Method for Detecting Surface Defects of Ceramic Tableware Based on Deep Learning

Shuo Ye¹,*, Lingzhen Sun², *
¹Huali College Guangdong University of Technology, Guangzhou 511325, China
²Guangzhou Huali Science and Technology Vocational College, Guangzhou 511325, China

*Corresponding author e-mail: sunlingzhen1983@163.com, *kame1986@126.com

Abstract. In view of the advantages of simple and high-precision detection methods for surface defects of ceramic tableware based on machine vision, the various links involved in such methods are reviewed. First, it summarizes the various imaging methods and common defect types on the surface of ceramic tableware; secondly, introduces and analyses the existing detection methods according to different mathematical modelling ideas; finally, summarizes the content the future research on the detection method of surface defects of ceramic tableware is prospected. It can be seen that the surface defect detection method of ceramic tableware based on machine vision has made great progress, but there is still room for improvement in the design of feature extraction algorithms, such as the feature extraction algorithm based on deep neural networks.

1. Introduction
With the advancement of computer real-time control technology, the degree of automation of ceramic tableware production system has been greatly improved. This not only reduces the manpower input of manufacturers, but also improves the production efficiency of the production line. However, the current competition in the ceramic tableware manufacturing industry is becoming more and more fierce. In addition, the special use of ceramic tableware makes the manufacturers put forward higher requirements on the appearance quality of the ceramic tableware produced. If the ceramic tableware produced has surface defects, such as black spots, yellow spots, oil stains, wrinkles, cracks, uneven surfaces or shape defects (such as discs that are not round), it will have an adverse effect on the overall sales of tableware, so, Ceramic tableware generally needs to be strictly inspected for the appearance quality of the tableware before leaving the factory to exclude defective products. However, at present, most of the surface quality of ceramic tableware is tested manually, which makes the detection effect often limited by the subjective consciousness of the inspectors and the degree of eye fatigue, resulting in the defect detection accuracy cannot be defined, and the overall inspection efficiency is low. The current popularity of machine vision inspection technology provides a new idea for the detection of ceramic tableware surface defects, that is, using industrial cameras to obtain tableware images, using image processing algorithms to analyse tableware images, and finally determine whether the tableware defect. Industrial machine vision inspection platforms generally vary greatly depending on the detected target and the site environment [1].

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd
Compared with the above detection methods, the surface defect detection method of ceramic tableware based on machine vision has obvious advantages. It can realize the integration of real-time monitoring, accurate judgment and detection device. With the advantages of convenience and speed, such methods have become the detection of surface defects of ceramic tableware one of the main development directions.

2. Algorithm introduction
The traditional machine learning method is based on the construction of a neural network. Although it has achieved good application results through the construction of a shallow neural network, it is unable to deal with the representation of too many dimensions and complex data. That said, as the number of layers continues to deepen, it not only means a lot of parameter tuning but also faces the risk of gradient disappearance. In 2006, Hinton and others of the University of Toronto proposed a new method for constructing deep neural networks, namely Deep Learning Neural Network (DLNN). The difference between deep learning and traditional deep neural networks is that it overcomes the reduction of deep error feedback caused by the construction of multi-layer neural networks, so that it can approximate a complex function by training and constructing a deep neural network. It has been widely used in image processing, pattern recognition and natural language processing.

2.1. Basic idea
Deep learning uses a layered idea to extract complex signals layer by layer. Assuming a network system \( S(S_1, S_2, \ldots, S_n) \) with layer \( n \), its inputs and inputs are \( I \) and \( O \), respectively, which can be expressed as:

\[
I \Rightarrow S_1 \Rightarrow S_2 \Rightarrow \ldots \Rightarrow S_n \Rightarrow O
\]

(1)

If the output \( O \) is equal to the input \( I \), then it can be said that there is no loss of information after this system change, that is, at any layer \( S_n \) is another representation of the original information \( I \). This layered abstract structure is usually carried out layer by layer Training, which advances layer by layer under the condition that the original information is fully expressed. Its unsupervised learning method can easily obtain the hidden features of the data, which is also one of the reasons why deep learning has become a popular research machine learning. Figure 1 shows the basic structure of deep learning [2].

![Figure 1. Basic structure of deep learning](image)

2.2. Common methods of deep learning
In 2007, Benjio proposed the concept of sparse coding, using the concept of base in linear algebra to improve the condition that the input and output of the automatic encoder are equal, so that it can be expressed in the form of base, that is \( O = a_1 \cdot \varphi_1 + a_2 \cdot \varphi_2 + \ldots + a_n \cdot \varphi_n \). Where \( a_n \) is the base and \( \varphi_n \) is the coefficient, so in the end, the problem of minimizing the difference between input and output is also
introduced, namely \( \text{Min} \ |(I - O)| \), where I is the input and O is the output. By solving this optimization problem, its base \( a_n \) and coefficient can be solved separately \( \varphi_n \), these coefficients and bases are another form of expression for input, which is also obtained through unsupervised learning.

