Surface defect detection and classification of Si$_3$N$_4$ turbine blades based on convolutional neural network and YOLOv5

Dahai Liao$^{1,2}$, Zhihui Cui$^{1,2}$, Xin Zhang$^{1,2}$, Jun Li$^{1,3}$, Wenjie Li$^{1,3}$, Zuoxiang Zhu$^{1,3}$ and Nanxing Wu$^{1,2}$

Abstract
Due to the influence of mechanical vibration, high temperature creep and other factors, Si$_3$N$_4$ turbine blades are prone to surface defects. Besides, traditional algorithms are incapable to detect and classify surface defects simultaneously. Aiming at solving these problems, an algorithm for defect detection and classification of Si$_3$N$_4$ turbine blades based on convolutional neural network is proposed. The detection and classification network of this algorithm is optimized based on YOLOv5 network, the PAN structure and FPN structure of YOLOv5 are replaced by BiFPN structure. We establish the dataset of Si$_3$N$_4$ turbine blades, which is expanded by data enhancement. For the purpose of achieving a higher level of feature fusion, the PAN and FPN structures of the Neck part are replaced by BiFPN structure. As a result, the accuracy of detecting and classifying the surface defects by this algorithm is as high as 97.4%, and the detection speed is as low as 16ms. This optimized algorithm is able to solve the problems of traditional detection methods such as heavy workload, long time consuming and low accuracy. The algorithm provides a feasible approach for the quality detection of Si$_3$N$_4$ turbine blades and has certain engineering application value.

Keywords
Defect detection, Si$_3$N$_4$ turbine blades, YOLOv5 algorithm, BiFPN structure, convolutional neural network

Date received: 8 October 2021; accepted: 2 February 2022
Handling Editor: Chenhui Liang

Introduction
Si$_3$N$_4$ turbine blades have excellent properties such as low coefficient of thermal expansion, heat resistance, corrosion resistance, and wear resistance. Particularly, they still have high strength and high hardness and small specific density at 1400°C$^{1,3}$ which can effectively reduce the weight of aircraft engine and fuel consumption. Therefore, they are widely used in aerospace and military industry.$^{4,5}$ Owing to the influence of mechanical vibration, high-temperature creep, and other factors, Si$_3$N$_4$ turbine blades are prone to surface defects, such as snowflake defects, pit defects, scratch defects, etc. Abnormal abrasive as a sharp head is the cause of the pit defects. As a blunt head, the upper grinding plate produces crackle defects.$^6$ The crystal structure of the material changes to produce snowflake

---

1School of Mechanical and Electronic Engineering, Jingdezhen Ceramic University, Jingdezhen, People’s Republic of China
2Laboratory of Ceramic Material Processing Technology Engineering, Jingdezhen, People’s Republic of China
3National Engineering Research Center for Domestic and Building Ceramics, Jingdezhen, People’s Republic of China

Corresponding author:
Nanxing Wu, School of Mechanical and Electronic Engineering, Jingdezhen Ceramic University, Jingdezhen 333403, People’s Republic of China.
Email: ldhydz20187@163.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
defects. Incorrect machining pressure and unbroken hard abrasive particles during finishing produce graze and scratch defects.\(^7\) These defects do harm to the service life and performance of Si\(_3\)N\(_4\) turbine blades.\(^8,9\)

Traditional silicon nitride turbine blade detection methods include manual detection method, which is inefficient and has strong subjective factors. Meanwhile, traditional algorithms cannot simultaneously detect and classify defects,\(^10\)-\(^12\) which reduces the production speed of products and is not conducive to improving product quality. Along with the image processing technology, nondestructive testing technology is gradually applied in the field of surface defect detection of silicon nitride turbine blades.\(^13\) Non-destructive testing technology, assisted by image processing,\(^14\) is unable to cause secondary damage to Si\(_3\)N\(_4\) turbine blades. It is of vital significance to study and design a detection and classification method for surface defects of Si\(_3\)N\(_4\) turbine blades.\(^15,16\) The new method plays an increasingly important part in improving the stability of product quality,\(^17\) reducing labor intensity of workers,\(^18,19\) and improving the level of industry automation.

