Mechanical Properties Prediction for Hot Roll Steel Using Convolutional Neural Network

Hao Xu[1], Zhiwei Xu[1], and Kai Zhang[1,2(✉)]

[1] School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430065, Hubei, China
zhangkai@wust.edu.cn

[2] Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan 430065, Hubei, China

Abstract. Prediction the mechanical properties is very important in many real-life industry fields. In this paper, we proposed an efficient convolutional neural network (CNN) to predict the mechanical properties of hot roll steel. In this study, 20,000 sets of data are collected from the hot roll factory, where 16,000 sets of data were used for training the CNN model, and 4,000 sets of data were used for testing the performance of the model. Compared with Support Vector Machine (SVM) and Artificial Neural Network (ANN), The experimental results have been demonstrated to provide a competitive and higher prediction accuracy.

Keywords: Convolutional neural network · Mechanical property prediction · Hot roll steel

1 Introduction

Hot roll steel is an important material which widely used in many real-life fields. These areas have different requirement for hot roll steel mechanical properties, so it is important to predict the mechanical properties of alloy steel accurately which are tensile strength (TS), yield strength (YS) and elongation (EL).

Traditionally, mechanical properties prediction is carried out by destructive testing, which is costly and time consuming. As the rolling process is complicated and final mechanical properties of steel determined by many parameters, including the chemical composition and the process parameters [1], it is extremely hard to express the relationships by mathematical model [2]. In the previous studies, scholars have widely use metallurgical mechanism models and statistical models to predict mechanical properties [3]. Due to the complexity and dynamic in steel

Supported by the National Natural Science Foundation of China (Grant Nos. U1803262, 61702383, 61602350).

L. Pan et al. (Eds.): BIC-TA 2019, CCIS 1160, pp. 565–575, 2020.
https://doi.org/10.1007/978-981-15-3415-7_47

manufacturing process [4], the model structure is needed too many experimental trials, which increases the cost and production time [5].

In the existing literature, the Artificial Neural Network (ANN) and Support Vector Machine [6] (SVM) have become widely used to predict the mechanical properties of hot roll steel. Liang et al. [7] used laser-induced breakdown spectroscopy combined with support vector machine to classify steel. Chou et al. [8] used a three-layer feedforward ANN with Taguchi particle swarm to optimize the chemical composition of steel bars and improve the mechanical properties. The methods based on ANN and SVM can demonstrate good performance when the data set is small, but they cannot achieve high precision when dealing with massive data. With the development of the steel industry, the processing is more complex, the relationship between input parameters and mechanical properties is more complicated, those methods might be powerless.

Recently, deep learning [9] has made a major breakthrough in the field of machine learning, especially CNN has demonstrated strong performance in the field of image recognition [10]. It uses the local connection and weight sharing to reduce the number of parameters, not only to extract local features from complex data but also to be insensitive to noise and has good model expression ability [11]. Based on these advantages of CNN, we consider that it can be used to solve hot roll steel mechanical properties prediction problem.

In this paper, a CNN-based method is proposed to predict the mechanical properties of hot roll steel. This method represents chemical composition and processing parameters as one-dimensional vector, and employs CNN-based model to extract features contained by the vector. The rest of the paper is organized as follows. Section 2 introduces the problem considered in this research. Section 3 presents the proposed one-dimensional CNN-based method. Section 4 describes the experimental settings. Section 5 concludes the study.

2 Background

2.1 Hot Rolled Processing

The hot rolled processing is one of the parameters affecting the mechanical properties. As shown in Fig. 1, after reheating, roughing rolling, finishing rolling, laminar cooling and down coiling, the steel slab becomes a coil of a thin sheet.

