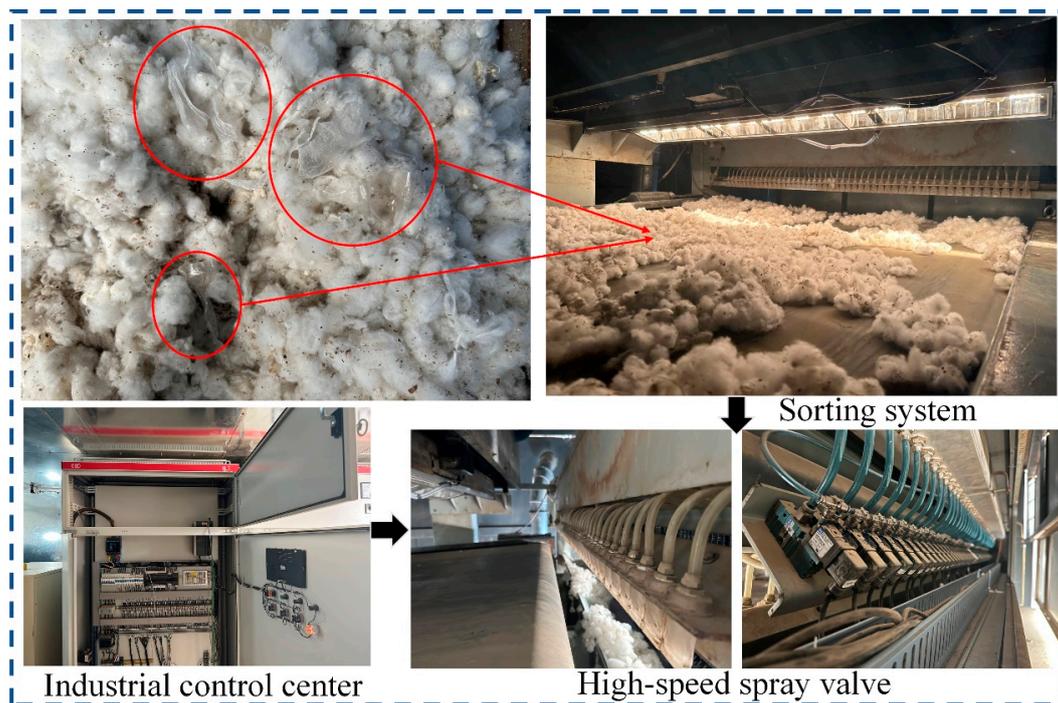


Figure 15. Hyperspectral sorting experiment.

Table 10. Application of test results.

Quantity of Trials	Quantity of Films in Cotton	Quantity of Removal Films	Removal Accuracy
1	132	128	96.97%
2	112	109	97.32%
3	98	95	96.94%
4	146	142	97.26%
5	157	152	96.82%
6	104	100	96.15%
7	128	125	97.66%
8	168	163	97.02%
9	84	82	97.62%
10	113	109	96.46%
Sum	1242	1205	97.02%

### 3.5. Summary of Discussions and Results

This chapter focuses on three main tasks: collecting laboratory data, conducting tests on the algorithm, and comparing the visualized recognition results with the experimental results. The recognition effects of LeNet, AlexNet, and VGGNet neural networks combined with LDA, PCA, and ICA dimension reductions are compared and analyzed. Finally, the feasibility of the proposed optimal model is verified for practical applications.

## 4. Conclusions

Based on hyperspectral images and the deep learning intelligent recognition algorithm, a novel intelligent recognition method for seed cotton mixed with colorless and transparent film is proposed in this paper. The main research topics include the construction of hyperspectral classification systems, dimensionality reduction for hyperspectral data processing, construction of algorithmic recognition models, and the practical application sorting tests.

(1) The basic principles of hyperspectral imaging are studied and a hyperspectral classification system is designed for the intelligent classification of seed cotton mixed with the film. The system can obtain 288 hyperspectral data bands with a resolution of

384 pixels  $\times$  600 pixels and a spectral range of 1000–2500 nm, which provides an excellent data basis for the recognition of film in seed cotton.

(2) LDA, PCA, and ICA are utilized to reduce the dimension of hyperspectral data to settle the problems of high latitude, large amounts of data, and redundant information of hyperspectral data. Experimental results suggest that LDA and PCA generalized better than ICA. LDA is the best method when the dimensionality reduction is the same as that of the three. PCA dimensionality reduction is more advantageous when computer performance is satisfied.

(3) An algorithm is successfully completed for hyperspectral image recognition of film in seed cotton. Based on the convolutional neural network architectures of LeNet, AlexNet, and VGGNet, the network model is constructed for hyperspectral image recognition applications in the seed cotton film domain.

