International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 7450-7463

https://doi.org/10.70517/ijhsa463639

Classification and identification of spring steel tempered flexural body based on the combination of deep residual network and convolutional neural network

Ziqiang Luo[1], Jiajun Fu[1], Zhili Hu[1], Yicheng Gong[2,*] and Zhilong Luo[3]
[1] School of Mechanical Engineering, Wuhan University of Science and Technology, Wuhan, Hubei, 430081, China
[2] Hubei Province Key Laboratory of Systems Science in Metallurgical Process, Wuhan University of Science and Technology, Wuhan, Hubei, 430065, China
[3] Shiyan Aotie Thermal Engineering Technology, Shiyan, Hubei, 442001, China

Corresponding authors: (e-mail: so1711111198@163.com).

Abstract As a key material for mechanical manufacturing, the microstructure of spring steel is closely related to its mechanical properties. Traditional identification methods have limited accuracy and are difficult to efficiently and accurately classify and identify microstructures such as tempered flexural, which restricts the optimization of spring steel properties. In this study, a deep residual network and a convolutional neural network are combined to construct a classification and identification model of spring steel tempered flexural microstructure. By establishing the SS-3000 dataset containing 5 classes of tissues and 600 images of each class, a migration learning strategy was used to optimize the model training process and validate the model performance on the TEST-1000 dataset. In the experimental design, images were normalized to a fixed interval and normalized, and hyperparameter settings such as batch size 32, learning rate 0.0005 and 500 training cycles were used. The results show that the average recognition accuracy of the SE-Resnext101 metallographic tissue recognition model based on migration learning for various types of microstructures of spring steel reaches 98.2%, and the recognition accuracy and recall of tempered flexural reaches 97.82% and 99.04%, respectively, which is significantly better than that of the traditional GLCM+SVM algorithm and other deep learning models. In the comparison experiments, the recognition accuracy of the model for tempered flexural is 5.69%, 7.19% and 5.43% higher than that of VGG19, AlexNet-TL and MobileNetV2-TL, respectively. The study confirms that the combination of deep residual network and convolutional neural network can effectively extract the microstructure features of spring steel, which provides reliable technical support for the study of the relationship between material properties and microstructure.

Index Terms Spring steel, Tempered flexural, Deep residual network, Convolutional neural network, Migration learning, Image recognition

I. Introduction

In recent years, with the rapid development of the economy, China vigorously develops the manufacturing industry, making the iron and steel industry become a pillar industry in China, which has a lasting and long-term impact on the national economy and life. Steel has a wide range of applications, and in many fields, especially in the material has an irreplaceable role [1]. Among them, spring steel is the main material for manufacturing springs, and its more varieties of alloying elements are also very complex, and there are more than a dozen grades that are often used [2], [3]. As the working environment of spring steel is quite harsh, long-term work in the cyclic bending, torsion and other alternating stress conditions, but also often subjected to tensile, compression, impact and fatigue corrosion and other roles, the performance requirements are extremely stringent, but also sometimes have to withstand a very high level of sudden load in a short period of time [4]-[7]. Therefore, all mechanical properties of springs must meet high requirements, including high ductility and toughness, as well as good hardening ability [8], [9]. At the same time, spring steel also needs to have good hardenability, not prone to decarburization, and more excellent surface quality, which is required in the process of spring steel [10], [11].

The performance of spring steel is obtained through a reasonable heat treatment process, the high elastic limit and toughness of tempered quartz can meet the performance requirements of the spring [12]. To obtain the tempered martensitic organization must be obtained by cooling the material at a certain rate after complete austenitization, and then tempering treatment at a certain temperature to obtain tempered martensite [13]-[15]. Therefore, one of the ways to develop high-performance spring steels is to optimize the heat treatment parameters to further improve the strength, plasticity, and fatigue resistance of spring steels with the existing spring steel elements and compositions [16]-[18]. Based on this, identifying the spring steel tempered quenchantite class to

determine good quenching production parameters and excellent tempering production parameters has an important impact and significance on the development of spring steel in China [19], [20].

As an indispensable functional material in the mechanical industry, spring steel is widely used in transportation, industrial equipment and national defense construction and other fields. In the actual service process, spring steel is subjected to long-term tensile, compression, torsion and other complex stresses, which puts forward harsh requirements on its mechanical properties. The microstructure of the material is to determine the basis of its macroscopic properties, accurate identification of the microstructure characteristics of spring steel for an in-depth understanding of its performance mechanism, optimize the heat treatment process and predict the service behavior of the material is of great significance. The traditional microstructure identification of materials mainly relies on the experience of professionals, which is not only inefficient, but also subjective and uncertain. Especially for the complex organization morphology such as tempered quartzite in spring steel, its characteristic subtle differences are difficult to be accurately distinguished by the naked eye, which restricts the systematic study of the relationship between spring steel properties and organization. In recent years, computer vision and deep learning technologies have made important breakthroughs in the field of materials science, providing new ideas for automated and intelligent identification of material microstructures. Deep residual network, with its powerful feature extraction capability and the advantage of solving the degradation problem of deep network, combined with the excellent performance of convolutional neural network in the field of image processing, lays the technical foundation for the accurate identification of spring steel microstructure.

Based on this background, this study proposes the combination of deep residual network and convolutional neural network for the classification and identification of spring steel microstructures such as tempered flexural. Firstly, the microstructure images of spring steel in different heat treatment states are collected by scanning electron microscope, and a dataset containing five types of tissues, namely, ferrite/pearlitic, martensite, martensite/bainite, martensite/residual austenite, and tempered dendrite, is established; secondly, the SE-Resnext101 metallographic tissue recognition model based on migration learning is designed to make full use of the general feature extraction capability of the pre-trained model in order to adapt to the special needs of material microstructure recognition; finally, the proposed method is comprehensively evaluated by comparing with traditional methods and other deep learning models to verify the effectiveness and superiority of the method in spring steel microstructure recognition by comparing the performance indexes such as recognition precision, recall rate and F1 score.

II. Overview of spring steel

II. A.2.1 Spring steel

Spring as an important part of the machinery industry, is widely used in transportation, industrial equipment and national defense construction and other fields, mainly connecting, shock absorption buffer, sealing, guiding and other roles. Spring steel as the manufacture of springs and a variety of elastic components of the material, in the service process for a long time by the tensile, bending and impact loads, and the working environment is poor. Therefore, the comprehensive performance of spring steel has high requirements [21]-[23].

