International Journal of Information Technologies and Systems Approach
Volume 17 • Issue 1 • January-December 2024

A Research on Character Feature Extraction for Computer Vision and Pattern Recognition

Xiaoyuan Wang
https://orcid.org/0009-0000-2745-6516
Basic Experiment and Training Center, Hefei University, Hefei, China

Hongfei Wang
Sinosoft Company Limited, Beijing, China

Jianping Wang
College of Electrical Automation, Hefei University of Technology, Hefei, China

Jingjing Ge
Basic Experiment and Training Center, Hefei University, Hefei, China

Haiyan Dong
Basic Experiment and Training Center, Hefei University, Hefei, China

ABSTRACT

Character feature extraction is a key area in computer vision and pattern recognition. Traditional methods often rely on manually designed extractors, which struggle with capturing complex structures and abstract features in character images, limiting their performance. The training and tuning of these models require considerable computational resources and time, reducing efficiency. This paper explores and compares various character feature extraction methods. It integrates two-dimensional wavelet decomposition with grid-based statistical and structural features. A detailed design of wavelet coarse and fine grid feature vectors is presented, starting with the construction and extraction of wavelet coarse grid feature vectors, followed by the finer grid feature vectors. The wavelet fine grid features demonstrate stronger specificity and discrimination than the coarse grid features. Experimental validation on 108 character samples yielded a 97.4% success rate, confirming the practicality and effectiveness of the proposed feature extraction method.

KEYWORDS

Feature Vectors Research, Character Feature, Extraction

INTRODUCTION

The world has entered the era of big data with the rapid development of the internet and the widespread adoption of information technology, bringing both unprecedented opportunities and challenges, particularly in Natural Language Processing (NLP). Character feature extraction is a fundamental task in NLP that plays a crucial role in applications like text classification, sentiment analysis, and machine translation. Initially, feature extraction relied on manually designed methods such as TF-IDF (TF-IDF is a method that measures the importance of words in a specific text by combining term frequency and inverse document frequency, commonly used in text classification and

DOI: 10.4018/IJITSA.366037

retrieval) and the BoW (Bag-of-Words model) model, which, despite their effectiveness in some tasks, struggled with complex linguistic phenomena. The advent of machine learning and deep learning has led to the evolution of these methods from rule-based approaches to data-driven ones, which leverage large-scale data to learn richer language patterns with improved generalization. However, the handling of vast data volumes presents challenges, including efficient storage and processing, meaningful feature extraction, and minimizing the impact of noise and bias, which have resulted in ongoing research in the areas of data preprocessing, feature extraction, and model optimization.

In the field of character recognition, the choice of feature extraction methods is crucial as it directly affects the accuracy and efficiency of recognition. This paper aims to analyze and compare existing character feature extraction methods, evaluating their overall performance and the statistical characteristics of inherent fields in characters. Building on prior laboratory research, we propose wavelet coarse grid features and wavelet fine grid features, which offer advantages in statistical and structural features, respectively. The latter is particularly optimized for confusing characters in license plates. This paper first briefly introduces various feature extraction methods, and then, leveraging the properties of two-dimensional wavelet decomposition, elaborates on the construction and extraction algorithms for wavelet coarse and fine grid feature vectors. Finally, the effectiveness and feasibility of these extraction methods are validated through the analysis of specific examples.

Character feature extraction is one of the core tasks in character recognition. The commonly used methods in this research area include shape-based feature extraction, texture-based feature extraction, and statistical feature extraction. Shape-based feature extraction methods include contour analysis, edge detection, and geometric feature extraction. Texture-based feature extraction methods mainly use texture statistics, the wavelet transform, and filtering techniques. Statistical feature extraction methods use statistical features to describe the distribution and shape of characters.

In the character recognition field, the choice of feature extraction methods is crucial as it directly affects the recognition accuracy and efficiency. The aim of this paper is to analyze and compare existing character feature extraction methods, evaluating their overall performance and the statistical characteristics of inherent fields in characters. Building on Based on existing research, we propose coarse grid and fine grid wavelet features. The latter is particularly optimized to deal with confusing characters that appear on license plates. This paper first briefly introduces various feature extraction methods. Second, it leverages the properties of two-dimensional wavelet decomposition, and elaborates on the construction and extraction algorithms for coarse and fine grid wavelet feature vectors. Last, the effectiveness and feasibility of these extraction methods are validated through the analysis of specific examples.

ANALYSIS OF THE FEATURE EXTRACTION METHOD

First, the fundamental goal of feature extraction is to find a more effective character feature vector in order to achieve character recognition (Cheng et al., 2023). Character features can be broadly divided into two categories: structural features and statistical features This paper stresses that the feature extraction should adhere to the following criteria:

1) Reliability. Similar objects should have relatively close feature values. Consider the example of green and ripe apples, which although have significant color differences, but are still both apples. Hence, color is not an ideal feature for distinguishing apples of different ripeness levels.
2) Discriminability. Objects from different categories should have distinctly different feature values.
3) Limited quantity. The complexity of pattern recognition systems increases rapidly as the number of features increases. This is particularly important because the numbers of samples required for training the classifiers and testing the results grow exponentially as the number of features increases. In some cases (Xu et al., 2023), it may even be impossible to obtain enough samples to train the classifiers. Furthermore, adding features with noise or high correlation can actually

degrade the classification ability of the classifiers, especially when the training set has a limited size.

4) Independence. The selected features should be unrelated to one another. Although highly correlated features can be combined (e.g., through mean subtraction) to reduce noise, they generally should not be used as individual features.

Methods for Extracting Structural Features

The feature extraction methods are divided into two categories: structural feature extraction and statistical feature extraction. In structural feature extraction, determination of the structural information represented by primitivesis of key importance. At present, the structural features acquisition mainly relies on skeletons, contours, strokes, etc. The skeleton of characters is obtained through thinning methods, which include iterative peeling and direct acquisition. Structural features based on skeletons include feature points, such as endpoints, intersections, turning points, etc. (Zhang et al., 2023). The image thinning quality plays an important role in the extraction of skeleton-based features. As the existing thinning algorithms cause changes of varying degrees in the topological structure, such as Y-shaped forks, spikes, broken lines, etc., the subsequent recognition should have a greater flexibility. At present, skeletons are usually only used for differentiation within small class sets or as an additional discriminative criteria (Waluyo, Isnanto, & Rochim, 2023).

Contours can also reflect the structure of character images, and there are roughly two ways to describe the contours: the first one uses arcs and structural points to construct the image, while the second one measures the distance between standardized borders and contours. Compared to skeletons, contours contain more precise positional information and have a lower computational complexity than thinning (Chopparapu, & Seventline, 2023). However, they can be affected by stroke width and discontinuities, and are more suitable for environments with better image quality and relatively consistent writing. According to certain research work(Qulub et al., 2023), locally significant features with topological significance can be directly extracted from images, such as pivots, holes, and inner and outer concave-convex points within local images.