If the regularization limit of the L1 norm is introduced in the previous formula, then \( \text{Min} |I - O| + u \bullet (|a_1| + |a_2| + ... + |a_n|) \) can be derived. Its purpose is to express the input signal using a linear combination of a set of bases, and the number of bases in this conversion expression must be far smaller than the size of the signal input itself; this method is called sparse coding.

The restricted Boltzmann machine model is a bipartite graph, so in the case of \( V \), the hidden layer nodes are independent of each other, that is, \( p(h \mid v) = p(h_v \mid v) = p(h_v \mid v) \). Similarly, in the case of the hidden layer \( h \), all the visible nodes are independent, and because all the sums satisfy the Boltzmann distribution, when \( h \) is input, the hidden layer \( h \) can be obtained through \( p(h \mid v) \), and after the hidden layer \( h \) is obtained, the visible layer can be obtained through \( p(h \mid v) \). By adjusting the parameters, the visible layer obtained from the hidden layer has the same result as the original visible layer. Then it can be said that the hidden layer is another expression of the visible layer, so the hidden layer can also be used as the feature of the input data of the visible layer.

Assuming that the time series samples are Dimensional vectors, K continuous data samples can be expressed as \( (x_1, y_1), ..., (x_k, y_k) \in \mathbb{R}^n \bullet R \), and the linear function is also set to \( f(x) = \omega \bullet x + b \), and the optimization problem is converted to the minimum value \( R(\omega, \xi, \xi^*) = \frac{1}{2} \omega \bullet \omega + C \sum_{i=1}^{k} (\xi + \xi^*) \), and its constraints are:

\[
\begin{align*}
    f(x_i) - y_i &\leq \xi^* + e, i = 1, 2, ..., k \\
    f(x_i) - y_i &\leq \xi + e, i = 1, 2, ..., k \\
    \xi, \xi^* &\geq 0, i = 1, 2, ..., k
\end{align*}
\]

Among them, \( 1/2 \omega \bullet \omega \) can improve the generalization ability of the support vector machine, \( C \sum_{i=1}^{k} (\xi + \xi^*) \) is the penalty term, and \( e \) is the insensitive loss function. The same problem becomes the problem of solving convex quadratic optimization. The Lagrange function is introduced:

\[
L(\omega, b, \xi, \xi^*, a, a^*, \gamma, \gamma^*) = \frac{1}{2} \omega \bullet \omega + C \sum_{i=1}^{k} (\xi + \xi^*) \\
- \sum_{i=1}^{k} a_i |\xi_i + e - y_i + f(x_i)| \\
- \sum_{i=1}^{k} a_i^* |\xi_i + e - y_i + f(x_i)| - \sum_{i=1}^{k} (\gamma_i \gamma_i^* - \xi_i \xi_i^*)
\]

\[ (3) \]

3. Visual image features

3.1. Visual image features

Colour is the most basic characteristic attribute of an image, and the colour attribute of an image is mainly described by the colour value of each pixel in the image in the colour space. The colour histogram feature is currently the most commonly used colour attribute description method. This feature has good stability for the position, size and rotation of the target object in the image. For the statistical distribution of image colour values, colour histograms in different colour spaces may have different performances. At present, the commonly used colour spaces in image processing are RGB, HSV, Lab, YIQ and YCbCr.
Because Lab colour space is more in line with the human eye's visual perception of true colours, it is often used to describe the colour information of images.

The colour histogram is the statistics of the global colour distribution of all pixels in the image. This feature extraction process for pixels is computationally efficient, but it lacks a description of the spatial positional relationship of pixel colour values. To this end, Stricker et al. Developed the colour moment feature. If the distribution of colour values in the image is regarded as a probability distribution, the colour feature can be described by first-order moments and second-order distances. The first-order colour moment is the colour of the image. The mean value describes the overall lightness and darkness of the image; the second-order colour moment is the standard deviation of the colour values, which mainly describes the discrete distribution of these colour values. This feature calculation is also more efficient, the feature vector is relatively simple, and the colour distribution can be obtained without quantizing the colour value, to a certain extent, avoiding the loss of colour information. Both the colour histogram and the colour moment are descriptions of the image's colour features. The main advantages are its efficient calculation process and low feature dimensions, so it is more suitable for the needs of target detection in robot vision.