II. A. 1) Classification of spring steel

(1) Classification by production and processing methods

According to the production and processing methods, spring steel can be divided into hot-rolled spring steel and cold-drawn spring steel. Hot rolled spring steel can be annealed after hot rolling to obtain good fatigue resistance, commonly used in large-size parts that require good plastic toughness. Cold-drawn spring steel can withstand higher loads, commonly used in small cross-sectional area and a larger force on the manufacture of small-sized structural parts.

(2) According to the external size classification

According to the size classification, spring steel can be divided into flat steel, wire, bar and steel wire, etc.. Flat steel is divided into right-angle flat steel and rounded flat steel. Flat steel width between 50-130mm, thickness between 6-60mm, mainly used in the production of automotive steel plate spring. Wire diameter of 5-40mm, bar diameter of 12-90mm, steel wire diameter less than 5mm.

(3) Classification by chemical composition

Classified by chemical composition, spring steel can be divided into carbon spring steel and alloy spring steel. Carbon spring steel carbon content in 0.6-0.9%, typical representative grades are 65, 70 and 85. Although the carbon spring steel has high strength, but the hardenability and decarburization resistance is poor, so it is generally used in the manufacture of agricultural machinery parts that do not require high performance.

Alloy spring steel is in the carbon steel based on the addition of Cr, Ni, Nb, V and Mo and other alloying elements of spring steel, its carbon content of 0.4 ~ 0.9%. Compared with carbon spring steel, alloy spring steel has more excellent comprehensive performance.

II. A. 2) Spring steel property requirements

Spring steel in service due to long-term torsion, tension and impact and other stresses, and the service environment is relatively harsh. Therefore, spring steel is required to have excellent overall performance, such as hardenability, decarburization properties, mechanical properties, anti-hydrogen delayed fracture properties and fatigue properties.

(1) Hardenability

Spring steel hardenability is the spring steel after quenching to obtain the ability of martensite, the advantages and disadvantages of hardenability determines the maximum size of the spring can be produced. When the spring steel hardenability is good, after quenching and tempering heat treatment of spring steel organization on the whole is more uniform, so that it has a better overall performance.

(2) Decarburization properties

Decarburization is steel in the heat processing and heat treatment process, the carbon elements in the steel and the environment of the atmosphere reaction, resulting in the steel surface of the carbon content of the phenomenon of reduction according to the decarburized layer of morphology, decarburization can be divided into full decarburization and partial decarburization. Decarburization of the spring steel surface will reduce the hardness and quality of the surface, resulting in a decrease in the fatigue life of spring steel. Therefore, in order to avoid the adverse effects of decarburization on the fatigue performance of spring products, the study of the decarburization properties of spring steel is particularly important.

(3) Mechanical properties

With the demand of modern society for the development of the automobile industry, in order to ensure the safety, stability and reliability of spring steel in service, spring steel is required to have a high design stress. Therefore, spring steel should have high strength, good plastic toughness and high flexural strength ratio.

(4) Anti-hydrogen delayed fracture performance

Hydrogen delayed fracture refers to the role of hydrogen, steel is subjected to less than the limit of its strength of the stress, the phenomenon of sudden fracture without warning. In the research and development of high-strength spring steel, in order to improve the anti-hydrogen delayed fracture performance of spring steel, attention should be paid to the study of its hydrogen delayed fracture.

(5) Fatigue resistance

Fatigue failure is one of the main failure modes of spring steel, anti-fatigue performance is an important indicator of the evaluation of the merits of spring steel □. Studies have shown that non-metallic inclusions in spring steel is one of the important reasons that affect fatigue performance and cause fatigue failure of spring steel.

II. B.Overview of spring steel heat treatment processes

Through heat treatment, the crystal structure of spring steel can be changed to refine its grain, thus increasing its hardness and strength. This can increase the bearing capacity and wear resistance of spring steel, so that it is not easy to deformation or fracture in the working process. Spring steel needs to have a certain degree of toughness to ensure that the impact of external forces will not occur when the brittle fracture. Through heat treatment, you can adjust the organizational structure of spring steel, so that it has good toughness and impact resistance. Spring steel is mainly used to withstand deformation force, in the process of long time use need to have good elasticity and fatigue resistance, heat treatment can change the crystal structure of spring steel, so that it has better elasticity and fatigue resistance.

II. C.Factors affecting the heat treatment process of spring steel

(1) heating temperature and holding time

The choice of heating temperature directly affects the organization and properties of spring steel. Too low a heating temperature may lead to the spring steel can not reach the required phase transition temperature, and thus can not obtain the ideal organizational structure. While too high a heating temperature is likely to cause grain growth, grain boundary diffusion and precipitates dissolved, resulting in a reduction in the hardness and strength of spring steel. Holding time is another important influence in the heat treatment process. Holding time is to reach the desired heating temperature, the material will remain at the temperature of the time. The length of the holding time directly affects the degree of completion of the tissue transformation and phase change in spring steel. Too short holding time may lead to the organization structure is not completely transformed, thus affecting the performance of spring steel, while too long holding time may cause grain growth and excessive phase transformation, resulting in a reduction in the hardness and strength of spring steel.

(2) cooling rate and cooling medium

The choice of cooling rate directly affects the organization and properties of spring steel. Rapid cooling can make the spring steel grain refinement, increase the hardness and strength of the material. This is because rapid cooling can prevent the growth of the grain and the line of phase transition, thus forming fine grains and high strength of the organizational structure. On the contrary, slow cooling will lead to grain growth and phase transformation of the line, so that the hardness and strength of spring steel reduced.

The choice of cooling medium directly affects the organization and performance of spring steel. Commonly used cooling media include water, oil and gas, etc. Different cooling media have different cooling rates and heat transfer characteristics, which have different effects on the organization and properties of spring steel. For example, water cooling rate is fast, can obtain high hardness and high strength organization. Oil cooling rate moderate, can obtain moderate hardness and strength.

(3) Tempering temperature and time

Tempering temperature refers to the quenched spring steel heated to a certain temperature after cooling process, the choice of tempering temperature directly affects the hardness and toughness of spring steel.

Usually, the higher the tempering temperature, the lower the hardness of spring steel, the higher the toughness. This is because the high temperature tempering can make the spring steel in the residual stress is released, grain recrystallization and growth, thus forming a larger grain and lower hardness. On the contrary, low temperature tempering can make the spring steel to maintain a higher hardness, but lower toughness. The length of tempering time also has an effect on the properties of spring steel. Typically, the longer the tempering time, the lower the hardness and higher the toughness of the spring steel. This is because a long time tempering can make the spring steel in the residual stress is more fully released, the grain to get a greater degree of recrystallization and growth, thus forming a larger grain and lower hardness. Conversely, a short tempering period allows the spring steel to maintain a higher hardness but a lower toughness.