Statistical Feature Extraction

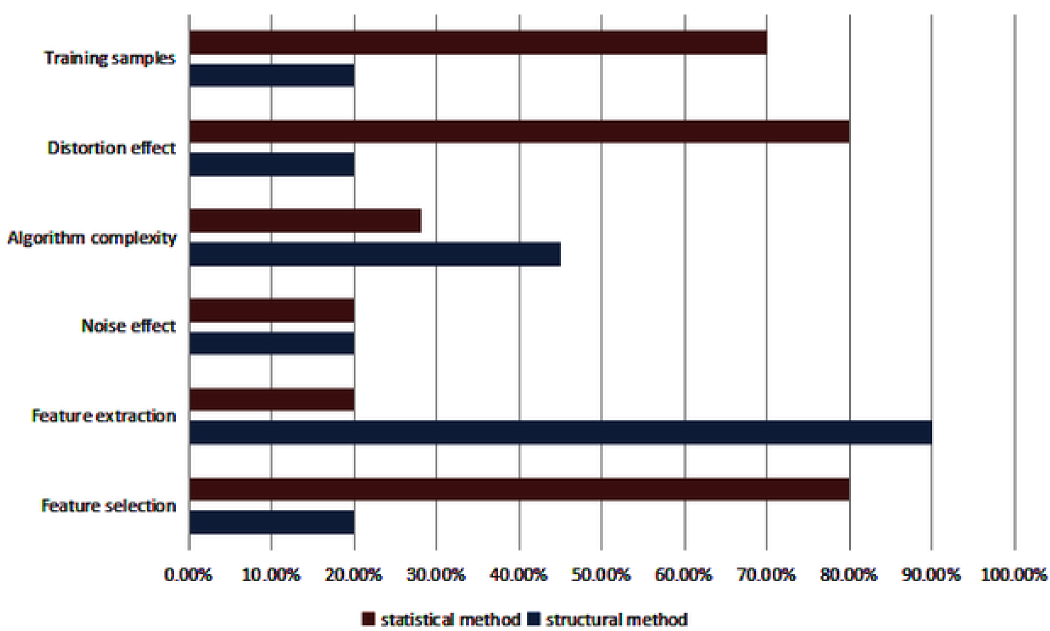
Statistical feature extraction obtains the most relevant information from raw data for classification, minimizing intra-class differences and maximizing inter-class differences. Statistical features can be divided into global and local features, where the former transform the entire character image into coefficients, which are used as the image features. Global statistical feature extraction methods mainly include the Fourier transform, Karhunen-Loeve (KL) transform, Cosine, Sine, Gabor transform, Hadamard transform, Rapid transform, Hough transform, algebraic features, moment features, coarse peripheral features, stroke density features, etc. The Fourier transform is a widely used mathematical transformation that is widely adopted due to its fast transformation method (DFFT) and a clear theoretical description (Y. V. Guntara, 2023). Moment features are widely used and researched because they have translation, rotation, and scale invariance, which match the human visual characteristics. The KL transformation concentrates information on the least number of dimensions of feature vectors. When the number of dimensions extracted after the KL transformation is fixed, it results in the minimization of the variance before and after transformation (Marcato et al., 2023).

Local features refer to transforming images using specific window positions and sizes, and mainly include local grayscale features, projection features, and directional line element features. Local grayscale features, also known as coarse grid features, are obtained by dividing standardized images into fixed or elastic grids. Subsequently, the average grayscale or the number of target pixels in each grid are counted to obtain feature vectors with a dimension equal to the number of grids (Akbar et al., 2023). Projection features involve projection of standardized images in the M and N directions to obtain two B-dimensional feature vectors, and have a low computational complexity with good discriminability for coarse classifi cation. Directional line element features and their variations are one of the most widely used and effective feature types in current systems. These features divide characters into certain grids and classify adjacent black points along different directions within each grid into several categories (Achmad et al., 2023). The number of directional line elements along various directions within a local area reflects the direction of strokes in that area. It comprehensively represents character information and exhibits a good temporal performance. However, ordinary directional line element features only consider fixed grids, and slight displacements near the grids may cause feature mutations (N. V. Gubanov, 2023). To effectively overcome deformation problems in character recognition, the fuzzy mathematical concepts were used to introduce fuzzy directional line element features for characters (Riadi et al., 2023). Additionally, elastic grids for stroke direction feature extraction were proposed in the aforementioned reference to effectively address issues such as unstable stroke positions and local glyph deformation caused by different writing styles.

Analysis and Comparison of Statistical and Structural Methods

As aforementioned, the features' stability will directly affect the performance of character recognition systems. Therefore, extraction and selection of features is one of the prerequisites and key steps for character recognition (Ardelean et al., 2023). The statistical and structural feature extraction methods have their own advantages and disadvantages, which are illustrated in Fig. 1 and described as follows:

Figure 1. Statistical method and structural method performance comparison diagram



1) One of the main advantages of the structural feature method is its ability to describe the structure of characters, breaking down character patterns into strokes, stroke segments, and radicals. It can effectively integrate geometric and structural knowledge during the recognition process, thus obtaining recognition results with relatively high reliability. However, it has a high computational

complexity, which makes it difficult to characterize line segments with unclear ownership or cause error-prone encoding (Siraji et al., 2023). Experiments have shown that it is difficult for any single feature to perfectly represent any pattern (Mayatopani et al., 2023).

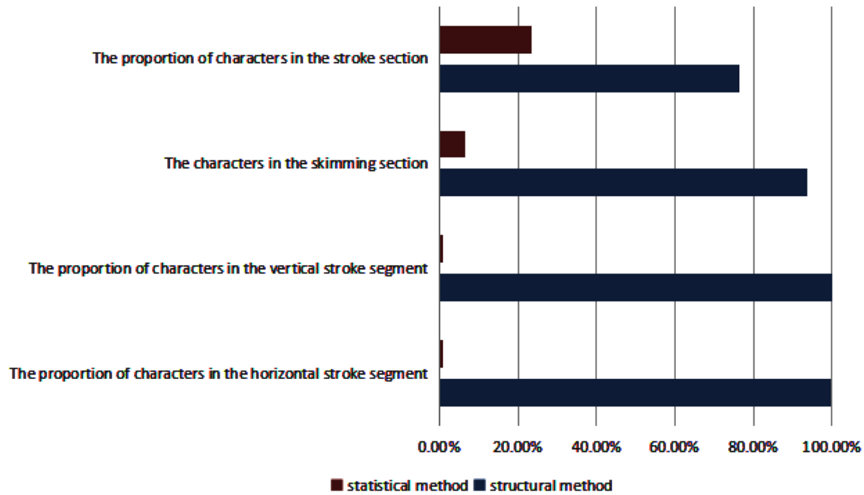
2) Once a certain feature is determined using a statistical feature method, the extraction algorithm is simple, easy to train, and can achieve relatively high recognition rates on a given training set. However, feature selection is the most difficult aspect because the morphology of characters is highly variable with no clear standards, which makes it extremely challenging to separate thousands of character patterns (Naushad et al., 2023). This difficulty can increase the dimensionality of feature vectors, which increases the difficulty of processing. Inappropriate features not only have lower utility in pattern classification but can also affect the separability of pattern categories. Therefore, it is crucial to find relatively stable features in a large amount of unstable data. As it is difficult to find the most important features in practical problems or measure them due to conditional limitations, the feature selection and extraction task becomes one of the most difficult ones in building pattern recognition systems (Dryuchenko, 2023). Therefore, organically combining the advantages of the two aforementioned methods is a topic worthy of in-depth research.