In the manufacturing processes, a series of complex microstructure have been changed. First, reheating process provide uniform austenite grain. Then, the roughing and finishing processes refine austenite by dynamic and static recrystallization. These grains determine the mechanical properties of the steel. In this process, the furnace temperature (FT), the roughing rolling temperature (RRT), the finishing rolling temperature (FRT) and the coiling temperature (CT) have the greatest influence on the mechanical properties of hot roll steel. According to China High Strength Low Alloy Structural Steel Standard GB/T1591-20608, most of the chemical components of alloy steel consist of carbon (C), manganese (Mn), silicon (Si), phosphorus (P) and sulfur (S). The effects of chemical composition on mechanical properties is another important parameter. The additions of

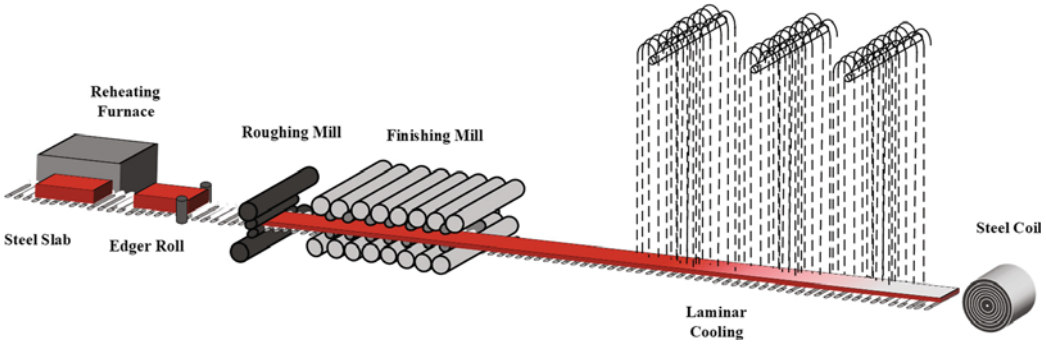

Fig. 1. Hot rolled processing.

some alloying elements affect ferrite transformation and change the mechanical properties of alloy steel [12]. The presence of microalloying elements generally control the grain size and have a significant impact on the strength. These mechanisms can be very complicated. For example, the major chemical component of steel bar is carbon, which determines mechanical strength [13]. When the ratio of C in the steel is below 0.8Wt%, the YS and TS of the steel increase dramatically with the increases of C content, but the EL of steel decreases. S will reduce the hot workability and strength of steel and P will reduce the plasticity and toughness of the steel [14].

2.2 Mechanical Properties of Hot Roll Steel

The mechanical properties of alloy steel are TS, YS, and EL. TS is defined as the maximum tensile stress that the steel can withstand before breaking, and the YS is defined as the maximum stress of the steel can withstand before plastic deformation begins. EL is defined as the percentage of stretched length to the original length after the steel is broken. Figure 2 shows how the YP, TS and EL of hot roll steel are related.

Chemical elements and processes together determine the mechanical properties of steel, the different combination of chemical components and processes parameters complicate those properties. It is difficult to express the relationship. An efficient method of predicting mechanical properties is needed.

3 Proposed Algorithm

This section introduces the proposed prediction method based on CNN for mechanical properties of hot roll steel.

3.1 CNN Prediction Model

CNN is widely used in the field of images, and the input to the network is a two-dimensional matrix at most. In order to adapt to the one-dimensional characteristic of hot rolled steel data, one-dimensional CNN-based architecture

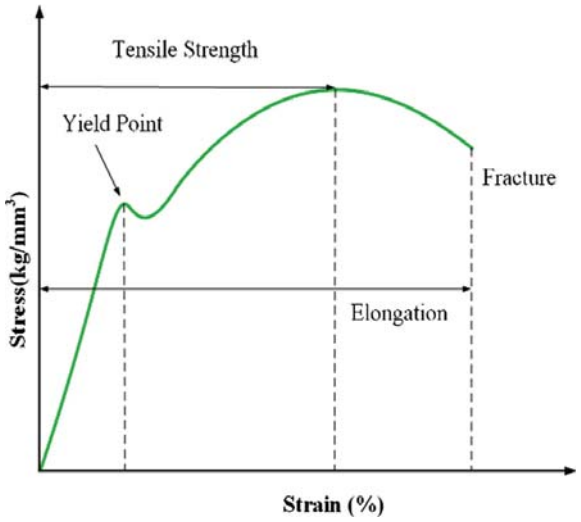

Fig. 2. Relationship between tensile strength, yield strength and elongation of hot roll steel.

is used which applies 1D arrays instead of 2D matrices for both kernels and feature maps.

As shown in Fig. 3, this model consists of three parts, model input, feature extraction part, prediction part. Each of the parts is explained below. The feature extraction is composed of several feature extractors, which is stacked by convolutional layer, batch normalization, nonlinear activation and pooling layer. The prediction part contains two fully connected (FC) layers. The predicted value will be output in the final layer.