II. D.Hardening and Tempering Processes and Isothermal Quenching Processes

The difference between the conventional quench and temper process and isothermal quenching is that the conventional quench and temper process cools to room temperature during quenching, giving martensite in the quenched state. Whereas in isothermal quenching process, it is not quenched to room temperature and less martensite is obtained. Although the few martensite obtained are mostly large angular grain boundaries, which are not favorable for the strength of the steel. However, large-angle grain boundaries can hinder crack extension during tensile impact and improve toughness and plasticity to some extent.

Under the traditional process, the spring steel organization changes according to the tempering temperature. Usually medium-temperature tempered quartzite is the product of martensite tempered at a medium temperature of about 400 ℃, tempered quartzite has a high elastic limit, yield strength and fatigue strength, and has a certain degree of toughness. The nature of its organization is composed of carbide organization and ferrite mechanical mixture. In order to distinguish between the flexural phase obtained from the transformation of supercooled austenite, it is called tempered flexite.

In this paper, the microstructure and mechanical properties of 51CrMnV steel are studied comparatively for the conventional quenching and tempering heat treatment process and isothermal quenching and heat treatment process. By comparing the organization and properties with those of the traditional process, it provides a reference for the application of the lower bainite organization in the heat treatment of springs. In order to better compare with the reality of the process, to exclude more interference from external factors, the test steel used in this section in the finished product of the leaf spring sampling. Conventional heat treatment process schematic shown in Figure 1.

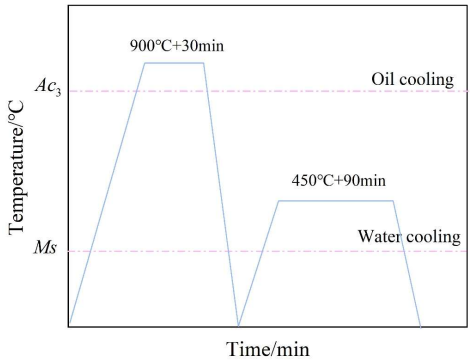

Figure 1: Traditional heat treatment process schematic

By observing the metallographic organization of 51CrMnV after quenching and tempering heat treatment, it can be seen as a single tempered flexural (T), and these organizations are tempered flexural. After tempering treatment, the organization is obviously refined, the tempered quartz body is uniformly small, and the slat feature of the tempered quartz body is obvious. After high temperature quenching temperature martensite, medium temperature tempering, tempering can not eliminate the martensite morphology characteristics. Tempered flexural, as a tempering product, has a martensite-like lath-like organization.

Mechanical properties of 51CrMnV spring steel with bainite and tempered quartzite organizations were studied. Microstructural characterization showed that the lower bainite and tempered quartzite differed in microstructural features, including slat size, dislocation structure and density, and size and distribution of carbide precipitates.

51CrMnV spring steel was treated with tensile test and impact test to determine the strength and toughness index of the steel, and the test results are shown in Table 1.

Table 1: The effect of heat treatment technology on steel ability

Processing process	Tensile strength(MPa)	Elongation (%)	Impact work (J)	Strong plastic product(GPa)
Quenched and tempered	1385.6	9.89	23.5	14.653
Austempering	1467.2	14.02	32.1	20.311

The stress-strain curves under different processes are shown in Fig. 2. Compared with the traditional quenching and tempering heat treatment process, the tensile strength, elongation, and impact work of 51CrMnV spring steel are improved after isothermal quenching and heat treatment, indicating that 51CrMnV spring steel has higher strength, hardness, and plasticity after isothermal quenching and heat treatment. Provide analytical support for the subsequent classification and identification of spring steel microstructure.

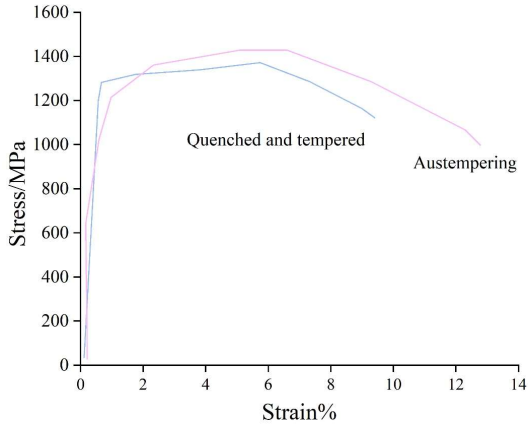

Figure 2: Stress - strain curve under different processes

III. Identification results of metallographic organization classification of spring steel

III. A. Deep learning based classification and recognition algorithm

III. A. 1) Deep residual networks

Deep residual network (ResNet) is an efficient deep learning model, which is important for solving the problem of training difficulty encountered by deep neural networks in classification tasks.

Ⅰ Forward Computation

This subsection focuses on the six main components involved in the forward computation process in a residual network: the input layer, the convolutional layer, the pooling layer, the base residual block, the global average pooling, and the Softmax classifier, as well as the associated computations involved.

The input layer of the neural network is the first layer of the neural network and the entrance to the whole neural network. It is mainly responsible for receiving and processing the input data and characterizing the raw data for further processing in the subsequent layers of the neural network.

The main role of the convolutional layer is to extract the features of the image and to solve practical problems by learning the features extracted from the image. The formula for convolution operation is given below:

$$output_{i,j} = \sum_{k=0}^{n}\sum_{l=0}^{n} input_{k,l} * \ker nel_{i-k,j-l} \tag{1}$$

where $input_{k,l}$ denotes the pixel value of the k th row and l th column of the input image. $\ker nel_{i-k,j-l}$ denotes the weights of the $i-k$ rows and $j-l$ columns of the convolution kernel. n denotes the size of the convolution kernel. $output_{i,j}$ denotes the output of the convolutional layer, which represents the result of the convolution operation for the i th row and j th column of the input image.

The formula for the convolutional layer is:

$$y(i,j,l) = \mathrm{sum}(x(i+p, j+q, k) * w(p,q,k,l)) + b(l) \tag{2}$$

where the sum operation is a cumulative summation over all k dimensions within the convolution kernel size, and p, q are the sliding steps of the convolution kernel. After computing the convolution, an activation function is also typically used to nonlinearly transform the output of the convolution.