CHARACTER FEATURE ANALYSIS

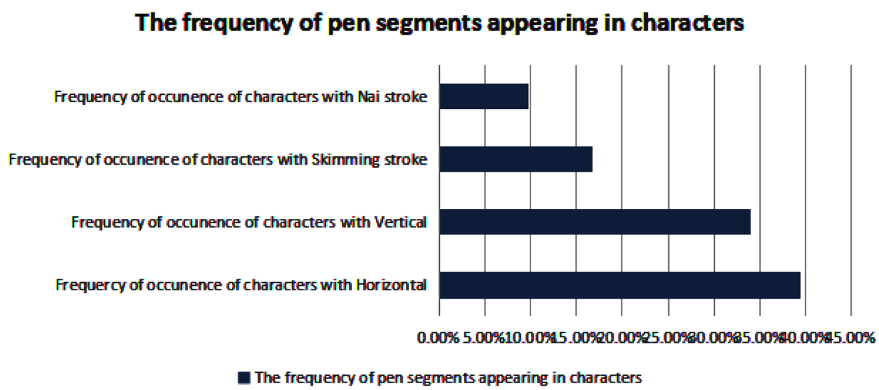
The preceding discussion serves as a reference for analyzing feature extraction standards. It is known that the Chinese characters are composed stroke-by-stroke, where each stroke is formed by writing continuously in a certain direction. A stroke typically represents a single character (such as " ", " ", etc.), but most Chinese characters are composed of multiple strokes. Basic strokes can be viewed as stroke segments, while compound strokes can be decomposed into stroke segments based on inflection points, where a stroke segment is defined as a collection of pixels having a grayscale value of one in the same direction (Hong et al., 2023). Although difficulties exist in automatic extraction of each type of stroke from characters using image processing techniques, it is possible to extract the four basic strokes of vertical, horizontal, left-falling and right-falling strokes. However, at present, there are issues with ambiguous stroke attribution and encoding errors.

Statistical data indicate that vertical, horizontal, left-falling, and right-falling strokes are the most widely used in Chinese characters. It was found from a survey of 6763 international two-level Chinese characters that those with vertical strokes, horizontal strokes, left-falling strokes and right-falling strokes account for 99.85%, 99.8%, 76.5% and 93.5%, respectively. This indicates that these few stroke forms constitute a variety of complex structures through different relationships, forming thousands of different characters. Additionally, there are significant differences in the occurrence frequency of these four stroke types in characters, with vertical strokes, horizontal strokes, left-falling strokes, and right-falling strokes accounting for 33.94%, 39.5%, 9.78% and 16.77%, respectively.

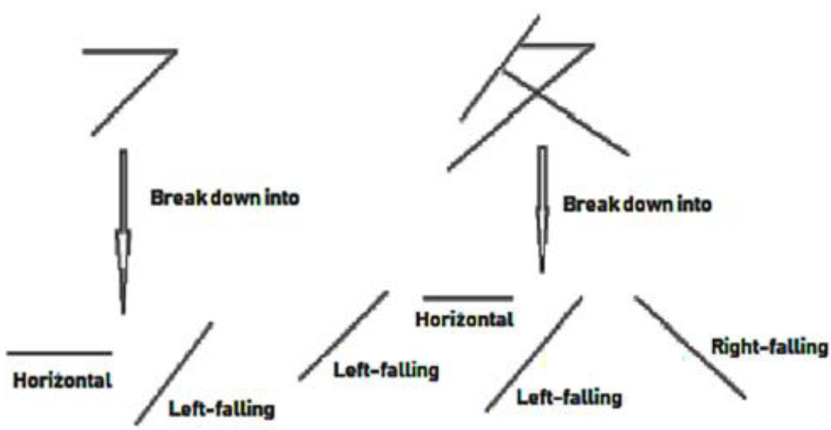
Figure 2. Analysis of the proportions and frequencies of basic strokes and stroke segments in Chinese characters and illustrative diagram of composite stroke decomposition



(a) Proportional distribution of basic strokes in Chinese characters



(b) Frequency of basic strokes appearing in Chinese characters



(c) Schematic diagram of compound strokes decomposed into basic stroke segments

English letters and numbers can also be considered to be composed of single or multiple strokes because character images are generally represented as digital raster images (Malik et al., 2023). Therefore, certain English characters, such as P, O, Q, etc., consist of arc-shaped stroke segments that can be classified as left-falling and right-falling strokes. Thus, English letters and numbers also include vertical, horizontal, left-falling and right-falling strokes. Based on this discussion, we can conclude the following two points:

1) The structure of characters can be described using four basic stroke segments: vertical, horizontal, left-falling, and right-falling strokes;
2) The stroke length can be used for quantitative description of characters.

The wavelet method is widely used in image and signal processing, enabling multi-scale analysis and applications like image compression, denoising and edge detection (Fukagata, 2023). In the pattern recognition field, wavelet transforms provide rich information that is useful for feature extraction and selection (Sari & Bantun, 2023), and for fine-tuning different parameters to improve the algorithm performance (Adekunle et al., 2023). The two-dimensional wavelet transform effectively decomposes strokes, avoiding encoding errors while combining left-falling and right-falling strokes (Katili et al., 2023).

Fine grid wavelet features aim to describe the local structure and texture information of images more precisely compared to coarse grid features, enhancing feature discriminability (Yadav & Yadav, 2023). These feature capture more subtle variations, making them more effective for character recognition tasks.

Therefore, this paper proposes fine grid wavelet features to deal with ambiguous characters(Saikhu, A., Setyadi & Hariadi, 2023). We analyze the structure of these characters and select effective fine grid statistical information that aligns with the feature descriptions obtained from the two-dimensional wavelet transform (Rameshbabu & Ramakrishnan, 2023). This extraction method reduces the data volume, enhances the distinguishability of ambiguous characters, and elevates the overall effectiveness of wavelet grid features.

RESEARCH ON WAVELET GRID FEATURE EXTRACTION METHODS——THEORETICAL PART

Construction and Design of Wavelet Grid Features

Based on the above statements, we can segment the normalized structural feature distribution map into appropriate grids. Subsequently, we can calculate the feature values of each grid to generate the feature vectors of image characters (Zeydabadinezhad, Horn, & Mahmoudi, 2023). As mentioned earlier, the length of character strokes can quantitatively describe the characters, while the grayscale statistical features of character images can provide an alternative description of the character stroke length information. Therefore, these features represent the dimensions and the numerical values reflect the character feature vectors.

In Definition 1, we assume that the grayscale levels of character images range from 1 to x, and the number of pixels with grayscale value h is equal to yh. Therefore, the total number of pixels can be represented as

$$B = \sum_{h=1}^{x} y_1 \, h = 1,2,\cdots ,x \; ; \tag{1}$$

The mean grayscale value of the image is given as:

$$A = \sum_{h=1}^{x} hQ_h \, h = 1,2,\cdots,x \; ; \tag{2}$$

The probability of each grayscale value is:

$$Q_h = \frac{y_h}{B} h = 1,2,\cdots,x \; ; \tag{3}$$

The mean grayscale value of the image characters reflects the trend of the overall grayscale values of the image. Essentially, it is a quantitative calculation of the wavelet feature vectors of the image characters, which represents the statistically "expected" grayscale value.

