Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 1111683, 11 pages
https://doi.org/10.1155/2022/1111683

Research Article

Evaluation Method of Product Shape Features based on Multidimension Spatial Data Mining

Feilong Liu [1] and Miao Li [2]

[1] *School of Hubei Engineering University, Xiaogan, Hubei 432000, China*
[2] *School of Yellow River Conservancy Technical Institute, Kaifeng 475000, Henan, China*

Correspondence should be addressed to Feilong Liu; peterlfl@hbeu.edu.cn

Received 5 July 2022; Revised 26 August 2022; Accepted 30 August 2022; Published 19 September 2022

Academic Editor: Gengxin Sun

Analysis of product shape design has attracted more and more attention of researchers because it can make products better meet perceptual needs. Based on the theory of spatial data mining, this study proposes an evaluation method of a product shape design scheme. The method proposed in this study not only obtains the shape design of the overall and local features of the product, which solves the problem of the insufficient utilization of spatial data by the analysis method. During the simulation process, the model obtains the product shape design and appeal from the online evaluation spatial data, which can integrate the product perceptual knowledge in the spatial data, which greatly reduces the manual operation steps and the required time for the degree of data utilization. The experimental results show that after the obtained data are filtered to extract the feature words, the weight of the feature words is calculated by the TF-IDF method, the number of neighbors is increased from 1 to 30, the interval is 5, and the vectorized representation of the spatial data is constructed. The similarity between the calculation sentences of the data mining method is 89.7%, which effectively improves the support function and design efficiency of spatial data mining for product design.

1. Introduction

In recent years, innovative design has been raised to an unprecedented height, and enterprises have invested funds in product innovation research, and the innovative design of product visual interface is an indispensable part of it [1]. However, the style inheritance and creativity of brand products are often a pair of contradictions. How to carry out the creative design in an orderly manner and critically inherit the original design style in the creative design of products has become an industrial design at home and abroad today, which is to be solved in academia and even in corporate practice [2–4]. This study discusses the definition and mathematical expression method of point, line, surface, and body elements in the visual interface of product modeling. For a specific product, the point, line, and surface elements in its shape design are often numerous and complicated [5]. How to grasp the visual interface in the main feature elements in the design, and accurately extract them in the visual interface graphics [6], has become one of the main technical problems of the inheritance theoretical analysis and calculation of the shape design [7].

In terms of product design research, Pan and Zhang [8] took sustainable product design as the research object, built the relationship between product clusters and environmental impact through mining and analysis by building a product tree and environmental impact database [9], and established a product life cycle. A periodic automatic evaluation system is set to assist in sustainable product design [10]. Ait Issad et al. [11] regard complex products as big data systems, use a decision tree algorithm to reduce the dimension of high-dimensional design problems, and significantly reduce the design space, and the anti-collision design problem of passenger products was verified [12]. Mezaal and Pradhan [13] used the method of data mining to study the Universal Design Criterion and summarized the relationship between the design process into raw materials, elements, and signals. The required functional basis is summarized, and the function realization flow chart is constructed based on this, and the Apriori algorithm is used to generate association rules to optimize the selection of functions, shapes, and parameters [14]. Dou et al. [15] took digital cameras as the research object, and digital design features and function configuration are classified, used the decision tree algorithm and Apriori algorithm to obtain the relationship between features and configuration parameters and consumer image and demand, and used this to provide support and assistance for product design and sales. Alsrehin et al. [16] derived a dimensionless shape design loss function for the shape design features of small and large shapes and established an improved multivariate shape design loss function by using the principal component analysis method, which extended Artiles, and fully consider the correlation between the various shape design features, making it more realistic [17–19]. Although this method considers the difference in the influence of each principal component variable on the results [20] and retains the original shape design fluctuation information, it does not consider the influence of the shape design feature variance [21].

Taking the feature elements of product modeling visual interface as the starting point, this study studies the analysis and evaluation method of product visual interface and its software development method. The study analyzes the definition method of each feature of point, line, surface, and body of the product visual interface, and establishes the mathematical expression method of each feature element. The semantic difference method is used for the evaluation of perceived usability, which converts people's subjective evaluation of product styling into quantitative evaluation values, and the category set is obtained through cluster analysis. Then, we select the dimension of the shape design parameters through the maximum information coefficient, build a Kansei engineering expert system based on the spatial data mining processing and classification algorithm, extract the multi-level shape design parameters for the spatial data of the product's perceptual demand, and input it into the Kansei engineering expert system to output and design characteristic parameters, so as to describe and generate a new design scheme, and realize the product image design method driven by the big data of shape design.

2. Structural Design Method of Spatial Data Mining

2.1. Visualization of Spatial Data. In the three-dimensional spatial data visualization form, the scale of one dimension is much smaller than that of the other two dimensions, and the form will show the tendency of appearance [22–24]. A surface can be thought of as a line moving trajectories or it can be composed of an infinite number of lines. The shape of the surface is extensible and has the overall visual characteristics of the sense of expansion.

$$
\begin{aligned}
&\{\text{persit}(i+1), \text{persit}(i+2), \text{persit}(i+3), \ldots, \text{persit}(i+n-1)\},\\
&\{(\text{setx}(1), \text{sety}(1)), (\text{setx}(2), \text{sety}(2)), (\text{setx}(3), \text{sety}(3)), \ldots, (\text{setx}(i-1), \text{sety}(i-1))\}.
\end{aligned}
\tag{1}
$$

Because the surface elements in the product modeling visual interface all have outlines that are visually perceived. Therefore, contour lines can be used to represent surface elements. The coordinate set of each end point on the contour line of the polygon face element will be able to uniquely determine the face element, that is, the face element can be represented by the coordinate set of these points.

$$
\begin{aligned}
&\begin{cases}\text{foret}((x, y)|x + y) < 1,\\ \text{foret}((x, y)|\max(x) - \max(y) < x - y) < 1,\end{cases}\\
&\sqrt{\max(x)) - \max(y)} + \sqrt{\min(y) - \max(x)}\\
&+ \sqrt{\max(x) + \max(y)} = 1 - x - y.
\end{aligned}
\tag{2}
$$

The product data information of spatial data visualization mainly includes the product's rating of the watched video or purchased product, the time of browsing the page, the number of times the page was clicked, and the length of the product watching the video, which are all important recommended raw data. The recommendation system of collaborative filtering uses statistical methods to search for similar products of the target product and also uses the scores of similar products to a certain item to predict the score of the target product for the item, and finally selects the top N neighbor products with high similarity. The indicators are clustered after statistical analysis of the corresponding data. The obtained category set provides data support for the construction of the product modeling usability evaluation model.