(4) The combination test of the hyperspectral data dimension reduction algorithm (LDA, PCA, ICA) and the CNN model (LeNet, AlexNet, VGGNet) is completed. The experimental results illustrate that when the computer performance is satisfied, AlexNet-PCA-12 can achieve the best cost-to-performance ratio for both recognition and dimensionality reduction, and the recognition accuracy of the algorithm can reach 98.07%; the overall removal rate of film is 97.02% with the data of several sorting experiments in Aksu, Xinjiang.

On the whole, considering the influence of environmental factors such as light, humidity and dust in the practical application sorting tests, data under different environmental variables should be collected to further improve the generalization of the model. However, the research has potential applications in various fields, including but not limited to tea stalks removal, fruit and vegetable flaw separation, and pesticide residue detection in agricultural products. Further research can explore the use of photoelectric separation technology to enhance agricultural development.

**Author Contributions:** Conceptualization, Y.Z. and Q.L.; methodology, Q.L. and X.Y.; software, Q.L.; validation, Q.L. and X.Y.; formal analysis, Y.Z., Q.L. and X.Y.; investigation, Z.L. and Q.L.; resources, Q.L.; data curation, Q.L.; writing—original draft preparation, Q.L.; writing—review and editing, Q.L., Y.Z. and L.Z.; visualization, Q.L.; supervision, X.Y. and Z.L.; project administration, Z.L.; funding acquisition, L.Z. and Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research is supported by the Key Research and Development Projects of the Xinjiang Uygur Autonomous Region: Research on Key Technologies of Automatic Recognition of Foreign Fibers in Machine-picked Long-Staple Cotton (project number: 2022397193).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yang, W.; Li, D.; Zhu, L.; Kang, Y.; Li, F. A new approach for image processing in foreign fiber detection. *Comput. Electron. Agric.* **2009**, *2*, 68–77. [[CrossRef](#)]
2. Whitelock, D.P.; Armijo, C.B.; Gamble, G.R.; Hughs, S.E. Survey of seed-cotton and lint cleaning equipment in US roller gins. *Eng. Ginning* **2007**, *11*, 128–140.
3. Zhang, H.; Wang, Q.; Li, Y.; Liu, Y.; Jia, D. Electrostatic separation motion analysis of machine-harvested cotton and residual film based on CFD. *J. Comput. Methods Sci. Eng.* **2020**, *2*, 771–783. [[CrossRef](#)]
4. Li, D.; Yang, W.; Wang, S. Classification of foreign fibers in cotton lint using machine vision and multi-class support vector machine. *Comput. Electron. Agric.* **2010**, *2*, 274–279. [[CrossRef](#)]
5. Guo, L.; Yu, Y.; Yu, H.; Tang, Y.; Li, J.; Du, Y.; Chu, Y.; Ma, S.; Ma, Y.; Zeng, X. Rapid quantitative analysis of adulterated rice with partial least squares regression using hyperspectral imaging system. *J. Sci. Food Agric.* **2019**, *2*, 5558–5564. [[CrossRef](#)]