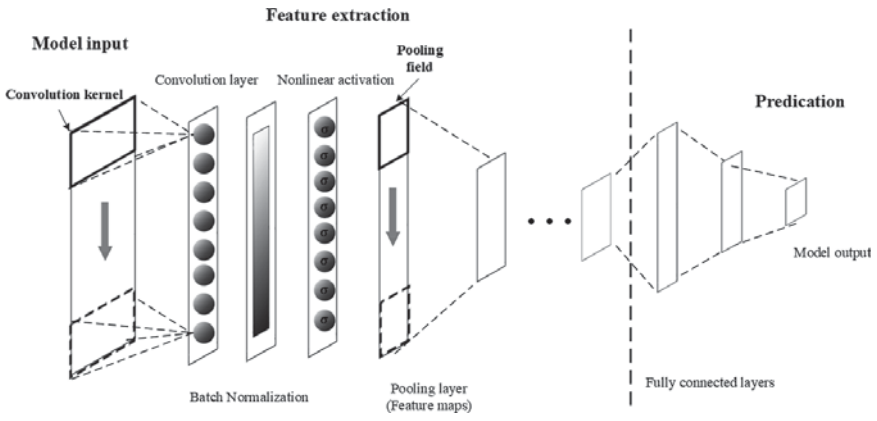

Fig. 3. Structure of CNN-based prediction model for mechanical properties.

First, model input includes 16 chemical compositions and 4 heat treatment process parameters. In order to eliminate the largely distinct scales in different fields, the raw data need to be normalized. If the raw data is $X = (x_1, x_2, x_3 \ldots x_{n-1}, x_n)$, the normalization equation is:

$$x_i^* = \frac{x - \min}{\max - \min} \tag{1}$$

where x_i^* is the normalized value of the input parameter x_i, and min and max are the minimum and maximum values in the data samples respectively.

Second, the feature extraction part is combination of convolutional layer, batch normalization, nonlinear activation layer and pooling layer.

The convolutional layer is the core part of the CNN model. It consists of a set of linear filters that convolute the input data. The convolution operation is as shown in the formula:

$$(h_k)_{ij} = (W_k * x)_{ij} + b_{ij} \tag{2}$$

where k is the index of the kth feature map in the convolutional layer, (i,j) is the index of the pixel point, x is the input data, and W and b are the weight parameter and the offset parameter of the kth feature map, respectively. $(h_k)_{ij}$ is the output value of the kth feature map. And in one-dimensional convolution, j is usually set to 1.

Batch Normalization [15] is added after each convolutional layer, which can speed up the training of the network and avoid gradient explosion. First, the distribution of the input data is normalized to a distribution with the mean of 0 and variance of 1, as follows:

$$\hat{x}^{(k)} = \frac{x^k - E\left(x^k\right)}{\sqrt{\mathrm{Var}\left(x^k\right) + \varepsilon}} \tag{3}$$

where $x^{(k)}$ represents the kth dimension of the input data, $E(x^k)$ represents the average of the dimension, and $\sqrt{\mathrm{Var}\left(x^k\right) + \varepsilon}$ represents the standard deviation. Then set two learnable variables γ and β, and use these two learnable variables to restore the data distribution learned from the previous layer, as follows:

$$y^{(k)} = \gamma^k \hat{x}^{(k)} + \beta^{(k)} \tag{4}$$

Rectified linear unit (ReLU) is used as the nonlinear activation function, which can prevent the problem of gradient vanished and gradient explosion in the neural network during training. Let max denote the function to select the larger value between x and zero, the ReLU activation function can be expressed as:

$$\mathrm{Relu}(x) = \max(0, x) \tag{5}$$

The pooling layer is a down-sampling layer, which can not only reduce the network scale of the CNN, but also identify the most prominent features of input layers. The maxpooling method is used in the proposed CNN model by selecting the maximum value in the pooling field.

Finally, the predicted value is obtained by feature extractions part to gain the features from raw data and applying fully connected layers to process the feature information.

