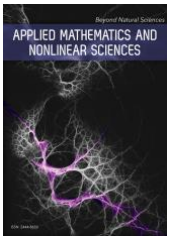
sciendo Applied Mathematics and Nonlinear Sciences, 9(1) (2024) 1-15



Applied Mathematics and Nonlinear Sciences

https://www.sciendo.com

A Study on the Application of Image Analysis Techniques to Autocoders in Civil and Hydraulic Engineering

Yusi Yu[1,†]

1. College of Civil Engineering, Xiangtan University, Xiangtan, Hunan, 411100, China.

Submission Info

Communicated by Z. Sabir
Received March 8, 2024
Accepted June 3, 2024
Available online July 9, 2024

Abstract

With the continuous updating of computer vision and image analysis technology, image processing, as well as analysis technology has become an important auxiliary means in the analysis of civil engineering and water conservancy projects. In this study, an image analysis model is constructed by defining the autoencoder and its expansion relation, combined with a convolutional neural network. On this basis, an engineering detection model is built by using a sparse-stacked autoencoder. The structural sub-transmission characteristics of civil engineering and water conservancy projects were investigated. The image separation optimization was carried out by using Beer Lambert's law, and finally the engineering structure extraction and recognition model based on deep learning was formed. Then, the performance of the model is examined. The average value of the repetition rate is higher than 80% in the brightness adjustment, rotation, and scaling operation change test. The Dice and IoU indexes are higher than 90%, and the HD distance is less than 27mm, so the overall performance is excellent. The practical application of civil engineering and water conservancy engineering has a fantastic performance, with a relative error of no more than 2%. The method in this paper has excellent stability and practical effect. It proposes an improvement method for optimizing the image analysis method in civil engineering and water conservancy engineering.

Keywords: Autoencoder; Convolutional neural network; Beer Lambert's law; Engineering detection; Civil engineering and water conservancy.
AMS 2010 codes: 97P20

†Corresponding author.
Email address: xtusun@qq.com

ISSN 2444-8656
https://doi.org/10.2478/amns-2024-1789

1 Introduction

With the development of computer technology in recent years, there has been a great improvement in terms of the level of hardware and software development, and the computing power and the speed of processing graphics have been greatly improved [1]. At this stage, the civil engineering inspection technology is not very perfect due to the relevant standards caused by the backwardness of the inspection method, so there is an urgent need to introduce advanced technology to standardize the appropriate standards so that its rapid and dynamic development [2-4].

Digital image analysis technology is a technology that gradually emerged in the late 1980s, the use of computers to analyze and process images to extract or describe the required digital information [5-6]. The prosperity of the technological digital market also provides support for the development of digital image analysis and processing technology [7]. Technicians in various industries increasingly favor digital images due to their inherent ease of storage, richness of information, and digitization characteristics [8-9]. In the field of civil engineering, which may face long periods of out-of-home operations, the application of digital image analysis-based techniques in engineering appears to be more convenient and efficient [10-11]. Autoencoders are used in several applications for degradation and anomaly detection and this architecture can be well used in civil engineering systems [12-13]. Auto-encoders are a subgroup of neural networks that contain a two-level structure: encoder and decoder [14]. The encoder reduces the dimensionality of the input data RM and projects it to one or more layers in the potential space RN, N<M [15]. Decoder, which projects the signal back to its original space [16].

Literature [17] used image analysis in air structure analysis of asphalt mixtures to obtain information about the void structure across all layer boundaries, solving the problem of not being able to distinguish between the appearance of different void structures and exclude layer boundaries when determining a single total void content of an asphalt layer by weighing and measuring methods in the past. Literature [18] addressed the problem of reduced durability of starting structures due to dry period cracking of concrete by using image analysis techniques to measure the width of shrinkage cracks, further evaluating the crack width of the degree of shrinkage cracking, and measuring and analyzing the crack area. Literature [19] used the image analysis method to estimate the tension of remote sling with modal frequency as a parameter, and by comparing with the battle force measured by vibration method, the difference between the modal frequency obtained by finite element calculation of sling and the measured modal frequency was effectively reduced. Literature [20] states that remote camera and drone acquisition based on computer vision technology provides a remote and non-contact solution for evaluating the condition of civilian infrastructure, transforming the captured image or video data into actionable information. Literature [21] investigated the application of particle image velocimetry and laser scattering techniques in reflecting the internal deformation characteristics of slopes, which yielded experimental results that help explain the sliding mechanism at the soil-rock interface. Literature [22] proposed a quantitative analysis method of coarse aggregate angle based on image technology to analyze the effect of different coarse aggregate angles on the performance of asphalt pavements using an aggregated imaging system.

In this study, the advantages and disadvantages of three conventional autoencoders are first discussed, and the sparse constraint term is added to the autoencoder model to establish the sparse autoencoder as the base structure. The image analysis model for civil engineering and hydraulic engineering was constructed using a convolutional neural network. On this basis, the stacked autoencoder is used to build the network on the neural tree complex composed of several basic sparse autoencoders and then optimized by combining the structural transfer characteristics of civil engineering and water conservancy with the image separation method. Finally, the model performance test will be applied

to actual engineering projects to study the specific performance and practical value of the method presented in this paper.

2 Deep learning-based engineering structure extraction and recognition model

2.1 Autoencoders and their extended relationships

Autoencoder (AE) is an unsupervised machine learning algorithm that can extract feature representations from data. It has been widely used in data degradation and feature extraction. An autoencoder network usually consists of a three-layer structure, i.e., There is an input layer, a hidden layer, and an output layer.

