INTELLIGENT DETECTION OF RUTILE SINGLE CRYSTAL GROWTH MORPHOLOGY BASED ON COMPUTER VISION UNDER HIGH TEMPERATURE

Weilun ZHAO[1]*, Yan WANG, Xiaoguo BI

Rutile single crystal is an excellent birefringent optical crystal with broad application prospects in the fields of precision detection and 5G communication equipment development, whose artificial preparation requires high temperature and oxygen-rich environment. To achieve the visual intelligent control of its growth, this paper designs a computer vision detection model of rutile single crystal morphology under high temperature. Feature analysis of rutile growth morphology images is achieved through convolutional neural network establishment, thus allowing image segmentation and feature point location during the dynamic growth process of rutile single crystals. The model can detect parameters such as crystal growth edge, melting boundary edge, melting cap edge, melting cap tangent angle, etc. According to the crystal growth morphology, it controls the given parameters in real time, effectively prevents the overflow of liquid crystal, thus improving the crystal growth success rate and crystal quality.

Keywords: rutile; computer vision; EfficientNet convolutional network; artificial intelligence

1. Introduction

High temperature oxide single crystals play an important role in scientific research and modern social life thanks to its special physical and chemical properties. For instance, rutile (TiO_2) single crystal with great birefringence and high refractive index is used in spectral prisms and polarizing devices, such as optical isolators, optical circulators, beam splitters, etc. So far, all the above-mentioned devices used in optical communication adopt yttrium vanadate (YVO_4) crystals, while high-end products require rutile (TiO_2) single crystals [1]. At present, there are three main ways to produce rutile single crystal: float zone method, dry pot pulling method, and flame melting method. Floating zone method boasts the advantages of short growth period for crystal material, no crucible and easy observation of crystal growth. Nonetheless, stability of the melting zone is maintained through surface tension and gravity, making it difficult to control the melt stability, so float zone method is not the best method for rutile growth [2,3].

[1]* Senior Engineer, Automation College, Shenyang Institute of Engineering, China, e-mail: 453364546@qq.com

Dry pot pulling method demands a vacuum environment or protection of inert gas, while rutile growth demands high temperature environment and oxygen-rich atmosphere, so pulling method is inapplicable to rutile growth [4]. Flame melting method, also known as the Werner method, is proposed by the French scientist Werner at the end of the 19th century based on improvement of the predecessor's crystal preparation method. It was first used to produce sapphire and ruby. Dr. CH Moore of National Lead Company in the United States et al. first prepared rutile single crystals using flame melting method. So far, flame melting method is still the main approach for preparing commercial rutile crystals. Dr. Rangxiu Zhongzhu of Japan designed and manufactured the "Central Pillar Type Flame Melting Single Crystal Furnace" to produce rutile single crystal with high-level crystal size and quality. The biggest improvement lies in the structure optimization of the three-tube burner, which enables stable temperature distribution of crystal growth interface. M. Kestigian also used a similar three-tube burner to prepare transition metal-doped rutile crystals [5-7]. J.A. Adamski modified the three-tube burner for hydrogen-oxygen combustion to prepare rutile single crystal with greater diameter, which overcomes the problems of horizontal temperature field fluctuation and turbulence in Moore's synthesis of rutile [8,9]. The structure of this burner lays the foundation for all present three-tube burners. In recent decades, with the continuous progress of basic disciplines (such as physics, chemistry) and preparation technology, the theory of crystal growth control enjoys rapid development in both research means and research levels. Nevertheless, in the control of crystal growth preparation, poor controllability of crystal growth kinetics and high crystal defect density are common.

The temperature in the rutile single crystal growth furnace is as high as 2000 degrees Celsius, making it difficult to measure the crystal growth parameters through sensors. The current mainstream crystal growth control method is to observe the crystal growth rate and growth shape with the naked eye and then manually set the control parameters [10,11]. Hence, errors are inevitable in manual observation and control, leading to structural integrity issues in grown crystal, such as high dislocation density, high stress, and non-uniformity, which impairs the volume and yield of crystal growth. Crystals with growth failure are shown in Fig. 1.