3.2 CNN Optimization and Evaluation Metric

The prediction of the CNN are the mechanical properties of hot roll steel, and the mean squared error (MSE) is employed to measure the distance between predictions and actual mechanical properties. Thus, minimizing MSE is taken as the loss function of the CNN. MSE can be written as:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{6}$$

where $\hat{y}_i$ represents the predicted value, y_i represents the actual value, and N represents the number of samples in the data set.

Four indicators are adopted as the evaluation metrics to assess the prediction capability comprehensively. Mean square error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and coefficient of determination (R^2) are used as evaluation of the model. MAE represents the average of the absolute error, MAPE is a measure of prediction accuracy of a forecasting method. R^2 explains how much of the variability of a factor can be caused or explained by its relationship to another factor.

4 Results and Discussions

This section consists of two part. First, the experimental steel data is introduced. Second, experiments are performed to demonstrate that the proposed CNN model can predict for mechanical properties, and its training results are compared with Artificial Neural Network (ANN) methods, and Support Vector Machine (SVM).

4.1 Data Description

In this paper, 20,000 sets of data are collected from the hot roll factory, where 16,000 sets of data were used for training the CNN model, and 4,000 sets of data were used for testing the performance of the model. The data is shown in the Table 1 below. The input data consists of sixteen chemical components and four hot roll process parameters. The output consists of three mechanical properties including TS, YS and EL.

4.2 Comparison Results

In order to test the performance of the proposed algorithm, ANN and SVM are chosen for comparison. ANN represents the traditional neural network and attempts to learn features through hidden layers. SVM find a hyperplane to divide the sample space of the data set into different samples. From Table 2, compared with the SVM and ANN methods, the proposed CNN model has achieved good results, and the evaluation metric on the test set are better than the other two algorithms. This shows that the proposed CNN model can effectively extract the features that affect the mechanical properties of hot rolled steel and has good generalization ability.

Table 1. Steel dataset.

Parameter	Unit	Minimum	Maximum	Mean
C	Wt%	0.0051	0.1936	0.149
Mn	Wt%	0.2362	1.3696	0.4674
Si	Wt%	0.0063	0.2994	0.1944
P	Wt%	0.0062	0.0364	0.0172
S	Wt%	0.0012	0.025	0.013
Cu	Wt%	0.0124	0.0826	0.0405
Al	Wt%	0.0003	0.5654	0.0067
Als	Wt%	0.0002	0.5652	0.0059
Ni	Wt%	0.0012	0.0673	0.0168
Cr	Wt%	0.0082	0.0942	0.0296
Ti	Wt%	0.0001	0.0691	0.0015
Mo	Wt%	0.0012	0.0184	0.0053
V	Wt%	0.0012	0.0056	0.0014
Nb	Wt%	0.001	0.0194	0.0011
N	Wt%	0.0008	0.0653	0.003
B	Wt%	0.0004	0.0031	0.0002
FT	°C	1188	1291	1242.6607
RRT	°C	976	1142	1049.3915
FRT	°C	200	1056	752.5484
CT	°C	630	934	774.2208

4.3 Hyperparameters Optimization

The hyperparameters are the important factors that should be considered cautiously which include convolutional kernels size, polling size and depth of the CNN, when implementing the structure of a CNN.

First, there is no rules for the selection of hyperparameters generally. Based on the parameter settings of VGG Net, which has achieved second place in the 2014 ILSVRC, we select convolutional filter of size (3,1) and max pooling of size (2,1). Xavier is adopted as the weight initialization method, the batch size is set to 128, and parameters in the CNN model are updated by Adam optimizer. The initial learning rate of the proposed CNN model is set as 0.001.

Second, the depth of the CNN should not be too big or too small, which can make the model difficult to converge or overfit. The structure of are shown in Table 3, where each convolutional layer is followed by a pooling layer, and the numbers represent quantities of convolutional filters in the layer. Table 4 shows the results of the CNN with different number of depths, TS, YS and EL represent the three mechanical performance predicted by the model. The Depth-2 achieves the best result. When the depth is too small, the relationship between the input

Table 2. Prediction performances of the CNN and other algorithms.