Based on the innovations of YOLOv4 and YOLOv5 algorithm in data enhancement, CNN(convolutional neural network), adaptive anchor frame calculation and adaptive image scaling, scholars at home and abroad have done in-depth research on YOLO (You Only Look Once) algorithm. Malta et al.\(^20\) proposed a task assistant model based on convolutional neural network. The data set of automobile engine parts was established. With training the neural network, using augmented reality glasses to identify the architecture of the system. The algorithm is able to accurately detect the parts in the real-time video stream; Neuhauser et al.\(^21\) put a kind of surface defect of extruded aluminum based on CNN classification and detection method. Using camera to continue extrusion profile, distinguishing the flawless surface defect and surface through YOLOv5 network. The defect surface was marked and classified. It has shown that this method has the advantages of fast detection speed, high classification accuracy, and high labeling classification accuracy. As a result, the method meets the needs of industrial production equipment in the detection of surface defects. Based on a single neural network, Redmon et al.\(^22\) proposed a target detection system, namely YOLO network, whose essence is to target detection problem into a regression problem. Given the input image, the target’s frame and its classification were directly regressed from multiple positions of the image. While ensuring the detection, the speed had been significantly improved. Zhang et al.\(^23\) proposed a bearing fault mode recognition method based on convolutional neural network. Combined with the ensemble empirical mode decomposition method, the adaptive decomposition was carried out. The eigen function components were screened, network images were extracted by CNN. The classification accuracy of bearing fault modes was improved by 4.26%. Based on the research of surface defect detection and classification algorithm by the scholars home and abroad, the surface defect detection and classification of Si\(_3\)N\(_4\) turbine blade was realized by improved YOLOv5 algorithm.

In order to accurately detect the surface defects of Si\(_3\)N\(_4\) turbine blades and improve the efficiency of defect classification, a defect detection and classification method based on CNN and YOLOv5 after optimizing is proposed. Deep CNN is used to extract semantic features of defects on the surface of Si\(_3\)N\(_4\) turbine blades. The PAN structure and FPN structure of YOLOv5 are replaced by BiFPN structure. Images are input into the network for training and verification. Images in the test set are used to test the trained network. The method we proposed has the advantages of high accuracy and recognition speed. It has certain guiding significance to improve the detection and classification of surface defects of Si\(_3\)N\(_4\) turbine blades.

The structure of the article is as follows. Firstly, Chapter 2 introduces the Si\(_3\)N\(_4\) turbine blade surface defect detection and classification platform, including Chapter 2.1 on the construction of detection environment and Chapter 2.2 on the fabrication of surface defect data set of Si\(_3\)N\(_4\) turbine blades. Besides, Chapter 3 describes the surface defect image processing of Si\(_3\)N\(_4\) turbine blade, Chapter 3.1 introduces the principle of data augmentation, and Chapter 3.2 is about the data set division and annotation process. What’s more, Chapter 4 is about the detection and classification method for surface defects of Si\(_3\)N\(_4\) turbine blades, Chapter 4.1 introduces defect detection and classification process, Chapter 4.2 describes the preparation of YOLOv5 algorithm network partition, Chapter 4.3 introduces the check for the model network structure, Chapter 4.4 describes the improvement of detection model, and Chapter 4.5 mainly introduces the BiFPN structure. In addition, Chapter 5 is the exposition of loss function calculation and algorithm performance analysis, Chapter 5.1 mainly introduces the loss of the model, Chapter 5.2 describes the model training process, and Chapter 5.3 is about the model evaluation. Next, Chapter 6 is the experimental results and analysis. Last but not least, Chapter 7 is conclusion.

**Si\(_3\)N\(_4\) turbine blade surface defect detection and classification platform**

**Detection environment setup**

Si\(_3\)N\(_4\) turbine blade surface cleanliness of sample will be about image quality, affect the accuracy of experimental results.\(^24\) For the purpose of getting high quality defect images, we build Si\(_3\)N\(_4\) turbine blade surface
defect detection and classification system. This system, shown in Figure 1, is mainly composed of PC workstations, stereo microscope, industrial camera, image acquisition card, support platform, coaxial light, circular shadowless light, and camera obscura. Based on the characteristics of silicon nitride turbine blades, the light source device we used are UD115W LED circular shadow light source and C20W coaxial light source. Mer2-2000-19u3m is used for industrial cameras, HN-2520-20M-C1/1x lens is used for stereoscopic microscope. The operating system of PC workstation is Windows10, equipped with Nvidia GTX1080Ti GPU and Intel i7-7700 CPU. Compared with other defect detection systems, this system performs much better.