Fig. 1. Sample of rutile single crystal with growth failure

To achieve computer vision detection of rutile single crystal growth and intelligent control of parameter setting, this paper incorporates computer vision and artificial intelligence technology to the control system of rutile single crystal flame melting furnace, thus establishing the EfficientNet convolution neural network for rutile single crystal morphology feature detection under high temperature. Currently, the more popular convolutional neural networks include AlexNet, VGGNet, ResNet, DenseNET, etc. AlexNet convolutional network [12-13], proposed by A.Krizhevsky, I.Sutskever, et al. in 2012, has an 8-layer network structure with about 60 million parameters, which uses ReLU as the activation function for faster convergence. Also, the method of Max pooling was proposed. The network won the championship of the image classification competition in 2012, and subsequent new networks are derived based on this network through technical improvements. VGG Net [14-15] proposed by K. Simonyan and Zisserman in 2014 inherits and develops the idea of ZFnet. A simple and effective structural design principle is proposed, so that multiple layers of small-size filters (3×3) replace large-size filters (11×11). This method greatly reduces the number of parameters while maintaining convolution effect of large-size filters, but its disadvantage of great depth (number of layers) of the convolution kernel results in increased computational cost [16]. In 2019, Google released the EfficientNet network to uniformly reduce all dimensions of the model via compound coefficients, which greatly improves the model accuracy and efficiency.

This paper introduces an attention mechanism for further improvement of EfficientNet and establishes a visual recognition network for the classification of high temperature crystal oxide feature parameters [17-20]. The network is capable of extracting different levels of features to automatically combine low hierarchical features into high-level features. By virtue of this automatic feature extraction function, the network can gain a good feature representation from the original image pixels, thus allowing feature parameter target classification of rutile single crystal growth under high temperature [21-22]. With greatly improved accuracy and shortened response time, the network supports automatic and precise control of various process parameters in crystal growth [23-25].

2. Establishment of Convolutional Neural Network for Crystal Morphological Object Classification

2.1 Architecture of Visual Recognition System

In this paper, the video of 2000 hours in the process of crystal growth is collected and decomposed into 200000 crystal morphology images, of which 150000 images are used as the training dataset of convolutional neural network and 50000 as the testing dataset. According to the main characteristics that affect the success rate of crystal growth, the data sets are divided into six categories, and

the annotations are melt boundary, melting cap, growth striation, solid crystal, liquid crystal and overflow. The network model is shown in Fig. 2.

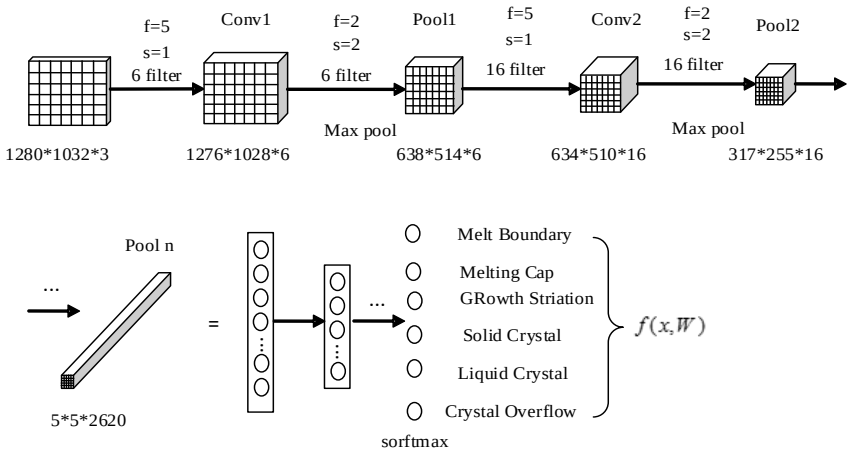


Fig. 2. Convolutional Neural Network for Object Classification of Crystal Parameters