We define the following to qualitatively and quantitatively describe the wavelet grid feature vectors:

Definition 2: By dividing the image area of dimensions x*y into l regions as needed using a set of uniform or non-uniform imaginary grid lines,

$$C=\{A_h\}=[A1\ A2\ A3\ A4\ \ldots\ldots\ldots],h=1,2,\ldots\ldots\ldots l \tag{4}$$

is calculated for each grid, which is then referred to as the grid feature vector of the image character, where h=1, 2, ..., l, are calculated in the order of column first followed by the row. The grid feature vector of the character image enhances the distinguishability of similar image character features, essentially making them "independent" and suitable for use in the fine quantitative calculation processing of character wavelet feature vectors.

The methods used to calculate the average grayscale value and the grid feature vectors of character images are applied to these structural subgraphs, forming the image character wavelet grid feature vectors, which are defined as follows:

Definition 3: The K-scale wavelet decomposition subgraphs of the image are represented as follows: El………Smooth subgraph; Fl………Horizontal subgraph; Ul……… Vertical subgraph; Il………Diagonal subgraph. The corresponding grid feature vectors are calculated for each subgraph, which are then arranged and combined in the following order: El, Fl, Ul, Il;………; El,Fl,Ul,Il. For the lth scale wavelet decomposition subgraph, we obtain:

$$G_l = \left\{C_{E_l}\, C_{F_l}\, C_{U_l}\; C_{I_l}\right\} \tag{5}$$

Therefore, the K-scale wavelet decomposition of the image provides the following wavelet grid vector:

$$G_{l.h} = \begin{bmatrix} A_{E_l(1)} & A_{E_l(2)} & A_{E_l(3)} & \cdots & \cdots & A_{E_l(h)} \\ A_{F_l(1)} & A_{F_l(2)} & A_{F_l(3)} & \cdots & \cdots & A_{F_l(h)} \\ A_{U_l(1)} & A_{U_l(2)} & A_{U_l(3)} & \cdots & \cdots & A_{U_l(h)} \\ A_{I_l(1)} & A_{I_l(2)} & A_{I_l(3)} & \cdots & \cdots & A_{I_l(h)} \end{bmatrix} \tag{6}$$

where l represents the level of wavelet decomposition, and h denotes the image grid. We set the wavelet decomposition level to three by trial-and-error after multiple experiments and tuning, because the character recognition accuracy at this decomposition level is high, and the computational complexity remains moderate. The wavelet feature vector of the image is denoted by Gl,h, with every four elements grouped to correspond to the statistical values of four structural sub-images. This grouping is similar to the way the human eye summarizes the character forms, such as vertical lines, horizontal lines, strokes and curves.

In this paper, the Morlet wavelet and the cubic B-spline wavelet are used as the basis functions. The former is suitable for capturing the high-frequency details and local oscillation features, making it particularly suitable for handling characters with complex shapes. On the other hand, the cubic B-spline wavelet focuses on low-frequency information, which allows it to effectively capture the overall contours and smooth parts of characters, and makes it especially useful for recognizing blurred or low-resolution characters.

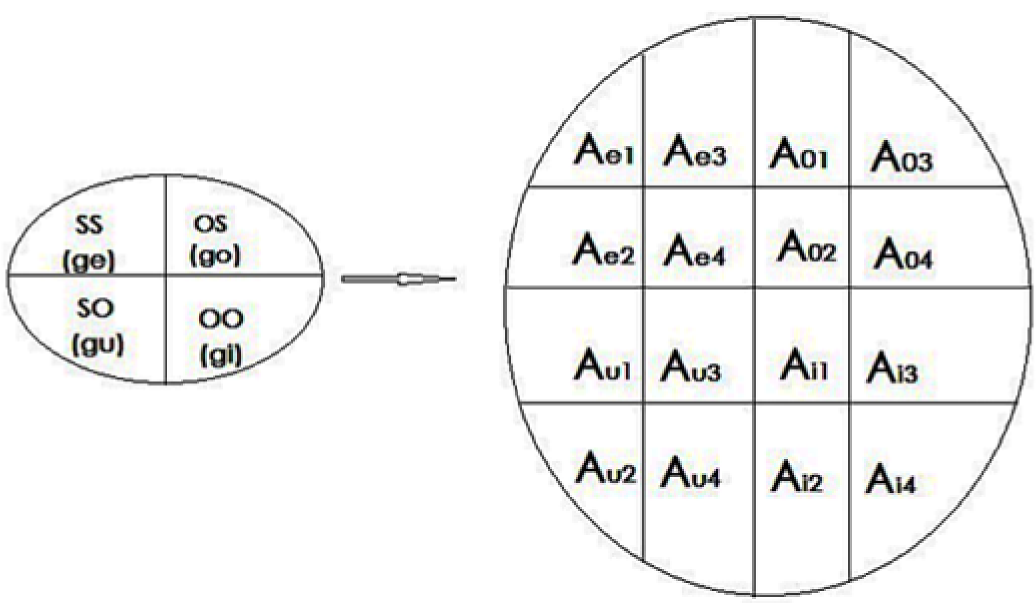
Construction and Extraction of Coarse Grid Wavelet Feature Vectors

For l=1, the image is uniformly divided into four regions using a grid method. A one-dimensional wavelet feature vector G1,4 consisting of 4*4 elements is obtained based on the aforementioned wavelet feature vector construction rules.

$$G_{1,4} = \begin{bmatrix} A_{E_1(1)} & A_{E_1(2)} & A_{E_1(3)} & A_{E_1(4)} \\ A_{F_1(1)} & A_{F_1(2)} & A_{F_1(3)} & A_{F_1(4)} \\ A_{U_1(1)} & A_{U_1(2)} & A_{U_1(3)} & A_{U_1(4)} \\ A_{I_1(1)} & A_{I_1(2)} & A_{I_1(3)} & A_{I_1(4)} \end{bmatrix} \tag{7}$$

Denoting E1 = β, F1 = γ, U1 = φ, and I1 = ∮, the above feature vectors are arranged in the following order:

Figure 3. 2*2 wavelet grid feature vectors



The coarse grid wavelet feature vector of the character image arranged according to the sequence given in (7) is obtained as follows: $\mathbf{E}=A_4 = [A_{e1} A_{e2} A_{e3} A_{e4} A_{f1} A_{f2} A_{f3} A_{f4} A_{u1} A_{u2} A_{u3} A_{u4} A_{i1} A_{i2} A_{i3} A_{i4}]$

where A_{hl} (h=e,o,u,i;l=1,2,3,4) represents the average grayscale value of the grid pixels. Table 1 provides a set of character wavelet feature vector data.

