Spatiotemporal early prediction of rock damage in rock engineering based on infrared radiation monitoring technology

Gao, Qiangqiang; Ma, Liqiang; Liu, Wei; Khan, Naseer Muhammad; Wang, Xiuzhe; Ni, Yanxiao; Yu, Kunpeng; Alarifi, Saad S.

DOI
10.1016/j.engfracmech.2025.110811

Publication date
2025

Document Version
Final published version

Published in
Engineering Fracture Mechanics

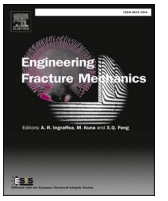
Citation (APA)
Gao, Q., Ma, L., Liu, W., Khan, N. M., Wang, X., Ni, Y., Yu, K., & Alarifi, S. S. (2025). Spatiotemporal early prediction of rock damage in rock engineering based on infrared radiation monitoring technology. *Engineering Fracture Mechanics*, *315*, Article 110811. https://doi.org/10.1016/j.engfracmech.2025.110811

Engineering Fracture Mechanics 315 (2025) 110811

Contents lists available at ScienceDirect

Engineering Fracture Mechanics

journal homepage: www.elsevier.com/locate/engfracmech



Spatiotemporal early prediction of rock damage in rock engineering based on infrared radiation monitoring technology

Qiangqiang Gao [d], Liqiang Ma [a,b,c,d,*], Wei Liu [d], Naseer Muhammad Khan [d,g], Xiuzhe Wang [f], Yanxiao Ni [e], Kunpeng Yu [d], Saad S. Alarifi [h]

[a] Xinjiang Key Laboratory of Coal-bearing Resources Exploration and Exploitation, Xinjiang Institute of Engineering, Urumqi 830023, China
[b] Xinjiang Engineering Research Center of Green Intelligent Coal Mining, Xinjiang Institute of Engineering, Urumqi 830023, China
[c] Key Laboratory of Xinjiang Coal Resources Green Mining (Xinjiang Institute of Engineering), Ministry of Education, Urumqi 830023, China
[d] School of Mines, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China
[e] School of Foreign Studies, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China
[f] Faculty of Civil Engineering and Geoscience, Delft University of Technology Stevinweg 1, Delft 2628CN, Netherlands
[g] Sustainable Advanced Geomechanical Engineering, National University of Sciences and Technology, Risalpur 23200, Pakistan
[h] Department of Geology and Geophysics, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

ARTICLE INFO

Keywords:
Coal and rock damage prediction and early warning
Spatio-temporal prediction
Spatio-temporal feature extraction
Infrared radiation monitoring

ABSTRACT

This study introduces the Spatio-Temporal Attention Enhanced Encoder-Decoder Damage Prediction Network (STAE-EDDPNet), an innovative deep learning model designed to enhance the predictive capabilities of coal-rock damage infrared temperature fields, which is crucial for the safe production in rock engineering and mining engineering. STAE-EDDPNet integrates a spatio-temporal attention mechanism, significantly improving the capture of complex nonlinear spatio-temporal information in rock infrared radiation. Compared with baseline models such as 3DCNN, ConvLSTM, and EDDPNet, STAE-EDDPNet demonstrated superior performance in both single-step and multi-step forecasting tasks. Test set results show that its predictive accuracy is 25.56% higher than 3DCNN, 5.69% higher than ConvLSTM, and 0.19% higher than EDDPNet. The study also found that the characteristics of brittle failure rock data significantly affect model training and predictive performance, providing a direction for future data collection and experimental design improvements. The introduction of STAE-EDDPNet not only promotes the application of infrared monitoring technology in the field of safety monitoring but also provides valuable reference for rock damage early warning.

1. 1、Introduction

Rock is a complex mechanical medium formed through long-term geological processes, containing natural defects such as microcracks and pores at various scales [1,2]. The instability failure of rock is the process of initiation, expansion, aggregation and penetration of internal micro-cracks, which is the result of rock damage accumulation. This process determines the destabilizing failure course of rock materials and their ability to be sustainably used [3–5]. Therefore, the prediction and quantitative characterization of rock damage evolution process under stress is of great significance for the stability monitoring and disaster warning of mining engineering and underground space engineering.

* Corresponding author.
E-mail address: ckma@cumt.edu.cn (L. Ma).

https://doi.org/10.1016/j.engfracmech.2025.110811
Received 13 October 2024; Received in revised form 29 December 2024; Accepted 8 January 2025
Available online 10 January 2025

Nomenclature

Symbols

ITF	Infrared temperature field
C	Number of channels
M, N	Row and column indexes of the infrared temperature matrix, respectively
I	Infrared temperature field sequence
i_t, f_t, C_t, O_t, H_t	Input gate, forget gate, cell State, output gate, hidden state
STAE-EDDNet	Spatio-temporal attention-enhanced encoder-decoder damage prediction network
AIRT	Average infrared radiation temperature (℃)
$AIRT_R(v)$	Average infrared radiation temperature of the reference specimen in the vth frame
A_i	Sample number
g_v^L,g_v^R	Infrared thermogram sequences of the rock loaded specimen and the reference specimen, respectively
$g_v^{'}$	Temperature matrix sequence of the rock infrared thermograms after denoising of the thermal background
u	Index of the frame after denoising of the multi-frame cumulative average
s	Size of the window
v	Matrix frame index
m, n	Rows and columns indexes, respectively
k, l	Maximal rows and columns frame indexes, respectively
g(u)	Infrared radiation temperature matrix sequence after multi-frame cumulative average denoising
U	Characteristic domain of rock infrared temperature field data
MAE	Mean absolute error
RMSE	Root mean square error
R	Pearson's correlation coefficient
$\widehat{y}_t$	Predicted rock ITF at moment t
y_t	Real rock ITF at moment t
$\overline{\widehat{y}_t}$,$\overline{y}_t$	Temperature averages of the predicted and real temperature fields, respectively

In recent years, technologies such as acoustic emission (AE) [6], resistivity [7], computed tomography (CT) [6,7], ultrasonic waves [8], and X-rays [9] have made significant progress in rock damage monitoring [8–10]. However, these methods have certain limitations in terms of sensitivity, real-time monitoring, and applicability to complex rock masses. For example, AE is sensitive to environmental noise, CT and X-ray technologies are expensive and lack real-time capabilities, while ultrasonic and resistivity methods face challenges in monitoring complex crack networks. In contrast, infrared radiation monitoring technology, by capturing changes in the surface temperature field of rocks, enables non-contact, real-time, and highly sensitive monitoring. It has been widely applied in studies on internal damage, fracture behavior, fatigue assessment, and damage evaluation of rock-like materials [11–15]. Some scholars have done a lot of research on the IR characteristics and failure precursors of rock materials under different loading conditions. Zhao et al. [16] studied the IR and acoustic emission characteristics and precursor characteristics of sandstones with different grain sizes, and found that the acoustic emission and AIRT spatial evolution of sandstones with different grain sizes are similar. Liu et al. [17] defined a new IR index, Infrared Energy Response of Damage, to describe the crack evolution state inside the rock, discovered the precursor of rock rupture, and established a new ontological model of rock damage. Franzosi et al. [18] studied the dynamic process of rock cooling by using infrared thermal imager and quantitatively evaluated the rock fracture state. Hao et al. [19] and He et al. [20] studied the IR rupture precursors of coal samples using the critical moderating theory. In recent years, researchers have put forward a variety of statistical quantitative indicators of IR to evaluate the characteristics of changes in IR of rock-like materials and to find the precursor of coal and rock fracture. For example, AIRT [21,22], extreme value (maximum and minimum temperature value) [23,24], variance [25], differential IR variance [26], infrared image entropy [27], IR b value [28], high temperature point scale factor [29], temperature distribution skewness [24,30], temperature distribution kurtosis [30], roughness [31] and other quantitative evaluation indexes. The damage state and fracture precursor of rock are determined mainly from the comparative analysis of IR information time series and stress state. At the same time, the spatial anomaly is assisted by infrared thermal image of indicator abnormal time[11,19,23,32,33].

However, the anomaly of infrared information on the rock surface is mainly caused by the temperature anomaly in some areas of the infrared thermal image. The description of quantitative indicators is only numerical statistics of the temperature field information, and the precursor characteristics are subjective to a certain extent, which cannot simultaneously take into account the historical spatiotemporal characteristics of the infrared thermal image sequence of the rock. Therefore, it is difficult to determine the spatial location and damage severity of rock surface damage with a single statistical feature. In addition, most of the existing studies are based on the temporal characteristics analysis of rock offline data, so it is difficult to predict the surface damage anomaly and the spatial location of the anomaly online. Therefore, if the two methods can be combined, that is, the time and space information of rock infrared information sequence can be considered at the same time, not only can the long-term monitoring of rock damage risk be realized, but also the advance prediction of rock surface damage spatial information can be realized.

With the rapid development of deep learning technology, deep learning forecasting methods have been widely used in nonlinear

spatio-temporal series forecasting tasks, including traffic flow forecasting [34], precipitation approaching forecasting [35,36], typhoon forecasting [37,38], video forecasting [39,40], automatic driving and so on. Shi et al. [36] firstly constructed ConvLSTM neural network units by combining convolutional neural network (CNN) with long short-term memory network LSTM, and then stacked ConvLSTM units using sequence-to-sequence architecture for precipitation proximity forecasting. And further creatively proposed the TrajGRU model [35]. Subsequently, the researchers proposed a series of improved models of ConvLSTM, which have been applied in many fields [41–46]. Xiao et al. [47] proposed a spatiotemporal deep learning model using ConvLSTM as the building block, which realized the accurate prediction of short and medium SST field. Yao et al. [48] proposed a spatio-temporal surface deformation spatial and temporal prediction convolutional long and short-term memory neural network based on temporal interferometric synthetic aperture radar (InSAR) for the prediction of surface deformation in mining areas.

The prediction of infrared temperature field (ITF) on rock surface is a typical spatiotemporal series prediction problem. Therefore, the deep learning method can be used to learn and characterize the complex characteristics of the ITF distribution on the rock surface and the spatio-temporal changes. In view of the current research status, it is urgent to predict the future ITF of rock damage by considering both time and space information of IR. Therefore, based on the infrared thermal image sequence of bearing rocks, this paper constructs the infrared spatiotemporal prediction data set of rocks, uses deep learning prediction models and methods, and combines the ITF information of rock surface to establish a high-precision prediction model of rock surface temperature field, and applies it to short-term prediction and prediction of regional temperature of rock surface. It can provide technical support for monitoring and early warning of surrounding rock stability of rock engineering.

The rest of the paper is organized as follows. We first introduce the model and methodology proposed in this paper in Section 2. Then, we present the acquisition and processing of the rock damage infrared radiation dataset in Section 3, and make an over-the-top multistep prediction of the infrared radiation information during the rock damage evolution in Section 4. Finally, we analyze the mechanism of rock damage infrared radiation information evolution as well as the over-advanced prediction mechanism in Section 5, and give the conclusions of this paper in Section 6.

2. Methodology

This section mainly defines the problem of IR prediction in the process of rock damage evolution, and introduces the structure and method of the IR temperature field prediction model proposed in this paper.