The working process of detecting surface defects of silicon nitride turbine blades in this system can be divided into the following steps:

1. Installing the debugging device and placing the Si3N4 turbine blade on the workbench.
2. The surface defect images of Si3N4 turbine blades are collected and transmitted to PC.
3. The defect detection and classification method based on CNN and optimized YOLOv5 network is used to detect and classify the surface defects of silicon nitride ceramic spheres.

**Fabrication of surface defect data set of Si3N4 turbine blades**

The foundation of defect detection and classification is the establishment of high-quality data set. The diversity and balance of sample types should be guaranteed when making the data set. In order to achieve satisfactory experimental results, it is necessary to select the data set that meets the certain standard. After training the CNN, the recognition results obtained have a certain orientation.

We have prepared 500 Si3N4 turbine blade experimental samples through experiment. Surface defect images are collected by the image acquisition system, and 1000 surface defect images are collected as the data set. Surface defects are detected by YOLOv5 network, five defects are mainly collected, namely pits, scratches, cracks, snowflakes, and grazes, the quantity of each defect is 200. The five surface defects are shown in Figure 2. The pit defects are showed partial tissue stripping, and the blade surface is uneven. The scratch defects are linear and shallow. The crack defects are annular or straight, the grain of which is obvious. The snowflake defects show large net white spots or pitting. The abrasion defects are banded and have no obvious depression.

**Surface defect image processing of Si3N4 turbine blade**

**Data enhancement**

With the deepening of CNN, the required learning parameters increase and the structure of learning model becomes more complex. When the data set is small, too many parameters will fit all the characteristics of the data set instead of the commonness between the data, resulting in over-fitting. In the field of Si3N4 turbine blade surface defect detection, for the lack of defect images, the database on a smaller scale. It is difficult to get a lot of defect images with labels. Small scale database leads to the overfitting phenomenon of the convolution neural network. Data enhancement technique is introduced, namely the original image data is enhanced data transformation. Adding new samples plays an important role in extending the dataset to five times of the original one. The enhanced images, shown in Figure 3, improve the generalization ability of the machine learning model. Data augmentation is a method to add new training data on the basis of the
original dataset. It is achieved by transforming the data of training set into a new data by using domain-specific technology. Create new defect images that belongs to the same category as the original images. Through resizing images and scaling pixel values, the dataset has been extended. As a result, representativeness of the defect dataset has been ensured.

Common data augmentation methods, used to increase the training set and improve the generalization ability, includes random cutting, image flipping, rotation, noise addition, affine transformation, etc. The data enhancement methods and purpose are shown in Table 1. The data enhancement methods adopted are as follows:

1. Horizontal or vertical flip: flip the surface defect images in horizontal or vertical dimensions.
2. Rotation: rotate the image by a certain angle, the common rotation is 90° or 180°.
3. Noise disturbance: the RGB of each pixel of the image is randomly disturbed, and the
commonly used noise mode is salt and pepper noise and Gaussian noise.

(4) Image color adjustment: by adjusting the image brightness, contrast, saturation to simulate different shooting environment.

**Data set partitioning and annotation**

Around 5000 surface defect images are obtained by data enhancement technology, and the defect images are divided into training set, verification set, and test set in a ratio of 7:2:1. What should be noted is that the ratio of the training set, validation set, and test set appear in any proportion. In the field of machine learning deep learning, it is very important to divide training set and test set reasonably. For machine learning, the typical ratio of training set to test set is 7:3 or 8:2. However, according to the characteristics of this data set, we have tried the ratio of 6:2:2 and 5:3:2, the training effect performs bad, the mAP and precision are low. Finally, the ratio has been changed to 7:2:1, the radio is unable to degrade the performance of the YOLOv5 algorithm we proposed. The sample distribution of data set is shown in Table 2.