2.2. Components of the Visual Interface. Starting from the generalized tolerance of the visual interface, the interaction relationship of the product shape design features at all levels is analyzed, and the product tolerance system model is established; using the historical data in the product manufacturing inspection process, the product shape design feature deviation based on the support vector nonlinear regression machine is established. Based on the model of the deviation relationship between the product shape design feature and the component shape design feature deviation, a Support Vector Nonlinear Regression (SVNR) model was constructed.

$$
\begin{aligned}
&\text{fin}(t + y(t))\mathrm{d}t\mathrm{d}y - \text{fin}(t - x(t))\mathrm{d}t\mathrm{d}x = \begin{bmatrix} t \\ y \end{bmatrix} - \begin{bmatrix} x \\ t \end{bmatrix},\\
&\text{seppher}(x, y) = \begin{cases}\text{sqrt}(\text{fin}(x) - \text{fin}(y)),\\ \sum \text{sqrt}(\text{fin}(x) - \text{fin}(y)),\end{cases}\\
&\begin{bmatrix} 1 - \cos\ j(x)\sinh(x) \\ 1 - \sin\ j(x)\cosh(x) \end{bmatrix} = \begin{bmatrix} \max(x, j) \\ \min(x, j) \end{bmatrix}.
\end{aligned}
\tag{3}
$$

Table 1: Description of feature extraction of spatial interface.

Description node	Feature point	Extraction type	Interface ratio
Performance	83.22	Data usability	0.55
Reviewed product	28.02	Data sample	0.94
Each sample	40.18	The purchase group	0.81
Number of errors	86.30	The experiment	0.04
Evaluation index	59.34	Extract product	0.60
Experiment invite	27.40	Corresponding data	0.23
Product evaluation	54.54	Space data	0.16

If the visual interface components are used to distinguish, qualitative inference-style perceptual engineering and forward-quantitative inference-style perceptual engineering are both based on the perceptual needs of designers or consumers to infer a design that conforms to the perceptual design. Inverse quantitative inferential sensibility engineering is to infer the sensibility of a design case.

2.3. Spatial Interface Feature Extraction. If conditions permit, a more scientific space interface feature tracking experiment can also be used to extract product feature elements. The test subject browses the whole pictures of several products by wearing the instrument, and the instrument will automatically record the subject's gaze track distribution and gaze dwell time, and obtain the attractive index of each characteristic part of the product, which is affected by the level of the index. The main features of the product styling style are extracted and finally the characteristics of the product styling that are most concerned by the subjects are statistically analyzed. These highly concerned modeling features are the extracted spatial interface features.

$$\operatorname{section}(\operatorname{cert}(i,j),i,j)=\begin{cases}\dfrac{\sqrt{\max(y,j)-\min(y,j)}}{y-j},\\[2ex] \dfrac{\sqrt{\max(x,j)-\min(x,j)}}{x-j},\end{cases}\qquad (4)$$

$$\begin{cases}\displaystyle\int\sqrt{\operatorname{fertin}(i,x)-1}\,\mathrm{d}x = x-i\mathrm{d}x,\\[2ex] \displaystyle\int\sqrt{\operatorname{fertin}(i,x)+\operatorname{fertin}(i,y)-1}\,\mathrm{d}x\mathrm{d}y = x-1.\end{cases}$$

In view of changing product performance requirements or adding new product requirements, based on the product tolerance system optimization design method and data mining, a product tolerance system redesign optimization model oriented to product performance requirements is established. The TF-IDF algorithm will reduce the weight value of high-frequency spatial data appearing in multiple documents as keywords, preventing the extraction of commonly used words with low importance, thus ensuring the extraction effect of the algorithm.

$$\begin{cases}\left\{\dfrac{1}{\operatorname{vert}(i)}-\dfrac{1}{\operatorname{vert}(j)}\right\}=1-ij,\\[2ex] \left\{\dfrac{\operatorname{vert}(i,j)}{\operatorname{vert}(i)\operatorname{vert}(j)}\right\}=1-\dfrac{i}{1-j},\end{cases}$$
$$\begin{bmatrix}\left(\operatorname{bert}(i-j)-\operatorname{bert}\left(\dfrac{1}{i+1/j}\right)\right)\\[2ex] \operatorname{bert}(i+j)+\operatorname{bert}\left(\dfrac{1/i}{1/j}\right)\end{bmatrix}=\operatorname{siergod}(i,j).\qquad (5)$$

However, in the product evaluation space data, because there is a certain degree of similarity between the reviewed products and the purchase groups, there are many similar cases to evaluate spatial data, and the TF-IDF algorithm tends to give lower weights to these spatial data, so that when using the TF-IDF algorithm, the weight value of the evaluation spatial data in Table 1 will be underestimated, which may easily lead to the inability to extract product keywords, to reflect real product experience.

The performance usability, that is, the completion time of the spatial interface feature task and the number of errors, is used as the evaluation index for testing, like task completion time, that is, the time taken from the start of the test product operation to the completion of the task. The evaluation indicators of usability are divided into subjective dimensions and there are two categories of objective dimensions, which correspond to perceived usability and performance usability. The operation is recorded as one error, and the experiment is repeated 4 times for each sample and each person to obtain the corresponding data.