The core operation of the pooling layer is to downsample the output of the convolutional layer and select a representative feature value from the output of the convolutional layer, common pooling operations are maximum pooling and average pooling. For maximum pooling, the maximum value within a custom size range is selected as the representative value. For average pooling, the average value within the customized size range is taken. The formula is expressed as follows: for maximum pooling, let $x_{i,j}$ denote an element in the input matrix, then the result of maximum pooling $y_{i,j}$ is:

$$y_{i,j} = \max(x_{i,j}, x_{i,j+1}, \cdots, x_{i+s-1,j+s-1}) \tag{3}$$

where s is the size of the pooling window. For average pooling, let $x_{i,j}$ denote an element in the input matrix, then the result of average pooling $y_{i,j}$ is:

$$y_{i,j} = \frac{1}{s^2} \sum_{k=0}^{s-1} \sum_{l=0}^{s-1} x_{i+k,j+l} \tag{4}$$

The pooling layer is part of a convolutional neural network and is responsible for reducing the spatial size of the input feature maps and decreasing the size of the feature maps while preserving key features. However, since some of the information may be lost during the pooling layer, the appropriate pooling operation needs to be selected based on the specifics of the problem.

The base residual block is the basic structural unit of deep residual networks, and its main role is to solve the problem that deep neural networks are difficult to train.

Softmax classifier is a commonly used classifier in deep learning, which is mainly used for multi-classification problems. The formula is as follows:

$$p_j = \frac{e^{a_j}}{\sum_{k=1}^{K} e^{a_k}} \tag{5}$$

where p_j denotes the probability that the sample belongs to the j th category, a_j denotes the score of the predicted j th category for the sample, and K denotes the total number of categories.

In order to train a Softmax classifier, a loss function is needed to measure the prediction error. The commonly used loss function is the cross-entropy loss with the following formula:

$$L = -\sum_{j=1}^{K} y_j \log p_j \tag{6}$$

where y_j is the value of the j th dimension of the true label, usually 0 or 1. The smaller the value of the cross-entropy loss, the smaller the gap between the classifier's prediction and the true label.

II Loss Function

In deep learning, the loss function is a metric to measure the difference between the model's predicted value and the true value, and it is also the key to optimize the objective function. The formula for calculating cross entropy is as follows:

$$CE(y, \hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{7}$$

where y denotes the true label, $\hat{y}$ denotes the label probability distribution predicted by the model, and C denotes the number of categories.

III Backpropagation

Backpropagation is one of the most important optimization algorithms in the CNN training process, which is mainly used to calculate the gradient of the neural network so as to achieve the optimal update of the model parameters.

Assume that L is the loss function and w is the parameters of the network, including weights and biases. a is the activation value in the network, and z is the weighted input value of the network. N is the number of neurons in the network and l is the number of layers in the network, excluding the input layer.

(1) Backpropagation from the output layer to the hidden layer

The relationship between the l th layer and the $l-1$ th layer in the process of forward propagation is:

$$a_i^{(l)} = \sigma\left(\sum_{j=1}^{N^{(l-1)}} w_{i,j}^{(l)} a_j^{(l-1)} + b_i^l\right) \tag{8}$$

where $w_{i,j}^{(l)}$ denotes the weight of the j th neuron in the $l-1$ th layer to the i th neuron in the l th layer, b_i^l denotes the bias of the i th neuron in the l th layer, and $\sigma(\cdot)$ denotes the activation function.

For the i th neuron of the output layer, the gradient of its corresponding loss function is:

$$\delta_i^{(L)} = \frac{\partial L}{\partial z_i^{(L)}} \tag{9}$$

According to the chain rule, the gradient of its corresponding previous layer is:

$$\delta^{(l)} = \frac{\partial l}{\partial z^{(l)}} = \left(\sum_{j=1}^{N^{l+1}} w_{j,i}^{(l+1)} \delta_j^{(l+1)}\right) \sigma'\left(z_i^{(l)}\right) \tag{10}$$

where $\sigma'(\cdot)$ denotes the derivative of the activation function, the commonly used activation functions are sigmoid function, ReLU function, tanh function, etc., and their derivatives can be obtained by some basic calculus.

For the i th neuron in the l th layer, the gradient of its weights is:

$$\frac{\partial L}{\partial w_{i,j}^{(l)}} = a_j^{l-1} \delta_i^{(l)} \tag{11}$$

The gradient of the bias is:

$$\frac{\partial L}{\partial b_i^{(l)}} = \delta_i^{(l)} \tag{12}$$

(2) Backpropagation from hidden layer to input layer

The backpropagation from the hidden layer to the input layer is similar to the backpropagation from the output layer to the hidden layer, which is calculated by the formula:

$$\delta^{(l)} = \frac{\partial l}{\partial z^{(l)}} = \left(\sum_{j=1}^{N^{l+1}} w_{j,i}^{(l+1)} \delta_j^{(l+1)}\right) \sigma'\left(z_i^{(l)}\right) \tag{13}$$

In this case, the gradient formula for weights and bias is also the same as the backpropagation from the output layer to the hidden layer.

Finally, the parameters of the network are updated by optimization algorithms such as the gradient descent method to minimize the loss function of the network.

III. A. 2) Convolutional Neural Networks

CNN is a special deep learning network for spatial data such as images, consisting of a stack of convolutional, pooling, fully connected and classifier layers. And it adopts the structure of weight sharing, which significantly reduces the number of training parameters and decreases the complexity of the model. Recursive gradient computation is performed using the differential chaining law (also known as the forward propagation algorithm), and the parameters are adjusted by gradient descent to achieve effective training of the entire model.

The structure of a typical convolutional neural network AlexNet model includes an input layer: the input is a pixel 3D matrix of the image, the length and width of the matrix indicate the size of the image, and the depth of the matrix indicates the color channel of the image, which is 3 for color images and 1 for grayscale images.

Convolutional layer: the input image is convolved with a linear convolutional filter to perform convolutional computation as shown in equation (14):

$$(h_k)_{ij} = (W_k * x) + b_k \tag{14}$$

where k is the index of the k th feature map in the convolutional layer. (i, j) is the index of the neuron in the k th feature map, and x denotes the input data. W_k and b_k are the trainable parameters (weights) of the linear filter and the bias of the neuron in the k th feature map, respectively. $(h_k)_{ij}$ denotes the output value of the neuron with position (i, j) in the k th feature map. The spatial 2-D convolution operation between the input data and the feature atlas is denoted by *. The convolutional layer analyzes each small region of the input image in depth, and the matrix of nodes processed by the convolutional layer becomes deeper, resulting in features with a higher level of abstraction.

Pooling layer: a nonlinear downsampling layer that takes the maximum or average value of each sub-region of the input data. Maximum pooling is often used to achieve small shift invariance of the feature maps.

Fully connected layer: Same as normal neural network layer, the layers are connected linearly as shown in equation (15):

$$y_k = \sum_l W_{k_l x_l} + b_k \tag{15}$$

where y_k denotes the output of the k th neuron. l_{x_l} is the output computed by convolution and pooling. W_{k_l} is the k_l th weight between x_l and y_k.