Construction and Extraction of Coarse Grid Wavelet Feature Vectors

Definition 4: Using a wavelet-transformed character image of size x*y, a set of hypothetical gridlines is considered to divide this image into x*y grids, so that each pixel occupies one grid. The fine grid wavelet characteristics can be defined as follows:

$$f = \sum \sum_{h=l_1}^{l_2} \sum_{k=b_1}^{b_2} A_t(h,k) \; l_1, l_2 = 1, \cdots \cdots x; b_1, b_2 = 1, \cdots \cdots y \; ; \tag{8}$$

In the above expression, t=e, o, u and i represent low frequency, horizontal, vertical and diagonal sub-images, respectively, and At(h,k) represents the grayscale average over the (h,k) grid in the t sub-image. The original image size should be scaled to 32*16 pixels to obtain one-scale wavelet decomposition using the biorthogonal wavelet syml. Based on the differences between the character structure and its wavelet components, the fine grid wavelet feature vector is obtained as follows: [Feature vector distinguishing between P and O, Q (left part of the vertical component)]

$$f = \sum_{h=6}^{12} \sum_{k=4}^{6} A_0\left(h,k\right) \tag{9}$$

Feature vector distinguishing between P and R (bottom-right part of the diagonal component):

$$f = \sum_{h=1}^{16} \sum_{k=1}^{4} A_u\left(h,k\right) \tag{10}$$

Feature vector distinguishing between O and Q (bottom-right part of the diagonal component):

$$f = \sum_{h=1}^{16} \sum_{k=1}^{4} A_i\left(h,k\right) \tag{11}$$

Feature vector distinguishing between E and F (bottom horizontal component right-bottom part):

$$f = \sum_{h=1}^{8} \sum_{k=5}^{8} A_e(h,k) + \sum_{h=1}^{8} \sum_{k=5}^{8} A_i\left(h,k\right) \tag{12}$$

It can be observed from the above that the fine grid wavelet characteristics reflect the differences in local structures between similar characters, which are the most distinguishable by the human brain.

ANALYSIS OF COMPUTATIONAL COMPLEXITY AND COMPARISON WITH EXISTING METHODS

This section examines the computational complexity of the proposed algorithm and compares it with those of the existing methods. The overall complexity of the algorithm is O(N\logN + I·K·N), where N represents the total number of pixels in the image, K is the number of clusters, and I is the number of iterations. It exhibits significant advantages over traditional K-means algorithms and direct wavelet transforms-based image processing methods in handling complex image features and improving the clustering accuracy.

REASONING BEHIND THE CHOICE OF THE PROPOSED ALGORITHM

In the proposed algorithm, first, the grayscale mean (A) of an image that reflects its overall brightness level is calculated by determining the weighted average of each grayscale value, and the probability of each grayscale value (Qh) is derived from the frequency of its occurrence. Second, the image is divided into multiple uniform or non-uniform regions and the local grayscale mean (Ah) for each region is calculated, which helps in capturing local features and image details. Third, a multi-scale wavelet transform is applied to the image to obtain subgraphs at different scales, including smooth (E), horizontal (F), vertical (U), and diagonal (I) subgraphs. The grayscale mean and other statistics of these subgraphs are used to construct the feature vectors that describe the structural information of the image at various scales and directions. Fourth, the local standard deviation (σ) of grayscale values is calculated within each region as follows:

$$\sigma_h = \sqrt{\frac{1}{n_h}\sum_{i=1}^{n_h}\left(y_{h,i} - \mu_h\right)^2} \tag{13}$$

where $y_{h,i}$ is the grayscale value of the i-th pixel in region h, and μ_h and n_h are the grayscale mean and number of pixels in region h, respectively. The local standard deviation is used in the proposed algorithm as it reflects the texture complexity and dispersion degree of different regions in the image. It is crucial for image segmentation and feature extraction, and helps in the identification of edges, textures and other local features.

Last, the algorithm includes an optimized initial center selection for clustering based on K-means++ initialization. (The relevant reference for K-means++ initialization is Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 1027–1035.) It improves the accuracy and stability of the clustering results by probabilistically selecting initial centers that are likely to produce better clustering outcomes. This step identifies different feature regions in the image. The proposed algorithm has significant applications in image processing and pattern recognition, effectively enhancing the image features extraction and analysis.

Research on Wavelet Grid Feature Extraction Methods——Experimental Part

By analyzing and comparing the comprehensive performance of existing character feature extraction methods, the inherent statistical data features of the characters, and the directionality of image wavelet decomposition, we propose the following:

Construction and Design of Wavelet Grid Features

That the two-dimensional wavelet transformation of images has directional selectivity. Therefore, for characters that have directional strokes, it is unnecessary to perform complex stroke preprocessing to obtain sub-images in various directions: horizontal sub-images, smooth sub-images, diagonal sub-images, and vertical sub-images(Alruwais et al., 2023) .

Focusing on the recognition target, i.e., the license plates, Fig. 4 shows the wavelet decomposition of digits, Chinese characters, and letters.For a clear visual demonstration, the standard printed characters are chosen for decomposition(Gondim, Bompan, & Haach, 2022).

Figure 4. Schematic diagram of single scale wavelet decomposition of characters

Original Image Smooth Sub-image (SS) Horizontal Sub-image (OS) Vertical Sub-image (SO) Diagonal Sub-image (OO)

(a) Wavelet decomposition of the digit "2" at a single scale

Original Image Smooth Sub-image (SS) Horizontal Sub-image (OS) Vertical Sub-image (SO) Diagonal Sub-image (OO)

(b) Wavelet decomposition of the Chinese character "安" (An)

Original Image Smooth Sub-image (SS) Horizontal Sub-image (OS) Vertical Sub-image (SO)

(c) Wavelet decomposition of the character "E"

It can be observed that the two-dimensional wavelet transform components exhibit directionality, which is consistent with human visual characteristics. We handle four different sub-images separately for the processed images: the OS sub-image mainly provides vertical details and is insensitive to horizontal details; the SO sub-image highlights the horizontal details (Meng et al., 2023); while the OO block is more sensitive to diagonal information, especially in the direction of left-falling and right-falling strokes for character images.

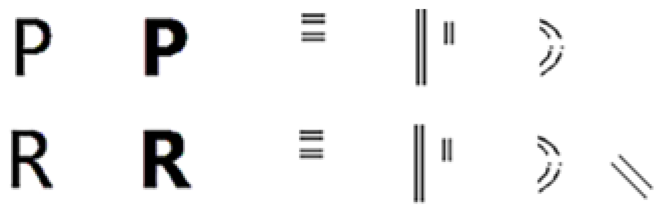
Construction and Extraction of Coarse Grid Wavelet Feature Vectors

We choose characters from standard templates to describe the directional visual characteristics of wavelets more clearly. The data in Table (X) reveal the following:

1) Different characters generally differ significantly in grid feature vectors under structural sub-images. Therefore, the wavelet feature vectors between characters have large vector distances, which provide a good foundation for character recognition.
2) It is not obvious to attribute a category to easily confused characters. Examples of easily confused character pairs are “P” and “R”, “O” and “Q”, and “E” and “F”. Thus, feature vectors with strong discriminative power need to be further proposed for easily confused characters.