2.2 Using principal component analysis for improvement of classification model

In the network structure of the traditional neural network, after the collected image information is input to the base station computer, the target can normally be classified based on geometric features. How to extract the geometric edge feature points of the obstacle target is a key link in the target classification. Normally, feature point is extracted based on the image gray value around the feature point. The computer detects the pixel value of the area around the pixel area of the candidate target feature point. If the candidate target feature point has sufficient coordinate points with pixel values greatly different from the gray value of the candidate point, the candidate point can be judged as the pixel coordinate of a feature point.

$$N = \sum_{x \forall (circle(p))} \left| I(x) - I(p) \right| > \varepsilon_d \tag{1}$$

Where:

$I(x)$: Grayscale at any point on the circumference;

$I(p)$: Grayscale of the center point;

ε_d: The set threshold.

If N is greater than the algorithm setting value (Fast-9 to 16 according to the default algorithm), the default value is 9, then the point is judged as a feature point.

The feeding method of vibrating screen is adopted in the system. After each screen vibration, a certain amount of titanium dioxide powder can enter the combustion furnace and be mixed with hydrogen and oxygen for combustion. The powder can be melted to the liquid by flame and attached to the semi molten seed crystal, and the crystal in the combustion furnace can grow continuously.

Fig. 3. Picture before blanking

Fig. 4. Pictures during blanking

The prerequisite for judgment of crystal growth health is to accurately identify the physical edges of each part of the crystal. In real-time processing of image data, feature extraction is required for images collected in each frame, that is, convolution operation is needed on the input image through multiple convolution kernels, so that convolution result can be synthesized into high-dimensional output feature image. At present, if the classical convolutional neural network is to achieve good learning effect, a large number of convolution kernels are required for convolution calculation, which will generate abundant parameter matrices. The multivariate large dataset composed of a large number of convolution kernels will provide rich information for the study of image feature extraction, which, however, also increases the data collection workload to a certain extent. More importantly, in many cases, correlation exists between many convolution kernels. Separate analysis of each convolution kernel is often isolated without sufficient utilization of data information. However, if we blindly reduce the convolution kernels, a lot of useful information will be lost, leading to incorrect conclusions.

In this paper, the attention mechanism is introduced into the convolution neural network. Due to the relatively large number of visual features of the image and the large overlap of information reflected by the feature parameters within a certain range, it is possible to select as few indicators in the image information as possible to reflect more image features of the original target according to the actual needs of crystal feature recognition. Because the image of crystal growth in the high-temperature furnace must have a high resolution to highlight its growth characteristics, the image data usually has a high dimension, and the data contains high redundant components. In order to improve the efficiency of the subsequent classification algorithm and reduce the possibility of redundancy and over fitting, this paper uses PCA to reduce the dimension of image data features and combines EfficientNet classification model to extract the relevant features of image local advantages. The network structure is shown in Figure 5.