2.1. Problem statement on ITF prediction in rock damage

The goal of IR monitoring and prediction during rock damage evolution is to use previously observed IR thermogram sequences to predict multi-step IR thermograms (i.e., temperature fields) for future damage states of rocks. From a machine learning perspective, the problem is essentially a spatio-temporal sequence prediction problem. Therefore, the prediction of ITF during rock rupture can be regarded as a spatio-temporal sequence prediction task, i.e., based on the temporal and spatial characteristics of the previous ITF modeling to predict the subsequent ITF, according to the predicted temperature distribution characteristics of the temperature field to achieve the purpose of predicting the hazardousness of the damage degree of the rock early warning.

Firstly, the tensor $I \in R^{C\times M\times N}$ is used to denote the ITF sequence, where M and N denote the row and column indexes of the infrared temperature matrix, respectively, and C denotes the number of channels. If we use $L = \{I_{t-i+1}, I_{t-i+2}, \ldots, I_{t-1}\}$ to denote the i-frame ITF acquired experimentally, $S = \{I_t, \ldots, I_{t+j-1}\}$ to denote the j-frame infrared thermal temperature field generated by the model prediction, and $\overline{S} = \{\overline{I}_t, \ldots, \overline{I}_{t+j-1}\}$ to denote the j-frame ITF acquired experimentally (i.e., the real value of the prediction target). Then, the ITF prediction problem can be modeled as a pixel-level regression problem from $\overline{S}$ to S. The main objective is to minimize the error between $\overline{S}$ and S, which can be expressed as:

$$\overline{I}_t, \ldots, \overline{I}_{t+j-1} = \underset{I_t,\ldots,I_{t+j-1}}{\operatorname{argmax}} p\left(I_t, \ldots, I_{t+j-1} \middle| I_{t-i+1}, \ldots, I_{t-1}\right) \tag{1}$$

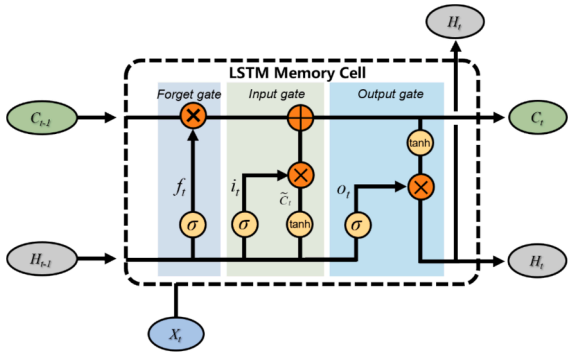


Fig. 1. LSTM structure diagram.

where p denotes the conditional probability. The above equation indicates that we use the observed IR temperature field of the first i frames to predict the IR temperature field of the future j frames.

2.2. LSTM

Hochreiter et al. proposed Long Short-Term Memory Network (LSTM) in 1997, which is similar to RNN and includes input, hidden and output layers, the most important feature in the LSTM memory unit is the use of gate structures to realize the retention, forgetting as well as updating of sequential information [49]. The structure of LSTM is shown in Fig. 1. There are three gate structures in the LSTM unit, which are forgetting gate, input gate and output gate. The input gates are mainly used to process and adjust the input information appropriately, the forgetting gates are mainly used to control the retention and forgetting of the historical information, and the output gates are mainly used to process the output information of the hidden layer. The formula for the inner gate function of the LSTM is:

$$f_t = \sigma\left(W_{xf}X_t + W_{hf}H_{t-1} + b_f\right) \tag{2}$$

$$i_t = \sigma(W_{xi}X_t + W_{hi}H_{t-1} + b_i) \tag{3}$$

$$\widetilde{C}_t = \tanh(W_{xc}X_t + W_{hc}H_{t-1} + b_c) \tag{4}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \widetilde{C}_t \tag{5}$$

$$o_t = \sigma(W_{xo}X_t + W_{ho}H_{t-1} + b_o) \tag{6}$$

$$H_t = o_t \circ \tanh(C_t) \tag{7}$$

f_t denotes the result after the information passes through the forgetting gate at moment t. ft is mapped to the range between 0 and 1 after the sigmod activation function, 0 means all discarded, 1 means all retained. i_t denotes the result after the input information passes through the input gate. $\widetilde{C}_t$ denotes the state of the cell of the current input, and C_t is the state of the cell at the current moment. O_t denotes the result after the input information passes through the output gate at moment t. H_t is the output result of the LSTM cell at moment t. H_{t-1} is the output of the previous moment, and X_t is the input at moment t. *Tanh* is the hyperbolic tangent function. The symbol $\circ$ denotes the Hadamard operation between matrices. W and b are the learnable parameters of the model, which are the weights and biases of the network, respectively.

2.3. ConvLSTM

ConvLSTM is a combination of convolutional filter and LSTM. Shi et al. [36] changed the feedforward methods in the directions of input-to-gate and gate-to-gate in the LSTM cell from pointwise multiplication to convolution operation, which enables ConvLSTM to process data with spatial continuity, such as images, videos, etc., and capture long-term dependencies at the temporal level in these data, thus realizing long time series of prediction. The basic structure of ConvLSTM is similar to that of LSTM, as shown in Fig. 2. All of its computations are realized by convolutional operations. The input gates, forget gates and output gates in the ConvLSTM cell are formulated as follows:

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \tag{8}$$

$$f_t = \sigma\left(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f\right) \tag{9}$$

$$\widetilde{C}_t = \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \tag{10}$$

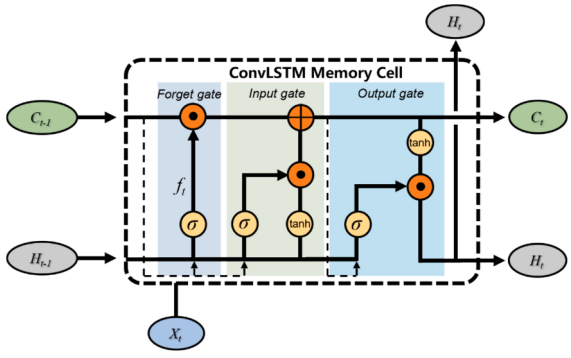


Fig. 2. Schematic diagram of ConvLSTM structure.

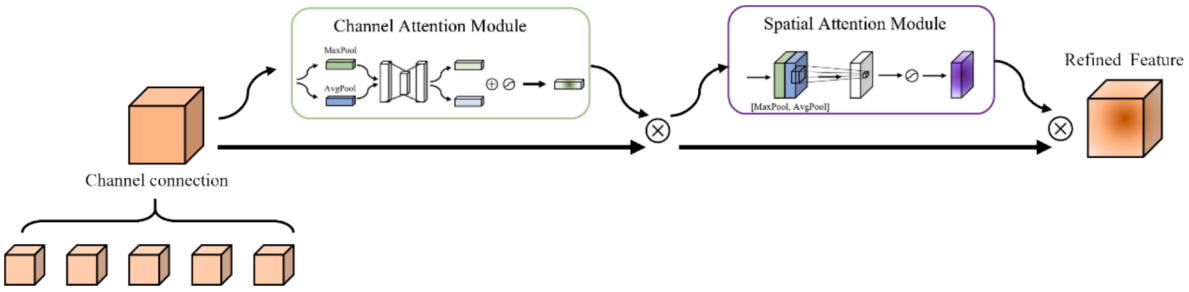


Fig. 3. Schematic diagram of CBAM structure.

$$C_t = f_t \circ C_{t-1} + i_t \circ \widetilde{C}_t \tag{11}$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o) \tag{12}$$

$$H_t = o_t \circ \tanh(C_t) \tag{13}$$

where * denotes a convolutional operation. Unlike LSTM, the connections between individual gates and inputs in ConvLSTM become convolution operations, as do the states.

2.4. Attention mechanisms

CBAM is a lightweight general-purpose module that can be seamlessly integrated into the model with negligible computational overhead [50], as shown in Fig. 3. CBAM enhances the feature representation of the network through two sequential modules (the channel attention module and the spatial attention module). This allows the model to not only identify which channels are important, but also to further localize to important regions within each channel.

2.5. Prediction model of ITF of rock

Rock ITF data contain high-dimensional spatio-temporal information, and an effective spatio-temporal fusion network structure is a prerequisite for realizing ITF prediction and warning. In this paper, a coding and decoding network model based on ConvLSTM units is designed (Fig. 4). The model integrates a spatio-temporal attention mechanism to enhance the performance of the model, which can

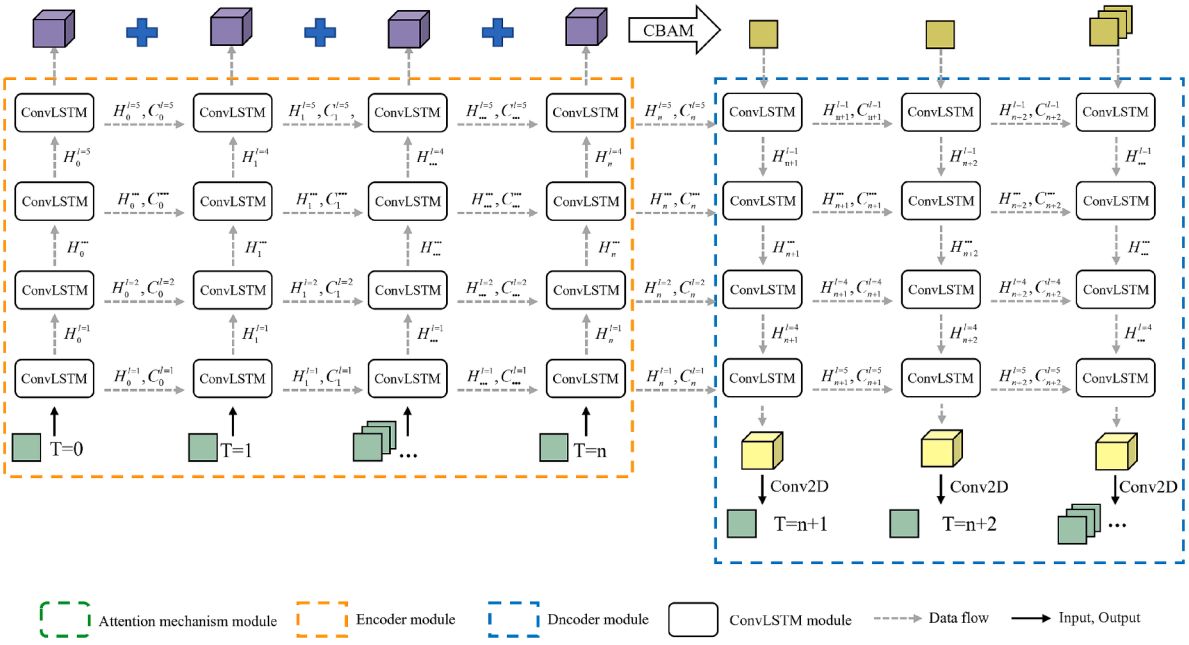


Fig. 4. STAE-EDDPNet model architecture diagram.

capture the complex nonlinear spatio-temporal information in the rock infrared information, and is used for the spatio-temporal prediction of infrared radiation information during the rock damage evolution. In this paper, we refer to it as the spatio-temporal attention-enhanced encoder-decoder damage prediction network (STAE-EDDNet).

a) Encoder: The STAE-EDDPNet proposed in this paper consists of two parts: an encoder and a decoder, see Fig. 4. For a given sequence of rock infrared radiation temperature matrices X_1, X_2, …, X_t, the encoder's task is to encode the input temperature matrix sequence into a hidden state with sequential spatiotemporal features. The encoder consists of a stack of multi-layer ConvLSTM units, where each layer further abstracts and encodes the input data, extracts spatio-temporal features at different levels, and transmits them to the next layer and the next time step. This recursive spatio-temporal structure can extract spatio-temporal features carrying the ITF of the rock to a large extent long and decode them for prediction.

b) Decoder: In line with the number of ConvLSTM layers in the encoder, the decoder also consists of multiple layers of ConvLSTM cells stacked together, while its task is to decode the output of the encoder to generate the predicted numerical matrix of the rock ITF. It is worth noting that the initial input to the decoder is obtained by processing the output of the encoder through the spatio-temporal attention mechanism. This is shown in Fig. 3 and Fig. 4.