2.4. Interactive Mode of Product Shape Design. In order to obtain the set of interactive perception usability categories of product shape design, the average value of the shape design evaluation obtained above is processed, and the 5-level category is obtained through isometric mining analysis. The higher the level, the more obvious the shape design. According to the above, the collected 50 product shape design samples were randomly sorted, the first 45 samples were used as training samples to construct the evaluation model, and the last 5 samples were used as verification samples. Taking aesthetics as an example, the evaluation results of aesthetics of 45 samples were obtained through SD questionnaire analysis.

$$
\begin{aligned}
&\text{cerberi}\{\text{cer}(i, j), \text{beri}(i, j) | \text{convert}(i) - \text{convert}(j), i, j < 1\},\\
&\begin{bmatrix} \text{ebersiler}(i) + \text{ebersiler}(j) \\ \text{ebersiler}(i) - \text{ebersiler}(j) \end{bmatrix}\\
&\prod_{i,j} |\text{sert}(\text{diserct}(i) - \text{diserct}(j) - 1)| - \prod |\text{sert}(i - j - 1)| = si - sj.
\end{aligned} \tag{6}
$$

In the visual interface, the point is a relatively small visual element, which plays the role of marking the position in the space, can attract people's attention, and has the characteristics of concentration and cohesion. Therefore, a point element can be understood as a point in geometry, that is, it does not have a size, but only represents a position. To compare the difference of point elements on the visual interface of similar products of a certain brand, it is to calculate the difference in the position center of the point. Although the part has a different shape on the right half, the shape of the left half has not changed. Therefore, we use the left half as the base to establish the coordinate system.

$$
\begin{aligned}
&\sum_{i,j} \text{omigaer}(i + j) - \sum_{i,j} \sqrt{\text{omigaer}(i + j)} - \sum_{i,j} \sqrt[3]{\text{omigaer}(i + j)} = 1 - \sqrt{(i + j)}\\
&\forall \text{terst}(\text{omigaer}(i), i) - \text{omigaer}(i) < i, \exists (1 - \text{dert}(\text{omigaer}(i) - i)) - i < \sqrt{\text{omigaer}(i) - 1}.
\end{aligned} \tag{7}
$$

First, the quantification of product shape is realized through the semantic description of product shape factors and the evaluation of product shape; the data of the survey statistics are tested for the level of agreement of the target attribute level to complete the data processing. The classification list of evaluation items of the method determines the constraints of the product from 60 independent aspects and summarizes 60 evaluation items, which basically cover all the basic evaluation items. The category set is obtained through actual testing and cluster analysis. The obtained category set provides data support for the construction of the product modeling usability evaluation model.

3. Construction of Evaluation Model for Product Shape Design Scheme

3.1. Spatial Data Mining Hierarchy. Spatial data mining mainly analyzes the correlation between items and then aggregates the item lists with the highest correlation together to form an item-based item neighbor group. Then, we use the list of products in the shopping basket to select the most similar products, finally form a recommended list, and send the recommended results to the products. Finding the nearest neighbor is the core part of the recommender system, and it plays a decisive role in the shape design of the recommender system, which greatly improves the effectiveness of the recommendation results and increases the loyalty of the product. The more consistent the rules of the algorithm are with the actual rules, the more accurate the prediction of the data in Figure 1 will be, and the better the recommendation effect will be.

(1) Divide the data into pieces and start the program on the cloud. (2) According to the deployment method of MapReduce and the master/slaves method, the master is responsible for task scheduling and detecting the working status of the workers. (3) In the map stage, the worker reads the data in the split, inputs it to the Map function in the form of a key-value pair, and caches the output key-value pair in memory. (4) The cached key-value pair will be periodically written to the disk, the key-value pairs of the same keyID will be put together, and the storage address of the key-value pair will be passed to the master record, while performing Reduce function case. (5) The reduce phase is to take the key-value pairs output by the Map function in the previous process as input to the Reduce function, and output a new result, that is, the result of mining.

For the big data of product review text, text preprocessing is realized by text deduplication, empty text deletion, etc., word segmentation, part-of-speech tagging, keyword extraction, and other means to extract the perceptual image vocabulary for the product. The obtained category set provides the data support in Table 2 for the construction of the product modeling usability evaluation model.

Taking a spatial data mining time as the starting node, traverse the product design points of storage and processing within the range that the system can afford. Under extremely ideal conditions, such as when the software and hardware conditions allow and the time limit is long enough, this

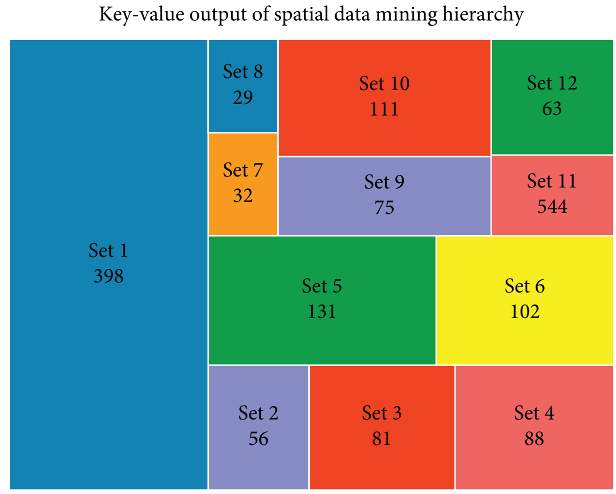

Figure 1: Hierarchical key-value output of spatial data mining.

Table 2: Collection evaluation of product styling categories.

Collection evaluation code	Product styling description
Import numpy as np	Taking a spatial $i + j$
From matplotlib import pyplot as plt	Call back to the main program
Fig = plt.figure()	Data mining time $\vert$sert$(i - j - 1)$
Ax = plt.axes(projection = “3d”)	The system can afford
$X = y$ = np.arrange(start = −4, stop = 4, step = 0.1)	Of the diserct(i) reduce phase
X, y = np.meshgrid(x, y)	As the starting node convert(i)
Separate = (distance, distance)	Within the range that d$w(y)$
Plt.figure(explode = separate)	Completing a mapreduce process
Plt.pie(sizes, autopct = “%1.1f%%”)	Of storage and processing
Plt.title(“separation distance = 0.8”)	The master will $1 - x$
Plt.show()	The product design points

strategy can capture a large-scale product design point set, which is almost consistent with the real network data.

3.2. Product Shape Image Design. In this stage, a semantic difference experiment was conducted by combining 4 target product shape image design words and 50 representative product shape design samples, and a fifth-order SD questionnaire was made for statistical analysis. 20 students majoring in industrial design were invited to participate in the survey. Four target profile designs of 50 samples were evaluated. During the investigation, each subject scored and evaluated each sample based on the shape design space data. If the line element is a curve, the comparison is relatively difficult. Using an approximation method of approximating a curve with an infinite number of small straight line segments to represent the curve, the calculation problem of the curve can be transformed into the calculation problem of the polyline. Perceptual appeal text big data, overall image extraction, partial image extraction, image parameterization, and other methods are to realize the parametric expression of perceptual image appeal. Obviously, the denser the selected points, the closer the polyline formed by the point set is to the shape of the curve.