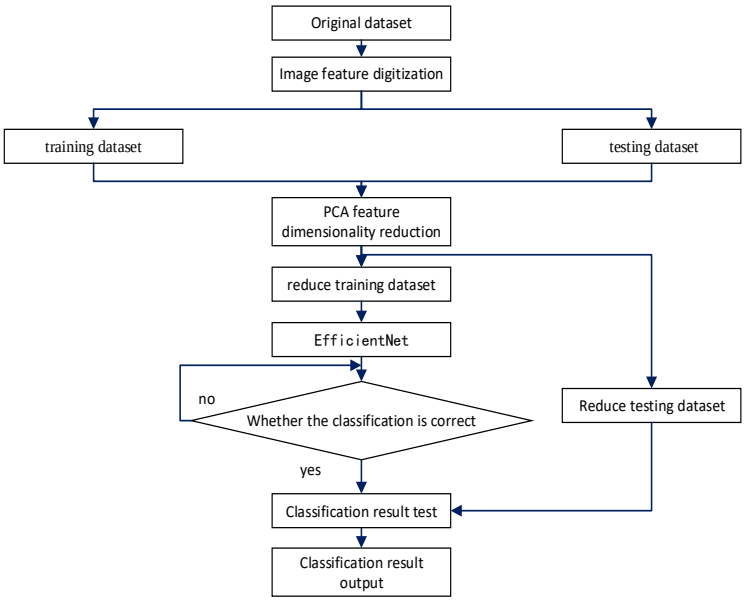


Fig.5. Feature Classification Process Based on PCA Efficient Net

Therefore, in this paper, principal component analysis method is taken to reduce the number of convolution kernels, thereby reducing the feature image dimension. While reducing the indexes to be analyzed, the loss of information contained in the original indexes is minimized, so that the image feature data can be comprehensively analyzed. Considering the certain correlation between the convolution kernels, we can modify the closely related convolution kernels into new convolution kernels in minimum number, so that these new convolution kernels are irrelevant. In this way, fewer comprehensive indexes can be used to represent the various types of information reflected in image features.

Suppose the visual image captures P-dimensional random vector:

$$x = (x_1, x_2, ..., x_p)^T \tag{2}$$

Where, each vector has n features:

$$x_1 = (x_{11}, x_{12}, ..., x_{1n})^T$$

Then the feature of the P-th vector is:

$$x_p = (x_{p1}, x_{p2}, ..., x_{pN})^T$$

Linearly combine the p features in x ,suppose:

$$F_1 = a_{11}x_1 + a_{21}x_2 + ...a_{p1}x_p = a^T x \tag{3}$$

Then,

$$F_p = a_{1p}x_1 + a_{2p}x_2 + ...a_{pp}x_p = a_p{}^T x \tag{4}$$

Where:

$$a_i = [a_{1i}, a_{2i} ... a_{pi}]^T$$

From the meaning of principal component analysis, we can see that:

(1) If the included information does not coincide, then the features of $F_i, i = 1,2,3...n$ should not be included in $F_j, j = 1,2,3...p$, that is, F_i is irrelevant with F_p, $\text{cov}(F_i, F_j) = 0$, namely, the covariance is zero;

(2) F_1 contains the most information, so there is need to maximize the F_1 variance as much as possible, that is, to compress the information in F_1.

From formula (10), the mathematical expectation of the random vector x can be calculated:

$$Ex = (Ex_1, Ex_2, ...Ex_p)^T = \mu \tag{5}$$

Let

$$\mu_1 = Ex_1, \mu_2 = Ex_2, ..., \mu_p = Ex_p$$

The covariance matrix of x can be expressed as:

$$D(x) = E(x - E(x))(x - E(x))^T \tag{6}$$

Where

$$D(x) = \begin{bmatrix} \text{cov}(x_1, x_1) & \text{cov}(x_1, x_2)... & \text{cov}(x_1, x_p) \\ ... & ... & ... \\ \text{cov}(x_p, x_1) & \text{cov}(x_p, x_2)... & \text{cov}(x_p, x_p) \end{bmatrix} \tag{7}$$

Let

$$\text{cov}(x_i, x_j) = \delta_{ij} \tag{8}$$

Define the random variable $y = (y_1, y_2, ..., y_p)^T$, then the covariance matrix of x and y is:

$$\text{cov}(x, y) = E(x - E(x))(y - E(y)^T \tag{9}$$

According to formula (7), the variance matrix is:

$$D(x, y) = \begin{bmatrix} \text{cov}(x_1, y_1) & \text{cov}(x_1, y_2)... & \text{cov}(x_1, y_p) \\ ... & ... & ... \\ \text{cov}(x_p, y_1) & \text{cov}(x_p, y_2)... & \text{cov}(x_p, y_p) \end{bmatrix} \tag{10}$$