The working process of an autoencoder is mainly divided into two parts: encoding and decoding. The encoding part mainly learns the hidden features of the input data and transforms the data input to the network into encoded information output after mapping, and the decoding part reconstructs the original input data by inverse mapping the newly learned features. Assuming x is the input vector, h is the extracted hidden layer and y is the output vector, the encoding process and decoding process of the autoencoder can be represented by the following equation:

$$h = f\left(W_1 x + b_1\right) \tag{1}$$

$$y = f\left(W_2 h + b_2\right) \tag{2}$$

Where W_1 represents the weight matrix of the mapping function when the input vector is mapped to the hidden layer during the encoding process of the input layer and the hidden layer, b_1 represents the bias of the function mapping during the encoding process, W_2 represents the weight matrix when the hidden features learned by encoding are reconstructed using the inverse mapping function for input vector reconstruction during the decoding process of the hidden layer and the reconstruction layer, and b_2 represents the bias generated during the decoding process. $f\left(\cdot\right)$ is a nonlinear activation function, and a sigmoid function is usually used as the activation function.

In the training process, the autoencoder is trained on the parameters by maximizing the reconstruction of the input data, so the training objective of the autoencoder is the minimum value obtained by the error $loss\left(x, y\right)$ between the original input data and the reconstructed data output after the encoding and decoding process, i.e:

$$J = \arg\min\left[loss\left(x, y\right)\right] \tag{3}$$

When the reconstruction error is minimized, it means that the hidden layer units contain enough data features while reducing the input data loss due to mapping. By constantly updating the weights to minimize the reconstruction error, the autoencoder can extract deeper and more abstract data features, which can be used as inputs to higher-level autoencoders to generate even deeper features, in addition to being used for deeper data analysis.

1) Stack autoencoder

A deep autoencoder is built on top of the AE to extract deeper features from the data. The stack autoencoder is a deep network structure formed by stacking multiple AEs, which contain one input layer and multiple hidden layers, and the model structure is represented in Figure 1.

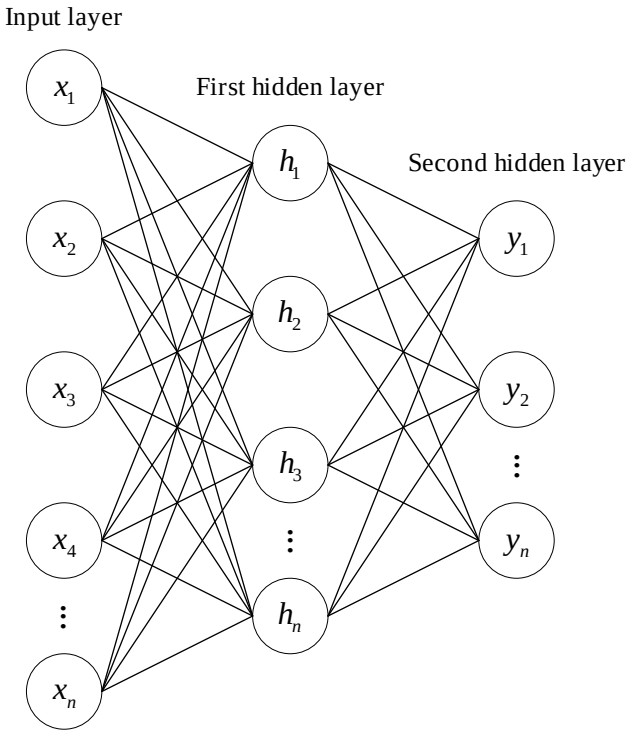


Figure 1. Model of stacked auto-encoder

2) Sparse autoencoder

The sparse autoencoder makes the AE network effective in extracting the features of the input data even when the number of nodes in the hidden layer is high by increasing the sparsity of the hidden units during training. In autoencoders, the number of nodes in the hidden layer tends to be smaller than the number of nodes in the output layer, which makes the autoencoder more inclined to learn the internal laws of the input data and obtain a dimensionality reduction representation of the data similar to principal component analysis. In order to increase the number of nodes in the hidden layer and obtain effective information, sparse autoencoders impose sparsity constraints on the entire autoencoding network by adding sparse penalty terms.

In a sparse autoencoder network, assuming that the hidden layer activation function is a sigmoid function, the output of an "active" node in the hidden layer of the network is 1, and the output of an "inactive" node is 0. Based on this, the KL dispersion is introduced to measure the similarity between the average activation outputs of the nodes in a hidden layer and a set sparsity of 1, i.e., the KL dispersion of the nodes in a hidden layer. The KL dispersion is introduced to measure the similarity between the average activation output of a hidden layer node and the set sparsity ρ, i.e:

$$KL\left(\rho \| \hat{\rho}\right) = \rho \log \frac{\rho}{\hat{\rho}} + \left(1-\rho\right) \log \frac{1-\rho}{1-\hat{\rho}} \tag{4}$$

Where $\hat{\rho}_j = \frac{1}{m}\sum_{i=1}^{m}\left[a_j^{(2)}\left(x^{(i)}\right)\right]$, $a_j^{(2)}(x)$ denote the activation of the hidden neuron j of the self-coding network with input x.