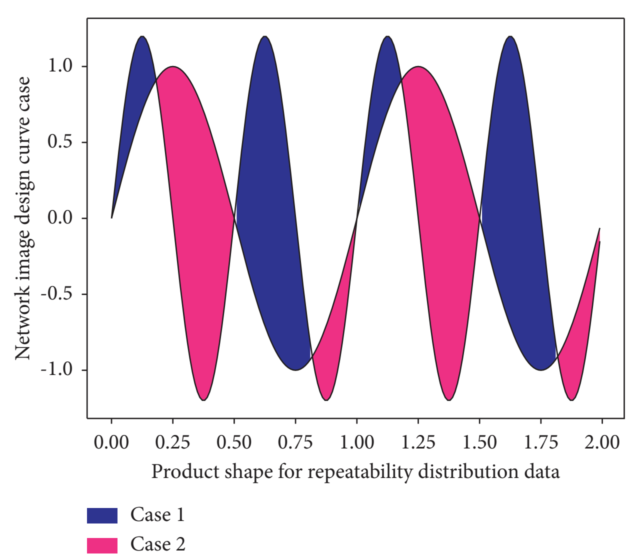

Figure 2: Product shape image design curve distribution.

URL sets can be created using the method shown in Figure 2, which is to link from one product design point to other product design points through hyperlinks. However,

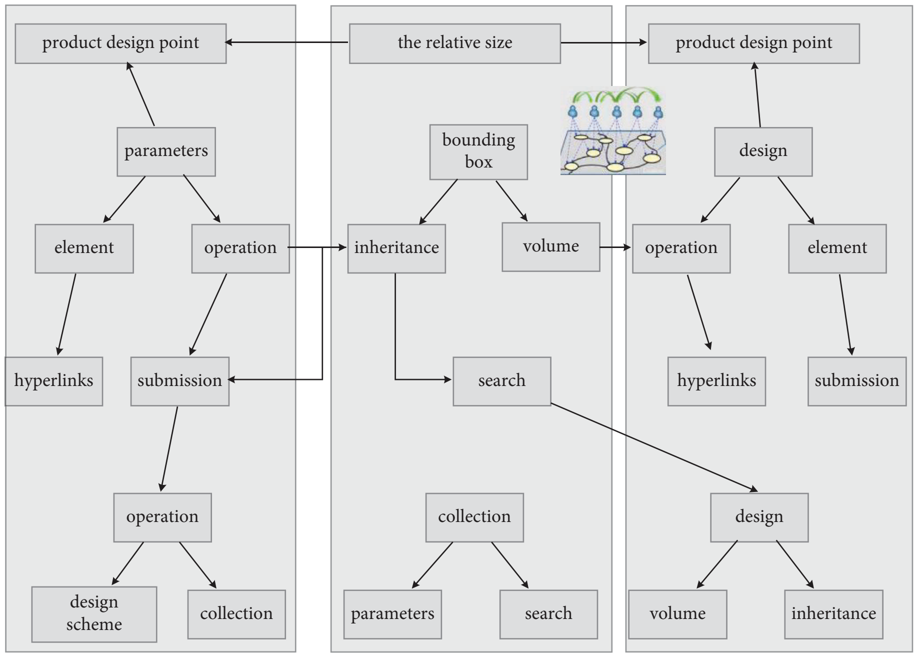

Figure 3: Product design point link topology.

in actual operation, it is not realistic to crawl all the product design points, so in the case that the automatic crawling cannot be crawled, it can be completed in the form of submission by the webmaster, that is, the webmaster submits and requests to the search engine which included and added to the URL collection.

The inheritance of the volume element can be calculated by comparing the minimum bounding box of each partition in Figure 3. The main idea of the calculation is to use the bounding box parameters of the shape to analyze and calculate the minimum bounding box volume of the shape and the inheritance of the shape. The inheritance of the shape is not only the inheritance of the saturation of the shape and the relative size but also the inheritance of the topological relationship between the parts.

3.3. Design Scheme Evaluation Threshold Calculation. The essence of the design scheme evaluation threshold algorithm is to build the relationship between user preference and product semantics, so as to further filter the product semantics, and finally get the semantic word set describing the object. This process is to match users' preferences based on product semantics, which is an evaluation problem. Since user preference is a continuous variable, it can be solved by regression tree technology. To analyze the user's emotional cognitive process, focusing on the physiological and emotional aspects of perceptual image cognition; the method mainly analyzes the user's description text of the product or invites the user to take a specific psychological test to obtain the user's tendency in the dimension of the perceptual image represented by vocabulary.

Because the directly crawled product design point data is very primitive, it contains all the information of the product design point, including some useless information, such as page name or link title and project name, etc., as well as some special characters that cannot be displayed, etc. Whether the subsequent operations will take up the analysis time or not has practical significance, it is filtered out first in this step. In order to solve this problem, the filtering algorithm in Figure 4 is designed based on the MapReduce framework, which can significantly speed up the process.

Subsequently, the relationship between product design elements that describe product characteristics is analyzed. The minimum word frequency is set to 20 to filter out the influence of some low-frequency spatial data and typos, and a total of 1065 product design elements and their correlation matrix are obtained. Then, the relationship between spatial data is analyzed using the maximum spanning tree algorithm, and 146 spanning trees are obtained. The product design elements in each tree are regarded as the same group, and the spatial data in some groups are shown in the figure.

The determination of the inheritance threshold should firstly collect adjectives of product styling features through various media and establish a corresponding regular morphological semantic set according to the results of the questionnaire; then we select the object semantic words describing each styling feature from the regular morphological semantic set according to the preference algorithm, select the optimal semantic word for each feature object by means of a fuzzy algorithm. Based on the optimal semantic words of each corresponding feature element in the product modeling visual interface of the two generations, a ten-order scale will be scored, and the inheritance value of each feature element will be obtained through comparison and calculation.

The RandomForestClassifier function of the sklearn package in python is also used to construct the mapping

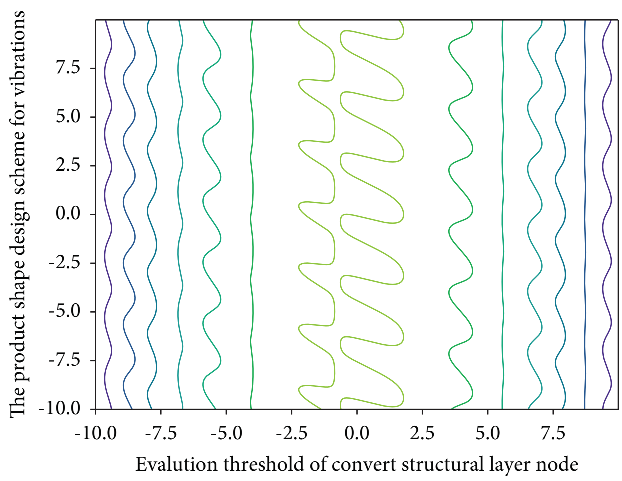

FIGURE 4: Analysis of design scheme evaluation threshold.