Let r_{ij} be the correlation coefficient matrix. From formula (11), there is:

$$r_{ij} = \frac{\text{cov}(x_i, x_j)}{\sqrt{\text{var}(x_i)} \cdot \sqrt{\text{var}(x_j)}} \tag{11}$$

Suppose:

$$\text{cov}(x_i, x_i) = \delta_{ii}, \text{cov}(x_i, x_j) = \delta_{ij} \tag{12}$$

Formula (11) can be transformed into:

$$r_{ij} = \frac{\delta_{ij}}{\sqrt{\delta_{ii}} \cdot \sqrt{\delta_{jj}}} \tag{13}$$

The vector form of formulas (1)-(2) is:

$$F = [F_1, F_2, ...F_P] \tag{14}$$

That is

$$F = [F_1, F_2, ...F_P] = a_{ij}{}^T x_{ij} \tag{15}$$

Then,

$$\max \operatorname{var}(F) = \operatorname{var}(a_{ij}{}^T x_{ij}) \tag{16}$$

According to the definition of variance, there is:

$$\operatorname{var}(a_{ij}{}^T x_{ij}) = a_{ij}{}^T E(x_{ij} - E(x_{ij}))(x_{ij} - E(x_{ij})^T a_{ij} \tag{17}$$

$E(x_{ij} - E(x_{ij}))(x_{ij} - E(x_{ij})^T$ is a symmetric matrix, the symmetric matrix must be similarly diagonalized, that is, there is an orthogonal matrix u, which meets $u \cdot u^T = E$.

Let

$$u = [u_1, u_2, ...u_p]$$

Then

$$u^T E(x_{ij} - E(x_{ij}))(x_{ij} - E(x_{ij})^T u = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & 0 & \lambda_P \end{bmatrix} \tag{18}$$

Where: $\lambda = [\lambda_1, \lambda_2, ...\lambda_p]$ is the eigenvalue of the covariance matrix.

Thus:

$$E(x_{ij} - E(x_{ij}))(x_{ij} - E(x_{ij})^T = \sum_{i=1}^{p} \lambda_i u_i u_i{}^T \tag{19}$$

Then,

$$\max \operatorname{var}(F) = a_{ij}{}^T \sum_{i=1}^{p} \lambda_i u_i u_i{}^T a_{ij} \tag{20}$$

By simplification, there is:

$$\max \operatorname{var}(F) = \sum_{i=1}^{p} \lambda_i (a_{ij}{}^T u_i)(a_{ij}{}^T u_i)^T = \sum_{i=1}^{p} \lambda_i (a_{ij}{}^T u_i)^2 \tag{21}$$

If:

$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$

Then:

$$\sum_{i=1}^{p} \lambda_i (a_{ij}{}^T u_i)^2 \leq \lambda_1 (a_{ij}{}^T (u_1, u_2, ..., u_p)(a_{ij}{}^T (u_1, u_2, ..., u_p))^T \tag{22}$$

Expand formula (28), there is

$$\sum_{i=1}^{p}\lambda_i(a_{ij}{}^T u_i)^2 \le \lambda_1 a_{ij}{}^T(u_1,u_2,...,u_p)(u_1,u_2,...,u_p)^T a_{ij} \tag{23}$$