The training set is used to learn the sample data set and estimate the parameters in the model, so that the model can reflect the reality and predict the future or other unknown information. The main function of the validation set is to adjust the hyperparameters of the model and evaluate the capability of the model. The test set evaluates the generalization capability of the model.29

Data augmentation, also known as data enhancement, is the process of making a limited amount of data produce value equivalent to more data without adding substantial data. For small data set, the neural network will definitely overfit with proper training. Since the training set is not enough to reflect the characteristics of the data set, the test accuracy is obviously low. In order to avoid the over-fitting phenomenon, data augmentation technology is adopted to expand the data set. Not only can data augmentation improve training efficiency and model classification performance but also enhance the abilities of model positioning and generalization.

In target detection, it is necessary to mark the original image and determine the detection target. LabelImg image annotation tool is used in this system to mark each defect in the image of Si$_3$N$_4$ turbine blade. The main function of this tool is to define the type and quantity of annotation and record the location information of the annotation object. The steps of drawing defect location mania are as follows: (1) observe the defect image, (2) select the corresponding defect type, and (3) draw the position box at the defect. The integer ID corresponding to each defect image is 0-pit defect, 1-scratch defect, 2-crack defect, 3-snowflake defect, and 4-graze defect.

### Detection and classification method for surface defects of Si$_3$N$_4$ turbine blades

#### Defect detection and classification process

Based on CNN and YOLOv5 network, the PAN structure of Neck part and FPN structure are replaced by BiFPN structure to obtain the defect detection and classification method of the system. The algorithm structure has been greatly improved. As for the situation that the pit, the snowflake, the scratch, the graze, and the crack defects occur simultaneously, the algorithm has the capacity of detecting and classifying the defects. The optimized YOLOv5 have the ability of keeping good performance when different defects appear simultaneously. The detection process is shown in Figure 4. The defect detection process is as follows:

1. With data enhancement method, the defect images are collected, and these images are

| Dataset     | Pits | Scratches | Grazes | Snowflakes | Cracks | Total |
|-------------|------|-----------|--------|------------|--------|------|
| Training set| 695  | 683       | 675    | 720        | 727    | 3500 |
| Verification set | 221  | 196       | 213    | 189        | 181    | 1000 |
| Test set    | 89   | 96        | 103    | 106        | 106    | 500  |

Figure 4. Flow chart of defect detection and classification.
labeled and classified, and the optimized YOLOv5 algorithm is used to train the defect images.

(2) According to Backbone network, different image granularity is aggregated to form CNN with image characteristics. Based on Neck network, images features are combined, which contributes to transferring the images to the prediction layer. Image features are predicted by Head network. LabelImg image annotation tool is used to generate the defect marking box and predict the defect category.

(3) Test and reason the learning model well trained, and analyze the results we has obtained through the defect detection algorithm.

The preparation of YOLOv5 algorithm network partition

The YOLOv5 algorithm is a single-stage target detection algorithm. Considering the characteristics of small surface defects and many kinds of defects, the network structure of YOLOv5 algorithm has been analyzed. In order to improve the training speed and detection accuracy of the algorithm, the main improvement ideas are as shown in Figure 5. It is the improvement of YOLOv5 that contributes to comprehensively improve the performance.

The input module. In the stage of model training, the improvement ideas are proposed. Compared with the previous model, the functions include data augmentation, adaptive anchor frame calculation, and adaptive image scaling has been added. As a result, the data set is amplified. The surface defect images are adjusted adaptively. The anchor frame is drawn according to the area of surface defect.

The Backbone part. As a classification network with excellent performance, the Backbone part is composed of the Focus structure, the BottleNeckCSP structure, and the SPP structure. As the reference network, Focus adopts interlaced sampling splicing structure. The input images are clipped through the slice operation. The CSP structure divides the original input into two branches and performs convolution operation to halve the number of channels. Therefore, the model has the ability to learn more features.
The neck part. In order to obtain higher semantic information, the SPP structure selects the feature layers of the last three shapes for the next operation. The input feature layers are maximized. The pooled results are stacked and convolved for three times. With the SPP structure, the most significant context features can be isolated. The PANet has only one top-down path or one town-top path. On the basis of PAN structure, an extra edge is added for the case that the input and output nodes are at the same level. More features are fused without increasing cost. As a consequence, this part has the ability of obtaining the information in the pictures and extracting features.