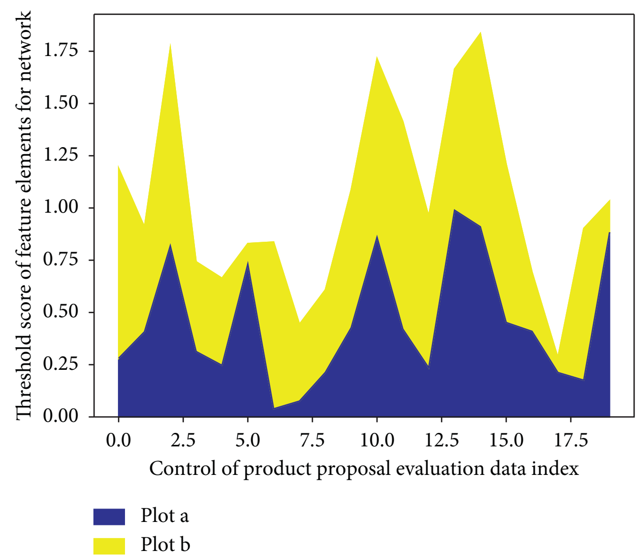

FIGURE 5: Threshold score of feature elements for scheme evaluation.

relationship, and the out-of-bag data is used for reliability verification. After verification, the classification accuracy of the local features of each product is shown in Figure 5. That is, selecting the corresponding method to extract specific knowledge according to the text mining target; pattern evaluation is to evaluate the effect of pattern mining, if it meets the requirements, save the result of the pattern mining, if it does not meet the requirements, return to the above improvement in a certain area.

The ratio is greater than 0.6, the product header design features many categories, and the corresponding mapping relationship has a low correct rate, which is in line with the expected results, and the mapping relationship is generally effective. Both the Dl and DZ functions are used for 2D graph comparison, but are calculated in different ways. The DZ function is to calculate the distance between any two points on the outline of the figure, and the Dl function is to calculate the distance from a fixed point on the outline of the figure to all other points. For more complex graphics, it takes too long to calculate with the DZ function, and the research in this study only needs to quantify the shape initially, so the Dl function is selected for calculation.

4. Application and Analysis of the Evaluation Model of Product Shape Design Scheme

4.1. Preliminary Convolution Processing of Spatial Data. The data convolution processing system will directly integrate the JDK. You can use the java-version command to check whether the system has been installed. If not, download the JDK compressed package from the Oracle official website to decompress it, and then configure the following environment, add the following statement in/etc/profile, the decompression path in this article is/opt/java, and the configuration environment needs to be modified according to the actual situation. The file path is/etc/hosts, and input the IP address and hostname of each host respectively.

It may group the comment space data of different products, and then perform word segmentation and part-of-speech tagging for the comment space data of each product sample in Figure 6, only keep the adjectives in the comments, and analyze the statistical characteristics of each adjective through the TextRank algorithm to obtain the space in data importance parameters, and then arrange the importance parameters from high to low, extract and quantify N adjectives with strong importance according to the number threshold, and the extracted shape design is expressed as the following. However, in the product evaluation text, because there is a certain degree of similarity between the products being reviewed, and the buying groups also have a certain similarity, there are many identical evaluation words in the review texts of different products, the algorithm tends to give a lower weight to these words. Nouns are used as candidate words for the same feature, and finally, irrelevant spatial data are manually deleted, so as to obtain different expressions of the same product feature.

Then, the relationship between the obtained spatial data is simplified by the maximum spanning tree algorithm to obtain the product design element grouping in Figure 7. Due to the low frequency of occurrence of some spatial data, word segmentation errors or typos in the spatial data will interfere with the construction of the product feature tree. Therefore, when calculating the relationship between words, a word frequency threshold is set, and only a space higher than the threshold is used. The data were included in the maximum spanning tree analysis. In addition, the spatial data mining process selects samples from the original data set through the Bootstrap method to train each decision tree. The training samples of each decision tree are independent of each other and have strong generalization ability. When Bootstrap selects the training data of each decision tree, about 36.8% of the data will be reserved as out-of-bag data to verify the validity of the spatial data mining processing classification model.

Spatial data mining constructs word vectors that incorporate semantic relationships as the basis for analyzing

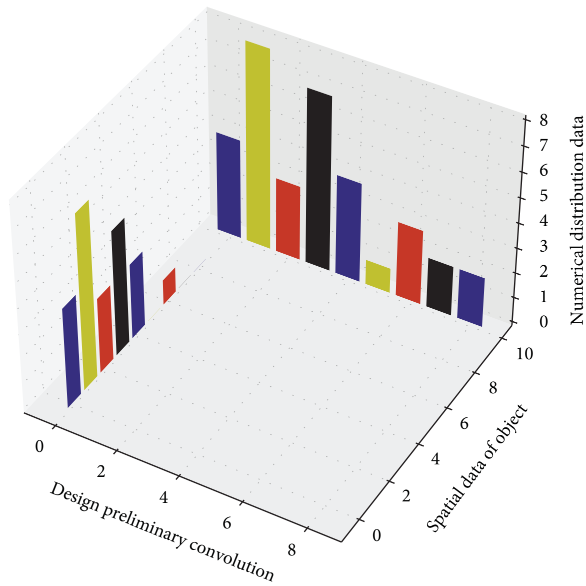

Figure 6: Preliminary convolution numerical distribution of spatial data.