The coarse grid wavelet feature vectors do not have strong discriminative power for easily confused characters. This is because the four-grid uniform division of wavelet decomposition sub-images may lead to vectors corresponding to different stroke features in similar characters to be allocated to the same grid. Consider the example of “P” and “R” shown in Fig. 8: In the second and fourth grids of the low-frequency, horizontal, vertical and diagonal sub-images, the corresponding grid regions of the two images have similar pixel distributions. According to distance measurement standards, the vector feature distance between these two types of characters is relatively small from both the overall and individual grid perspectives. Therefore, it is difficult to determine their category as they belong to ambiguous samples.

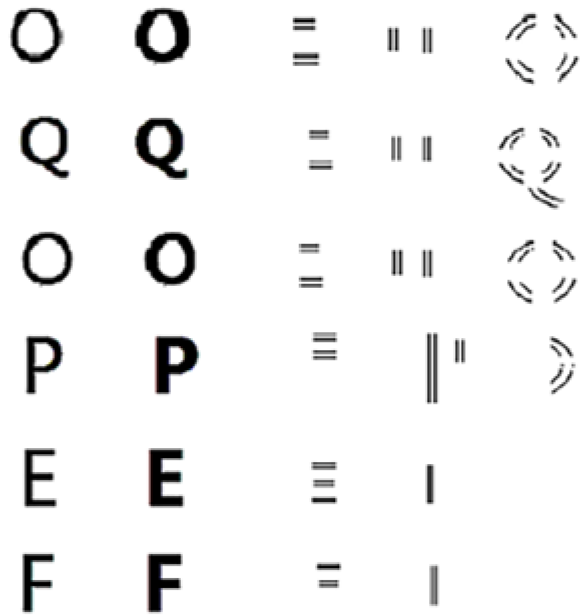
Figure 5. Grid decomposition schematic of characters “P” and “R”



Construction and Extraction of Coarse Grid Baud Eigenvector are Carried Out

Based on the character structure and wavelet decomposition features, we propose fine grid wavelet features with a strong specificity. The similarity of style distribution leads leads pixels to be distributed similarly in similar characters causing interference among them. The character “P” has an additional vertical stroke on the left compared to “O” and “Q,” which is reflected in the wavelet transform and results in more pixel contents on the left side of “P.” The distinction between “O” and “Q” lies in the higher pixel count in the lower right portion, which can be differentiated using the wavelet diagonal component. There is no confusion between “E” and “F” in the coarse-level classification because “E” has a horizontal stroke beneath it, which is visible in the horizontal sub-image. Although both “E” and “F” have significant pixel distributions in the lowest horizontal sub-image, there is a significant difference between their horizontal distributions, which allows the grid statistical information to be used as a distinguishing feature. Fig. 6 depicts wavelet grid decomposition diagrams of several types of characters:

Figure 6. Illustration of network decomposition of characters O, Q, O, P, E and F



This paper proposes fine grid features that decompose the character images into different frequency components and consequently, capture both global shapes and local details across multiple scales. Although high-frequency details may be lost even in low-resolution conditions, the overall contours and shapes of the characters can still be clearly extracted from the low-frequency components of the wavelet decomposition. Furthermore, similar characters can be distinguished using local features obtained through higher-level decompositions, which help in maintaining a high recognition performance. The fine grid features divide the image into detailed grids, extracting pixel distribution information from each local region. These local features provide sufficient information for classification even in low-resolution images. For example, local textures and structural features can still retain subtle differences although the overall shape may be blurred, which enhances the recognition accuracy.

The combination of wavelet decomposition and fine grid features can be used to mitigate the information loss caused by image downsampling. Fine grid wavelet features extract rich details from high-resolution images, especially edges, strokes and local textures, while low-frequency components preserve the global contours of the characters.

In addition, the scale and diversity of the dataset also affect the character feature extraction performance (Guo et al., 2023). Future research work should focus on addressing the challenges in image character feature extraction to improve the character recognition accuracy and robustness.

Data Sources and Feature Extraction Methods in this Study

This study explores character or signal feature extraction methods by gathering data from various sources, including sensor data, human trials, and simulated environments. Sensor data are collected using devices such as image and acoustic sensors, while human experiments involve participants adding input characters via tablets or keyboards. Data are also simulated for testing the algorithms under specific conditions. The wavelet transform is used to decompose the signals into coarse (low-frequency) and fine (high-frequency) grids, capturing different character characteristics to extract features including amplitude, phase, and statistical features such as mean and variance. Subsequently, dimensionality reduction techniques like the PCA or LDA are applied to refine these vectors, removing redundancy and preparing them for further analysis or classification.

Limitations and Future Research Directions

The challenges in character recognition primarily stem from the reliance on fine grid wavelet features. Despite capturing local structural differences, the accuracy of these features often decreases when either the shape or position of the character changes significantly. Additionally, the diversity of fonts and character sizes negatively impacts the feature extraction effectiveness, which limits the model's applicability across various character types.

To address the aforementioned issues, future research will focus on enhancing the algorithm's robustness to variations in character morphology, size and position. Issues such as character tilt and rotation can be tackled using advanced image processing techniques to ensure high recognition accuracy in complex scenarios. Moreover, the dataset can be extended by adding more data samples to improve the model's generalization ability, making it applicable to a wider range of real-world situations. Last, optimizing the computational efficiency is crucial to achieving efficient and accurate character recognition in practical applications. These improvements will help overcome the limitations of the proposed method and advance the development of character recognition technology.

Discussion

1) Test the described image processing and feature extraction algorithm, design a simulation environment and choose appropriate parameters

(1) Experimental Environment

Table 1. Hardware and software configuration requirements for image processing tasks

Configuration	Description
Hardware Configuration	
Processor	Multi-core CPU, such as Intel i7 or AMD Ryzen 7, to ensure sufficient computational power.
Memory	At least 16GB RAM to handle large-scale image data.
Storage	Solid-state drive (SSD) to improve data read/write speed.
Image Processing Accelerator (optional)	GPUs, such as the NVIDIA RTX series, to accelerate computationally intensive tasks.
Software Configuration	
Operating System	Windows 10 or Linux (e.g., Ubuntu 20.04).
Programming Language	Python 3.8+.
Required Libraries	NumPy, OpenCV, PyWavelets, scikit-learn, Matplotlib, etc.

(2) Experimental Methods and Parameter Settings

It is crucial to select appropriate datasets and parameters to test the image processing and feature extraction algorithms. The test datasets may include standard image processing datasets such as the MNIST, CIFAR-10 and ImageNet, covering both grayscale and color images. Image resolutions should span multiple scales, such as 28x28, 64x64 and 128x128 pixels, to comprehensively test the algorithm's performance. It is typically necessary to convert color images to grayscale and normalize pixel values between the [0, 1] range in the image preprocessing phase.

Commonly used wavelet transforms are the Daubechies wavelets, e.g., db1 and db2, and one to three levels of decomposition can be chosen to capture features at different scales. Based on the image size and resolution, uniform or non-uniform methods can be chosen to divide the image into multiple regions of size such as 4x4 or 8x8 pixels. The number of clusters can be chosen based on the dataset and actual needs, e.g., K=3, 5 and 10, and K-means++ initialization can be used to optimize the initial centers. Maximum iterations, e.g., 300 iterations, and tolerance, e.g., 1e-4, should be set to ensure the convergence.