Texture is also a very important feature attribute in visual images. Texture is a regular distribution structure of pixels in an image, but there is currently no uniform definition of image texture. Local binary mode is a simple but highly discriminative texture feature extraction algorithm. The core idea of LBP is to represent the local structure in the image by using a 0/1 binary string, and the image content is formed by the combination of these regular texture structures. By comparing the size relationship between the centre pixel of the image and its neighbours, 0 or 1 is used to represent this size relationship, and the texture structure near the centre pixel can be represented by this 0/1 binary string. Then the histogram distribution of these local texture structures in the image is counted to obtain the final LBP feature vector. Therefore, the LBP feature is to analyse the texture structure of the grayscale image, which has grayscale invariance and illumination invariance. In order to improve the invariance of this feature to the target rotation, there is also a rotation-invariant LBP feature that further improves the LBP feature. In addition, the texture structure distribution of these binary patterns in the image is not uniform. To this end, they proposed the concept of "equivalent patterns" to reduce the length of feature vectors [3].

As shown in Figure 2, SIFT can be divided into two parts: key point detection and key point feature description. In the retrieval process of key points, first use the scale space model of the image to simulate the formation process of the image in the human eye retina. Therefore, the image in the scale space has more complete image information. The Gaussian difference operator is used to construct the scale space, and the extreme point is defined and detected in the scale space as the potential key point position. In addition, in order to improve the stability of key points and remove some possible noise points, these key points need to be further screened. In the feature description process of the key point, firstly, the coordinate rotation is performed by calculating the main direction of the feature of the point. Before the feature description, the neighbourhood of the key point is spatially rotated, so that the extracted SIFT has rotation invariance; then the gradient direction histogram in the area is calculated, and the histogram statistics weighted by the gradient weight are finally formed 128-dimensional SIFT feature vector. Therefore, in the process of image feature description, SIFT features have invariance to rotation, scale, illumination, and translation, and also have certain anti-interference ability for affine transformation or 3D rotation of objects in the image.
3.2. SLIC super pixel segmentation algorithm

Image super pixel segmentation is to divide a large number of pixels in the image into a small number of integral super pixels. This segmentation is also a preliminary segmentation process for image content, using basic K-means algorithm or conditional random field methods Collect pixels with similar properties into a super pixel. The super pixel segmentation in the image will also greatly reduce the number of nodes in the image. For example, a grayscale image with a size of 300 * 400 contains 120,000-pixel nodes, and super pixel segmentation can generate super pixel nodes by manually specifying the number of nodes greatly reduces the number of nodes. This is very important for the target object detection process based on the visual attention mechanism. When a pixel is used as a graph node and the relationship between pixels is a graph model for the connected edge of the graph node, the process of image object detection is the graph model learning process. At this time, the size of the graph model is directly affected by the number of pixels, and when the graph model is too large, the efficiency of target detection will also decrease.

4. Deep learning machine vision detection method for surface defects of ceramic tableware

The general process of the surface defect detection method of ceramic tableware based on machine vision can be simply summarized as: imaging the surface of the ceramic tableware to obtain the surface of the ceramic tableware based on machine vision The general process of the surface defect detection method of ceramic tableware based on machine vision can be simply summarized as: The surface is imaged to obtain image data of the surface of the ceramic tableware, and then various machine vision methods are used to detect the defect area of the image. Figure 3 shows an example of surface defects of ceramic tableware based on machine vision [4].
According to different mathematical modelling ideas, the existing surface defect detection methods of ceramic tableware based on machine vision can be roughly divided into: gradient-based detection methods, cluster-based detection methods, frequency-domain analysis-based detection methods, matrix decomposition Detection method and detection method based on machine learning. The following will introduce and briefly analyse various methods.

4.1. Detection method based on gradient features

The basis of this kind of method is: the brightness of the surface defect area of the ceramic tableware and the rest are obviously different, and the boundary between the two has a high gradient. The following is a brief introduction to some representative work in such methods. Use the gradient characteristics of the image to sharpen and smooth the image of different gradient areas in the image. Since the gradient value of the defect edge area is higher, the defect area is sharpened first, and the flawless area with a lower gradient value is taken Smoothing, so this method can effectively strengthen the defects while suppressing noise [5]. The diffusion model in this method is characterized by Gray scale and gradient to adjust the diffusion coefficient equation, which is an adaptive smoothing process. Only the pixels with low Gray scale and high gradient in the defect area will produce high diffusion coefficient. Using this model to smooth the suspected defect area can retain the original Gray level of the complete area. By subtracting the diffusion image from the original image, micro cracks can be obtained the obviously enhanced difference image can be segmented by simple binary threshold segmentation and morphological operation. The algorithm block diagram is shown in Figure 4.