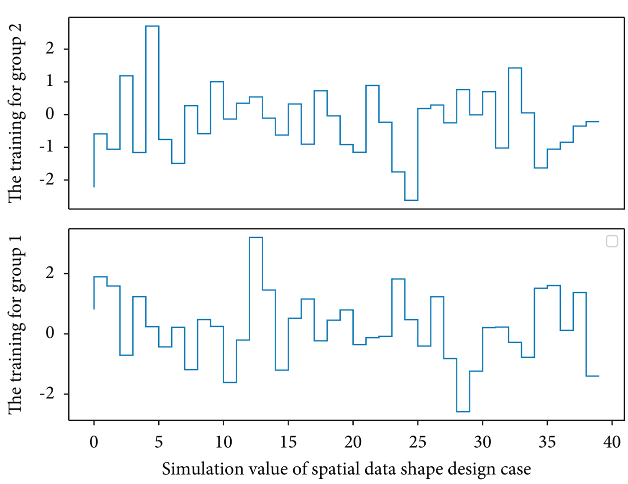

Figure 7: Extraction of spatial data shape design representation.

the relationship between spatial data and spatial data parameterization; extracting adjectives and their importance parameters in spatial data through keyword extraction algorithms as the overall image of the product. When extracting perceptual images for local features of products, it is first necessary to identify the description objects of the sentences, that is, to find the nouns in the text that describe the local features of products. The shape design is parameterized by word vectors, and the dimension of the shape design parameters is reduced by combining principal component analysis, so as to obtain the parametric expression of the global and local images.

4.2. Evaluation and Simulation of Product Shape Design Scheme. The development content of the product shape design software mentioned in this study mainly includes the related investigation of sensibility engineering and the inheritance calculation of the product shape visual interface. The software programming in this study is realized by the mixed programming of VB.NET and MATLAB. The specific realization process of product outline design contour extraction is: (1) grayscale the image; (2) use the maximum class variance method to select the appropriate closed value, segment the grayscale image, and generate the corresponding binary image; (3) use the eight-neighborhood method of hollowing out the interior points to find all closed areas in the image, and mark their boundary contours; (4) save the pixel coordinates of the marked contour points in the TXT file to prepare for subsequent processing.

Although the dimension of the Distributed Representation word vector obtained by the word2vec tool is lower than that of the One-hot Representation word vector, the single dimension is still greater than 100, which is still too high compared to the sample size of the product design research in Figure 8. Therefore, the dimension of the shape design space data is reduced by factor analysis, which reduces the dimension of the word vector involved in the calculation of the shape design parameters while retaining the semantic difference information between the spatial data. When analyzing the semantic relationship, the semantic relationship between words and nouns is divided into two aspects: similarity and dissimilarity. Since word vectors can reflect the similarity of semantic relations of words, the similarity between nouns is determined by calculating the cosine similarity between word vectors. Commonly used feature selection methods include variance screening, correlation coefficient, hypothesis testing, and mutual information. Since the purpose of dimension selection in this study is to obtain the dimension of shape design parameters that is more correlated with product feature parameters, and the type of correlation between perceptual image parameters and design feature parameters cannot be predicted, the method of calculating mutual information is used to filter perceptual imagery parameter dimension.

In order to improve the accuracy of Figure 9, some attribute words such as product attributes, words expressing related fields, and some idioms can be selectively reserved, so as to maximize the retention of meaningful words for the sentence, so that the remaining word for the meaning of the composed sentence is closest to the original text, which makes the final mining effect better. First, we take each point as the bottom mining, then calculate the distance between each mining, and then form the mining in the upper level by merging the most similar mining, when a certain termination condition is reached or all data points are merged into one mining sometimes ends.

4.3. Example Application and Analysis. This survey invited 10 students majoring in industrial design as subjects to select 10 adjectives from all the collected product image semantics. The selected adjectives must be suitable for describing the sample of this survey. In the process, the selected adjectives should have as little overlap as possible, and the selection

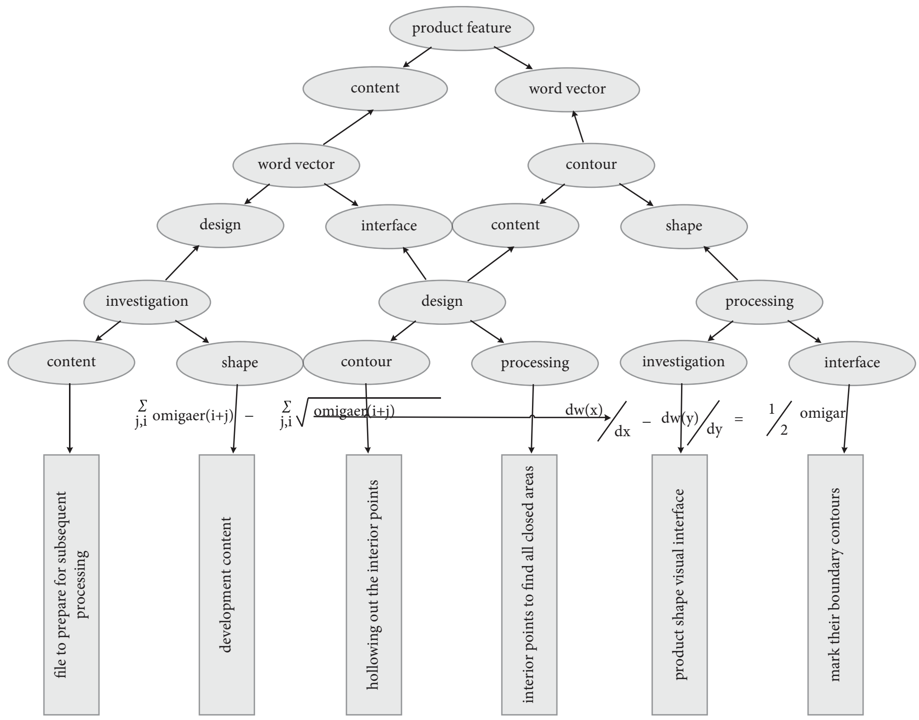

FIGURE 8: Structure optimization of product shape design scheme.