Then, adding the KL scatter as a regular term constitutes a sparsity penalty term to the loss function, which can constrain the sparsity of the whole self-coding network, i.e.:

$$J_{spare}(W,b) = J(W,b) + \beta\sum_{j=1}^{s_2} KL(\rho \square \hat{\rho}) \tag{5}$$

Where β is the weight of the sparse penalty term.

3) Denoising autoencoder

The denoising autoencoder solves the constant function risk in AE by randomly introducing noise to the input data, which necessitates AE to recover and denoise the network input. The denoising autoencoder can obtain a good characterization of the input data, i.e., the hidden layer features of the input data are robustly obtained from the noisy input, and the obtained features can be used to recover their corresponding noiseless original inputs to enhance the robustness of the network.

2.2 Image analysis model based on convolutional neural network

Convolutional neural networks (CNNs) are a type of feed-forward neural network that uses convolutional computation and has a deep structure. Convolutional neural network is nowadays recognized in the field of deep learning and widely used in the field of image processing due to its simple network structure, parsimonious network parameters, good application background and excellent experimental results.

The convolutional layer scans and calculates the local features of the image through the convolutional kernel function, and its output matrix is called the feature map. To extract the internal features of the image is its primary function. Convolution is an important integral transform method in mathematical analysis. It can be described as connecting a convolution kernel to a localized region of the previous layer, multiplying the scanned matrices in that region, and then working on the summation. The convolution operation is usually denoted by *, i.e.:

$$Y = X*Z \tag{6}$$

The pooling layer is a downsampling operation whose main purpose is to minimize the dimension of the data output from the convolutional layer. The pooling process is shown in Figure 2. The pooling layer can maximize the dimensionality of the data and maintain the local invariance of the features, greatly avoiding the occurrence of overfitting phenomena.

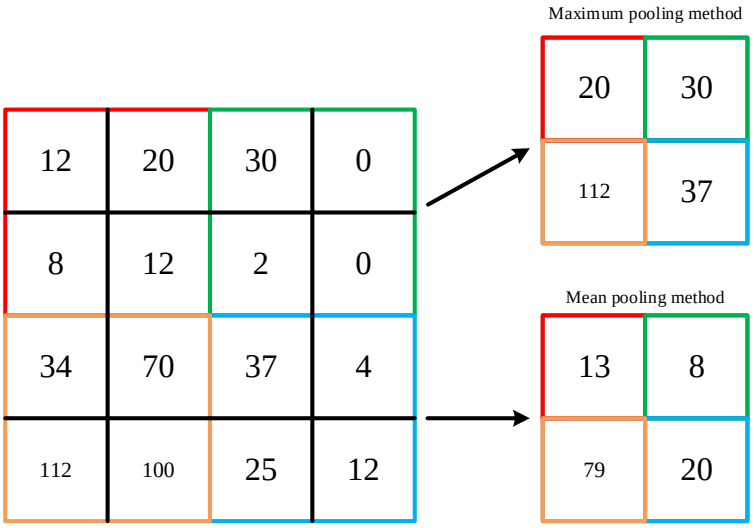


Figure 2. Pooled layer sampling process

Fully connected layer: The fully connected layer is located at the end of the convolutional neural network and connects the feature maps processed by the convolutional layer to the pooling layer for output. According to the different requirements of the target task, the fully connected layer will adjust the connected nodes accordingly and transmit the data to the output layer to obtain the final output result.

The training process of a convolutional neural network is mainly divided into two stages: forward propagation and back propagation. Assume that the input data of a given hyperspectral image is $O=\{(x_1,y_1),\cdots,(x_i,y_i),\cdots(x_n,y_n)\}\in X\times Y$ where $X=[x_1,\cdots,x_i,\cdots x_n]\in R^{m\times n}$ is a $m\times n$ -dimensional input sample and $Y=[y_1,\cdots,y_i,\cdots,y_n]^{\mathrm{T}}\in R^n$ is the label of the input sample. The convolutional layer maps the input through k trainable filter banks and biases to obtain the corresponding k -dimensional convolutional features, where the size of the filter bank is $l\times l\times q$. After inputting X into the initialized convolutional neural network, the output features obtained by the i th convolutional layer can be represented by the output feature map z^c:

$$z^c=f\left(\sum_{i=1}^{q}W_i*x^i+b_i\right) \tag{7}$$

Where W_i is the filter bank of layer i, i.e., the feature map of the convolutional kernel layer. x^i is the input data of layer i, b_i is the trainable deviation parameter of layer i. f() is the neuron nonlinear excitation, and * is the convolutional computation. The final output feature map of layer i is obtained from the neuron nonlinear excitation function, and the weights in the same feature map are shared.

The data will enter the pooling layer for downsampling after the operation of convolutional layer and the downsampled feature map is obtained z^p. The calculation formula of pooling layer is shown in the following equation:

$$z^p=f\left(down\left(z^c\right)+b_{ci}\right) \tag{8}$$

Where $down(\cdot)$ is the pooling layer subsampling function and b_{ci} is the bias of the i rd pooling layer.

In this way, after feature extraction in multiple convolutional and pooling layers, the data enters the fully connected layer to obtain the actual output of the data training sample O, i.e.:

$$O = f_\theta(X) \tag{9}$$

Where θ is the set of parameters in the convolutional neural network.