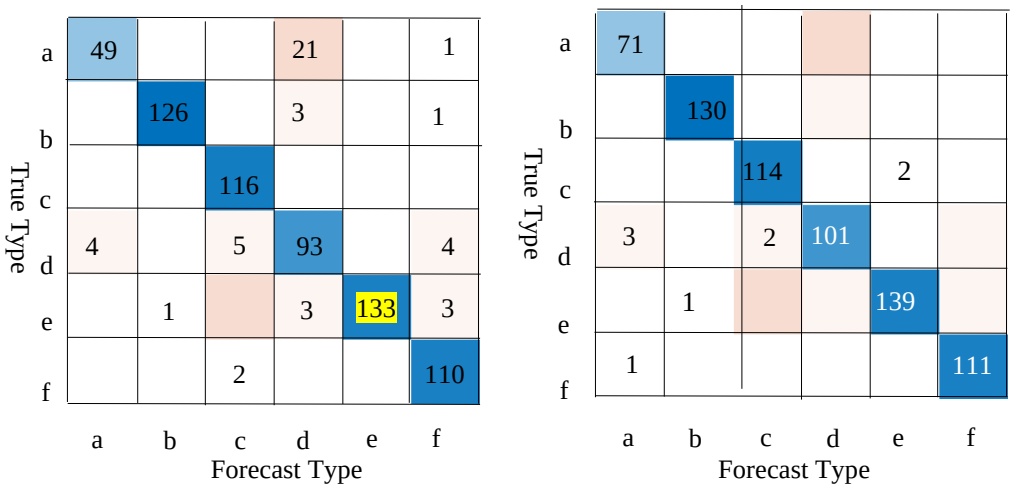
u is known to be an orthogonal matrix which meets $u\cdot u^T = E$, then

$$\sum_{i=1}^{p}\lambda_i(a_{ij}{}^T u_i)^2 \le \lambda_1 a_{ij}{}^T a_{ij} = \lambda_1 \tag{24}$$

As it can be seen from formula (24), the maximum variance is highly correlated with the maximum eigenvalue, and the principal component of x_1 - x_p is a linear combination of coefficients with the eigenvectors of the covariance matrix, and the variance of each principal component is equivalent to the feature root corresponding to the covariance. The principal component is the convolution kernel with great correlation difference. By reducing the number of convolution kernels through principal component analysis, it is possible to increase the model running speed while maintaining the image recognition accuracy, reduce the GPU computation amount, thereby achieving further model optimization.

3. Experiment and result analysis

Improved model takes 35 minutes to train after 1000 iterations, with training speed increased by 5%. The mini-batch accuracy also reaches 100%, with mini-batch loss error dropping to 3.17×10^{-6}. As the number of model iterations increases, a single iteration roughly consumes the same amount of time. Comparison of recognition accuracy of crystal characteristic parameters before and after optimization is shown in Fig. 6.



a: Melt Boundary; b: Melting Cap; c: Growth Striation; d: Solid Crystal; e: liquid Crystal; f: Overflow

Fig. 6. Comparison of recognition accuracy of crystal characteristic parameters before and after optimization

Comparison of recognition rates before and after optimization is shown in Table 1.

Table 1

Image coordinates corresponding to feature points (part)

Coordinate calibration of feature points after image optimization			
Image type	Accuracy before optimization	Optimized accuracy	Accuracy improvement value
Melt Boundary	92.4%	94.6%	2.2%
Melting Cap	99.2%	99.2%	0%
Growth Striation	81．1%	98.2%	17.1%
Solid Crystal	77.5%	100%	22.5
Liquid Crystal	100%	98.6%	-1.4%
Crystal Overflow	92.4	100%	7.6%

The comparison of the convergence curve of the loss function before and after the improvement of the model is shown in Fig. 7.