In Fig. 4, H and C represent the hidden state and time memory unit of the ConvLSTM unit, respectively, with superscripts and subscripts representing layers and timestamps, respectively. The number of encoder and decoder steps in the model is consistent with the input step and the predicted step, respectively. This design allows the STAE-EDDPNet model to handle input data of various lengths and to generate the desired predictive data.

In the STAE-EDDPNet model, the initialization of both encoder and decoder is 0 initialization. This is because we want the network to start processing data without any prior information. In addition, the initialization state of all layers of the decoder's first time step is provided by the state of the encoder's last time step. This is because we want the decoder to be able to take advantage of the features and information of the input data extracted by the encoder to generate more accurate predictions.

3. Data acquisition and preprocessing

This section describes in detail the infrared radiation monitoring experiments during rock rupture carried out in this study, as well as the methods of obtaining and processing the experimental data.

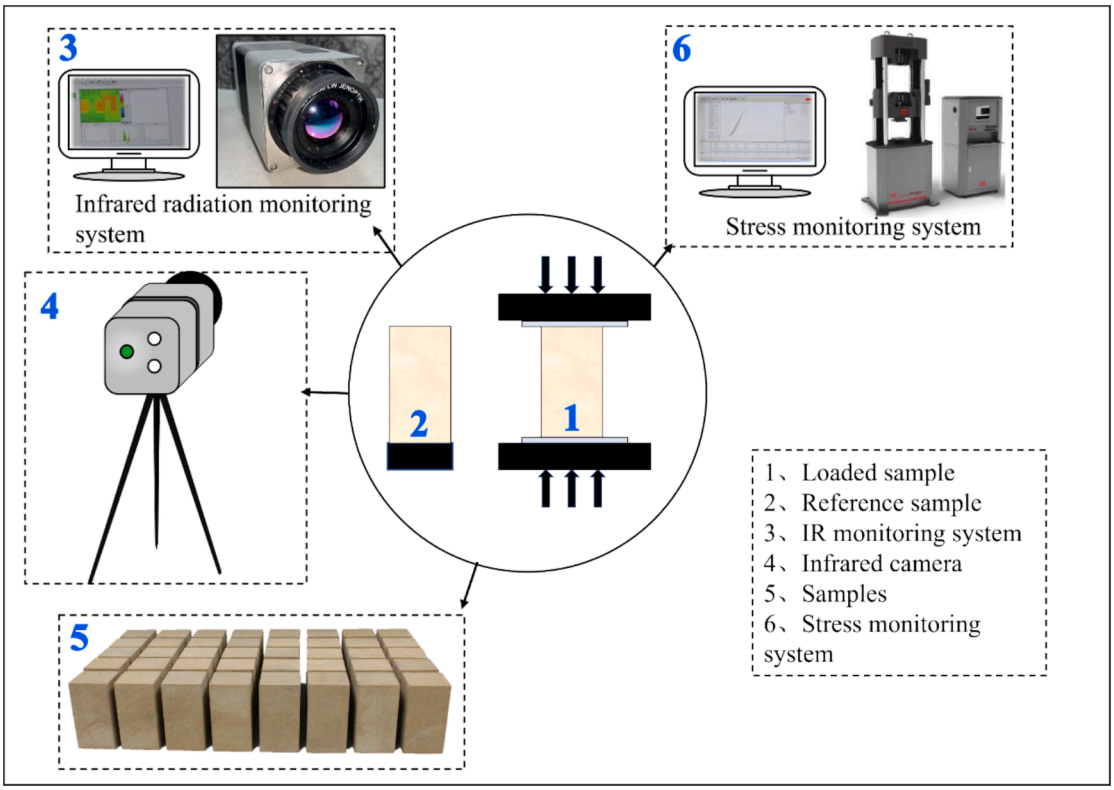


Fig. 5. Schematic diagram of the experiment.

3.1. Infrared radiation monitoring experiments on loaded rocks

3.1.1. Experimental equipment and materials

Fig. 5 illustrates the schematic of the experimental setup. The loading apparatus employed in the experiment is an MTS electro-hydraulic servo universal testing machine with a maximum load capacity of 1000kN. The infrared camera used is the VarioCAM HD head 880 high-definition thermal imaging camera from German manufacturer InfraTec, featuring a temperature range of −40 °C to 1200 °C, a resolution of 1240 × 768 pixels, a temperature sensitivity of 0.02 °C, a maximum frame rate of 30 frames per second (fps), and a measurement wavelength band of 7.5–14 μm.

The experimental material used in this experiment is sandstone common in mining engineering, geotechnical engineering and other engineering fields, and the sandstone samples used are obtained from the whole rock of a quarry in Linyi, Shandong Province. Processing the sample size of 50 mm × 50 mm × 100 mm cuboid standard specimen, two end surfaces with sandpaper grinding, so that the end surface unevenness error is less than 0.05 mm, the height of the two sides of the non-parallel error is less than 0.03 mm, a total of 40 pieces, numbered A_i (i = 1, 2, …,40). Before the start of the experiment, the experimental material is put into the laboratory 24 h in advance, so that the rock sample is consistent with the ambient temperature of the laboratory. The basic physical and mechanical parameters of all samples are shown in Table 1. The average peak stress is 59.06Mpa, the average elastic modulus is 4.54Gpa, and the average loading time is 525.52 s. Among them, one sample failed to load due to operation error during the experiment, and it was numbered A_{40}.

3.1.2. Experimental process

Set the parameters of the instrument before the experiment. Set the loading rate of the loading device to 0.2 mm/min. The acquisition rate of infrared radiation monitoring equipment is set to 30 frames /s. Place the reference sample about 10 cm next to the loaded sample. The reference sample is at the same level as the loading sample. The infrared thermal imager is about 1 m away from

Table 1
Physical and mechanical parameters of samples.

Sample numbering	Peak stress (Mpa)	Elastic modulus (Gpa)
A_1	29.77	3.40
A_2	59.79	4.46
A_3	58.46	4.23
A_4	58.04	4.43
A_5	49.93	3.61
A_6	58.97	4.96
A_7	57.32	5.31
A_8	60.36	5.52
A_9	48.26	3.04
A_{10}	53.70	3.22
A_{11}	51.68	2.89
A_{12}	58.41	4.19
A_{13}	58.08	4.07
A_{14}	53.04	4.07
A_{15}	58.36	4.57
A_{16}	39.08	2.24
A_{17}	50.88	2.77
A_{18}	58.37	4.33
A_{19}	45.94	3.25
A_{20}	61.21	4.96
A_{21}	57.30	4.60
A_{22}	54.08	3.90
A_{23}	35.48	3.56
A_{24}	59.08	4.41
A_{25}	44.20	2.91
A_{26}	62.12	5.25
A_{27}	48.57	3.23
A_{28}	59.84	4.76
A_{29}	55.56	4.57
A_{30}	58.00	5.58
A_{31}	72.42	5.70
A_{32}	78.41	6.14
A_{33}	95.33	6.48
A_{34}	87.26	5.44
A_{35}	59.94	4.42
A_{36}	68.84	5.64
A_{37}	84.41	4.98
A_{38}	86.52	5.74
A_{39}	66.27	3.29
A_{40}		
Average value	59.06	4.54

the loaded sample. At the same time, before starting the experiment, 0.1kN force is applied to the loaded sample to make the press indenter fully contact with the rock. After the preparation is completed, the loading equipment and infrared radiation monitoring instrument begin to work at the same time. In order to reduce the influence of the environment on the experimental results, no one moved around during the whole experiment, and the doors and Windows in the laboratory were closed to reduce the air flow. During the whole experiment operation, the experimenter wore heat insulation gloves and kept quiet during the loading process.

3.2. Rock ITF prediction data preprocessing and data set production

3.2.1. Reconstruction of rock ITF

The infrared radiation information released during the rock damage process can be detected by the detector of the thermal imager and recorded as the infrared thermal image sequence. As shown in Fig. 6, in this paper, the infrared thermograms of the rocks in the entire thermogram area are resampled and processed into a sequence of temperature matrices of the loaded specimen $f_p^L(\mathrm{m},n)$ and a sequence of temperature matrices of the reference specimen $f_p^L(\mathrm{m},n)$ using matlab2022.

3.2.2. Rock ITF data set denoising

During the experiment, various noises can interfere with the effective information of the infrared radiation from the rock. For example, the surrounding environment and equipment radiation, the temperature drift caused by the correction of non-uniformity of the whole column of infrared focal plane of the uncooled thermal imaging camera. In order to minimize the influence of noise on the experimental results, Sun et al. creatively set up a control specimen during the experiment and found that the average infrared radiation temperature (AIRT) of the loaded specimen and the reference specimen have a strong linear correlation, which led to the proposal of the background thermal noise correction model [51]. In addition for static infrared image sequences, for the characteristics of strong inter-frame correlation and weak time domain correlation of the effective signals, the time domain noise effect on infrared temperature can be effectively reduced by utilizing the multi-frame cumulative averaging denoising method [52].

In this paper, based on the above analysis as well as previous studies, the background thermal noise is first removed from the ITF data using a reference specimen, and then it is subjected to multi-frame cumulative averaging to remove the temporal noise, which is calculated as follows.

$$AIRT_R(v) = \frac{1}{k}\frac{1}{l}\sum_{m=1}^{k}\sum_{n=1}^{l} g_v^R(m,n) \tag{14}$$

$$g_v'(m,n) = g_v^L(m,n) - AIRT_R(v) \tag{15}$$

$$g(u) = \frac{1}{s}\sum_{(u-1)s+1}^{us} g'(u)u = 1,2,\ldots,[v_{\max}/u] \tag{16}$$

where $\mathrm{AIRT_R(v)}$ denotes the average infrared radiation temperature of the reference specimen in the vth frame; g_v^L and g_v^R denote the infrared thermogram sequences of the rock loaded specimen and the reference specimen, respectively; and g_v' denotes the sequence of the temperature matrix of the rock infrared thermograms after denoising of the thermal background. u is the index of the frame after denoising of the multi-frame cumulative average, and s is the size of the window. In this paper, the acquisition rate of the thermal imaging camera is 30 frames/s. Considering the amount of data computation and the need for analysis, the cumulative window in this paper is set to 30, that is, the final ITF sequence per second is obtained. v denotes the matrix frame index, m, n are the rows and columns indexes, respectively, and k and l are the maximal rows and columns frame indexes, which are 200 and 100, respectively.

