Method for State Recognition of Egg Embryo in Vaccines Production Based on Support Vector Machine

Yan-wei XU¹,², Meng-hui CHEN² and Tan-cheng XIE²,*

¹School of Mechanical Engineering, Tianjin University, Tianjin 300072, China
²School of Mechatronics Engineering, Henan University of Science and Technology, Luoyang 471023, China

*Corresponding author

Keywords: Egg embryo, Pattern recognition, Support vector machine, Computer image processing.

Abstract. The method for state recognition of egg embryo in biological vaccines production based on computer image processing and support vector machine was researched. Firstly, the median filtering method was applied to eliminate noises in the images of egg embryo that were classified by the requirements of practical production, then the threshold segmentation method was used to segment the images, then the blood vessels and black blocks in the egg embryo images were selected as major characteristics to extract. The decision tree classification model with structure of binary tree was built on support vector machine, and the model was trained with the toolkit LIBSVM, the penalty factor C=2, RBF kernel parameter $\gamma=3.0512e-05$ and the precision of cross validation accuracy is 98.913%. Finally, 100 egg embryos were randomly selected as the test samples to do pattern recognition, and compared with manual test result.

Introduction

The method for eggs embryonic breeding viruses formed by short incubation of special fertilized eggs is a mature, safe and effective to cultivate influenza vaccine. Before the original strain of the diluted strain is injected into the chicken embryo allantoic cavity, and the state of the egg embryo body must be strictly checked and filtered to remove weak embryo and dead embryo unsuitable for vaccine culture. Otherwise, in the process of virus after the end of the vaccine culture, once mixed with the necrotic embryo egg, then the same batch of the virus would be polluted and abandoned, but also may cause major medical security problems [1].

The researchers have carried out research on the quality of egg testing [2-5], to a certain extent, to promote the development of egg quality testing technology. But most of the work still needs to be done manually at the time of modernized incubation of egg processing plants and bioengineering companies. To identify egg embryo body based on computer image processing technology, not only can improve the level and efficiency of egg embryo body detection technology, reduce the cost of vaccine preparation, but also can effectively solve the problem of poor efficacy, low efficiency, high labor intensity, and subjective and objective factors in the process of artificial detection of egg embryo body.

Image Processing of Egg Embryo Body

According to the actual production requirements of a biological engineering Limited company, researchers selected the incubation of 6 days or so of the egg embryo body, used LED light source and color area array CCD industrial camera, and collected the egg embryo body image of different status. According to the actual production classification standard, the embryo was divided into four types: living embryo, weak embryo, polluted embryo and dead embryo. According to the actual production classification standard, the egg embryo body was divided into four types: live embryo, weak embryo, polluted embryo and dead embryo (as shown in Figure 1). The weak embryo, dead embryo and
polluted embryo were the embryo bodies that needed to be screened. The blood vessels were many and thick in live embryo image, and blood vessels were few and thin in live embryo image. There were obvious irregular black blocks in the interior of the pollution embryo, and there were no obvious features such as blood vessels and black patches in the dead embryo.

Method for median filter was used to reduce noise in the image. Color image vector median filtering was to take the average value $\bar{X}(i, j)$ of all vectors $X(i, j) = [R(i, j), G(i, j), B(i, j)]^T$ in a given window, then calculated the distance from the vector $X(i, j)$ to average vector $\bar{X}(i, j)$ in the window, and took the vector with the smallest distance as the output value of the center pixel of the window. Assuming that the size of a color image set $X(i, j)$ was $M \times N$. The algorithm for using a window for median filtering is as follows

1. To calculate the average vector $\bar{X}(i, j) = [\bar{R}(i, j), \bar{G}(i, j), \bar{B}(i, j)]^T$ in the window by solving the average value $R$, $G$ and $B$.

   $$
   \bar{R} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} R(i, j) / n^2 
   $$

   $$
   \bar{G} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} G(i, j) / n^2 
   $$

   $$
   \bar{B} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} B(i, j) / n^2
   \tag{1}
   $$

2. To calculate distance $S_{ij}$ from vector to average vector, compare the size of $S_{ij}$, and get the smallest $S_{\min}$.

   $$
   S_{ij} = \| X(i, j) - \bar{X}(i, j) \| 
   $$

3. Using the corresponding pixel $X_{\min}$ of $S_{\min}$ as the vector of the window to replace the window center pixel vector.