(3) Experimental Process and Results Analysis:

During the testing process, the image dataset is first loaded and preprocessed, including grayscale conversion and normalization. Next, the algorithm is implemented according to the chosen parameters, including the mean grayscale calculation, region division, the wavelet transform, feature vector construction, and clustering analysis. The computation time and memory usage are recorded at each step of the algorithm's execution, and intermediate results such as sub-images are saved after the wavelet transform, feature vectors, and clustering results.

Finally, the algorithm's accuracy such as the classification accuracy and feature extraction effectiveness are evaluated, the performance is compared under different parameter settings, such as the computation time, memory usage, convergence speed, etc., and the algorithm is compared with existing methods to identify its strengths and weaknesses. The process is concluded by summarizing the experimental results, analyzing the algorithm's performance, and proposing possible improvements and directions for future research.

2) The results are compared with the existing literature

In the field of character recognition, the recognition accuracy and efficiency depend on the choice of feature extraction methods. Feature extraction involves transformation of the raw image data into feature vectors that can be processed using pattern recognition algorithms. These vectors should effectively represent the intrinsic features of characters and simultaneously exhibit robustness against noise and deformation. This paper analyzes and compares existing character feature extraction methods and evaluates their overall performance and the statistical properties of character. Furthermore, it describes the algorithms for constructing and extracting coarse and fine grid features based on the two-dimensional wavelet decomposition. The experimental results show that the former have high accuracy and lower computational costs in large-scale character recognition, while the latter can effectively reduce the error rates when dealing with confusing characters. Future research can further optimize the feature extraction methods or integrate them with deep learning techniques such as the Convolutional Neural Networks (CNNs) to enhance their robustness and accuracy. Furthermore, specific character sets can be optimized according to particular application scenarios, such as license plate recognition.

The purpose of the proposed design is to fully exploit the powerful capabilities of fine grid wavelet features, making character feature extraction more comprehensive and targeted. The effectiveness of the proposed feature extraction method is verified by experiments performed on 108 actual characters, with recognition accuracy of 97.4%. The experiments prove the feasibility and effectiveness of our feature vector extraction method in practical applications, providing strong support for research in the field of character recognition.

CONCLUSION

(1) This paper analyzed several feature extraction methods and compared them. It combined the two-dimensional wavelet decomposition algorithm of images, using coarse and fine grid wavelet features;

(2) This paper provided details concerning the construction and design of coarse and fine grid wavelet feature vectors. It proposed to construct and extract coarse grid wavelet feature vectors first, followed by the construction and extraction of fine grid feature vectors. The latter features type exhibited stronger specificity and discrimination than the former type.

Character feature extraction is a crucial step in character recognition and has considerable significance in the digital age. Traditional feature extraction methods can improve the performance and effectiveness of feature extraction to a certain extent (Wang et al., 2023). In the future, we need to further study the robustness, generalization ability, and adaptability of character feature extraction methods to cope with the changing application scenarios and demands.

Further research work on character feature extraction can include the following aspects: First, traditional image processing techniques can be used to propose more efficient feature extraction models for enhancing the discriminability and generalization of features. Second, transfer learning and weakly supervised learning can be employed to reduce the dependency on labeled data and improve the model's adaptability. Third, attention mechanisms and global information can be integrated to further enhance the feature extraction performance. Last, key features selection or application of the dimensionality reduction algorithms, such as the PCA, can help in the elimination of redundant information and compress the feature dimensions. In this way, the computational complexity can be reduced without significantly affecting the recognition accuracy, which will enable real-time processing.

In summary, character feature extraction is an important component of character recognition technology that has both theoretical and practical values. It is believed that the continuous development and improvement of technology can further improve the character feature extraction technology and its application in future research.

AUTHOR NOTE

The authors declare no conflicts of interest.

The data used to support the findings of this study are included within the article.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Correspondence concerning this article should be addressed to Xiaoyuan Wang, Basic Experiment and Training Center, Hefei University, Hefei, Anhui, 230601, China. Email: hfuuw20231022@126.com.

PROCESSING DATES

12, 2024

This manuscript was initially received for consideration for the journal on 09/13/2024, revisions were received for the manuscript following the double-anonymized peer review on 11/24/2024, the manuscript was formally accepted on 12/04/2024, and the manuscript was finalized for publication on 12/17/2024

REFERENCES

Adekunle, T. S., Adeleke, T. A., Sunday, O., Ebong, G. N., & Bamisaye, T. A. (2023). A framework for robust attack detection and classification using rap-densenet. *ParadigmPlus*, *4*(2), 1–1. DOI: 10.55969/paradigmplus.v4n2a1

Akbar, S., Sitohang, B., Pardede, J., Amal, I., Yunastrian, K., Ahmada, M., & Prameswari, A. (2023). Folk Games Image Captioning using Object Attention. [Rekayasa Sistem dan Teknologi Informasi]. *Jurnal RESTI*, *7*(4), 758–766. DOI: 10.29207/resti.v7i4.4708

Alruwais, N., Obayya, M., Al-Mutiri, F., Assiri, M., Alneil, A. A., & Mohamed, A. (2023). Advancing retinoblastoma detection based on binary arithmetic optimization and integrated features. *PeerJ. Computer Science*, *9*, e1681. DOI: 10.7717/peerj-cs.1681 PMID: 38077613

Ardelean, E. R., Coporîie, A., Ichim, A. M., Dîn oreanu, M., & Mure an, R. C. (2023). A study of autoencoders as a feature extraction technique for spike sorting. *PLoS One*, *18*(3), e0282810. DOI: 10.1371/journal.pone.0282810 PMID: 36893210

Cheng, R. R., Zhao, X. L., & Zhou, H. J. (2023). Research on Chinese Font Style Transformation Based on Font Features and Multi-Scale PatchGAN. *Yunnan Daxue Xuebao. Ziran Kexue Ban*, *45*(6), 1228–1237.

Chopparapu, S., & Seventline, J. B. (2023). An efficient multi-modal facial gesture-based ensemble classification and reaction to sound framework for large video sequences. Engineering, Technology &. *Applied Scientific Research*, *13*(4), 11263–11270.

Dryuchenko, M. A. (2023). An adaptive image steganography algorithm based on the use of non-cryptographic hash functions for data extraction. *Komputernaia Optika*, *47*(3), 415–425. DOI: 10.18287/2412-6179-CO-1215

Fukagata, K. (2023). Reduced order modeling of fluid flows using convolutional neural networks. *Journal of Fluid Science and Technology, 18*(1), JFST0002-JFST0002.

Gondim, R. M. L., Bompan, K. F., & Haach, V. G. (2022). Cross-correlation for feature extraction applied to ultrasonic test for stress evaluation in concrete. *Revista IBRACON de Estruturas e Materiais*, *16*(4), e16402. DOI: 10.1590/s1983-41952023000400002

Gubanov, N. V. (2023). Software implementation of the system for learning to write Chinese characters. *Herald of Dagestan State Technical University Technical Sciences*, *50*(2), 58–66. DOI: 10.21822/2073-6185-2023-50-2-58-66

Guntara, Y. V. (2023). Pengenalan Kepribadian Melalui Tulisan Tangan Menggunakan Convulutional Neural Network Dengan LS Classifiers. *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi, 14*(2), 151-167.