After this, the second phase of convolutional neural network training i.e. back propagation phase is entered. In this phase, the residuals between the actual outputs of the training samples O and the actual sample labels of the data Y are computed. Its objective function can be expressed as:

$$J = \min_\theta \|O - Y\|_2^2 \tag{10}$$

After multiple feature extraction stages, in order to train the convolutional neural network, stochastic gradient descent with backpropagation can be applied to optimize its functionality.

2.3 Engineering detection model based on sparse stacked autoencoder

2.3.1 Sparse Stacked Autoencoder Detection Model Construction

This paper creates a sparse autoencoder model by adding sparse constraint terms to the autoencoder model (AE). This is because the number of neurons in the generally hidden layer is generally larger, sometimes even more than the number of neurons in the input layer, and the hidden structural information in the data is uncovered by adding some other constraints to the autoencoder network, which is expanded as:

$$\Im_{AE}(\theta) = \frac{1}{N}\sum_{k=1}^{N}(L(x(k), d_{\hat{\theta}}(e_{\hat{\theta}}(x(k))))) + \beta\|W\|_2^2 + \alpha\sum_{j=1}^{n} KL(\rho \square \hat{\rho}_j) \tag{11}$$

In the third term of the above equation n denotes the number of units in the hidden layer, j denotes the j rd neuron in this hidden layer, $KL(\rho \square \hat{\rho}_j)$ describes the KL distance between the average activity $\hat{\rho}_j$ of the hidden layer neurons on the training set and the pre-set activity ρ of the hidden layer.

The specific calculation formula is as follows:

$$KL(\rho \square \bar{\rho}_j) = \rho \log\frac{\rho}{\bar{\rho}_j} + (1-\rho)\log\frac{1-\rho}{1-\bar{\rho}_j} \tag{12}$$

The stacked autoencoder used in this paper is carried out in a neural tree complex composed of several underlying sparse autoencoders, where the output of the prequel network is used as the input to the next layer of the network, and the later networks are constructed analogously. The main consideration is the two-layer sparse autoencoder model, which is simple for SSAE.

Similar to the training of a single-layer autoencoder model, the training of a multilayer Indo-Active Encoder Heapsux model also finds an optimal set of parameters by minimizing the wakamats between

the reconstructed image and the original image $\theta=(W,b_h,b_x)$. Based on this optimal set of parameters, the SSAE model is able to obtain the mapping or transformation relation equation $f:R^{'}\to R^{'}$, which transforms the input pixel-by-pixel image into a new feature representation $h^{(2)}=f(x)\in R^{\delta_{n+1}}$. Using the SSAE model, each training image block $x(k)$ can be transformed into a dimensionally oriented feature representation, i.e., the output features of the second hidden layer. All the training image blocks can be re-represented as $\{h^{|2|}(k),y(k)\}_{i=1}^{N}$, where $k\in\{1,2,3,.....N\},\{h^{(2)}(k),y(k)\}$ represents a set of images corresponding to a small high-dimensional feature table and its corresponding class label. It is important to point out that the class label information is not utilized during the training of the SSAE model Y. After all the high-dimensional features of the images have been extracted, the training dataset $\{h^{(2)}(k),y(k)\}_{k=1}^{N}$ is fed into the final output layer.

The classifier used in this paper's method is the Softmax classifier, or SMC for short, which is a supervised model expressed in the form of a logistic regression function:

$$f_{W^{(3)}}(z)=\frac{1}{1+\exp(-W^{(3)T}z)} \tag{13}$$

In the above equation towards is the activation function expressed in parameter $f_{W^{(3)}}$, Z is the representation of the features extracted using the SSAE model, and parameter $W^{(3)}$ of the SMC model is obtained using the energy cost equation of the minimization formula:

$$J(W^{(3)})=-\frac{1}{N}\left[\sum_{i=1}^{N}y(i)\log f_{m^{(3)}}(h^{(2)}(i))+(1-y(i))\log(1-\log f_{W^{(3)}}(h^{(2)}(i)))\right] \tag{14}$$

The objective equation above is optimized by the gradient descent method so that we get parameter $W^{(3)}$ of the model.

After SSAE extraction of features and training of the classifier, the detector for detecting engineering structures is obtained in order to apply the trained detector to the detection of civil engineering and water conservancy project structures, and taking into account the huge number of engineering structures, spread, and irregularity, this paper decides to use the sliding window method, i.e., selecting the appropriate square window for the size of the structure in the image to be detected, so as to be able to contain The entire engineering structure.

2.3.2 Structural transfer characteristics of civil engineering and hydraulic engineering

The study of structural problems in civil and hydraulic engineering is based on idealized mechanical models that reflect their main mechanical properties. In fact, some seemingly complex structures can be simplified or even reduced to a single degree of freedom system, and satisfactory results can still be obtained. The single-degree-of-freedom model can uncover many essential phenomena of system vibration, and its research method is the basis of structural vibration research. Here, we take the single degree of freedom system as the object of our study and analyze its frequency response characteristics and vibration structure transfer characteristics.

The equation for the structural dynamic equilibrium of the single degree of freedom system is:

$$m\ddot{X} + c\dot{X} + kX = F(t) \tag{15}$$

Where m、c、k is the mass, damping and stiffness, $\ddot{X}$, $\dot{X}$ and X are the acceleration, velocity and displacement vectors, and F is the load vector.

The system produces an acceleration response $\ddot{X}$ under the action of F. The acceleration derivative is defined by the mechanical impedance method:

$$Z_a = \ddot{X} / F \tag{16}$$

In a physical sense, the conductance represents the amount of motion produced by a system under a unit excitation.