6. Ma, J.; Sun, D.W. Prediction of monounsaturated and polyunsaturated fatty acids of various processed pork meats using improved hyperspectral imaging technique. *Food Chem.* **2020**, *2*, 126695. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, M.; Li, C.; Yang, F. Classification of foreign matter embedded inside cotton lint using short wave infrared (SWIR) hyperspectral transmittance imaging. *Comput. Electron. Agric.* **2017**, *2*, 75–90. [[CrossRef](#)]
8. Jiang, Y.; Li, C. mRMR-based feature selection for classification of cotton foreign matter using hyperspectral imaging. *Comput. Electron. Agric.* **2015**, *2*, 191–200. [[CrossRef](#)]
9. Zhang, R.; Li, C.; Zhang, M.; Rodgers, J. Shortwave infrared hyperspectral reflectance imaging for cotton foreign matter classification. *Comput. Electron. Agric.* **2016**, *2*, 260–270. [[CrossRef](#)]
10. Zhao, Z.Q.; Zheng, P.; Xu, S.; Wu, X. Object detection with deep learning: A review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *2*, 3212–3232. [[CrossRef](#)]
11. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* **2018**, *2018*, 7068349. [[CrossRef](#)] [[PubMed](#)]
12. Morales, G.; Sheppard, J.W.; Logan, R.D.; Shaw, J.A. Hyperspectral dimensionality reduction based on inter-band redundancy analysis and greedy spectral selection. *Remote Sens.* **2021**, *2*, 3649. [[CrossRef](#)]
13. Jia, S.; Zhao, Q.; Zhuang, J.; Tang, D.; Long, Y.; Xu, M.; Zhou, J.; Li, Q. Flexible Gabor-based superpixel-level unsupervised LDA for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *2*, 10394–10409. [[CrossRef](#)]
14. Kang, X.; Xiang, X.; Li, S.; Benediktsson, J.A. PCA-based edge-preserving features for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *2*, 7140–7151. [[CrossRef](#)]
15. Lupu, D.; Necoara, I.; Garrett, J.L.; Johansen, T.A. Stochastic Higher-Order Independent Component Analysis for Hyperspectral Dimensionality Reduction. *IEEE Trans. Comput. Imaging* **2022**, *2*, 1184–1194. [[CrossRef](#)]
16. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *2*, 279–317. [[CrossRef](#)]
17. Ni, C.; Li, Z.; Zhang, X.; Zhao, L.; Zhu, T.; Wang, D. Online sorting of the film on cotton based on deep learning and hyperspectral imaging. *IEEE Access* **2020**, *2*, 93028–93038. [[CrossRef](#)]
18. Firat, H.; Asker, M.E.; Bayindir, M.I.; Hanbay, D. Spatial-spectral classification of hyperspectral remote sensing images using 3D CNN based LeNet-5 architecture. *Infrared Phys. Technol.* **2022**, *2*, 104470. [[CrossRef](#)]
19. Jiang, B.; He, J.; Yang, S.; Fu, H.; Li, T.; Song, H.; He, D. Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues. *Artif. Intell. Agric.* **2019**, *2*, 1–8. [[CrossRef](#)]
20. Zhao, J.; Pan, F.; Li, Z.; Lan, Y.; Lu, L.; Yang, D.; Wen, Y. Detection of cotton waterlogging stress based on hyperspectral images and convolutional neural network. *Int. J. Agric. Biol. Eng.* **2021**, *2*, 167–174. [[CrossRef](#)]
21. Zebari, R.; Abdulazeez, A.; Zeebaree, D.; Zebari, D.; Saeed, J. A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction. *J. Appl. Sci. Technol. Trends* **2020**, *2*, 56–70. [[CrossRef](#)]
22. Qin, X.; Wang, S.; Chen, B.; Zhang, K. Robust Fisher Linear Discriminant Analysis with Generalized Correntropic Loss Function. In Proceedings of the 2020 Chinese Automation Congress (CAC), Shanghai, China, 6–8 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 7117–7121.
23. Wu, S.X.; Wai, H.T.; Li, L.; Scaglione, A. A review of distributed algorithms for principal component analysis. *Proc. IEEE* **2018**, *2*, 1321–1340. [[CrossRef](#)]
24. Ye, M.; Ji, C.; Chen, H.; Lu, H.; Qian, Y. Residual deep PCA-based feature extraction for hyperspectral image classification. *Neural Comput. Appl.* **2020**, *2*, 14287–14300. [[CrossRef](#)]
25. Ghosh, A.; Barman, S. Application of Euclidean distance measurement and principal component analysis for gene identification. *Gene* **2016**, *2*, 112–120. [[CrossRef](#)] [[PubMed](#)]
26. Luo, Z. Independent Vector Analysis: Model, Applications, Challenges. *Pattern Recognit.* **2023**, *138*, 109376. [[CrossRef](#)]
27. Sajjad, M.; Yusoff, M.Z.; Yahya, N.; Haider, A.S. An efficient VLSI architecture for FastICA by using the algebraic Jacobi method for EVD. *IEEE Access* **2021**, *2*, 58287–58305. [[CrossRef](#)]
28. Huang, J.T.; Li, J.; Gong, Y. An analysis of convolutional neural networks for speech recognition. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, Australia, 19–24 April 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 4989–4993.
29. Chen, Z.Q.; Li, C.; Sanchez, R.V. Gearbox fault identification and classification with convolutional neural networks. *Shock. Vib.* **2015**, *2015*, 390134. [[CrossRef](#)]
30. Ranjbarzadeh, R.; Jafarzadeh, G.S.; Bendeche, M.; Amirabadi, A.; Ab, R.M.N.; Baseri, S.S.; Aghamohammadi, A.; Kooshki, F.M. Lung Infection Segmentation for COVID-19 Pneumonia Based on a Cascade Convolutional Network from CT Images. *BioMed Res. Int.* **2021**, *2021*, 5544742. [[CrossRef](#)] [[PubMed](#)]
31. Sun, M.; Song, Z.; Jiang, X.; Pan, J.; Pang, Y. Learning pooling for convolutional neural network. *Neurocomputing* **2017**, *2*, 96–104. [[CrossRef](#)]
32. Basha, S.H.S.; Dubey, S.R.; Pulabaigari, V.; Mukherjee, S. Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing* **2020**, *2*, 112–119. [[CrossRef](#)]

33. Janocha, K.; Czarnecki, W.M. On loss functions for deep neural networks in classification. *arXiv* **2017**, arXiv:1702.05659. [[CrossRef](#)]
34. Kanezaki, A. Unsupervised image segmentation by backpropagation. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1543–1547.
35. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaria, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *2*, 1–74.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.