Guo, F., Li, X. H., Liu, W. T., & Tang, J. (2023). Rapid panorama image stitching algorithm based on parameter regression. *Journal of Communication*, (9), 36–47.

Hong, E., Jeon, S., Kim, N., Kim, H. W., Kang, H., Moon, K., & Woo, J. (2023). Convolutional kernel with PrCaMnOx-based resistive random-access memory for neuromorphic image processing. *AIP Advances*, *13*(1), 015318. DOI: 10.1063/5.0133846

Katili, I. F., Soeleman, M. A., & Pramunendar, R. A. (2023). Character Recognition of Handwriting of Javanese Character Image using Information Gain Based on the Comparison of Classification Method. [Rekayasa Sistem dan Teknologi Informasi]. *Jurnal RESTI*, *7*(1), 193–200.

Malik, H., Naeem, A., Hassan, S., Ali, F., Naqvi, R. A., & Yon, D. K. (2023). Multi-classification deep neural networks for identification of fish species using camera captured images. *PLoS One*, *18*(4), e0284992. DOI: 10.1371/journal.pone.0284992 PMID: 37099592

Marcato, M., Tedesco, S., O'Mahony, C., O'Flynn, B., & Galvin, P. (2023). Machine learning based canine posture estimation using inertial data. *PLoS One*, *18*(6), e0286311. DOI: 10.1371/journal.pone.0286311 PMID: 37342986

Mayatopani, H., Handayani, N., Septarini, R. S., Nuraini, R., & Heriyani, N. (2023). Implementation of Self-Organizing Map (SOM) Algorithm for Image Classification of Medicinal Weeds. [Rekayasa Sistem dan Teknologi Informasi]. *Jurnal RESTI*, *7*(3), 437–444. DOI: 10.29207/resti.v7i3.4755

Meng, Y. B., Liao, Y. M., Tan, F., & Wang, X. H. (2023). Simulation of natural resource land feature extraction under remote sensing image fusion. *Jisuanji Fangzhen*, (9), 162–166.

Qulub, M., Hammad, R., Irfan, P., & Yuliana, Y. (2023). Improvement of Spelling Correction Accuracy in Indonesian Language through the Application of Hamming Distance Method. *Journal of Applied Informatics and Computing*, *7*(2), 271–277. DOI: 10.30871/jaic.v7i2.5926

Rameshbabu, S., & Ramakrishnan, S. (2023). Machine learning approach for diagnosis and prognosis of cardiac arrhythmia condition using a minimum feature set and auto-segmentation-based window optimisation. *Elektronika ir Elektrotechnika*, *29*(5), 51–61. DOI: 10.5755/j02.eie.34357

Riadi, I., Fadlil, A., & Putra, I. J. D. P. (2023). Batik pattern classification using naïve bayes method based on texture feature extraction. *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika, 9*(1).

Rizal, A., Hadiyoso, S., Aulia, S., & Wijayanto, I. (2023). Triwiyanto, & Said, Z. (2023). Multiscale hjorth descriptor on epileptic EEG classification. *Journal of Electrical and Computer Engineering*, (1), 4961637.

Saikhu, A., Setyadi, A. T., & Hariadi, V. (2023). Linear and Non-Linear Spatio-Temporal Input Selection In Wireless Traffic Networks Prediction using Recurrent Neural Networks. [Rekayasa Sistem dan Teknologi Informasi]. *Jurnal RESTI*, *7*(6), 1332–1340.

Sari, J. Y., & Bantun, S. (2023). Contactless Biometric Based on Palm Vein Recognition Using Wavelet and Local Line Binary Patterns. [Rekayasa Sistem dan Teknologi Informasi]. *Jurnal RESTI*, *7*(6), 1300–1308. DOI: 10.29207/resti.v7i6.4530

Siraji, M. I., Rahman, A. A., Nishat, M. M., Al Mamun, M. A., Faisal, F., Khalid, L. I., & Ahmed, A. (2023). Impact of mobile connectivity on students' wellbeing: Detecting learners' depression using machine learning algorithms. *PLoS One*, *18*(11), e0294803. DOI: 10.1371/journal.pone.0294803 PMID: 38011194

Varish, N., Singh, P., Tugiti, P., Manikanta, M. H., Yedlapalli, B., Pappusetty, A., Thakkar, H. K., & Sharma, G. (2023). Color Image Retrieval Method Using Low Dimensional Salient Visual Feature Descriptors for IoT Applications. *Computational Intelligence and Neuroscience*, *2023*(1), 6257573. DOI: 10.1155/2023/6257573 PMID: 36873380

Waluyo, W. N., Isnanto, R. R., & Rochim, A. F.R. Rizal IsnantoAdian Fatchur Rochim. (2023). Comparison of mycobacterium tuberculosis image detection accuracy using CNN and combination CNN-KNN. [Rekayasa Sistem dan Teknologi Informasi]. *Jurnal RESTI*, *7*(1), 80–87. DOI: 10.29207/resti.v7i1.4626

Wang, D. H., Liu, C., Wang, J., Yao, K., & Zhang, Z. (2023). High-precision parallel principal skewness analysis algorithm and its application in remote sensing images. *Journal of Electronics and Information.*, (10), 3492–3501.

Xu, S. J., Du, M., Duan, Z. X., Li, M. H., & Han, J. Q. (2023). Multi-Scale Region Attention InfoGAN License Plate Recognition Network. *Journal of Xi'an Jiaotong University*, (8), 206–218.

Yadav, S., & Yadav, J. K. P. S. (2023). Automatic Cataract Severity Detection and Grading Using Deep Learning. *Journal of Sensors*, *2023*(1), 2973836. DOI: 10.1155/2023/2973836

Zeydabadinezhad, M., Horn, C. C., & Mahmoudi, B. (2023). Quantifying the effects of vagus nerve stimulation on gastric myoelectric activity in ferrets using an interpretable machine learning approach. *PLoS One*, *18*(12), e0295297. DOI: 10.1371/journal.pone.0295297 PMID: 38039299

Zhang, P. C., Tusongjiang, K. Y., Liu, P., Di, Q., & Li, Z. (2023). Automatic Meter Reading Method for Digital Instruments Based on YOLOv5 and Improved VGG-CTC. *Modern Electronics Technique*, (2), 107–112.

Xiaoyuan Wang, graduated from Hefei University of Technology in 2008 with a master's degree, is a senior experimentalist at Hefei College. Her research areas include electrical and electronic engineering, electrical automation, electromechanical integration, information and communication technology, etc. Two Chinese invention patents, and won the third prize in the first self-made experimental teaching instrument in Anhui Province in 2016.

Hongfei Wang, male, (Han ethnicity), Anhui native, Bachelor's degree, research area: software development. Obtained one Chinese invention patent and one utility model patent. Participated in writing 3 EI conference papers.

Jianping Wang, male (Han ethnicity), Professor, Anhui native, Ph.D., Ph.D. supervisor, Vice Dean of the School of Electrical and Automation Engineering at Hefei University of Technology, Director of the Institute of Intelligent Monitoring and Diagnosis, member of the school's Academic Committee, and director of the school's Academic Committee. Research areas: electrical and automation education and research.