To study the relationship between $\ddot{X}$ and F in the frequency domain is to analyze the acceleration frequency response characteristics of a single degree of freedom system. It is analyzed as follows, let the Fourier operator be L and the Fourier transform (FT) be:

$$X''(\omega) = L\ddot{X}(t) = \frac{1}{2\pi}\int \ddot{X}(t)e^{-i\omega t}dt \tag{17}$$

Similarly, the FT of F is $F(\omega)$.

By the integral property of operator L:

$$X'(\omega) = X''(\omega) / j\omega; \quad X(\omega) = -X''(\omega) / \omega^2 \tag{18}$$

From the linearity of operator L, the differential equation reduces to an algebraic equation:

$$(m - \frac{k}{\omega^2} + \frac{c}{j\omega})X^*(\omega) = F(\omega) \tag{19}$$

Convert to:

$$H(\omega) = \frac{X''(\omega)}{F(\omega)} = \frac{\omega^2}{\omega^2 m - k - j\omega c} \tag{20}$$

2.3.3 Image Separation Based on Sparse Nonnegative Matrix Decomposition

The civil and hydraulic engineering image is based on the principle that each earth structure differs in its degree of light absorption. Beer Lambert's Law (BLL) is satisfied by the degree of absorption and light it possesses, which applies to all electromagnetic radiation and light-absorbing substances, such as gases, solids, liquids, molecules, atoms, and ions. The equation is:

$$A_c = \log_{10}\frac{I_{o,c}}{I_{i,c}} \tag{21}$$

Where A_c denotes the absorbance at c, $I_{o,c}$ denotes the intensity of the incident light at c, which can be interpreted as the intensity of white light, and denotes the intensity of the transmitted light at

c, which can be interpreted as the pixel value displayed in the picture. Under the RGB color model space, the R-channel value I_R, the G-channel value I_G, and the B-channel value I_B of each pixel point are obtained from the camera. Because the relevant pixel value of each channel is related to the absorption concentration of a substance in a particular light, and this relationship is nonlinear according to the above. So it can be separated under the space of optical density, which is calculated as defined below:

$$A_c' = -\log_{10}\frac{I_{o,c}}{I_{i,c}} \tag{22}$$

Using equation (22) can convert the image under light intensity into an image under optical density, referred to as optical density image, in order to achieve separation of color images.

The target equation is:

$$\min_{A,S} f(A,S) = \frac{1}{2}\| X - AS\|_F^2 + \alpha\sum_{i=1}^{m}\|A(i,:)\|_1^2 + \beta\sum_{j=1}^{n}\|S(:,j)\|_1^2, s.t. \quad A,S \geq 0 \tag{23}$$

Where $A(i,:)$ denotes the i rd row row vector in matrix A, $a>0$ is a regularization parameter to regulate the estimation accuracy and sparsity balance of matrix A, and similarly $S(:,j)$ denotes the j th column column vector in matrix S, and $\beta>0$ is also a regularization parameter to regulate the estimation accuracy and sparsity balance of matrix S. In this paper, the iterative least square algorithm of Eq. (24)) is used to optimize the objective equation of Eq. until its convergence:

$$\min_{A\geq 0}\left\|\sqrt{\beta}e_{1\times k}\right\| s - \left\| \begin{matrix} x \\ 0_{1\times n} \end{matrix}\right\|_F^2 \tag{24}$$

In the above equation $e_{1\times k}\in \square^{1\times k}$ is a row vector whose element values are all 1's and $0_{1\times n}\in \square^{1\times n}$; is a row vector whose element values are all 0's:

$$\min_{S\geq 0}\left\| \begin{pmatrix} s^T \\ \sqrt{\alpha}e_{1\times k} \end{pmatrix} A^T - \begin{pmatrix} X^T \\ 0_{1\times m} \end{pmatrix}\right\|_F^2 \tag{25}$$

In the above equation, $e_{1\times k}\in \square^{1\times k}$ is the row vector whose element values are all 1, $0_{1\times n}\in \square^{1\times n}$; is the row vector whose element values are all 0. The formula uses minimizing the L-1 paradigm of the columns of the $S\in R^{k\times n}$ matrix to achieve the purpose of sparsity, and similarly uses minimizing the L-I paradigm of the rows of the $A\in R^{m-1}$ matrix to achieve the sparsity of the day, and at the same time in the process of solving, attention should be paid to ensure that the knowledge of the matrix A and the matrix S, each element of the matrix is greater than or equal to 0, in the iterative process of realizing the results of each time the In the realization iteration process, after each result, the elements in the matrix less than 0 are set to 0, which satisfies the condition of non-negativity.

3 Empirical analysis based on the application of civil and hydraulic engineering

3.1 Effectiveness of detection model based on sparse stacked autoencoder

3.1.1 Comparison experiment of image feature extraction

In order to verify the effectiveness of the civil engineering image matching effect in this paper, the method of this paper is compared with three commonly used feature point extraction algorithms (SIFT, SUSAN, Harris) in a comparison experiment, specifically examining their stability under geometric distortion, i.e., luminance change, rotation change, and size change. The experimental data was obtained from images of civil and hydraulic engineering structures collected in S. The total number of pictures for the city is 245. In comparison experiments, the repeatability of feature points in the changed samples and the original samples is used as an evaluation index. Specifically, brightness adjustment, rotation, and scaling are performed on an image. Then, three feature extraction algorithms are used to extract feature points at the same time under the same level of change. Then, the set of feature points extracted by each method under different levels of change is matched with the set of feature points of the original image. The ratio of the recurrence of feature points detects the stability of the algorithms. The average value is taken from several experiments to obtain more reliable test results. The test results for brightness change are displayed in Fig. 3. The repeatability is 56%, except for the SUSAN algorithm with a brightness of 75. The rest of the algorithms are higher than 65% between brightness 20 and 80. In particular, the repeatability of this paper's method is 89.29% on average, which is better for stability.