Classifier Layer (Output Layer): the last layer of the network which is used to calculate the probability that the input image belongs to a particular class.The most widely used classifier in CNN is the softmax function. This function generates a real-valued vector between $(0,1)$ that represents the probability distribution that the input vector X belongs to the j th class, as shown in equation (16):

$$p(y = j \mid x; W, b) = \frac{\exp(X^T W_j)}{\sum_{k=i}^{k} \exp(X^T W_j)} \tag{16}$$

Loss function: to calculate the difference between the real category labels and the predicted category labels, the cross-entropy loss function is commonly used, as shown in equation (17):

$$L = -\sum_x P'(x) \log P(x) \tag{17}$$

where $P(x)$ is the minimum value between the unique heat coding and the predicted classification probability calculated by the softmax classifier.

The gradient descent method is utilized to update the weights at each step of the backpropagation process so that the loss function is gradually reduced, and the update method is shown in equation (18):

$$W_{ij}^{new} = W_{ij}^{old} - \eta \frac{\delta L}{\delta W_{ij}} \tag{18}$$

where η denotes the learning rate. δ denotes the gradient. Dropout method is used in the training process to randomly ignore or discard some neurons in the network from the network structure during the training phase to reduce the overfitting of the CNN model and improve the generalization ability of the model.

III. A. 3) Deep learning based metallographic organization recognition of spring steel

(1) Three important steps are required to study the recognition of metallographic organization of spring steel based on convolutional neural network:

STEP1: The spring steel metallographic tissue image dataset needs to be established to complete the training and validation of the model.

STEP2: The size of the spring steel metallographic tissue images in the acquired dataset is not fixed in size, while the input image size of the constructed model must be fixed, so the data preprocessing of the images fed into the model must be performed to obtain a standardized image dataset.

STEP3: Convolutional neural network based on deep learning in the field of image recognition has been proposed many classical models and algorithms, on this basis, the study to establish a model and method based on the convolutional neural network of the spring steel metallurgical tissue recognition, comparative analysis of the performance of the modeled model, to determine the optimal recognition model.

(2) The convolutional neural network model takes the metallographic tissue image of spring steel as input, and the output result is the probability value that the metallographic image belongs to a certain category, and the recognition result of the model is the category with the highest probability value. The process of convolutional neural network modeling is shown in Figure 3.

(3) Based on the convolutional neural network spring steel metallographic training modeling, the images are preprocessed as follows when input to the convolutional neural network:

The spring steel metallographic image dataset is preprocessed before input to the neural network by normalizing the image to the interval $[0,1]$, denoted as:

$$f(x) = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{19}$$

In the formula, $f(x)$ is the pixel value of a point x on the image, and $x_{\min}$ and $x_{\max}$ represent the minimum and maximum values of the interval to which x belongs.

The normalization process makes the dataset image distributions more similar, which can reduce the training time of the model, and finally the image is subjected to a mean reduction divided by standard deviation operation. I.e:

$$f(x) = \frac{x - mean}{std} \tag{20}$$

where $mean$ and std are the mean and standard deviation values of the image, respectively, in Eq.

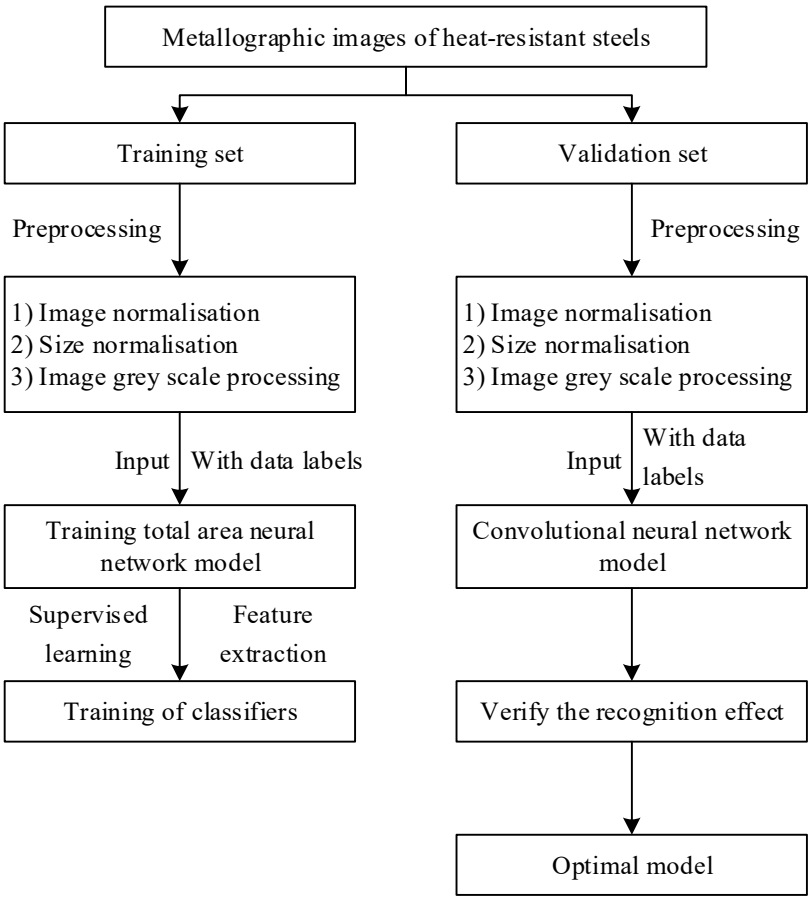

Figure 3: The process of the convolution neural network modeling

(4) Model hyperparameter settings

Batch Size (Batch Size), set the size of Batch Size to 32.

Learning Rate (Learning Rate), set Learning Rate to 0.0005.

Training Cycle Count (Epoch), set Epoch to 500.

Optimizer: Adams, Adaptive Learning Rate with wide adaptation. Input image size: 224×224.

(5) SE-Resnext101 metallographic tissue recognition model based on migration learning

The SE module is divided into a compression phase and an excitation phase, which can adaptively emphasize the most useful feature channels, thus reducing feature noise and improving model performance. The compression phase compresses the height and width (H×W) of the feature matrix to 1 through a global average pooling operation, which reduces the input feature map to a channel vector with global information. The excitation phase utilizes two fully connected layers to learn the weights of each channel. First the first fully connected layer acts as a dimensionality reduction to reduce the feature dimension to 1/4, after activation using the ReLU function. Secondly the feature dimensions are reduced through a fully connected layer, thus enabling the network to automatically scale meaningful features and suppress useless information and noise. Again the weight coefficients are normalized through the High-Sigmoid function, and finally the feature weighting operation, i.e., the new residual structure, is implemented by applying the weights from the output of the High-Sigmoid activation to each channel of the original feature map using Scale.