3.2.3. 3D spatiotemporal data construction

In the problem of infrared radiation temperature prediction on rock surface, the dimension of original input data cannot meet the input requirements of network model. The main reason is that the original input data cannot fully encode the spatiotemporal information of temperature field. In order to solve this problem, it is necessary to convert all inputs, unit outputs and unit states into three-

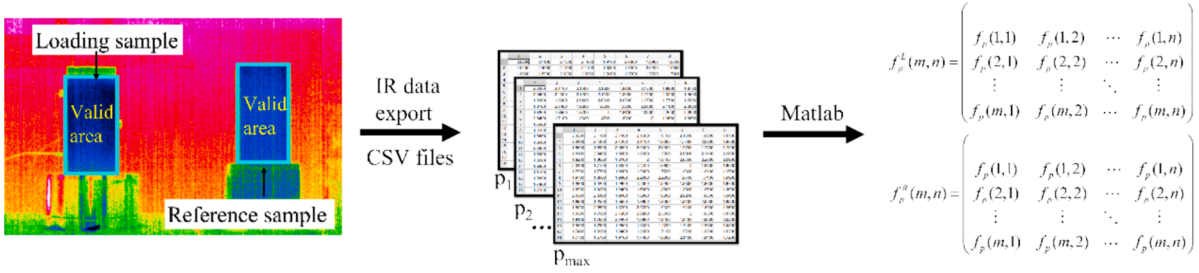


Fig. 6. Reconstruction of ITF of rock.

dimensional tensors $X \in U^{T\times W\times H}$, where U represents the characteristic domain of rock infrared temperature field data, and W and H represent the row and column spatial dimensions respectively. Since the rock loading time in this paper is relatively short, the time series data is segmented using overlapping sliding windows within the entire time series range of each rock sample to ensure the benchmark amount of training data, and the original single-frame infrared temperature matrix is converted into a continuous multi-frame time dimension vector in the spatial grid.

4. Experimental setup and results

In this section, we will conduct a series of experiments to evaluate the performance of the proposed model.

4.1. Dataset

In this paper, a total of 39 sandstone specimens were subjected to infrared radiation observation experiments during the bearing process. The data sets were randomly divided according to the ratio of 7:2:1 during model training. That is, the infrared radiation temperature field sequences of 27 sets of specimens were used as the training set, 7 sets as the validation set, and 5 sets as the test set. For each set of infrared radiation temperature field sequences in the training, validation and test sets, each frame of temperature field matrix is saved separately in chronological order. Then a sliding window is used to segmentally encode the ITF data of each set of specimens in chronological order. The final infrared temperature data set is formed. Where the size of the sliding window will be determined experimentally in 4.4.1.

4.2. Experimental setting

In this paper, we compare the proposed model with three other different models, namely 3DCNN, ConvLSTM[36] and the model without attention mechanism (EDDPNet). Where EDDPNet is the STAE-EDDPNet model in which the attention mechanism CBAM is removed and the initial input of the decoder is replaced by a zero matrix, and the other parameters are consistent with STAE-EDDPNet. The 3DCNN model consists of a three-layer 3DCNN layer as well as a two-layer fully connected layer. The number of output channels of the three 3DCNN layers are 32, 64, and 128, respectively. The number of ConvLSTM model layers is consistent with STAE-EDDPNet.

During model training and testing, the batch size was set to 8 and the OneCycleLR[53] learning rate tuning strategy was used with the maximum learning rate set to 0.003, as shown in Fig. 7. Adam stochastic gradient descent was also used to optimize the network model parameters. All models were implemented using pytorch framework as a backend and trained and tested on a single NVDIA GeForce RTX 4080 GPU.

4.3. Evaluation metrics

In order to evaluate the performance of the rock ITF prediction model, this paper adopts three indexes, namely, mean absolute error (MAE), root mean square error (RMSE) and Pearson's correlation coefficient (R), for a comprehensive quantitative evaluation of the predicted rock surface ITF errors. The definitions are as follows:

$$MAE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}|\hat{y}_t(i,j) - y_t(i,j)| \tag{17}$$

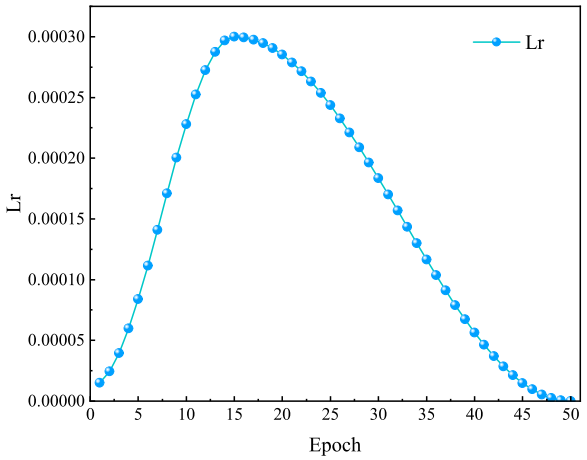


Fig. 7. Learning rate change curve.

$$RMSE = \sqrt{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[\hat{y}_t(i,j) - y_t(i,j)]^2} \tag{18}$$

$$R = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(\hat{y}_t(i,j) - \overline{\hat{y}_t})(y_t(i,j) - \overline{y}_t)}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(\hat{y}_t(i,j) - \overline{\hat{y}_t})^2}\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(y_t(i,j) - \overline{y}_t)^2}} \tag{19}$$

Where $\hat{y}_t$ denotes the predicted rock ITF at moment t. y_t denotes the real rock ITF at moment t. $\overline{\hat{y}_t}$ and $\overline{y}_t$ denote the temperature averages of the predicted and real temperature fields, respectively. m, n denote the rank indexes of the rock ITF, respectively.

4.4. Prediction results of infrared radiation temperature field of rock damage

4.4.1. Model layer number, hidden layer state number and sliding window

In this section, the single-step prediction strategy is adopted, and the following three types of experiments are carried out to determine the number of hidden layer states, the number of model layers and the sliding window in the model.

(1) **Number of hidden layer units:**the number of model layers is fixed to 3, and the number of hidden layer states of ConvLSTM units in STAE-EDDPNet is set to 80, 64, 48, 32, 24, 16, 12, and 8, respectively. The performance of different numbers of hidden layer states in the validation set is shown in Fig. 8. As can be seen from the table, the overall performance of the model in the validation set is as follows: as the number of hidden layer states goes from small to large, the model performance goes from strong to weak, and then becomes strong again. When the number of hidden layers is greater than 16, although the model indicators show that the model performance is gradually enhanced, however, the model has a weak overfitting phenomenon, which is manifested in the fact that the training error is decreasing, while the validation error is slightly increasing.

In addition, the size of the temporary tensor generated by the convolution operation during model computation directly affects the overall model storage space occupation. Too large number of hidden layer units will lead to GPU memory overflow. Therefore, taking into account the model's computational volume, space occupation computational efficiency and other issues, this paper sets the number of hidden layer states to 8.

(2) **Number of model layers:** based on the optimal number of hidden layer states from experiment (1), the number of hidden layer states is fixed to 8, and the number of model layers is set to 2, 3, 4, 5, and 6. The performance of different model layers in the validation set is shown in Fig. 9. The training time increases with the increase of the number of model layers, and the RMSE and MAE indexes are minimized when the number of model layers is 5. In th[6]is paper, the number of model layers is set to 5.

(3) **Sliding window:** Similarly, in order to determine whether repeated sliding Windows are used in the formation of data sets, according to the experimental results of (1) and (2), the number of model layers is 5, and the hidden layer unit is set to 8. Assuming the input step size is j, when s = j, it is a sliding window without repetition. When s < j, there are repeated sliding windows. The duplicate sliding window and the non-duplicate sliding window will affect the redundancy of input information and have a great impact on the model prediction performance.

In order to determine the optimal sliding window size for our dataset, we conducted a series of representative experiments.

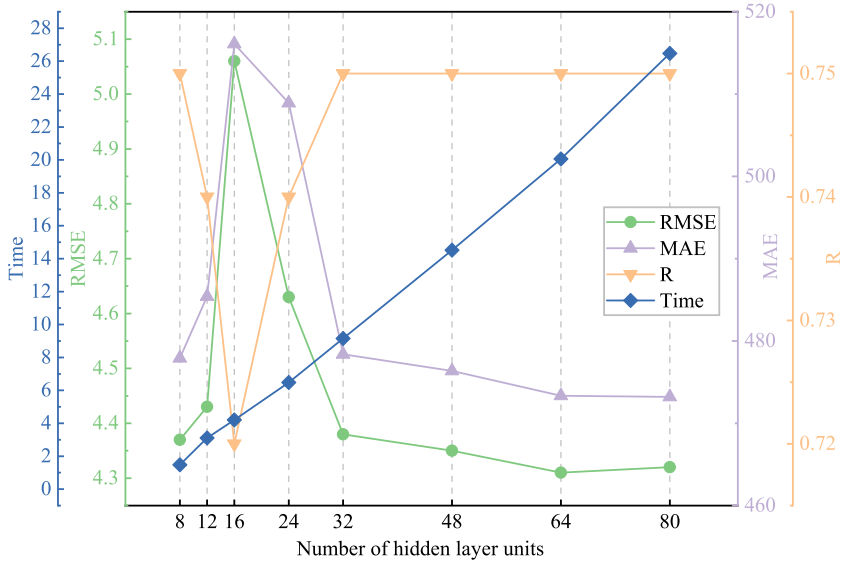


Fig. 8. Performance of the model in the validation set for different size hiding states.

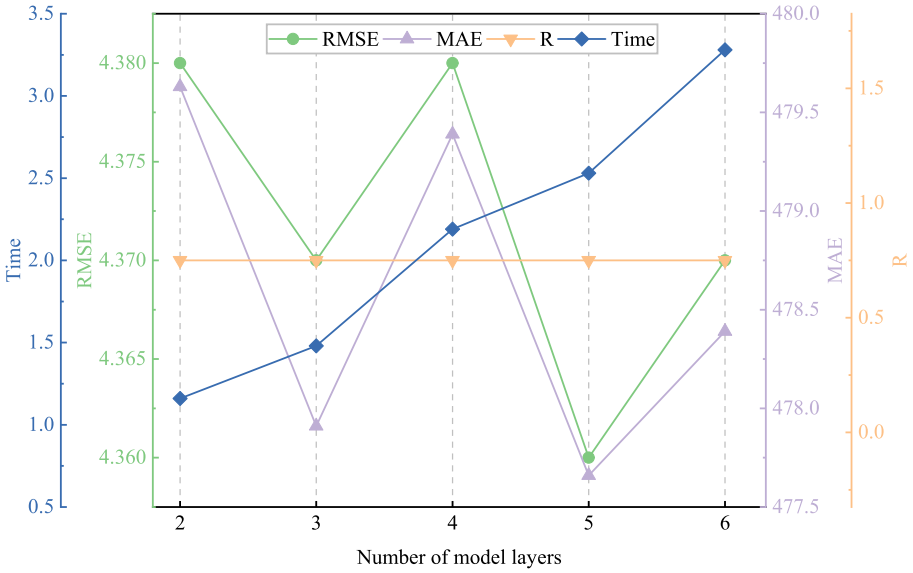


Fig. 9. Performance of models with different number of layers in the validation set.