![Figure 4. Detection method based on gradient features](image)

4.2. Cluster-based detection method

The core idea of this kind of method is to distinguish defective areas from non-defective areas by clustering algorithm, and finally use threshold segmentation and other algorithms to obtain binary images containing only defective areas. The following is a brief introduction to some representative work in such methods. Detection method of surface defects of ceramic tableware based on the maximum between-class variance. This method first uses a Gaussian filter to smooth the image; secondly, it performs edge positioning and image segmentation to separate individual battery blocks; then, the maximum inter-class variance method is used to threshold the image and divide the image into Defective foreground part and non-defective background part; Finally, Hough transform is used to perform straight line detection on the segmented binary image, and the difference between the resulting image and the original image after smooth pre-processing is used to obtain the defective part.

Literature [5] a new clustering algorithm is used to detect surface defects of ceramic tableware. This method uses defect-free images as training samples. In the training stage, a binary tree clustering algorithm is used to cluster the training samples. Specifically, a consistency measurement criterion based on principal component analysis (PCA) is proposed to evaluate each cluster. If a cluster has the lowest metric score, the fuzzy C-means (FCM) algorithm is used to split the cluster into 2 new cluster. In the
testing phase, the distance between the input data and the centre of each cluster is calculated to measure whether the input data contains defects. As shown in Figure 5.

![Figure 5](image)

**Figure 5.** The application of clustering algorithm in the detection of surface defects of ceramic tableware

4.3. Detection methods based on machine learning

This section mainly introduces the surface defect detection method based on machine learning. The main machine learning methods used include: support vector machine (SVM), independent component analysis and deep learning. The following is a brief introduction to some representative work in such methods. The paper proposes a detection method for surface defects of ceramic tableware based on support vector machine. In the training mode, the method obtains a set of samples marked with cracks and non-cracks in advance, and then uses a set of local descriptors to extract features from these samples, to use the obtained feature vector to represent the sample, and send the feature vector into SVM for training. Among them, the crack widths in the crack samples are different, making the sample set more representative. During online detection, the same local descriptor is used to obtain the feature vector of the input image, and then sent to the SVM to determine whether there is a crack [6]. The block diagram of the algorithm flow is shown in Figure 6.

![Figure 6](image)

**Figure 6.** Block diagram of the algorithm flow of the machine learning detection method
5. Conclusion
With strong feature extraction capabilities, deep learning-based ceramic tableware surface defect detection has become a research hotspot in this field. However, due to the small number of manual annotation samples and the relatively insufficient generalization ability of existing methods, this problem will be weak in the future. Label samples (only mark whether they contain defects or types of defects, but do not know the specific location of the defects, which can reduce the difficulty of manual labelling and facilitate the increase of the number of samples). Semi-supervised learning may be the next research direction. One of the main difficulties in detecting surface defects of current ceramic tableware is that when the brightness difference between the defect and the background area is small, the detection effect is relatively poor. It may be a research direction in the future to study the feature extraction algorithm of subtle defects or design a feature extraction framework based on deep neural networks.

References
[1] Berger, S., Stamminger, R., Wiebke M. Schünemann, & André Lipski. Development of a method for the analysis of microbial load reduction factors on dishes cleaned by hand and by machine. Tenside Surfactants Detergents, 52 (3) (2015) 206-212.
[2] Gradmann, R., Berthold, C., & Schüssler, Ulrich. Composition and colouring agents of historical Islamic glazes measured with epma and μ-xrd2. European Journal of Mineralogy, 27 (3) (2015) 701-701.
[3] Waraporn Pongmonjit, Kla Somtrakool, & Wilailak Ratanapeantamma. Conservation and export development of mon pottery in central thailand. International Journal of Academic Research in Business & Social Sciences, 5 (2) (2015) 5-17.
[4] Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. Machine learning for high-throughput stress phenotyping in plants. Trends in Plant science, 21 (2) (2015)110-124.
[5] Obermeyer, Z., & Emanuel, E. J. Predicting the future — big data, machine learning, and clinical medicine. New England Journal of Medicine, 375 (13) (2016) 1216-1219.
[6] Long, N, Gianola, DRosa, GJ, Weigel, KA, & Avendano, S. Machine learning classification procedure for selecting snps in genomic selection: application to early mortality in broilers. Developments in biologicals, 124 (6) (2015) 377-389.