The head part. The anchor frame mechanism of the output layer is the same as that of YOLOv4. Used to complete the output of target detection results, there are two branches in the output side, namely classification branch and regression branch. In order to further improve the detection accuracy of the algorithm, the major improvements have been put forward, including the loss function during training and NMS. The loss function of YOLOv5 algorithm consists of three parts: category loss, position loss, and confidence loss. It is through this section that feature graphs and prediction boxes can be output.

Check the model network structure

The structural network of YOLOv5 algorithm is mainly composed of Backbone, Neck, and Head, which is shown in Figure 6. The three parts are described as follows:

Backbone mainly includes Focus, CSP, and SPP. The function of Focus structure is to slice $608 \times 608 \times 3$ input images into $304 \times 304 \times 12$ feature images, and output $304 \times 304 \times 32$ feature images through a 32-channel convolution layer. The function of CSP structure is to split the feature graph into two parts, one part is used for convolution operation, the other is used for feature fusion with the result of convolution operation. SPP structures fuse multi-scale features by pooling different convolution kernels.

Neck is located in the middle layer between Backbone and Head, and adopts FPN structure and PAN structure, which is mainly used to generate feature pyramid. Feature pyramid is able to enhance the detection of objects with different scale of the model, so as to identify the same object with different size and scale.

Head, the last detection part, outputs feature maps of three scales for detection of small, medium, and large objects. Each grid contains three prediction boxes, in which the confidence of the object and the location of the object are contained. Through non-maximum suppression NMS, the repetitive and redundant prediction boxes are eliminated, and the information of the prediction box with the highest confidence is retained. As a result, target detection is completed.

Improvement of detection model

The key to improve the performance of the model is to fuse the features of different scales. In order to enhance the feature fusion of different layers, the FPN structure and the PAN structure are adopted in Neck part of YOLOv5 network, which is conducive to multi-scale prediction.

FPN structure establishes a top-down path for feature fusion. The fused maps with higher semantic information are used for prediction. As a result, the accuracy has been greatly improved. The feature map of the upper level has stronger semantic information, which is beneficial to object classification; while the feature map of the lower level has stronger location information, which is beneficial to object localization. Consequently, the structure improves the semantic information of the predicted feature graph, but loses a few location information. On the basis of FPN structure, a new path from bottom to top is created to transmit location information to the prediction feature graph, namely PAN structure. The function of the structure are as follows: making the prediction feature graph have high semantic information and location information, improving the accuracy of target detection task. It has been proved that the effectiveness of bidirectional fusion. Since the bidirectional fusion of PAN structure is relatively simple, the PAN structure is replaced by the more complex BiFPN structure. The three structural models are shown in Figure 7.

BiFPN structure

BiFPN structure is shown in Figure 7(c). The blue part is the top-down pathway, conveying semantic
information of high-level features. The red part is the bottom-up path, conveying the location information of low-level features. The purple part is a new edge added between the input node and the output node of the same layer. Compared with PAN structure, there are three major design changes of this structure:

1. Delete the node with only one input edge: in the case of a node with only one input edge without feature fusion, its contribution to blend different feature networks with different features will be small. The existence of the node does nothing to the network. Deleting the node is conducive to simplify the bidirectional network.

2. In order to incorporate more features without too much cost, in the case that the original input node and output node are in the same layer, an extra edge is added to BiFPN structures between the original input node and the output node.

3. There are two paths in PAN structure, one top-down path and one bottom-up path. Different from PAN structure, for the purpose of achieving higher level feature fusion, BiFPN structure treats each bidirectional path as a feature network layer, and loops the same layer several times.