Predicting	Evaluation	CNN	SVM	ANN
TS	MSE	**0.0032**	0.0071	**0.0032**
	MAE	**0.0231**	0.0312	0.0221
	MAPE	**0.0612**	0.0731	0.0693
	R^2	**0.8281**	0.792	0.8262
YS	MSE	**0.0015**	0.0032	0.002
	MAE	**0.0116**	0.0129	0.0123
	MAPE	**0.0453**	0.0488	0.0441
	R^2	**0.9025**	0.8299	0.8519
EL	MSE	**0.0015**	0.0032	0.002
	MAE	**0.0116**	0.0129	0.0123
	MAPE	**0.0453**	0.0488	0.0441
	R^2	**0.9025**	0.8299	0.8519

and the output cannot be completely extracted and when the depth is too large, the model shows poor generalization ability.

Table 3. Different depths for CNN.

Depth	Structures of CNN Model
Depth-1	32 conv→ 64 conv
Depth-2	32 conv →64 conv→128 conv
Depth-3	32 conv →64 conv →128 conv →256 conv

The details of the depth-2 CNN are listed in Table 5. It contains three convolutional layers, three pooling layers and two fully-connected layers. FC1 represent the first FC layer, FC2 represent the second FC layer, and the followed numbers indicate the number of neuron nodes in the FC layer. The denotation of Filter (3*1*64) means that it is a convolutional layer which filter size is 3*1 with 64 channels. Maxpool (2*1) denotes that it is a maxpooling layer with a 2*1 pooling field.

Figure 4 shows the actual values on the test set and the predicted values of the proposed CNN model by scatter diagrams of TS, YS and EL. These values are normalized so they are only in the range of 0 to 1.

Table 4. Results of CNN in different depths.

Depth	Evaluation	TS	YS	EL
Depth-1	MSE	0.0061	0.0025	0.0035
	MAE	0.0525	0.0302	0.0578
	MAPE	0.0921	0.0976	0.1132
	R^2	0.6241	0.8556	0.7921
Depth-2	MSE	**0.0032**	**0.0015**	**0.0023**
	MAE	**0.0231**	**0.0116**	**0.0216**
	MAPE	**0.0712**	**0.0653**	**0.0618**
	R^2	**0.8281**	**0.9025**	**0.8649**
Depth-3	MSE	0.0052	0.0032	0.0037
	MAE	0.0427	0.0296	0.0376
	MAPE	0.0834	0.0876	0.1032
	R^2	0.6724	0.8667	0.801

Table 5. Details of the CNN.

Layer	Name	Parameters
L1	Conv	Filter(3*1*64)
	Pool	Maxpool (2*1)
L2	Conv	Filter(3*1*128)
	Pool	Maxpool (2*1)
L3	Conv	Filter(3*1*256)
	Pool	Maxpool (2*1)
FC1	Fully-connected 1	1280
FC2	Fully-connected 2	256

The Fig. 4 demonstrates that the predicted values of the proposed model are in good agreement with the actual mechanical properties, which indicates that the proposed CNN model has a good effect.