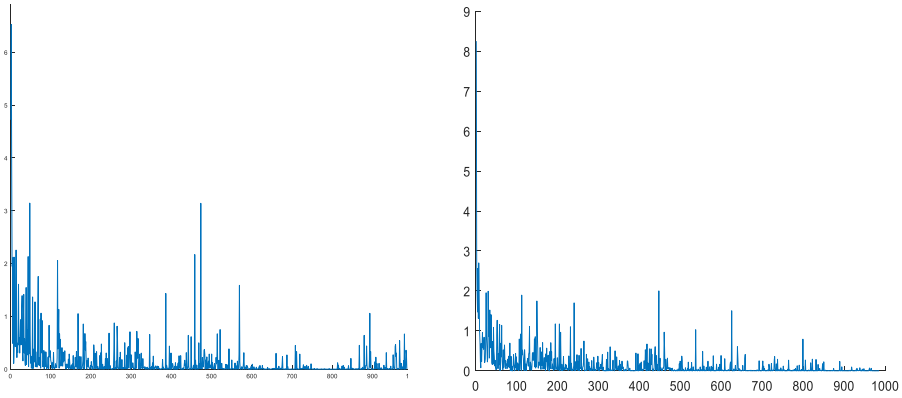


Fig. 7. Comparison of convergence curves of loss function before and after model improvement

From Fig. 3 to Fig. 4 and Table 1, by improving the EfficientNet model through the principal component analysis method, great improvement is achieved in its model performance compared with ordinary convolutional neural networks in terms of target recognition, accuracy, and response time.

5. Conclusions

In this paper, attention mechanism is introduced to optimize the EfficientNet network structure. By virtue of the visual feature dimensionality reduction algorithm of principal component analysis, we balance the network

depth and width, reduce the number of parameters and computation while improving the network performance. From image comparison, it can be seen that when common edge detection algorithm is chosen, due to the high image resolution and the relatively clear image texture, there are many image feature points. Without optimization via attention mechanism, all the feature edges will be displayed, thus prone to noise pollution. Moreover, the image does not capture the object details, the target object and the background image are not well understood, good recognition effect is impossible, making it more difficult to calibrate and detect feature objects. Through edge feature extraction optimized by the attention mechanism, it is possible to reduce the dimensionality of the input parameters, lower the non-feature interference, strengthen the recognition ability, coordinately control the depth, width and dimension of the image information to allow focus on more relevant areas of object details. Algorithm optimization of the attention mechanism allows more accurate and rapid visual target detection, calibration and classification of rutile single crystal growth features, accelerates the image data transmission speed, so that the system enables better real-time detection and control effects.

REFERENCES

[1]. *A.S. Kassiopeia, I.S. Andreas, Y.Q.Wu, et al.*, "Effects of intermediate energy heavy-ion irradiation on the microstructure of rutile TiO2 single crystal", in Journal of the American Ceramic Society., **vol. 101**, no. 9. 2018, pp. 4357-4366

[2]. *C.N. Li, X.G.Bi, X.D.Sun*, "Experimental study on optimization of gas parameters for growth of rutile single crystal by numerical control flame melting method", in Journal of Functional Materials., **vol. 50**, no. 7. 2019, pp. 07101-07110

[3]. *H.H.Chou, Y.H. Liou, M. Calatayud, et al.*, "Modelling rutile TiO2 nanorod growth preferences: A density functional theory study", in Catalysis Today., **vol. 356**. 2020, pp. 49-55

[4]. *K. Dai, O. Daichi, F. Tomoteru*, "Thickness Effects on Crystal Growth and Metal-Insulator Transition in Rutile-Type RuO2 (100) Thin Films", in Physica Status Solidi., **vol. 257**, no. 9. 2020, pp. 1-5

[5]. *G.B.Song, J.K.Liang, F.S. Liu, et al.*, "Preparation and phase transformation of anatase-rutile crystals in metal doped TiO2/muscovite nanocomposites", in Thin Solid Films., **vol. 491**, no. 1/2. pp. 110-116

[6]. *F. Schroff, D. Kalenichenko, J. Philbin*, "FaceNet: a unified embedding for face recognition and clustering", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, USA: 2015, pp. 815-823

[7]. *K.M. He, X.Y. Zhang, S.Q. Ren*, "Spatial pyramid pooling in deep convolutional networks for visual recognition", in IEEE Transactions on Pattern Analysis and Machine Intelligence., **vol. 37**, no. 9, 2015, pp. 1904-1916

[8]. *K. N. Sung, D. M. Chen, J. J. Xiong, et al.*, "Heterogeneous Computing Meets Near-Memory Acceleration and High-Level Synthesis in the Post-Moore Era", in IEEE Micro., **vol. 37**, no. 4. 2017, pp. 10-18.