Initially, with the input step size fixed at j = 5 and the prediction step size at 1, we evaluated the model's performance across datasets with varying sizes of sliding windows, as illustrated in Fig. 10. Observing the results, it becomes evident that increasing the sliding window size leads to reduced memory usage within the dataset, shorter training times, and higher computational efficiency. However, this comes at the expense of gradual degradation in model performance. Notably, when the sliding window size is set to 2, significant improvements in model performance are observed, while the training time is only about half that of a window size of 1. Therefore, in this study, we set the sliding window size to 2. This choice strikes a balance between computational cost and information redundancy, optimizing the overall performance of the model.

4.4.2. Predicting the spatio-temporal temperature field of rock damage in infrared

Space-time sequence prediction includes single-step and multi-step space–time sequence prediction. When j = 1 in Equation (1), it represents single-step prediction, i.e., predicting the next observation value based on historical space–time data. When j > 1, it represents multi-step prediction, i.e., predicting observation values for the coming period based on historical space–time data. In this paper, the model performance is evaluated from both the perspectives of single-step prediction and multi-step prediction.

4.4.2.1. Single-step prediction. This section evaluates the performance of STAE-EDDPNet and baseline models on single-step prediction

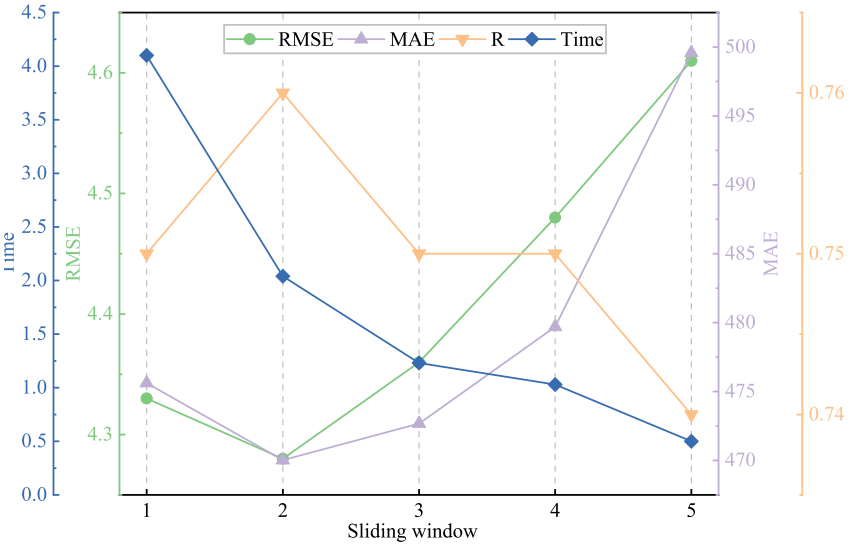


Fig. 10. Model performance in different sliding window data sets.

of rock damage status dataset. The performance of different models with varying input time step sizes is shown in Table 2.

According to the analysis of Table 2, increasing the input time step size has a dual impact on the model performance: the initial phase prediction accuracy and performance improve, but performance begins to decline after surpassing the optimal point. Meanwhile, training time and floating-point operations (FLOPs) show linear growth. Particularly, when the input time step size reaches 17, the STAE-EDDPNet model demonstrates the best performance in evaluation metrics such as RMSE, MAE, and R.

Furthermore, under the condition of an input step size of 17, we compared various baseline models. The results indicate that the performance of the 3DCNN model is the poorest, while ConvLSTM performs similarly to EDDPNet. It is noteworthy that the proposed STAE-EDDPNet model outperforms other baseline models in evaluation metrics such as RMSE, MAE, and R. Taking RMSE as an example, the model proposed in this paper improves the performance by 38.25 % over 3DCNN, 1.63 % over Covlstm, and 1.40 % over EDDPNet. This further highlights the significant role of the space–time attention mechanism module (CBAM) in model optimization.

4.4.2.2. Multi-step time prediction. In this study, we fixed the model's input time step size at 10 and examined the influence of different prediction time step sizes on the model's performance, as shown in Table 3. The consistent evaluation metrics of RMSE, MAE, and R indicate that as the prediction time step size increases, the model's performance gradually declines, primarily due to the increase in cumulative errors. Although the growth in prediction time step size leads to an increase in computation time and FLOPs, longer prediction steps are crucial for disaster early warning, especially in the progressive failure process of rocks. Despite the presence of error accumulation, longer prediction time step sizes still hold significant predictive value.

Under the condition of an input time step size of 10 and a prediction time step size of 5, the proposed STAE-EDDPNet model demonstrates significant performance advantages over other baseline models in multi-step prediction tasks. Taking the RMSE indicator as an example, the performance of STAE-EDDPNet is 29.18 % higher than that of 3DCNN, 7.12 % higher than that of ConvLSTM, and 0.41 % higher than that of EDDPNet. The 3DCNN model's performance on evaluation metrics is not satisfactory, and it consumes far more computational resources (FLOPs and training time) than other models, exhibiting lower efficiency. Compared to single-step prediction, the performance of the ConvLSTM model sharply declines in multi-step prediction. Further analysis reveals that STAE-EDDPNet surpasses its base model EDDPNet in evaluation metrics such as RMSE and MAE, thereby proving the effectiveness of the space–time attention mechanism module (CBAM) in enhancing prediction accuracy.

4.4.3. Spatial prediction accuracy assessment

To evaluate the spatial prediction accuracy of the model, we conducted testing on the test set and visualized the prediction results. The five test sets are samples A_4, A_{23}, A_{33}, A_{30}, and A_{39}. Table 4 displays the overall performance of the model on these test sets. The proposed model demonstrates the most outstanding performance across these test sets. Specifically, in terms of the key evaluation metric RMSE, the performance of STAE-EDDPNet is 25.56 % higher than that of the 3DCNN model, 5.69 % higher than ConvLSTM, and 0.19 % higher than EDDPNet, proving its superiority in spatio-temporal prediction accuracy.

Fig. 11 depicts in detail the stress curves of five rock samples in the test set along with the corresponding variations in average infrared radiation temperature (AIRT). Although the infrared radiation variations of each sample fluctuate over time, the overall trend shows temporal consistency. The variation in AIRT undergoes initial small-scale fluctuations, followed by accelerated growth, and ultimately exhibits a sudden change at the moment of rock rupture. This process is directly linked to the cumulative damage experienced by the rock under stress [17,24,54–56].

In this study, using sample A_{39} as an example, a qualitative analysis was conducted on the spatial prediction accuracy of different models for the ITF of rock damage at different loading times (three time points represented as a, b, and c in Fig. 11).

a) During the initial loading phase

In Fig. 12, we present a test instance (at time point a) during the initial loading phase before the peak load for sample A_{39}. During this loading stage, the rock undergoes initial compaction, where internal primary cracks and micropores are compacted, accompanied by partial frictional heating effects. The distribution of the ITF of rock damage is relatively uniform, with the damaged infrared temperature being relatively low. For this instance, the model receives 10 frames as input and predicts the subsequent 5 frames. In other words, it predicts the damage temperature field for the next 5 s based on the temperature field data from the previous 10 s. Upon

Table 2
The performance of models in single-step prediction under different input step sizes.

Models	RMSE (↓)	MAE (↓)	R (↑)	Time (min/epoch)	FLOPs/GB	Parameters/MB
STAE-EDDPNet (3–1)	4.33	475.73	0.75	1.26	1.81	28.86
STAE-EDDPNet (5–1)	4.28	470.02	0.76	2.04	2.72	28.86
STAE-EDDPNet (7–1)	4.27	469.83	0.76	2.46	3.63	28.85
STAE-EDDPNet (9–1)	4.31	471.78	0.76	3.22	4.53	28.86
STAE-EDDPNet (11–1)	4.28	468.66	0.76	4.02	5.44	28.86
STAE-EDDPNet (13–1)	4.26	466.43	0.76	4.43	6.34	28.86
STAE-EDDPNet (15–1)	4.24	464.54	0.76	5.23	7.25	28.86
STAE-EDDPNet (17–1)	**4.23**	**463.32**	**0.76**	**5.59**	**8.15**	**28.86**
STAE-EDDPNet (19–1)	4.27	467.73	0.76	6.34	9.06	28.86
STAE-EDDPNet (21–1)	4.28	469.01	0.76	7.05	9.96	28.86
3DCNN	6.85	761.63	0.25	4.26	10.60	94.31
ConvLSTM	4.30	470.81	0.76	5.32	7.70	14.43
EDDPNet	4.29	470.74	0.76	5.50	8.15	28.86

Table 3
The performance of the models in multi-step prediction with a fixed input time step size.

Models	RMSE (↓)	MAE (↓)	R (↑)	Time (min/epoch)	FLOPs/GB	Parameters/MB
STAE-EDDPNet (10–1)	**4.33**	**472.57**	**0.75**	**3.41**	**4.98**	**28.86**
STAE-EDDPNet (10–2)	4.54	494.05	0.74	4.07	5.45	28.86
STAE-EDDPNet (10–3)	4.63	505.26	0.72	4.26	5.89	28.86
STAE-EDDPNet (10–4)	4.74	517.80	0.71	4.52	6.34	28.86
STAE-EDDPNet (10–5)	4.83	528.86	0.70	5.10	6.79	28.86
STAE-EDDPNet (10–6)	4.91	538.89	0.69	5.34	7.25	28.86
STAE-EDDPNet (10–7)	5.07	552.22	0.68	5.54	7.70	28.86
STAE-EDDPNet (10–8)	5.07	556.60	0.67	6.12	8.15	28.86
STAE-EDDPNet (10–9)	5.14	563.69	0.67	6.37	8.60	28.86
STAE-EDDPNet (10–10)	5.16	567.23	0.66	6.58	9.06	28.86
3DCNN (10–5)	6.82	760.07	0.25	2.52	55.52	51.64
ConvLSTM (10–5)	5.20	573.17	0.66	3.32	4.53	14.43
EDDPNet (10–5)	4.85	531.69	0.70	5.03	6.79	28.86

Table 4
The model's performance in the multistep prediction in the test set.