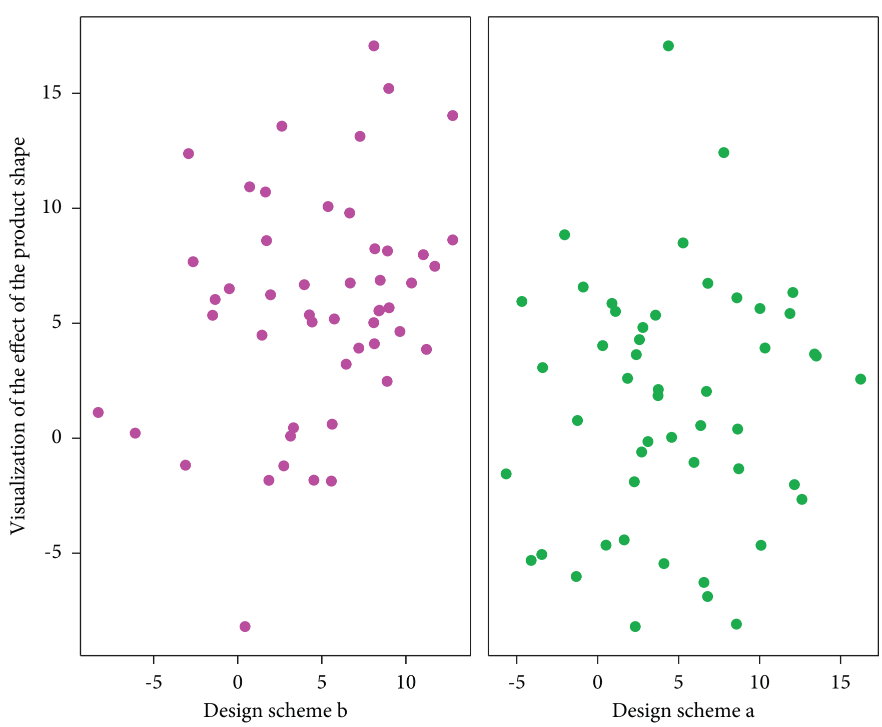

FIGURE 9: Visualization of the effect of product shape design scheme.

results of the subjects were analyzed using the frequency statistics method, and subjective predictions were appropriately added to screen out 8 product image semantics. In this way, a conventional image semantic set is formed. Due to the low frequency of occurrence of some words, most of them are word segmentation errors or typos in the text, which will interfere with the construction of the product feature tree. The reduce function is to summarize the Canopy output of the map function, output the mining analysis results, and save the results on HDFS.

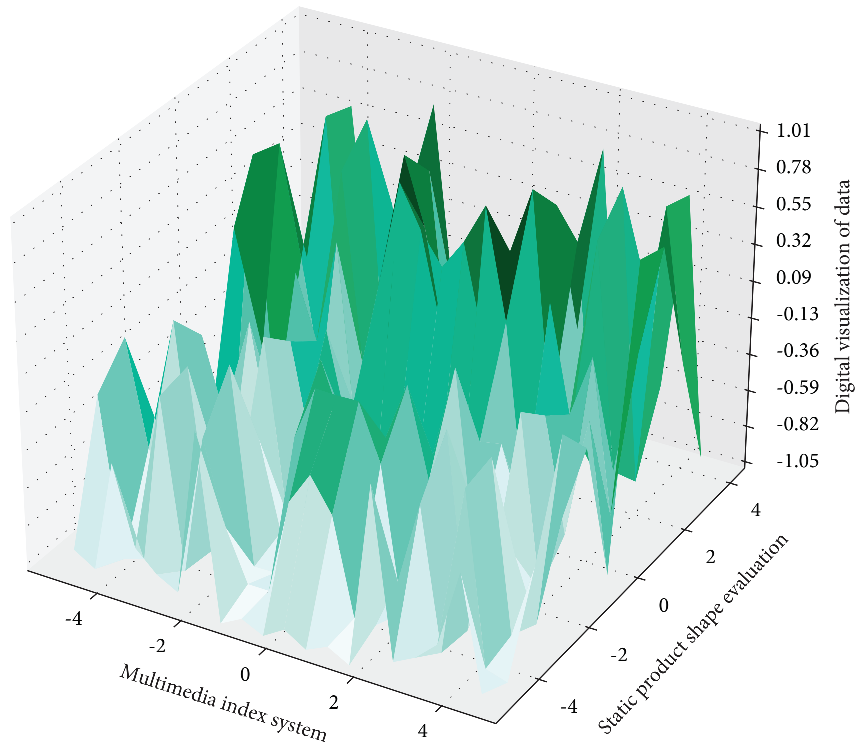

Figure 10: Three-dimensional distribution of product shape evaluation index system.

The form of the questionnaire is different depending on the method of product appearance evaluation. According to the dimensional differentiation method of product appearance evaluation, the form of the questionnaire is divided into a multi-sample questionnaire and a single-sample questionnaire. The production of the multi-sample questionnaire adopts the product appearance scoring method, which takes the attribute layer in the evaluation index system as the direct target, and directly classifies and evaluates the product appearance on the attribute layer; the production of the single-sample questionnaire adopts the product appearance ranking method. Therefore, when calculating the relationship between words, a word frequency threshold is set, and only words higher than the threshold are included in the maximum number of words. Through experiments, we found that the multi-sample survey method is based on the solid foundation of Vts, its data processing method is more delicate, the effectiveness and utilization of raw data are higher, and the horizontal relationship between each evaluation factor and each scheme can be completed. Compared with longitudinal comparison and analysis, however, because of the data collection and statistics for different schemes, the data collection work is complicated and the implementation is relatively difficult.

The level of Figure 10 provides the keyword information required for the upper-level description and defines a certain dependency between the keywords, which can usually define individualized static attribute rules at this level. The description layer is not only for specific products but also for specific resources. Arbitrarily it may pick out K objects as a cluster, and the object is the center of the cluster, calculate the distance between all the remaining objects and the K objects, the closest distance is divided into the corresponding cluster, and then calculate each average value of the cluster, the most central object is selected as the center of the new cluster, and the iteration starts at this time, and the remaining objects except the cluster center are calculated again, divided into new clusters, and then a new center is obtained, and repeat the execution, and finally get it if the center of the new cluster is consistent with the previous center or the distance is less than the set value, the iteration ends and the mining is completed.

5. Conclusion

Based on the concept of spatial data mining, this study studies the relationship between spatial data and product shape design. The product visual interface is decomposed into several characteristic elements. For the product shape design cycle, the model defines the loss of product shape design as the present value of the loss caused by scrapping the product after the investment because the product shape design features exceed the specified range and product shape design loss model of spatial data distribution; for shape design features that obey different distributions, a calculation method of product spatial data distribution density function based on tolerance requirements of shape design features is proposed. First, we obtain the big data of online product evaluation spatial data through web crawlers, use word2vec to vectorize the spatial data after word segmentation, divide the product shape design into two categories: overall product shape design and product partial shape design, finding adjectives related to the nouns in the group through syntactic relationship and counting the word frequency, and then obtaining the local image of each product. The keyword extraction algorithm and the keyword extraction method of the syntactic relationship extract the image and appeal, use the word vector model to represent it as a real vector and perform principal component analysis to reduce the dimension, and combine the importance parameters to calculate the corresponding shape design parameters, so hierarchical product shape design appeals to big data mining methods.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this study.