ResNeXt50 network model structure: first is the input and initial convolutional layer, where the network receives the input image and passes it through a common convolution, followed immediately by a maximum pooling layer, which is mainly used to roughly extract the features of the image and reduce the spatial dimensionality of the feature map. This is immediately followed by successively stacked residual modules Block1, Block2, Block3, and Block4, which are cyclically stacked 3, 4, 6, and 3 times, respectively. Finally, a global mean pooling layer is used, which

outputs each channel as a single mean, and a fully connected layer, which implements the identification or prediction required for the final task.The ResNeXt101 and ResNeXt50 networks are essentially similar in structure, except that the number of times that the four residual modules are stacked in a loop is deepened to 3, 4, 23, and 3 times, which deepens the network's layers from 50 to 101, and the goal is to deepen the network's layers from 50 to 101. The purpose of deepening the network from 50 to 101 is to express more complex features for better results.

For deep learning models, some low-level features such as points, lines, shapes, and texture information extracted at shallow layers are generalizable across different recognition tasks. Migration learning refers to the reapplication of weights trained in one task to another new task, and is used to improve learning by transferring generalized knowledge from a related task that has already been learned to a new task.

In deep learning models, the models are usually trained on the large dataset ImageNet (14 million images, 1,000 categories), allowing the network model to fully extract generic feature information. Using the migration learning method, the generic feature information is effectively migrated to small sample datasets, thus enhancing the performance of the model on small sample datasets.

III. B. Performance analysis

III. B. 1) Data set establishment

Compared with EBSD and TEM, SEM has become the most commonly used method for researchers to observe the microstructure of materials due to the advantages of simple sample preparation and large field of view.

In order to use the experimentally obtained SEM pictures for the study of the relationship between the microstructure of spring steel and its mechanical properties, the first thing to be done is to categorize and organize the microstructure pictures, and link them with the processing technology, processing conditions (pressure, temperature, holding time, etc.), temperature curve and mechanical properties curve of the material. Currently have collected all kinds of microstructure images totaling more than 600, mainly including spring steel forming process may occur in five types of organizations, ferrite / pearlite (F + P), martensite (M), martensite / bainite (M + B), martensite / residual austenite (M + A) and tempered quartzite. Material microstructure images are essential for the identification, classification, and quantitative characterization of spring steel microstructure, which in turn is of great significance for the study of the relationship between its microstructure and spring steel mechanical properties.

Accumulation of SEM images from experiments related to the spring steel forming process. These images are mainly of two resolutions, 1024×768 and 1024×943, with a magnification of 1000 times, and are derived from the SEMs of the university's sub-measurement center and the external cooperation unit, respectively.

Segmentation of SEM images is required in order to build a dataset of microscopic tissue images that can be used for feature extraction, machine learning and deep learning. Segmentation of the images and elimination of the non-compliant images from them resulted in the SS-3000 dataset: 5 classes, 600 images per class, and 224 × 224 in size per image.

At the end of the experiment, in order to explore the recognition accuracy of the model for each class of microtissue, a test set TEST-1000 was established for testing. The test set images are derived from the gallery established above, but are taken from outside the SS-3000 dataset, so they are processed in the same way as the SS-3000 dataset, all of which are 224×224 in size, with 200 sheets in each class, for a total of 1,000 sheets.

III. B. 2) Experimental results and analysis

The training results of SE-Resnext101 metallographic tissue recognition model based on migration learning are shown in Fig. 4. The training results of SE-Resnext101 metallographic tissue recognition model based on migration learning on the dataset SS-3000 are shown.

From the figure, it can be seen that the accuracy of the model is very close in the training set and validation set, which is basically higher than 96%, with better classification accuracy.

The change of loss function during the training process of SE-Resnext101 metallographic tissue recognition model based on migration learning is shown in Fig. 5.

It can be seen from the figure that in the first 10 iterations, the accuracy rate rises rapidly, and then gradually stabilizes, and finally stabilizes near 98%. The value of the loss function decreases rapidly and gradually stabilizes around 0.1.

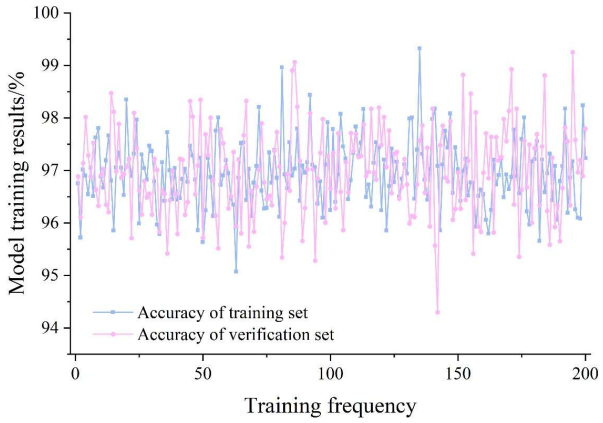

Figure 4: Identification model training results

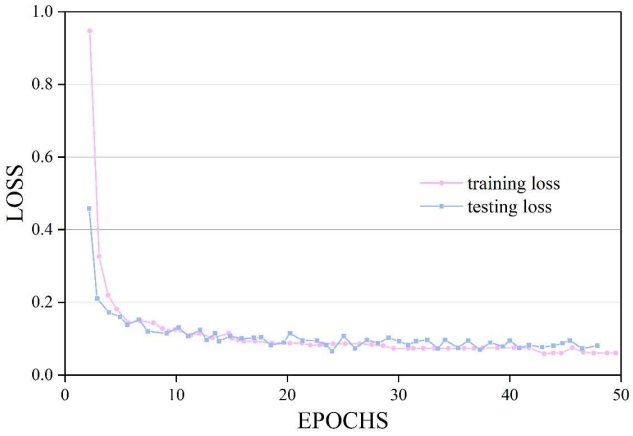

Figure 5: The loss function of the model training process

In order to explore the effectiveness of the SE-Resnext101 metallographic tissue recognition model based on transfer learning for each type of tissue, the tissue category prediction was performed for each image on the TEST-1000 dataset. The model prediction results are shown in Figure 6. The comparison algorithms include the GLCM+SVM algorithm (Gray level covariance matrix (GLCM) is a method for calculating the frequency of pixel pairs with the same gray level in an image), the SE-Resnext101 model, and the SE-Resnext50 model based on migration learning.