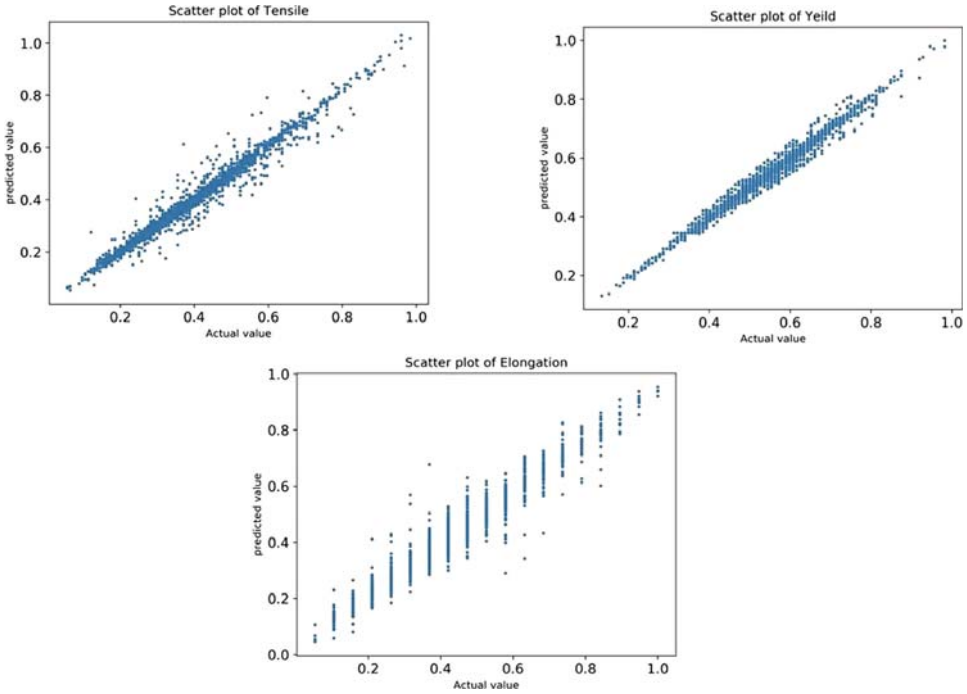

Fig. 4. Comparisons of predicted and actual mechanical properties

5 Conclusion

The proposed CNN model can predict the mechanical properties of steel through sixteen kinds of chemical composition and four kinds of hot rolling production processes of hot rolled steel. The raw data is processed by normalization to keep the data in the same range. The best CNN structures in predicting TS, YS, and EL are determined by comparison experiments, and the model prediction results are compared with the SVM model and the ANN model. The experimental results show that our algorithm provides the best accuracy than the traditional methods on 20,000 hot roll datasets. The results demonstrate the ability of the proposed CNN-based model to predict the mechanical properties of hot rolled steel.

References

1. Beghini, M., Bertini, L., Monelli, B.D., Santus, C., Bandini, M.: Experimental parameter sensitivity analysis of residual stresses induced by deep rolling on 7075–T6 aluminium alloy. Surf. Coat. Technol. **254**, 175–186 (2014)
2. Cheng, C.K., Tsai, J.T., Lee, T.T., Chou, J.H., Hwang, K.S.: Modeling and optimizing tensile strength and yield point on steel bar by artificial neural network with evolutionary algorithm. In: 2015 IEEE International Conference on Automation Science and Engineering (CASE), pp. 1562–1563. IEEE (2015)
3. Yun, X., Gardner, L., Boissonnade, N.: The continuous strength method for the design of hot-rolled steel cross-sections. Eng. Struct. **157**, 179–191 (2018)

4. Wu, Y., Ren, Y.: Prediction of mechanical properties of hot rolled strips by BP artificial neural network. In: 2011 International Conference of Information Technology. Computer Engineering and Management Sciences, vol. 1, pp. 15–17. IEEE (2011)
5. Mohanty, I., et al.: Online mechanical property prediction system for hot rolled IF steel. Ironmak. Steelmak. **41**(8), 618–627 (2014)
6. Guenther, N., Schonlau, M.: Support vector machines. Stata J. **16**(4), 917–937 (2016)
7. Liang, L., et al.: Classification of steel materials by laser-induced breakdown spectroscopy coupled with support vector machines. Appl. Opt. **53**(4), 544–552 (2014)
8. Chou, P.Y., Tsai, J.T., Chou, J.H.: Modeling and optimizing tensile strength and yield point on a steel bar using an artificial neural network with Taguchi particle swarm optimizer. IEEE Access **4**, 585–593 (2016)
9. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436 (2015)
10. Schmidhuber, J.: Deep learning in neural networks: an overview. Neural Netw. **61**, 85–117 (2015)
11. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
12. Esfahani, M.B., Toroghinejad, M.R., Abbasi, S.: Artificial neural network modeling the tensile strength of hot strip mill products. ISIJ Int. **49**(10), 1583–1587 (2009)
13. Saravanakumar, P., Jothimani, V., Sureshbabu, L., Ayyappan, S., Noorullah, D., Venkatakrishnan, P.G.: Prediction of mechanical properties of low carbon steel in hot rolling process using artificial neural network model. Procedia Eng. **38**, 3418–3425 (2012)
14. Hwang, R.C., Chen, Y.J., Huang, H.C.: Artificial intelligent analyzer for mechanical properties of rolled steel bar by using neural networks. Expert Syst. Appl. **37**(4), 3136–3139 (2010)
15. Li, X., Chen, S., Hu, X., Yang, J.: Understanding the disharmony between dropout and batch normalization by variance shift. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2682–2690 (2019)