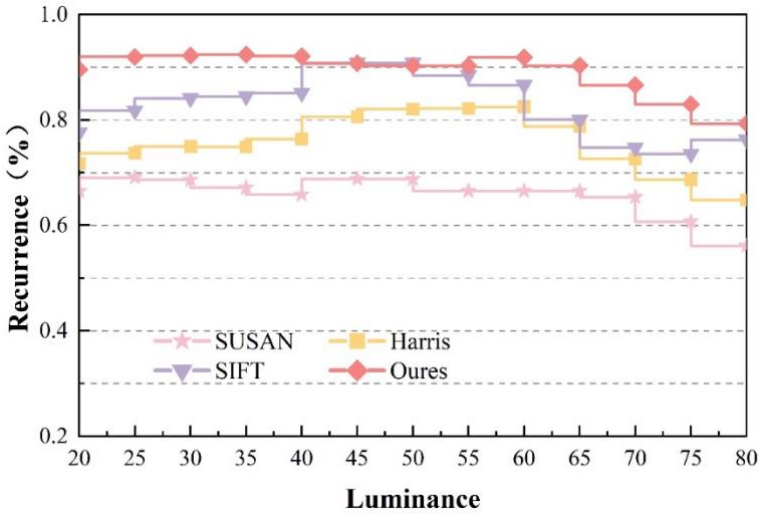


Figure 3. Test results of image brightness change

The repeatability curves of different algorithms when the rotation changes are shown in Fig. 4. The effect of SIFT and this paper's algorithm is not much different, and the repeatability basically stays above 60% in most of the rotational angles, which shows that both algorithms have better rotational stability. The SUSAN algorithm's effectiveness drops when the image experiences rotational changes, only remaining above 50%. When the image is rotated by 90°, the overlap of feature points extracted by all algorithms decreases, which is worth noticing. However, the algorithms in this paper still maintain good performance overall.

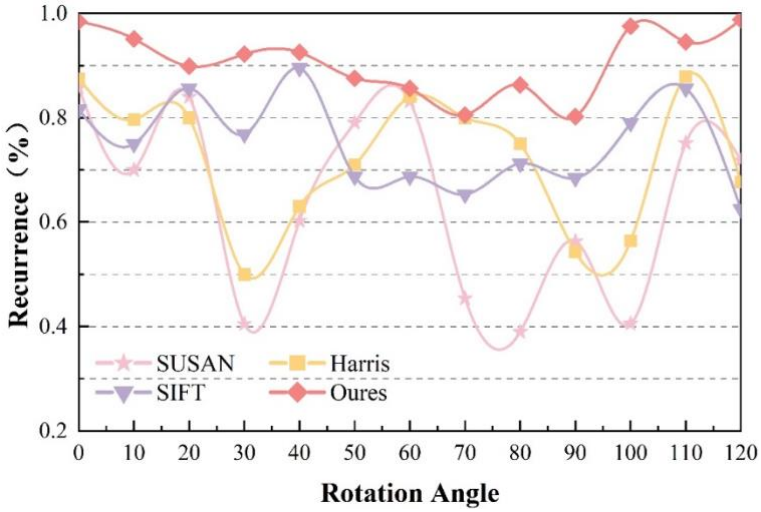


Figure 4. The rotation Angle of the different algorithm is the repetition of the rotation

The repeatability curves of different algorithms when the size changes are shown in Fig. 5. In the shift in image scale from 0.4 to 2.2 times, the algorithm in this paper is significantly more effective than the other algorithms, and the repeatability of the feature points is always above 70%, with an average repeatability of 82.47%. The performance of Harris and SIFT algorithms is comparable when the size of the image changes, and the SUSAN algorithm performs poorly. Harris and SIFT algorithms have poor adaptability to scale changes. Between 1 and 1.8 times, the effect of the algorithms decreases significantly without scale invariance, further proving the superiority and stability of the algorithms in this paper.