**Loss function calculation and algorithm performance analysis**

**Loss**

The loss function of the YOLOv5 algorithm consists of three parts: category loss, position loss, and confidence loss. Category loss and confidence loss are calculated with the method of cross entropy. And it is shown in formula (1)–(3):

\[
\text{Loss} = L_{\text{cla}} + L_{\text{GIOU}} + L_{\text{con}}
\]

\[
L_{\text{con}} = \lambda_{\text{obj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij} \left[ -\hat{C}_i \ln C_i - (1 - \hat{C}_i) \ln (1 - C_i) \right] + \lambda_{\text{nobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij} \left[ -\hat{C}_i \ln C_i - (1 - \hat{C}_i) \ln (1 - C_i) \right]
\]

\[
L_{\text{cla}} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} I_{ij} \left[ -\hat{C}_i \ln C_i - (1 - \hat{C}_i) \ln (1 - C_i) \right]
\]

Where, \(S^2\) represents the number of grids divided; \(B\) denotes the number of predicted boundary boxes per grid; \(I_{ij}\) represents whether the \(J\)TH prior box of the \(I\)TH grid has a target to be predicted; \(l_{ij}^{\text{obj}}\) denotes whether the \(J\)TH boundary box of the \(I\)TH grid has a target that does not need to be predicted; \(\lambda_{\text{obj}}\) and \(\lambda_{\text{nobj}}\) are the weight coefficient of mesh with or without target; \(C_i\) and \(\hat{C}_i\) are the confidence value of predicted target and actual target; \(C\) is the target category predicted by the boundary box; \(p_i(c)\) represents the prediction probability of the target category to which the \(I\)TH grid belongs when the target is detected; \(\hat{p}_i(c)\) represents the actual probability that the \(I\)TH network belongs to the target category when it detects the target. GIOU loss function is adopted for position loss. As for an important indicator to evaluate the performance of target detector. IoU stands for overlap degree and intersection ratio, which can be used to measure the similarity between real frame and predicted frame. The formula is shown in (4)–(6):

\[
\text{IoU} = \frac{A \cap B}{A \cup B}
\]

\[
\text{GIOU} = \text{IoU} - \frac{|A_c - U|}{|A_c|}
\]

\[
L_{\text{GIOU}} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} (1 - \text{GIOU})
\]
Where, $A$ denotes the area of the real box, $B$ stands for the area of the predicted box. IoU is directly used as the position loss. When there is no overlap between the prediction box and the real frame, IoU is 0. In this case, the distance between the two boxes cannot be reflected by IoU, and the gradient of the loss function is 0, which cannot be optimized. $A_c$ represents the area of the smallest closed box containing the prediction box and the real box, and $U$ represents the area of the prediction box and the real box. The larger the GIoU of the two bounding boxes is, the network will be optimized in the direction of high overlap between the prediction box and the real box.

**Model training**

In order to test the performance of the model in dataset, the established data set is input into the optimized YOLOv5 network for model training. Epoch is set to 300, batchsize to 16, and initial learning rate to 0.0001. Adam optimizer is selected for iterative optimization. The variation curves of each loss function in the training process are shown in Figure 8. It can be seen that with the increase of iterations, the neural network is gradually optimized and the positioning accuracy of predicted frame is getting higher and higher.

**Model evaluation**

The combination of the real defect category and the learner’s prediction category is divided into True Positive, False Positive, True Negative, and False Negative cases. A true example indicates that a defect is correctly detected, a false positive example indicates that no defect is incorrectly detected as a defect, a true negative example indicates that no defect is correctly detected, and a false negative example indicates that a defect is incorrectly detected as a defect.

Precision, recall, and mAP are used to evaluate the defect extraction ability of neural network. The accuracy rate refers to the proportion of the correctly identified samples to the correctly identified Si$_3$N$_4$ turbine blade surface defects samples and non-Si$_3$N$_4$ turbine blade surface defects samples. Recall rate refers to the

![Figure 8. Training and validation loss: (a) Loss(GIoU), (b) Loss(con), (c) Loss(cla), and (d) Loss.](image-url)
proportion of correctly identified samples of surface defects on Si$_3$N$_4$ turbine blades among all samples identified as surface defects on silicon nitride turbines. According to the relationship between the precision and recall, the precision-recall curve (P–R curve) can be drawn. Accuracy of all categories (AP) stands for the area of the area enclosed by the PR curve and the coordinate axes. Accuracy of all categories (AP) refers to the area of the zone enclosed by the PR curve and the coordinate axes. Map of all categories can be obtained by calculating AP of all categories and taking the average value. The calculation formula is shown in (7)–(10).