[9]. *K. Małgorzata, B. Alexander, T. Larysa*, "Synthesis of EMB-based Moore FSM with Splitting Set of Logical Conditions", in AIP Conference Proceedings. **vol. 1702**. 2015, pp. 1-4

[10]. *S. David*, "Mastering the game of Go with deep neural networks and tree search", in Nature, **vol. 529**, no. 7587. 2016, pp. 484-489

[11]. *M.D. Zeiler, R. Fergus*, "Visualizing and understanding convolutional networks", in Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland: Springer, 2014, pp. 818-833

[12]. *C.P. Shi, C. Tan, J. Zuo*, "Expression Recognition Based on Improved AlexNet Convolutional Neural Network", in Telecommunication Engineering, **vol. 60**, no. 9. 2020, pp. 1005-1012

[13]. *S. Sadhana, R. Mallika. A. Paul, et al.*, "An intelligent technique for detection of diabetic retinopathy using improved AlexNet model based convoluitonal neural network", in Journal of Intelligent & Fuzzy Systems., **vol. 40**, no. 4. 2021, pp. 7623-7634.

[14]. *X.L. Chen, X.W. Luo, Z.Y. Wu, et al.*, "An Improved Method Based on VGGNet for Refined Bathymetry from Satellite Altimetry: Reducing the Errors Effectively", in Geoscientific Model Development Discussions., 2022. pp. 1-16.

[15]. *D.X. Guo, K.Y. Chen, Y.Zhang, et al.*, "New template attack method for encryption chip based on VGGNet convolutional neural network", in Application Research of Computers, **vol. 36**, no. 9. 2019, pp. 2809-2855.

[16]. *Z. Wang, Z. Xu, D.Y. Zhang, et al.*, "A VGGNet-like approach for classifying and segmentingcoal dust particles with overlapping regions.", in Computers in Industry., **vol. 132**. 2021.

[17]. *Y. Pooja, M. Neeraj, R. Vinayakumar, et al.*, "EfficientNet convolutional neural networks-based Android malware detection", in Computers & Security. **vol. 115**. 2022.

[18]. *C.F. Zhao, W.Y.* Yao, *P. Tang, M.J. Yi*, "Arrhythmia Classification Algorithm Based on a Two-Dimensional Image and Modified EfficientNet", in Computational Intelligence & Neuroscience. 2022, pp. 1-10.

[19]. *T. Linh, Duong, T. Phuong*, "Automated fruit recognition using EfficientNet and MixNet", in Computers and Electronics in Agriculture. 2020 (C)

[20]. *L. T. Duong, P. T.Nguyen, Di Sipio C,et al.* " Automated fruit recognition using EfficientNet and MixNet", in Computers and Electronics in Agriculture,2020，171;105326

[21]. *M. Tan, Q. Le*, "EfficientNet: Rethinking model scaling for convolutional neural networks", in International Conference on Machine Learning, 2019, pp. 6105-6114.

[22]. *M.X. Tan*, "Rethinking Model Scaling for convolutional Neural Networks. Google. network, 2019.

[23]. *Y.Q. Jia, E. Shelhamer, J. Donahue*, "Caffe: convolutional architecture for fast feature embedding", in Proceedings of the ACM International Conference on Multimedia, Orlando, USA: 2014, pp. 675-678

[24]. *K. Cao, A.K. Jain*, "Latent orientation field estimation via convolutional neural network", in Proceedings of 2015 International Conference on Biometrics, Phuket, Thailand: 2015, pp. 349-356

[25]. *H.C. Shin, H.R. Roth, M.C. Gao*, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning", in IEEE Transactions on Medical Imaging., **vol. 35**, no. 5. 2016, pp. 1285-1298