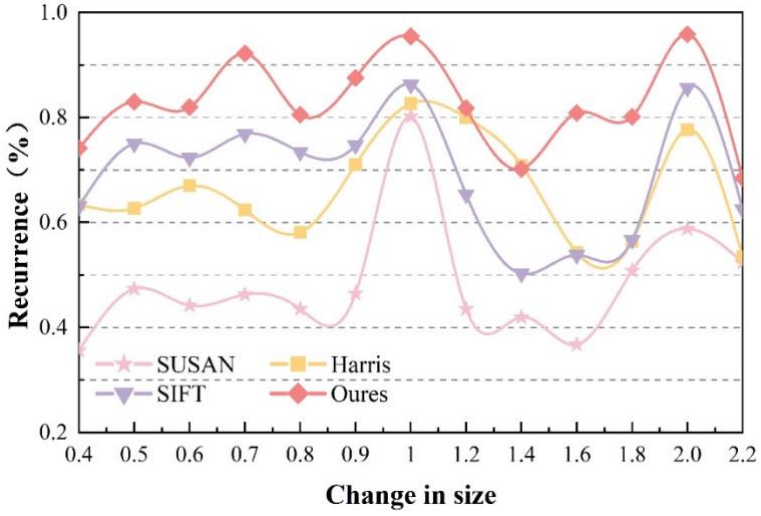


Figure 5. Different algorithms are repeated in size

3.1.2 Stacked Convolutional Self-Coding Model Performance Analysis

To fully utilize structural information and node features in the image data of civil engineering and hydraulic engineering, this paper proposes a stacked autoencoder model based on a graph convolutional neural network. The performance of the model is analyzed through experiments, and the features of the graph are used to disperse the information among the nodes so that each node can obtain the necessary information and update the features. To demonstrate the performance of the model in this paper, the designed model is compared with four frontier methods, 3D U-Net, 3D

ResNet, MSS U-Net, and 3D nnU-Net, on two datasets from KID1 and KID2. The three commonly used quantitative criteria for image segmentation are chosen to include the Dice coefficient, IoU coefficient, and HD distance. Dice is frequently used in the task of measuring segmentation quality by indicating the degree of overlap of the predicted regions and labels. IoU is another metric used to measure the degree of similarity between two sets, with higher Dice values implying more accurate segmentation results. Dice and IoU, although both are capable of rating the degree of similarity between the segmentation results and the true labeling similarity, cannot well assess the boundaries of segmentation results. Therefore, HD distance, which evaluates the distance between the segmentation prediction result and the surface of the actual result, is also used for comprehensive evaluation. The smaller the HD distance, the better the performance. The performance comparison of modeling methods for different approaches is shown in Table 1. Among the three evaluation metrics of Dice, IoU, and HD, the Dice and IoU metrics of this paper's method are higher than 90% on both data, and the HD distance is 22.548 mm and 26.48 mm, respectively, which is the best overall performance. Compared with the remaining four methods, HD indexes are reduced by 0.440-10.034mm and 20.867-41.279mm on the KID1 dataset and KID2 dataset, respectively, which illustrates that this paper's modeling method has the best performance among all the compared methods and is more effective for image edge segmentation.

Table 1. Performance comparison of model methods of different methods

Model	KID1			KID2		
	Dice	IoU	HD	Dice	IoU	HD
3D U-Net	0.965	0.933	28.569	0.876	0.782	67.759
3D ResNet	0.948	0.905	32.582	0.827	0.723	81.956
MSS U-Net	0.960	0.922	22.988	0.886	0.785	47.347
3D nnU-Net	0.963	0.929	31.07	0.903	0.823	48.185
Ours	0.968	0.933	22.548	0.925	0.904	26.48

3.2 Analysis of civil and hydraulic engineering applications

Soil structure crack is an important link in the quality inspection of civil engineering and water conservancy projects. The traditional detection method mainly relies on manual detection of crack scale and magnifying glass to detect the surface characteristics of the crack. The procedure is time-consuming and laborious, the detection results are greatly affected by human factors, and the accuracy is obviously insufficient. In order to test the practical effect of this paper's model in the application of civil engineering and water conservancy projects, this paper's model is applied to the detection of cracks in the soil structure of the earth dam project in M city for empirical analysis. The earth dam project is divided into 6 dam sections according to the location, which is labeled as B1~B6, and the error comparison test is carried out by comparing the measured value of the soil crack structure and the calculated value of this paper's model. The error comparison of the earth dam project in M city is shown in Table 2. It can be seen that the relative errors between the calculated values and the measured values of the stacked autoencoder modeling method based on the convolutional network in this paper are between 0.111~1.807% in general, and the relative errors do not exceed 2%. The accuracy of the model meets the engineering requirements, which proves that the modeling method in this paper has obvious advantages in the large-scale search and detection in the actual civil engineering and water conservancy projects and is suitable for the initial inspection of the project.

Table 2. Error comparison of dam engineering in M city

Earth dam position	Measured value	Calculated value	Relative error(%)
B1	0.879	0.887	0.911
B2	0.895	0.894	0.111
B3	0.123	0.124	0.813
B4	0.332	0.330.8	1.807
B5	0.635	0.641	0.945
B6	0.235	0.238	1.276

4 Conclusion

In this paper, on the basis of image analysis technology, the engineering detection model based on a sparse stacked ground hole encoder is constructed by using an autoencoder and convolutional neural network, and the following conclusions are drawn through the actual effect of civil engineering and water conservancy project application:

1) The average values of the repetition rate of the image feature extraction of the model in this paper are 89.29%, 90.54%, and 82.47% in the tests of brightness adjustment, rotation, and scaling operation changes, respectively, which are better than other methods. Among the three evaluation indexes of Dice, IoU, and HD, the Dice and IoU indexes of this paper's method are higher than 90% on both data. The HD distance is less than 27 mm. Compared to the other four methods, HD indexes are reduced by 0.440~10.034 mm and 20.867~41.279 mm on the KID1 dataset and KID2 dataset, respectively. It proves the effect of this paper's algorithm on image features, mentioning stability, with excellent performance.

2) This paper, based on the convolutional network stacked autoencoder model method, calculated value and measured the value of the relative error between the overall 0.111 ~ 1.807%. The relative error does not exceed 2%. It can meet the civil engineering and water conservancy project identification and detection of the accuracy requirements, indicating that this paper's modeling method in the detection and identification of civil engineering and water conservancy project has practical application value.