Models	RMSE (↓)	MAE (↓)
3DCNN	7.12	779.00
ConvLSTM	5.62	599.43
EDDPNet	5.31	560.67
STAE-EDDPNet	**5.30**	**559.94**

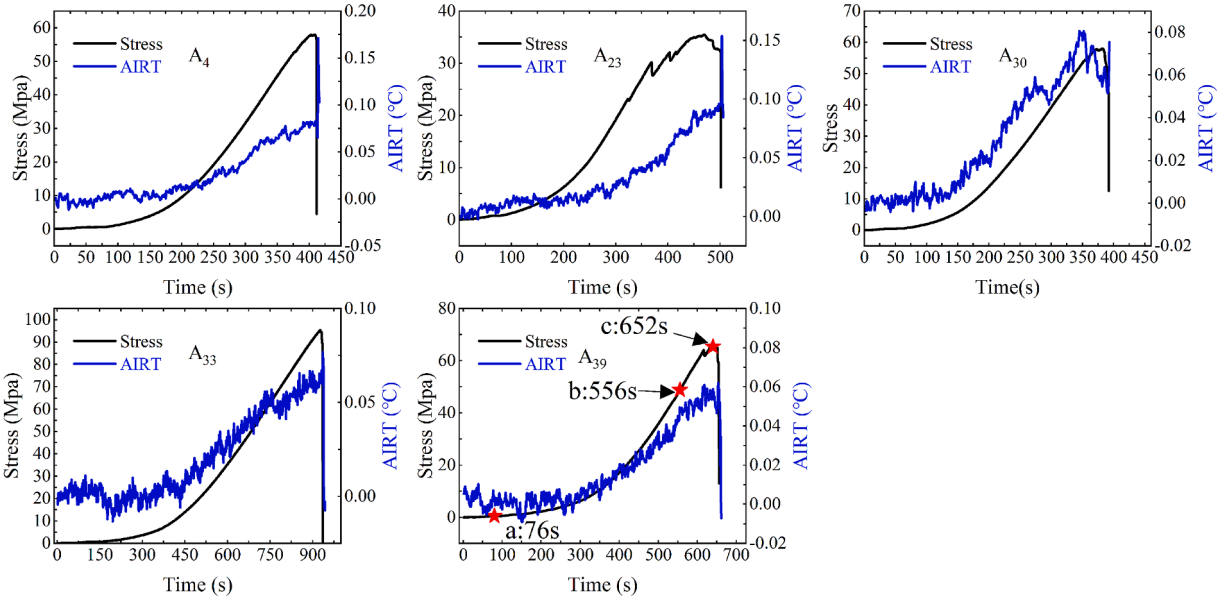


Fig. 11. The stress, AIRT, and time curves for each rock sample in the test set.

examining the prediction results, it's evident that the damage temperature field predicted by the 3DCNN model appears blurry with significant errors. Comparatively, ConvLSTM, EDDPNet, and STAE-EDDPNet exhibit better prediction performance. However, since the rock is in the initial loading phase and the ITF differentiation is not prominent, it's challenging to intuitively evaluate the performance of the three models from the graph. Therefore, using the RMSE metric as an example, the prediction accuracy of each model for each frame is evaluated.

Fig. 13 depicts the frame-by-frame evaluation metrics for the prediction instances at loading time points a and b for sample A_{39}. From panel (a) of the figure, it can be observed that the prediction accuracy of ConvLSTM is consistently lower than that of EDDPNet and STAE-EDDPNet for each frame. EDDPNet performs slightly worse than STAE-EDDPNet. The prediction indicators of the first to fifth frames in the initial loading phase show that the prediction accuracy of STAE-EDDPNet for the first frame is always greater than that of the subsequent frames, which is caused by error accumulation. On average, in the early stage of loading, the prediction accuracy of the STAE-EDDPNet model is 40.60 % higher than that of the 3DCNN model, 9.93 % higher than that of ConvLSTM, and 2.20 %

Fig. 12. The prediction results and error distribution plots of each model for the test instance corresponding to loading time point a of sample A_{39}.

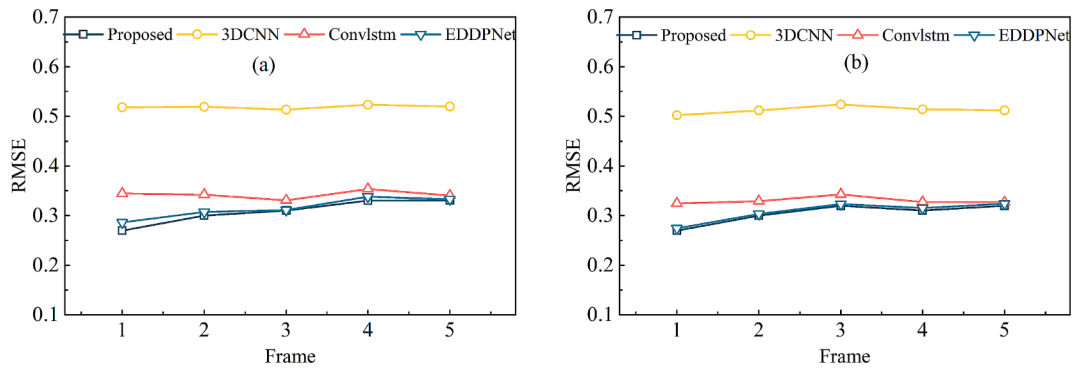


Fig. 13. The evaluation metrics for each frame of the prediction instances at loading time points a and b for sample A_{39}.

higher than that of EDDPNet.

b) During the mid-term loading phase

Fig. 14 illustrates a test instance of sample A_{39} during the mid to late stages of loading (loading time point b). From the input image in the figure, it can be observed that during the mid to late stages of loading, high-temperature regions appear in the ITF of rock damage. At this stage, influenced by the cumulative effects of thermoelasticity and frictional heating during the elastic deformation and plastic deformation stages, the surface temperature of the rock sample is higher than that during the initial loading, with the high-

Fig. 14. The prediction results and error distribution plots of each model for the test instance corresponding to loading time point b of sample A_{39}.

temperature regions located in the central part of the rock sample. The performance of each model is similar to that during the initial loading phase, with 3DCNN still unable to provide satisfactory prediction results. The performance of the other three models remains consistent with that during the initial loading phase. The same conclusion can also be drawn from Fig. 13 (b). Overall, in the mid-term loading phase, the advance prediction accuracy of the STAE-EDDPNet model is 40.71 % higher than that of the 3DCNN model, 7.92 % higher than that of ConvLSTM, and 1.32 % higher than that of EDDPNet.

c) During the final loading phase

Fig. 15 represents the last instance in the damage image sequence of sample A_{39}, which includes macroscopic cracks in the ITF of damage. At this loading time point, the rock sample gradually exhibits strain softening effects, with the load-bearing capacity of the sample gradually decreasing, entering the post-peak loading stage. Compared to the actual damage temperature field data, the prediction of the damage temperature field by the 3DCNN model is very poor, with not only blurry temperature field data but also many details not being reflected. Although the prediction results of the ConvLSTM model have improved, the errors are still significant. The prediction effects of EDDPNet and STAE-EDDPNet are similar, but the prediction results of STAE-EDDPNet exhibit more details. Additionally, it is noteworthy that macroscopic cracks appear in the real data at T = 13, 14, and 15, but these are not reflected in the effect images predicted by all models, as clearly shown in the error plots.

In Fig. 16, the frame-by-frame evaluation metrics for the prediction instance at loading time point c of sample A_{39} are presented. From Fig. 16, it can be observed that the performance of STAE-EDDPNet is consistent with what was observed in Fig. 15. None of the models can effectively extract the spatio-temporal evolution characteristics of macroscopic cracks, resulting in larger prediction errors as the prediction step increases. However, this cumulative error is not very evident in the prediction errors at loading time points a and b. This indicates that the models have weaker spatio-temporal feature extraction capabilities for the late loading stage. From the overall prediction results in the final loading phase, it is found that the advance prediction accuracy of the STAE-EDDPNet model is 40.49 % higher than that of the 3DCNN model, 26.58 % higher than that of ConvLSTM, and 1.28 % higher than that of EDDPNet.

This phenomenon is closely related to the dataset. In this study, when creating the dataset, the data was first subjected to multi-frame cumulative denoising and compressed to integer-second level. During the dataset creation process, the temperature field data of rock damage for the next 5 s was predicted based on the rock damage data for the previous 10 s, with a sliding window of 2. As a

Fig. 15. The prediction results and error distribution plots of each model for the test instance corresponding to loading time point c of sample A39.

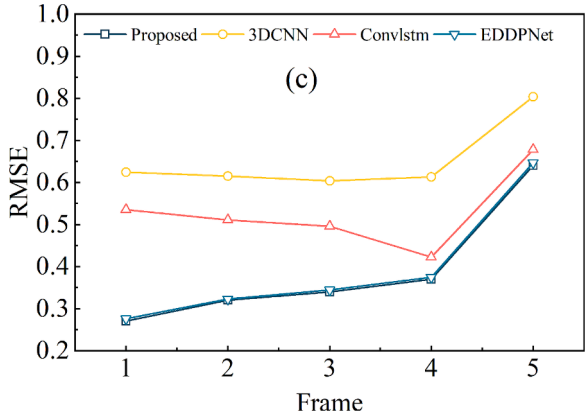


Fig. 16. The frame-by-frame evaluation metrics for the prediction instance at loading time point c of sample A_{39}.

result, the number of temperature field data containing macroscopic cracks in the input data to the model is relatively small because the time at which each macroscopic crack appears does not exceed 15 s. This can be explained by the stress curve in Fig. 11. From Fig. 11, it can be seen that the final rupture of the sample is a clear brittle failure. After the peak stress, the stress on the rock drops instantly, the rock loses its load-bearing capacity, macroscopic cracks appear on the surface, and the AIRT undergoes a sudden change. Therefore, these reasons lead to the model not effectively capturing the spatio-temporal evolution characteristics of macroscopic

cracks.

To further validate the above explanation, this study conducted error statistics on all instances in the test dataset. There are a total of 1440 instances in the test dataset, and each instance was predicted by the model, and the results were compared with the real dataset. Using RMSE as the evaluation metric, the results are shown in Fig. 17. It can be observed that the STAE-EDDPNet model exhibits a sudden increase in the prediction evaluation metrics for the late loading stage of each sample. This indicates that the proposed model has weak spatio-temporal feature extraction capabilities for the ITF of rocks in the late loading stage, while the prediction performance of the damage ITF in other loading stages is relatively stable. Therefore, if the prediction errors of rock damage in the post-peak deformation stage are not considered, the average prediction error of the proposed model will be significantly reduced, enabling effective extraction of the spatio-temporal features of rock damage ITF and advanced prediction.

5. Discussion

5.1. Ideas for modeling the prediction of ITF in rock damage

The infrared radiation temperature field sequence during the evolution of rock damage exhibits significant temporal and spatial correlations [52,57–60]. Temporal correlation is manifested in the dynamic changes over time in the infrared radiation energy transfer (heat transfer) received by the rock surface from internal voids and microcracks development (due to stress), mainly reflected in the trend of changes in the heat radiation information per unit time interval on the rock surface (as shown in Fig. 11). Spatial correlation is reflected in the differences in heat radiation across different areas of the rock surface, primarily related to the spatial positions of internal microcrack, microvoid expansion, and nucleation breakthrough in the load-bearing rock[57] (as indicated by the damage ITF inputs in Fig. 12, Fig. 14, and Fig. 15). Therefore, spatiotemporal advance prediction of the infrared radiation temperature field during the rock damage evolution process is theoretically feasible. Regarding the establishment of predictive models for the ITF on the rock surface, there are generally two approaches: theoretical physics-driven models and data-driven models.