Acknowledgments

This work was supported by the School of Hubei Engineering University.

References

[1] S. Lv, H. Kim, B. Zheng, and H. Jin, "A review of data mining with big data towards its applications in the electronics industry," *Applied Sciences*, vol. 8, no. 4, p. 582, 2018.

[2] P. Du, X. Bai, K. Tan et al., "Advances of four machine learning methods for spatial data handling: a review," *Journal of Geovisualization and Spatial Analysis*, vol. 4, no. 1, pp. 13–25, 2020.

[3] A. Dogan and D. Birant, "Machine learning and data mining in manufacturing," *Expert Systems with Applications*, vol. 166, Article ID 114060, 2021.

[4] G. Atluri, A. Karpatne, and V. Kumar, "Spatio-temporal data mining: a survey of problems and methods," *ACM Computing Surveys*, vol. 51, no. 4, pp. 1–41, 2019.

[5] C. Cocco, C. Rezende Freitas, A. C. Mourao Moura, and M. Campagna, "Geodesign process analytics: focus on design as a process and its outcomes," *Sustainability*, vol. 12, no. 1, p. 119, 2019.

[6] W. Chen, H. Hong, M. Panahi et al., "Spatial prediction of landslide susceptibility using gis-based data mining techniques of anfis with whale optimization algorithm (woa) and grey wolf optimizer (gwo)," *Applied Sciences*, vol. 9, no. 18, p. 3755, 2019.

[7] A. Mobasheri, H. Huang, L. C. Degrossi, and A. Zipf, "Enrichment of openstreetmap data completeness with sidewalk geometries using data mining techniques," *Sensors*, vol. 18, no. 2, p. 509, 2018.

[8] Y. Pan and L. Zhang, "A BIM-data mining integrated digital twin framework for advanced project management," *Automation in Construction*, vol. 124, Article ID 103564, 2021.

[9] H. N. Dai, H. Wang, G. Xu, J. Wan, and M. Imran, "Big data analytics for manufacturing internet of things: opportunities, challenges and enabling technologies," *Enterprise Information Systems*, vol. 14, no. 9-10, pp. 1279–1303, 2020.

[10] J. Zhu, Z. Ge, Z. Song, and F. Gao, "Review and big data perspectives on robust data mining approaches for industrial process modeling with outliers and missing data," *Annual Reviews in Control*, vol. 46, pp. 107–133, 2018.

[11] H. Ait Issad, R. Aoudjit, and J. J. P. C. Rodrigues, "A comprehensive review of Data Mining techniques in smart agriculture," *Engineering in Agriculture, Environment and Food*, vol. 12, no. 4, pp. 511–525, 2019.

[12] K. Bai, K. Li, N. B. Chang, and W. Gao, "Advancing the prediction accuracy of satellite-based PM2. 5 concentration mapping: a perspective of data mining through in situ PM2. 5 measurements," *Environmental Pollution*, vol. 254, Article ID 113047, 2019.

[13] M. R. Mezaal and B. Pradhan, "Data mining-aided automatic landslide detection using airborne laser scanning data in densely forested tropical areas," *Korean Journal of Remote Sensing*, vol. 34, no. 1, pp. 45–74, 2018.

[14] W. Yan, S. Lin, O. L. Kafka et al., "Data-driven multi-scale multi-physics models to derive process–structure–property relationships for additive manufacturing," *Computational Mechanics*, vol. 61, no. 5, pp. 521–541, 2018.

[15] J. Dou, A. P. Yunus, D. Tien Bui et al., "Evaluating GIS-based multiple statistical models and data mining for earthquake and rainfall-induced landslide susceptibility using the LiDAR DEM," *Remote Sensing*, vol. 11, no. 6, p. 638, 2019.

[16] N. O. Alsrehin, A. F. Klaib, and A. Magableh, "Intelligent transportation and control systems using data mining and machine learning techniques: a comprehensive study," *IEEE Access*, vol. 7, pp. 49830–49857, 2019.

[17] Y. Pan and L. Zhang, "Roles of artificial intelligence in construction engineering and management: a critical review and future trends," *Automation in Construction*, vol. 122, Article ID 103517, 2021.

[18] A. C. Benabdellah, A. Benghabrit, and I. Bouhaddou, "A survey of clustering algorithms for an industrial context," *Procedia Computer Science*, vol. 148, pp. 291–302, 2019.

[19] T. Wang, C. Ren, Y. Luo, and J. Tian, "NS-DBSCAN: NS-dbscan: a density-based clustering algorithm in network space," *ISPRS International Journal of Geo-Information*, vol. 8, no. 5, p. 218, 2019.

[20] Z. Salih Ageed, S. R M Zeebaree, M. Mohammed Sadeeq et al., "A survey of data mining implementation in smart city applications," *Qubahan Academic Journal*, vol. 1, no. 2, pp. 91–99, 2021.

[21] Z. Lan and Y. Liu, "Study on multi-scale window determination for GLCM texture description in high-resolution remote sensing image geo-analysis supported by GIS and domain knowledge," *ISPRS International Journal of Geo-Information*, vol. 7, no. 5, p. 175, 2018.

[22] T. Van Nguyen, J. Zhang, L. Zhou, M. Meng, and Y. He, "A data-driven optimization of large-scale dry port location using the hybrid approach of data mining and complex network theory," *Transportation Research Part E: Logistics and Transportation Review*, vol. 134, Article ID 101816, 2020.

[23] F. Tsung, K. Zhang, L. Cheng, and Z. Song, "Statistical transfer learning: a review and some extensions to statistical process control," *Quality Engineering*, vol. 30, no. 1, pp. 115–128, 2018.

[24] O. Ghorbanzadeh, H. Rostamzadeh, T. Blaschke, K. Gholaminia, and J. Aryal, "A new GIS-based data mining technique using an adaptive neuro-fuzzy inference system (ANFIS) and k-fold cross-validation approach for land subsidence susceptibility mapping," *Natural Hazards*, vol. 94, no. 2, pp. 497–517, 2018.