   The results of pretreatment of egg embryo image by median filter method was as shown in Figure 2. It can be seen that the median filter can effectively eliminate the impurity points in the image, reduce the ambiguity of the edge of the egg embryo image, and protect the edges and details of the image well.

This study carried out image enhancement processing for color brightness on the basis of the HSI model of the egg image after median filtering, saturation and hue. Color brightness $I$, hue $H$, and saturation $S$ are the 3 basic attributes of color, which information can be obtained by $R$, $G$, $B$ three color information obtained using formula.
In formula (3), \( F = \frac{2R-G-B}{G-B} \). R, G and B represent the stimulus values of red, green, and blue primaries.

**Extraction Image of Characteristics**

It can be seen from the image of the egg embryo that the blood vessel and the black block was the distinguishing features of living embryo, weak embryo, polluted embryo and dead embryo. This article studied the two characteristics of the blood vessel and the black block in the image of egg embryo body.

**The Feature Extraction of Blood Vessel Characteristics**

In order to extract the characteristics of blood vessel in the egg embryo image, the median filtering and image enhancement was performed firstly. Then, threshold of R, G and B component were read and analyzed respectively in the area of vascular image region and non vascular image region. The statistical results of R, G and B components corresponding to some pixels in the vascular region and the non-vascular region were shown in Table 1.

<table>
<thead>
<tr>
<th>Vascular region pixels point</th>
<th>Non vessel region pixels point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinate point (x, y)</td>
<td>R</td>
</tr>
<tr>
<td>(77, 110)</td>
<td>198</td>
</tr>
<tr>
<td>(78, 110)</td>
<td>197</td>
</tr>
<tr>
<td>(79, 110)</td>
<td>192</td>
</tr>
<tr>
<td>(80, 110)</td>
<td>193</td>
</tr>
<tr>
<td>(81, 110)</td>
<td>196</td>
</tr>
<tr>
<td>(82, 110)</td>
<td>197</td>
</tr>
<tr>
<td>(83, 110)</td>
<td>199</td>
</tr>
<tr>
<td>(84, 110)</td>
<td>196</td>
</tr>
<tr>
<td>(85, 110)</td>
<td>198</td>
</tr>
</tbody>
</table>

From Table 1, we can see that the fluctuation range of the R component threshold is 192–202, the fluctuation range of G component threshold is 27–44, and the fluctuation range of B component threshold is 38–49. The fluctuation range of the R component threshold of the non-vascular image is respectively 255, 210–218 and 162–169, the fluctuation range of the G component threshold is respectively 131–139, 58–63 and 21–24, and the fluctuation range of the B component threshold is respectively 72–80, 36–45 and 36–42. The R, G, and B thresholds of non-vascular images had no fixed fluctuation range, and compared with the R, G, and B component thresholds of the vascular images, there was a significant difference.

The fluctuation range of R, G, B component threshold was an important reference for the extraction of vascular features, but also needed to carry out color segmentation based on color histogram information, then extracted the vascular characteristics. The color histogram described the proportions of the different colors in the entire image, did not care about the spatial position of each color in the picture, and any egg embryo image can only give a corresponding color histogram. Through analyzing the color histogram of R, G and B component in egg embryo images, and combining with the fluctuation range of R, G, B vascular image component threshold in Table 1, the actual threshold value of R component in blood vessel image was 194–205, the actual threshold value of G component was 20–40, and the actual threshold of B component was 30–50. In this way, to color
the image of the egg embryo body was segmented by the threshold segmentation range of the image, and the image data was saved. Most of the contours of the blood vessel could be obtained, and the characteristic of blood vessel was extracted from the egg embryo image (shown in Figure 3).