Theoretical physics-driven modeling involves utilizing a series of physical equations associated with changes in the surface ITF during the evolution of rock damage. While this method is highly interpretable, the complexity, nonlinearity, and heterogeneity of rock damage make it challenging to establish physical equations for the infrared radiation changes during the damage evolution process. As a result, the physical mechanisms of infrared radiation changes during rock damage evolution are mostly qualitatively explained, making it difficult to establish physical equations for these changes. Therefore, this method is difficult to implement and requires further research into the physical mechanisms of infrared radiation changes during rock damage evolution.

Data-driven models, on the other hand, establish predictive models for the entire damage evolution process based on objective historical data of factors influencing rock damage and the infrared radiation during damage evolution. These models delve deep into large-scale datasets to discover complex patterns and trends hidden within, avoiding the need for complex internal mechanism derivation. This modeling approach, based on historical data and trends in rock infrared radiation, offers significant advantages in enhancing decision-making, reducing subjective biases, and deepening understanding of the problem.

5.2. The significance of advanced prediction of rock ITF

This study provides an innovative solution for the early identification of rock mass damage and disaster early warning, especially by utilizing the non-contact and real-time monitoring characteristics of infrared radiation technology, which offers effective technical support for safe production in geotechnical engineering. With the continuous development of geological disaster monitoring and early warning technologies, the application prospects of infrared radiation monitoring technology in practical engineering are becoming increasingly broad. By extracting the spatiotemporal features of infrared radiation temperature fields, this study can monitor the damage evolution of rock masses in real time and predict future damage development through data analysis. This provides an early warning mechanism to prevent the sudden occurrence of disasters, making it of significant engineering application value.

This technology has broad application scenarios in geotechnical engineering, not limited to tunnels and mines, but also extending to complex engineering environments, including:

- **Rock damage monitoring in underground tunnels and mines:** The stability of surrounding rock in tunnels and mines is critical to engineering safety. Infrared radiation technology can monitor real-time changes in the temperature field of tunnel or mine walls, detecting early signs of crack propagation caused by stress concentration, and providing early warnings of rock bursts or instability.
- **Monitoring the stability of coal pillars**: By monitoring the temperature changes in coal pillars with infrared radiation, it is possible to identify stress-concentrated damage areas in advance, preventing disasters caused by coal pillar instability.
- **Slope stability monitoring:** Slopes in open-pit mines or large excavations are prone to instability due to factors like rainfall and weathering. Infrared radiation monitoring technology can detect abnormal temperature fields on slopes in time, track the crack propagation process, and prevent landslides and collapses.
- **Health monitoring of dam structures:** Infrared radiation technology can be used to monitor the temperature field of dams, detecting leaks or crack propagation due to stress concentration or changes in water pressure. This provides early warning of dam instability and ensures reservoir safety.

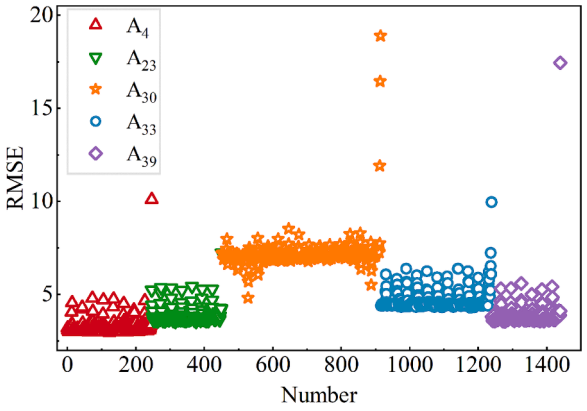


Fig. 17. Evaluation index distribution map of all test cases of each sample in the test set.

- **Monitoring underground reservoirs and energy storage facilities:** The long-term stability and sealing of rock are critical for underground reservoirs or energy storage facilities. Infrared radiation monitoring can capture signs of crack propagation and leakage in real time, ensuring the safety of these facilities.
- **Monitoring high-stress zones in deep mining:** High-stress zones in deep mines are prone to deformation and failure. By monitoring the temperature changes in these zones with infrared radiation, it is possible to predict damage and deformation in advance.
- **Rock damage monitoring induced by earthquakes:** In seismically active areas, infrared radiation technology can be used to monitor rock damage caused by earthquakes, providing post-earthquake assessments of rock stability and preventing secondary disasters.
- **Fracture monitoring during hydraulic fracturing:** In the hydraulic fracturing process of oil and gas extraction, infrared technology can monitor temperature changes during crack propagation, ensuring the safety and effectiveness of fracturing.

Through these wide-ranging applications, infrared radiation technology not only plays a crucial role in real-time monitoring but also predicts future disasters based on historical data, making it especially suitable for continuous monitoring of hazardous areas. Compared to other monitoring methods, infrared radiation monitoring offers significant advantages such as non-contact operation, high sensitivity, and wide applicability, playing a key role in early warning systems for safe production in complex geotechnical engineering environments. The widespread adoption of this technology can provide robust safety assurance for projects such as mines, tunnels, dams, and underground reservoirs.

5.3. Limitations and future research prospects

This study has preliminarily achieved advanced prediction of rock damage infrared temperature fields at the laboratory scale, with some positive results. However, there are still several limitations: First, the model's ability to extract spatiotemporal features of rock damage during the entire stress loading stage is not yet comprehensive, especially in handling complex crack evolution processes. To further improve model performance, future efforts will focus on optimizing the dataset through the collection of post-peak rock data and refining experimental designs to ensure the model can effectively capture damage features at all stages. On the other hand, from the prediction results of the early and middle loading stages, it is observed that as the prediction lead time increases, there is still a phenomenon of error accumulation. To address this, future research will incorporate adversarial neural networks and the fusion of multi-physics field features to further enhance the accuracy and robustness of advanced prediction.

Additionally, this study was conducted under a single rock type (such as a specific type of coal or sandstone) and a single loading path, without considering the complex conditions of different rock types, water content states, and loading paths. Different types of rocks, due to differences in their physical and mechanical properties, may exhibit different infrared radiation evolution characteristics, which limits the broad applicability of the model. Future research will expand to include experiments on multiple rock types, different water content states, and various loading paths to validate the universality and scientific validity of the advanced prediction method for infrared temperature fields. This will provide more reliable theoretical support and technical references for safety monitoring in production sites under complex stress environments and hydrogeological conditions.

In practical engineering applications, the implementation of the prediction model proposed in this study may face various challenges, such as environmental temperature fluctuations, sensor calibration and deployment, and differences in rock types. For instance, diurnal temperature variations and weather changes may introduce noise that affects the prediction accuracy of the model; sensors may experience data drift due to site conditions; and differences in rock composition and water content can significantly influence infrared radiation characteristics. Therefore, in future research, not only should data collection and model training processes be improved in laboratory settings, but algorithms for environmental temperature calibration, sensor drift correction techniques, and model adaptation methods based on transfer learning should also be developed to enhance the model's generalization ability and

applicability.

6. Conclusion

In this paper, the temporal and spatial characteristics of the ITF of rock damage are extracted and the single time step and multi-time step length are predicted. The main conclusions are as follows:

(1) A deep learning network, STAE-EDDPNet, was established for the first time for predicting the infrared temperature field of rock damage. Through experimental comparisons, the number of layers, the number of hidden units, and the sliding window size of the dataset were determined to be 5, 8, and 2, respectively. This model is a spatiotemporal attention-enhanced encoder-decoder prediction network, and the spatiotemporal attention mechanism module improved the network's predictive performance by 0.19 %.
(2) In both single-step and multi-step forecasting, the model proposed in this paper outperforms other baseline models in evaluation metrics such as RMSE and MAE. Taking RMSE as an example, in single-step forecasting, the model proposed in this paper achieved a 38.25 % improvement over 3DCNN, a 1.63 % improvement over ConvLSTM, and a 1.40 % improvement over EDDPNet. In multi-step forecasting, STAE-EDDPNet's performance was improved by 29.18 % compared to 3DCNN, by 7.12 % compared to ConvLSTM, and by 0.41 % compared to EDDPNet.
(3) In both single-step and multi-step forecasting, the model proposed in this paper outperforms other baseline models in evaluation metrics such as RMSE and MAE. Taking RMSE as an example, in single-step forecasting, the model proposed in this paper achieved a 38.25 % improvement over 3DCNN, a 1.63 % improvement over ConvLSTM, and a 1.40 % improvement over EDDPNet. In multi-step forecasting, STAE-EDDPNet's performance was improved by 29.18 % compared to 3DCNN, by 7.12 % compared to ConvLSTM, and by 0.41 % compared to EDDPNet.
(4) The dataset of rock damage ITF significantly influences the prediction results of damage. Due to the short loading time after the peak for brittle rock, there is limited information in the data regarding the post-peak evolution of rock damage. This results in a weaker ability of the model during the training phase to extract the spatiotemporal characteristics of the ITF during the post-peak loading stage, ultimately reducing the prediction performance of the ITF during the post-peak damage stage. In the future, it is necessary to further address these limitations through improvements in data collection, processing, and experimental design.
(5) The concept of advanced prediction of rock damage ITF based on data-driven theory was proposed for the first time, and the feasibility of this idea was verified. The study found that the application of infrared radiation monitoring technology in the advanced prediction of rock damage is feasible, which has a driving effect on the application of infrared monitoring technology in the fields of mining and rock engineering safety monitoring.

CRediT authorship contribution statement

Qiangqiang Gao: Writing – original draft, Validation, Software, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Liqiang Ma:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Wei Liu:** Writing – review & editing, Validation, Software, Formal analysis, Data curation. **Naseer Muhammad Khan:** Writing – review & editing, Validation, Formal analysis. **Xiuzhe Wang:** Writing – review & editing, Visualization, Investigation, Formal analysis. **Yanxiao Ni:** Writing – review & editing. **Kunpeng Yu:** Funding acquisition, Writing – review & editing. **Saad S. Alarifi:** Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This paper was supported by the Natural Science Foundation of China (52464015), the National Key Research and Development Program (2023YFC2907501), the Xinjiang Uygur Autonomous Region Key R&D Program Projects (2023B03009), the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX23_2811), the Graduate Innovation Program of China University of Mining and Technology (2023WLKXJ046), the National Natural Science Foundation of China (grant number 52404105). Also, this work was supported by the Researchers Supporting Project (RSP2025R496), King Saud University, Riyadh, Saudi Arabia.

Data availability

Data will be made available on request.