\[
P = \frac{TP}{TP + FP} \tag{7}
\]

\[
R = \frac{TP}{TP + FN} \tag{8}
\]

\[
AP = \frac{1}{N} \sum p(r) \tag{9}
\]

\[
mAP = \frac{1}{N} \sum \frac{AP_i}{N} \tag{10}
\]

### Experimental results and analysis

There are 5000 images using in the defect detection system to verify the algorithm, the values of each evaluation index are shown in Table 3. The recall rate of the optimized YOLOv5 model is 97.4%, which indicates that 97.4% surface defects are correctly detected and classified. Recall values of the first M positive samples are taken to calculate a maximum precision value. AP value is averaged for the obtained precision value. The map value is the average of all AP values. As for the function of AP and mAP, the AP value measures how well the learned model performs in each category. However, the mAP value measures how well the learned model performs across all categories.

Due to the defect area is larger than the others, the characteristics of snowflake defects are obvious and easy to identify, all evaluation indexes of snowflakes are high. There are similarities between the pits and snowflake defects. Since the pit defects are small in area, single in shape, and not obvious in characteristics, it is difficult to detect and classify pits accurately. Therefore, all the evaluation indexes of pits are relatively low. On account of the adaptive image scaling technology in the YOLOv5 algorithm, the image size of the input image does not need to be considered.

When it comes to the advantages of data augmentation, not only can this increase the amount of training data but also improve the generalization ability of the model. It is data augmentation that add noise data to improve the robustness of the model. For a dataset of 1500 defect images, the values of each evaluation index are shown in Table 4. The recall rate of the optimized YOLOv5 model is 95.6%, which indicates that 95.6% surface defects are correctly detected and classified.

Compared to the results of the data augmentation dataset, the mAP is lower and the comprehensive performance of the dataset of 1500 defect images is poor. As a result, for small sample datasets, data augmentation plays a significant role in improving the performance of the model.

The optimized YOLOv5 algorithm keeps good performance when different defects appear simultaneously. Different defect images in the dataset are fused together to form new representative surface defect images by image enhancement technology. The model is improved and trained, and the model training results are obtained as shown in Figure 9. As shown in the figure, when multiple defects exist in the surface defect image at the same time, the defects can still be marked. Therefore, it can be proved that for an image with multiple defects at the same time, the model can recognize and mark multiple defects.

As the function of this detection system is to detect and classify the defects of Si$_3$N$_4$ turbine blades, in order to fully evaluate the optimized YOLOv5 algorithm, the Faster RCNN algorithm and YOLOv5 algorithm are introduced for comparative experiments. Compared with the three algorithms, the precision rate, recall rate, average accuracy value, and detection speed are selected as evaluation indexes, the target detection and classification abilities of different detection algorithms are evaluated scientifically and reasonably. The comparison results are shown in Figure 10 and Table 5.

| Class      | Precision | Recall | AP | mAP |
|------------|-----------|--------|----|-----|
| All        | 0.974     | 0.974  | /  | 0.986 |
| Graze      | 1         | 0.928  | 0.994 | /   |
| Scratch    | 0.972     | 1      | 0.996 | /   |
| Crackles   | 0.99      | 1      | 0.996 | /   |
| Pit        | 0.907     | 0.943  | 0.949 | /   |
| Snowflake  | 1         | 1      | 0.995 | /   |

### Table 3. Validation results of model.

| Class      | Precision | Recall | AP | mAP |
|------------|-----------|--------|----|-----|
| All        | 0.919     | 0.956  | /  | 0.981 |
| Graze      | 1         | 0.909  | 0.986 | /   |
| Scratch    | 1         | 0.999  | 0.999 | /   |
| Crackles   | 0.842     | 0.833  | 0.975 | /   |
| Pit        | 0.755     | 1      | 0.915 | /   |
| Snowflake  | 1         | 1      | 0.995 | /   |

### Table 4. Model validation results of 200 defect image data sets.
surface defects of Si$_3$N$_4$ turbine blades, and the prediction box can select defects completely. YOLOv5 algorithm and Faster RCNN algorithm are poor in detecting small defects. Both algorithms can detect defects, but the prediction box cannot accurately select the defect box and its confidence is not as high as the optimized YOLOv5 algorithm.