For the identification results of spring steel tempered flexural, the migration learning-based SE-Resnext101 metallographic organization identification model proposed in this paper is 5.83%, 3.74%, and 2.13% more than the SE-Resnext101 model, the migration learning-based SE-Resnext50 model, and the GLCM+SVM algorithm, respectively. Model training using migration learning can further improve the feature extraction of the network model for microscopic tissues.

The SE-Resnext101 metallographic tissue identification model based on migration learning has a mean value of 98.2% for the identification of various types of microstructures of spring steel.

For network model comparison, the migration learning-based AlexNet network, the migration learning-based MobileNetV2 network and the VGG19 network model (the names of the three methods are abbreviated as AlexNet-TL, MobileNetV2-TL and VGG19, respectively, in the following) were used before reclassifying the metallograms, comparing the other network models with the paper's methods in terms of accuracy, precision, recall, F1 score and convergence speed.

The inverted residual structure and linear bottleneck structure introduced by MobileNetV2 network make the network classification accuracy significantly higher compared to MobileNetV1, and MobileNetV2 is a well-known lightweight network that can be compared with the present method. The VGG19 network model consists of 16 convolutional layers and 3 fully connected layers, both of which use 3×3 convolutional kernels and 2×2 maximum pooling, the VGG19 model is also commonly used for image recognition and is representative.

The recognition accuracies of each model on the metallographic map of spring steel are shown in Fig. 7. For the recognition accuracies of each model on the tempered flexural of spring steel, the recognition accuracies of the

VGG19, AlexNet-TL, and MobileNetV2-TL models are 92.13%, 90.63%, and 92.39%, which are lower than that of the proposed model of this paper of 97.82%, respectively.

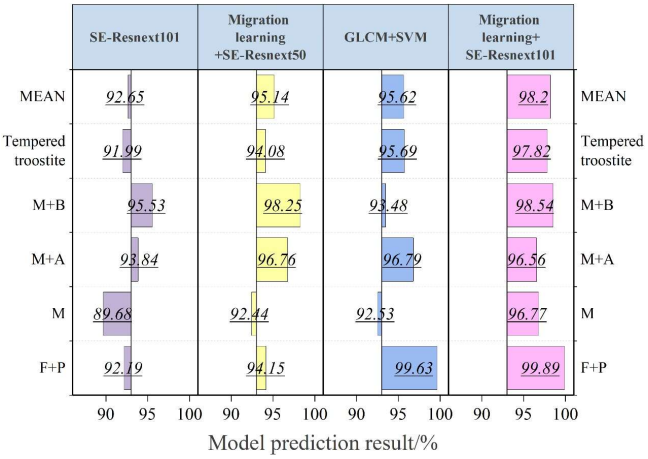

Figure 6: The prediction of the identification model

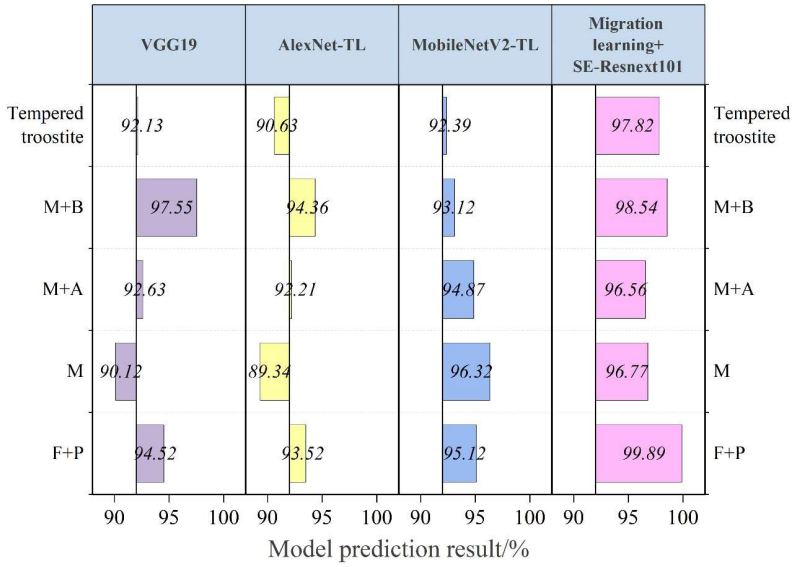

Figure 7: The exact rate of identification of the spring steel metallographic diagram

The recognition recall of each model for spring steel metallographs is shown in Table 2.The AlexNet-TL model has a low recall for all types of spring steel tissues. The SE-Resnext101 metallographic tissue recognition model based on migration learning generally has a recall rate higher than 93%, and the recognition recall rate for spring steel tempered flexural is as high as 99.04%.

Table 2: Recognition of the recall of the spring steel metallographic diagram

	VGG19/%	AlexNet-TL/%	MobileNetV2-TL/%	Migration learning+SE-Resnext101/%
F+P	91.25	85.63	90.11	92.53
M	88.36	86.25	89.42	94.25
M+A	98.63	97.21	94.56	99.21
M+B	96.52	84.25	92.36	99.16
Tempered troostite	98.25	94.36	93.28	99.04

The F1 scores of each model for the metallographic map of spring steel are shown in Table 3. Comprehensively comparing the recognition accuracy, recall, and F1 score of each model for the metallographic map of spring steel, the SE-Resnext101 metallographic tissue recognition model based on migration learning is the better method among the four models.

Table 3: F1 points for the metal phase diagram of the spring steel

	VGG19/%	AlexNet-TL/%	MobileNetV2-TL/%	Migration learning+SE-Resnext101/%
F+P	0.912	0.905	0.992	0.953
M	0.875	0.867	0.973	0.966
M+A	0.904	0.933	0.981	0.994
M+B	0.915	0.25	0.867	0.983
Tempered troostite	0.866	0.934	0.953	0.989

IV. Conclusion

A high-precision classification and recognition model for spring steel microstructures has been successfully constructed through a systematic study of the combination method of deep residual network and convolutional neural network. The SE-Resnext101 metallographic tissue recognition model based on migration learning achieves an average recognition accuracy of 98.2% for various types of spring steel microstructures on the TEST-1000 dataset, and exhibits a precision rate of 97.82% and a recall rate of 99.04% for tempered flexural tissues, with an F1 score of 0.989. The experimental comparison reveals that the model's recognition accuracy of spring steel tempered flexural is 5.83% higher than that of the traditional GLCM+SVM algorithm, and 3.74% and 2.13% more than that of the SE-Resnext101 model and the SE-Resnext50 model based on migration learning, respectively. In the comparative analysis of heat treatment processes, the 51CrMnV spring steel treated by isothermal quenching process exhibits a tensile strength of 1467.2 MPa and an elongation of 14.02%, which is 81.6 MPa and 4.13 percentage points higher than that of the traditional quenching and tempering process, respectively. The combination of deep residual network and convolutional neural network for the identification of spring steel microstructure lays the foundation for the accurate evaluation of material properties by effectively extracting the tissue features. This identification system is valuable for promoting the digital and intelligent characterization of spring steel microstructure, and provides technical support for the optimization of spring steel heat treatment process and performance improvement.