References

[1] Bobet A, Arson CF, Elsworth D, Nelson P, Tomac I, Modiriasari A. The role of rock mechanics in the 21st century. Geotechnical Fundamentals for Addressing New World. Challenges 2019:319–57.
[2] Meng F, Wong LNY, Zhou H. Rock brittleness indices and their applications to different fields of rock engineering: a review. J Rock Mech Geotech Eng 2021;13:221–47.
[3] Feng Q, Jin J, Zhang S, Liu W, Yang X, Li W. Study on a damage model and uniaxial compression simulation method of frozen–thawed rock. Rock Mech Rock Eng 2022;55:187–211.
[4] Shirole D, Hedayat A, Walton G. Damage monitoring in rock specimens with pre-existing flaws by non-linear ultrasonic waves and digital image correlation. Int J Rock Mech Min Sci 2021;142:104758.
[5] Zhang Y, Zhao G-F, Li Q. Acoustic emission uncovers thermal damage evolution of rock. Int J Rock Mech Min Sci 2020;132:104388.
[6] Zhao K, Ma H, Liang X, Li X, Liu Y, Cai R, et al. Damage evaluation of rock salt under multilevel cyclic loading with constant stress intervals using AE monitoring and CT scanning. J Pet Sci Engng 2022;208:109517.
[7] Jia P, Li L. Liu D-q, Wang X-s, Wang D-c. Insight into rock crack propagation from resistivity and ultrasonic wave variation. Theor Appl Fract Mech 2020;109:102758.
[8] Du Y, Xie M, Jiang Y, Chen C, Jia B, Huo L. Review on the formation mechanism and early warning of rock collapse. Metal Mine 2021;535:106–19.
[9] Du Y, Lu Y, Xie M, Jia J. A new attempt for early warning of unstable rocks based on vibration parameters. Bull Eng Geol Environ 2020;79:4363–8.
[10] Du Y, Xie M-w, Jiang Y-j, Li B, Chicas S. Experimental rock stability assessment using the frozen–thawing test. Rock Mech Rock Eng 2017;50:1049–53.
[11] Salami Y, Dano C, Hicher PY. Infrared thermography of rock fracture. Géotechnique Letters 2017;7:1–5.
[12] Sheinin VI, Blokhin DI, Maksimovich IB, Sarana EP. Experimental research into thermomechanical effects at linear and nonlinear deformation stages in rock salt specimens under cyclic loading. J Min Sci 2017;52.
[13] Wang CL, Lu ZJ, Liu L, Chuai XS, Lu H. Predicting points of the infrared precursor for limestone failure under uniaxial compression. Int J Rock Mech Min Sci 2016;88:34–43.
[14] Minh PL. Infrared thermovision of damage processes in concrete and rock. Eng Fract Mech 1990;35:291–301.
[15] Gao Q, Ma L, Liu W, Wang H, Ma Q, Wang X. Identification of damage states of load-bearing rocks using infrared radiation monitoring methods. Measurement 2025;239:115507.
[16] Zhao Y, Liu B, Ling C, Gao Y, Wang X, Danesh NN. Infrared radiation and acoustic emission characteristics of sandstone with different granularities under uniaxial compression. Bull Eng Geol Environ 2023;82.
[17] Liu W, Ma L, Gao Q, Spearing AJS, Wang Y, Cui R, et al. Fracture precursor recognition and damage quantitative characterization of stressed rock using infrared radiation. Rock Mech Rock Eng 2023.
[18] Franzosi F, Casiraghi S, Colombo R, Crippa C, Agliardi F. Quantitative Evaluation of the Fracturing State of Crystalline Rocks Using Infrared Thermography. 2023.
[19] Hao T, Li F, Tang Y, Zhao L, Wang Z. Infrared precursor of pre-cracked coal failure based on critical slowing down. Geomatics Nat Hazards Risk 2022;13:1682–99.
[20] Tian H, Li Z, Shen X, Zang Z, Song J, Zhang Q. Identification method of infrared radiation precursor information of coal sample failure and instability under uniaxial compression. Infrared Phys Technol 2021;119.
[21] Wu LX, Liu SJ, Wu YH, Wang CY. Precursors for rock fracturing and failure—Part II: IRR T-Curve abnormalities. Int J Rock Mech Min Sci 2006;43:483–93.
[22] Ma LQ, Sun H, Zhang Y, Zhou T, Li K, Guo JS. Characteristics of infrared radiation of coal specimens under uniaxial loading. Rock Mech Rock Eng 2016;49:1567–72.
[23] Shi WZ, Wu YH, Wu LX. Quantitative analysis of the projectile impact on rock using infrared thermography. Int J Impact Eng 2007;34:990–1002.
[24] Zhang K, Liu X, Chen Y, Cheng H. Quantitative description of infrared radiation characteristics of preflawed sandstone during fracturing process. J Rock Mech Geotech Eng 2021;13:131–42.
[25] Ma L, Zhang Y, Cao K, Wang Z. An experimental study on infrared radiation characteristics of sandstone samples under uniaxial loading. Rock Mech Rock Eng 2019;52:3493–500.
[26] Liqiang MA, Zhang D, Guo X, Sun H, Najeem A, Zhang Y, et al. Characteristics on the variance of differential infrared image sequence during coal failures under uniaxial loading. Chin J Rock Mech Engng 2017.
[27] Gao Q, Ma L, Liu W, Cao K, Khan NM, Zhao Z. Appraisal of rock IR law and damage precursor: insight infrared thermogram entropy. Infrared Phys Technol 2022;104443.
[28] Liu W, Ma LQ, Sun H, Khan NM. Using the characteristics of infrared radiation b-value during the rock fracture process to offer a precursor for serious failure. Infrared Phys Technol 2021;114:103644.
[29] Cao K, Ma L, Zhang D, Lai X, Zhang Z, Khan NM. An experimental study of infrared radiation characteristics of sandstone in dilatancy process. Int J Rock Mech Min Sci 2020;136:104503.
[30] Shan T, Li Z, Zhang X, Niu Y, Tian H, Zhang Q, et al. Infrared radiation and acoustic emission of damage evolution and failure precursory for water-bearing coal. Rock Mech Rock Eng 2022.
[31] Liu S, Wei J, Huang J, Wu L, Zhang Y, Tian B. Quantitative analysis methods of infrared radiation temperature field variation in rock loading process. Chin J Rock Mech Eng 2015;34.
[32] Huang J, Song L, Yu M, Zhang C, Li S, Li Z, et al. Quantitative spatial analysis of thermal infrared radiation temperature fields by the standard deviational ellipse method for the uniaxial loading of sandstone. Infrared Phys Technol 2022:104150.
[33] Hao T, Li F, Tang Y, Wang Z, Zhao L. Infrared Radiation Characterization of Damaged Coal Rupture Based on Stress Distribution and Energy. ACS Omega 2022;7:28545–55.
[34] Zhang K, He F, Zhang Z, Lin X, Li M. Graph attention temporal convolutional network for traffic speed forecasting on road networks. Transportmetrica B: Transport Dynamics 2020;9:153–71.
[35] Shi X, Gao Z, Lausen L, Wang H, Yeung D-Y, Wong W-k, et al. Deep learning for precipitation nowcasting: A benchmark and a new model. Adv Neural Inf Proces Syst 2017;:30.
[36] Shi X, Chen Z, Wang H, Yeung D-Y, Wong W-K, Woo W-c. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Advances in neural information processing systems. 2015;28.
[37] Wang X, Qin MJ, Zhang Z, Wang YY, Du ZH, Wang N. Typhoon cloud image prediction based on enhanced multi-scale deep neural network. Front Mar Sci 2023;9.
[38] Rüttgers M, Lee S, Jeon S, You D. Prediction of a typhoon track using a generative adversarial network and satellite images. Sci Rep 2019;9:6057.
[39] Villar-Corrales A, Karapetyan A, Boltres A, Behnke S. MSPred: Video Prediction at Multiple Spatio-Temporal Scales with Hierarchical Recurrent Networks. arXiv preprint arXiv:220309303. 2022.
[40] Zhou Y, Dong H, El Saddik A. Deep learning in next-frame prediction: A benchmark review. IEEE Access 2020;8:69273–83.
[41] Fan H, Zhu L, Yang Y. Cubic LSTMs for Video Prediction. Proceedings of the AAAI Conference on Artificial Intelligence 2019;33:8263–70.
[42] Gao Z, Tan C, Wu L, Simvp LSZ. Simpler yet better video prediction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2022. p. 3170–80.
[43] Wang Y, Gao Z, Long M, Wang J, Yu PS. PredRNN plus plus : Towards A Resolution of the Deep-in-Time Dilemma in Spatiotemporal Predictive Learning. 35th International Conference on Machine Learning (ICML). Stockholm, SWEDEN2018.
[44] Wang Y, Jiang L, Yang MH, Li LJ, Long M, Fei-Fei L. Eidetic 3D LSTM: A Model for Video Prediction and Beyond. International Conference on Learning Representations2019.

[45] Wang Y, Long M, Wang J, Gao Z, Yu PS. PredRNN: Recurrent Neural Networks for Predictive Learning using Spatiotemporal LSTMs. 31st Annual Conference on Neural Information Processing Systems (NIPS). Long Beach, CA2017.
[46] Wang Y, Zhang J, Zhu H, Long M, Wang J, Yu PS. Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition2019. p. 9154-62.
[47] Xiao C, Chen N, Hu C, Wang K, Xu Z, Cai Y, et al. A spatiotemporal deep learning model for sea surface temperature field prediction using time-series satellite data. Environ Model Softw 2019;120:104502.
[48] Yao S, He Y, Zhang LF, Yang W, Chen Y, Sun Q, et al. A ConvLSTM Neural Network Model for Spatiotemporal Prediction of Mining Area Surface Deformation Based on SBAS-InSAR Monitoring Data. IEEE Trans Geosci Remote Sens 2023;61.
[49] Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput 2019;31:1235–70.
[50] Woo S, Park J, Lee J-Y, Kweon IS. Cbam: Convolutional block attention module. Proceedings of the European conference on computer vision (ECCV)2018. p. 3-19.
[51] Sun H, Ma LQ, Adeleke N, Zhang Y. Background thermal noise correction methodology for average infrared radiation temperature of coal under uniaxial loading. Infrared Phys Technol 2017;81:157–65.
[52] Liu W, Ma LQ, Sun H, Khan NM. An experimental study on infrared radiation and acoustic emission characteristics during crack evolution process of loading rock. Infrared Phys Technol 2021;118:103864.
[53] Smith LN, Topin N. Super-Convergence: Very Fast Training of Neural Networks Using Large Learning Rates. Conference on Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications. Baltimore, MD2019.
[54] Li X, Bu J, Yang Z, Li H, Zhou J, Liu Z, et al. The law of infrared radiation generation and evolution of coal rock fracture under load. Infrared Phys Technol 2024:105338.
[55] Sun H, Zhu H-Y, Han J, Fu C, Chen M-M, Wang K. Energy and infrared radiation characteristics of the sandstone damage evolution process. Materials 2023;16.
[56] Hao T, Yuan M, Li F, Wang G. Quantitative analysis method for infrared characterization of coal damage under load based on image enhancement and pixel extraction. Geomatics Nat Hazards Risk 2023;14:2272574.
[57] Sun H, Liu XL, Zhang SG, Nawnit K. Experimental investigation of acoustic emission and infrared radiation thermography of dynamic fracturing process of hard-rock pillar in extremely steep and thick coal seams. Eng Fract Mech 2020;226.
[58] Wu LX, Liu SJ, Wu YH, Wang CY. Precursors for rock fracturing and failure - Part I: IRR image abnormalities. Int J Rock Mech Min Sci 2006;43:473–82.
[59] Yin S, Li Z, Wang E, Li X, Tian H, Niu Y. Infrared temperature evolution law and thermal effect mechanism of concrete impact failure. Journal of Building Engineering 2024;91:109592.
[60] Liu W, Ma L, Jaboyedoff M, Derron M-H, Gao Q, Bu F, et al. Extraction of the key infrared radiation temperature features concerning stress and crack evolution of loaded rocks. Int J Min. Sci Technol 2024.