Figure 3. Extraction of the vascular contour characteristic.

**Pattern Recognition Method of SVM**

**Support Vector Machines**

The basic idea of SVM method can be described in terms of circumstances. In the two types of optimal classification super-plane diagram shown in Figure 4, the solid point and the hollow point represented two kinds of training samples respectively, and H was the two kinds of classification lines which are not mistakenly separated. H1 and H2 were straight lines, which passed through the closest point from H and parallel to H in the two samples. It was the optimal classification line to separate the two types of samples without error, and make classification interval the greatest. The optimal classification function was as follows, which was obtained by applying the quadratic programming method.

\[
f(x) = \operatorname{sgn}(\sum_{i=1}^{n} \alpha_i^* y_i (x_i \cdot x) + b^*)
\]

In (4), \(\alpha_i^*\) is Lagrange multiplier; \(b^*\) is threshold.

![Figure 4. Schematic diagram of optimal classification of hyperplane.](image)

For the linear inseparable situation, the non-negative relaxation terms \(\xi_i\) and error penalty constant \(C\) needed to be introduced, the optimal classification function was the same as formula (4). For the nonlinear problem, it is the support vector machine that the nonlinear problem could be transformed into the linear problem in the high dimensional space by the kernel function, and then the optimal classification surface was obtained in the transformation space. The corresponding classification function became:

\[
f(x) = \operatorname{sgn}(\sum_{i=1}^{n} \alpha_i^* y_i K(x_i \cdot x) + b^*)
\]

The form of kernel function is the main factor affecting the performance of SVM classifier. Commonly used kernel functions are:

1) Linear kernel function
\[ K(x, x_i) = (x \cdot x_i) \]  
\[ K(x, x_i) = [(x \cdot x_i) + 1]^q \]  
\[ K(x, x_i) = \exp(-|x-x_i|^2/\gamma^2) \]  
\[ K(x, x_i) = \tanh[v(x \cdot x_i) + c] \]

**Training of Egg Embryo Body Classifier Based on SVM**

The LIBSVM toolkit was used as an experimental platform to train an egg embryo body classifier. Sixty eggs incubated for 6 days and tested by artificial detection, and the egg embryos were selected as the training samples, which included weak embryos, polluted embryos and dead embryos 15 each. The characteristic data of 60 egg embryo images were obtained by image acquisition, image preprocessing, image segmentation and feature extraction, and were converted into sample data of LIBSVM format. LIBSVM constructed an egg embryo body classifier model by studying the egg embryo sample data. The RBF function of the radial basis kernel is selected for the kernel function in the training process. The cross factor method is used to obtain the penalty factor \( C = 2 \), the kernel parameter \( \gamma = 3.0512 \times 10^{-5} \), and the accuracy of the cross validation is 98.913%. The total number of support vectors obtained by training of egg embryo samples was 16, and the number of support vectors for each type of live embryo, weak embryo, contaminated embryo and dead embryo was 6, 4, 3, 3.

**Conclusion**

We studied the recognition technology of egg embryo body status in vaccine preparation by image processing technology and support vector machine pattern recognition method. After the median image filtering and image enhancement of the egg embryo images collected by the CCD industrial camera, we used the threshold segmentation method to segment the blood vessel and the black block in the egg embryo image, and extracted the characteristics of the blood vessel and the black block. We constructed the decision tree classifier model based on the binary tree structure, and trained it with cross validation and quick selection of the LIBSVM toolkit. The penalty factor \( C \) was 2, RBF kernel parameter \( \gamma \) was 3.0512e-05, and the cross validation accuracy was 98.913%.

Finally, 100 eggs randomly selected as experimental samples carried out pattern recognition. Compared with the results of artificial test, the accuracy of SVM classifier was 100% for the weak embryo, polluted embryo and dead embryo, and 94.62% for the live embryo.

**Acknowledgement**

This work was financially supported by the National Natural Science Foundation of China (51305127), the youth backbone teachers training program (2016GGJS-057) and scientific research key project fund of the Education Department Henan Province of China (14A460018).
References