References

[1] Song, L. (2023). China: steel industry. In Encyclopedia of Mineral and Energy Policy (pp. 142-152). Berlin, Heidelberg: Springer Berlin Heidelberg.

[2] Moller, A. B., Tulupov, O. N., & Fedoseev, S. A. (2017). IMPROVEMENT OF SURFACE QUALITY OF ROLLED SECTION STEEL FOR SPRINGS MANUFACTURING. Journal of Chemical Technology & Metallurgy, 52(4).

[3] Hauserova, D., Dlouhy, J., & Kotous, J. (2017). Structure refinement of spring steel 51CrV4 after accelerated spheroidisation. Archives of Metallurgy and Materials, 62(3), 1473-1477.

[4] Tang, H. Y., Wang, Y., Wu, T., Li, J. S., & Yang, S. F. (2017). Characteristics analysis of inclusion of 60Si2Mn–Cr spring steel via experiments and thermodynamic calculations. Ironmaking & Steelmaking, 44(5), 377-388.

[5] Tang, H., Cai, S., Lan, P., Ma, Y., Wang, Y., & Wang, K. (2024). Effect of Cerium Content on Non–Metallic Inclusions and Solidification Microstructure in 55SiCr Spring Steel. Materials, 17(22), 5450.

[6] Masoumi, M., Centeno, D. M. A., & Echeverri, E. A. A. (2024). Tailoring the Microstructure using quenching and partitioning processing in a commercial Mn-Si-Cr spring steel to improve tensile properties. Arabian Journal for Science and Engineering, 49(11), 15121-15141.

[7] Murathan, Ö. F., & Kilicli, V. (2022). Effect of isothermal heat treatments under Ms temperature on the microstructures and mechanical properties of commercial high-silicon spring steel. Materials Testing, 64(8), 1112-1121.

[8] Lazarević, M. S., Nedić, B. P., Bajić, D. M., Đurić, S., & Maruš ić, L. (2023). Quality parameters of explosively welded spring steel and carbon steel. Tehnički vjesnik, 30(2), 530-537.

[9] Jiang, Y., Zou, J., Liang, Y., Yin, C., & Yang, M. (2021). Influence of carbides on the strain hardening behavior of 60Si2CrVAT spring steel treated by a Q&T process. Materials Science and Engineering: A, 823, 141695.

[10] Wang, H., Su, F., Wen, Z., & Li, C. (2023). Effects of Mn and Si on the ferrite decarburization of spring steel. Journal of Materials Research and Technology, 27, 363-371.

[11] Xia, B., Zhang, P., Wang, B., Li, X., & Zhang, Z. (2023). Effects of quenching temperature on the microstructure and impact toughness of 50CrMnSiVNb spring steel. Materials Science and Engineering: A, 870, 144856.

[12] Chen, K., Jiang, Z., Liu, F., Yu, J., Li, Y., Gong, W., & Chen, C. (2019). Effect of quenching and tempering temperature on microstructure and tensile properties of microalloyed ultra-high strength suspension spring steel. Materials Science and Engineering: A, 766, 138272.

[13] Bajželj, A., & Burja, J. (2022). Influence of austenitisation time and temperature on grain size and martensite start of 51CrV4 spring steel. Crystals, 12(10), 1449.

[14] Schröder, C., Wendler, M., Volkova, O., & Weiß, A. (2020). Microstructure and mechanical properties of an austenitic CrMnNiMoN spring steel strip with a reduced Ni content. Crystals, 10(5), 392.
[15] Gonçalves, V. R. M., Podgornik, B., Leskovšek, V., Totten, G. E., & Canale, L. D. C. F. (2019). Influence of deep cryogenic treatment on the mechanical properties of spring steels. Journal of Materials Engineering and Performance, 28, 769-775.
[16] Królicka, A., Radwański, K., Ambroziak, A., & Żak, A. (2019). Analysis of grain growth and morphology of bainite in medium-carbon spring steel. Materials Science and Engineering: A, 768, 138446.
[17] Zurnadzhy, V. I., Efremenko, V. G., Wu, K. M., Azarkhov, A. Y., Chabak, Y. G., Greshta, V. L., ... & Pomazkov, M. V. (2019). Effects of stress relief tempering on microstructure and tensile/impact behavior of quenched and partitioned commercial spring steel. Materials Science and Engineering: A, 745, 307-318.
[18] Wang, Y., Sun, J., Jiang, T., Yang, C., Tan, Q., Guo, S., & Liu, Y. (2019). Super strength of 65Mn spring steel obtained by appropriate quenching and tempering in an ultrafine grain condition. Materials Science and Engineering: A, 754, 1-8.
[19] Kim, S. H., Kim, K. H., Bae, C. M., Lee, J. S., & Suh, D. W. (2018). Microstructure and mechanical properties of austempered medium-carbon spring steel. Metals and Materials International, 24, 693-701.
[20] Li, Y., Wang, E., Zhang, L., Fautrelle, Y., Zhao, X., Guo, X., & Zhang, D. (2023). Microstructural evolution and mechanical properties of 60Si2CrVNb spring steel under quenching-tempering heat treatment process. Journal of Materials Research and Technology, 25, 6829-6842.
[21] Aliakbar Abedini,Habibollah Rastegari & SayyedMohammad Emam. (2024). Mechanical properties and work hardening behaviour of spring steel after the quenching-partitioning process. Canadian Metallurgical Quarterly,63(4),1169-1182.
[22] Yihao Ma,Minmin Dong,Peng Qi,Shaojie Zheng,Simon Yisheng Feng,Yunshan Zhang & Rongfu Xu. (2024). Effect of a new two-step austenitizing-Q&P process on the microstructure and mechanical properties of microalloyed 51CrMnV spring steel. Journal of Alloys and Compounds,1004,175864-175864.
[23] JiaxiChen,WanlinWang,LejunZhou,PeishengLyu,JieZeng & MinggangLi. (2024). Initial Solidification Behavior of Spring Steel Billet Near the Mold Corner During Continuous Casting Process. steel research international,95(10),2400136-2400136.