References

[1] Shen, H., Yan, J., Liu, D., & Liu, Z. (2020). A new method for determination of time-of-day breakpoints based on clustering and image segmentation. Canadian Journal of Civil Engineering(8).

[2] Poggiogalle, T. M., Meehan, C. L., Clarke-Sather, A. R., & Talebi, M. (2022). A digital image analysis technique for improved strain measurement in geosynthetic tensile testing. Geotechnical testing journal(3), 45.

[3] Tuan, N. M., Kim, Y., Lee, J. Y., & Chin, S. (2022). Automatic stereo vision-based inspection system for particle shape analysis of coarse aggregates. Journal of computing in civil engineering(2), 36.

[4] Wang, X., Wittich, C. E., Hutchinson, T. C., Bock, Y., & Kuester, F. (2020). Methodology and validation of uav-based video analysis approach for tracking earthquake-induced building displacements. Journal of Computing in Civil Engineering, 34(6), 04020045.

[5] Ivana Kumpová, Michal Vopálensk, Tomá Fíla, Kyt, D., Daniel Vavík, & Pichotka, M., et al. (2018). On-the-fly fast x-ray tomography using a cdte pixelated detector – application in mechanical testing. IEEE Transactions on Nuclear Science, PP(12), 1-1.

[6] Sudarsanan, N., Arulrajah, A., Karpurapu, R., & Amrithalingam, V. (2019). Digital image correlation technique for measurement of surface strains in reinforced asphalt concrete beams under fatigue loading. Journal of Materials in Civil Engineering, 31(8), 04019135.1-04019135.10.

[7] Obaidat, M. T., Ghuzlan, K. A., & Alawneh, M. M. (2017). Analysis of volumetric properties of bituminous mixtures using cellular phones and image processing techniques. Canadian Journal of Civil Engineering, cjce-2017-0085.

[8] Zhu, Y., Zhang, Y., Hussein, H. H., & Qu, S. (2021). Existing inverse analysis approaches for tensile stress–strain relationship of uhpc with treated steel fibers. Journal of Materials in Civil Engineering, 33(6), 04021118.

[9] Xiao, B., & Kang, S. C. (2021). Development of an image data set of construction machines for deep learning object detection. Journal of Computing in Civil Engineering, 35(2), 05020005.

[10] Chen, Ling, Wang, & Yuhong. (2017). Improved image unevenness reduction and thresholding methods for effective asphalt x-ray ct image segmentation. Journal of computing in civil engineering.

[11] Yang, G., Wang, K. C. P., Li, J. Q., Fei, Y., Liu, Y., & Mahboub, K. C., et al. (2021). Automatic pavement type recognition for image-based pavement condition survey using convolutional neural network. Journal of computing in civil engineering(1), 35.

[12] Jagadeesh, A., Ong, G. P., & Su, Y. M. (2019). Development of discharge-based thresholding algorithm for pervious concrete pavement mixtures. Journal of materials in civil engineering, 31(9), 04019179.1-04019179.12.

[13] Xiangqian, F., Shengtao, L., Xudong, C., Saisai, L., & Yuzhu, G. (2020). Fracture behaviour analysis of the full-graded concrete based on digital image correlation and acoustic emission technique. Fatigue & Fracture of Engineering Materials & Structures, 43.

[14] Bidgoli, M. A., Hajikarimi, P., Naderi, K., Golroo, A., & Pourebrahimi, R. (2020). Introducing adhesion-cohesion index to evaluate moisture susceptibility of asphalt mixtures using a registration image-processing method. Journal of Materials in Civil Engineering, 32(12).

[15] Chen, X., Chen, C., Cheng, X., Wu, C., Shi, Z., & Li, W. (2021). A study of loading rate effect fracture behavior of concrete based on digital image correlation and finite-element method:. The Journal of Strain Analysis for Engineering Design, 56(3), 161-172.

[16] Ahn, S. J., Han, S. U., & Al-Hussein, M. (2019). 2d drawing visualization framework for applying projection-based augmented reality in a panelized construction manufacturing facility: proof of concept. Journal of Computing in Civil Engineering, 33(5), 04019032.1-04019032.15.

[17] Tielmann, M. R. D., & Hill, T. J. (2018). Air void analyses on asphalt specimens using plane section preparation and image analysis. Journal of Materials in Civil Engineering, 30(8), 04018189-.

[18] Hemalatha, T., & Ramesh, G. (2019). Mitigation of plastic shrinkage in fly ash concrete using basalt fibres. Canadian Journal of Civil Engineering.

[19] Sung-Wan Kim, Jin-Hwan Cheung, Jae-Bong Park, & Sung-Ok Na. (2020). Image-based back analysis for tension estimation of suspension bridge hanger cables. Structural Control and Health Monitoring, 27.

[20] Spencer, B. F., Hoskere, V., & Narazaki, Y. (2019). Advances in computer vision-based civil infrastructure inspection and monitoring. Engineering, 5(2).

[21] Zhuang, W., Chi, L. I., & Xuan-Ming, D. (2019). Application of transparent soil model tests to study the soil-rock interfacial sliding mechanism. Journal of Mountain Science.

[22] Huyan, J., Li, W., Tighe, S., Zhang, Y., & Yue, B. (2020). Image-based coarse-aggregate angularity analysis and evaluation. Journal of materials in civil engineering, 32(6), 04020140.1-04020140.14.