As can be seen from the table, the precision, recall, and average accuracy values of the optimized YOLOv5 algorithm are higher than those of the other algorithms. It is indicated that compared with the other algorithms, the optimized YOLOv5 algorithm can better identify and classify defect images. In terms of defect detection speed, the optimized YOLOv5 algorithm is the fastest, as low as 16 ms, 69.6% Faster than the YOLOv5 algorithm and far faster than the Faster RCNN algorithm, which better meets the task requirements of defect detection and classification of silicon nitride turbine blades. Among these algorithms, the optimized YOLOv5’s mAP performs best, the YOLOv3 and the YOLOv5 both perform worst.

The shortcoming of the optimized YOLOv5 algorithm lies in its dependence on the size of data sets. Different from small sample deep learning algorithm, the YOLOv5 algorithm needs the support of large sample data set. For small data sets, the neural network will inevitably overfit with proper training. Because the training set is not enough to represent the problem, the test accuracy is obviously low. Because the training set is not enough to reflect the characteristics of the data set, the testing accuracy is obviously low. Thus, the model does not perform well. However, for large sample data sets, this model can achieve better results of defect recognition and classification.

### Conclusion

Due to mechanical vibration and high temperature creep and other factors, surface defects of Si$_3$N$_4$ turbine blades cannot be detected and classified by traditional algorithms at the same time. Therefore, a method for surface defects detection and classification of Si$_3$N$_4$ turbine blades based on optimized YOLOv5 network and CNN is proposed. According to the defect image characteristics of Si$_3$N$_4$ turbine blade surface, a defect detection system for Si$_3$N$_4$ turbine blade surface is built. The Neck part of YOLOv5 has been optimized, and the PAN and FPN structure are changed to a more complex bi-directional fusion BiFPN structure, which improved the detection accuracy and detection speed. It is shown that this method used to detect and classify surface defects of Si$_3$N$_4$ turbine blades makes an excellent performance. The defects are accurately and completely segmented. The classification accuracy is 97.4%, and the detection speed is as low as 16 ms, which meets the industrial requirements.

When it comes to the further research direction of the model, it lies in continuous improvement of the structure according to the latest proposed algorithm. Since the Transformer algorithm was proposed as a machine translation algorithm, it has gradually replaced feature extraction led by CNN and RNN. To be completely recursive, the idea of transformer’s algorithm is to use the attentional mechanism to handle the dependencies between inputs and outputs. Transformer is undoubtedly an overwhelming improvement over the recursive neural network. However, the ability of Transformer algorithm to obtain local information is not as strong as that of YOLOv5. There are also problems in location information coding. The future
research direction is to combine the YOLOv5 algorithm with the Transformer algorithm to obtain a training model with stronger adaptive ability.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work are supported by are supported by the National Natural Science Foundation of China [grant number: 51964022] and the Key Foundation of Jiangxi Province of China [grant number 2021A C B 204012] and Jiangxi Youth Foundation [grant number 2021A A B 214033].

Figure 10. The comparison diagram of model prediction results.

ORCID iD
Nanxing Wu https://orcid.org/0000-0001-8544-4417

References
1. Yu L, Yang Z and Zhang J. High-temperature mechanical properties of Si₃N₄ composite ceramics prepared by freeze gelation. J Chin Ceram Soc 2021; 49: 1439–1445.
2. Pan S, Zhang W, Zhang W, et al. Diagnostic model of coronary microvascular disease combined with full convolution deep network with balanced cross-entropy cost function. IEEE Access 2019; 7: 177997–178006.
3. Yang T, Wang L and Gu L. Processing defect analysis and nondestructive evaluation technology for Si₃N₄ bearing ball. Acta Armamentarii 2007; 38: 353–357.
4. Zhou L, Gong J and Li B. Image information restoration of automotive strip steel surface based on sparse representation. J Hunan Univ (Nat Sci) 2021; 48: 141–148.
5. Ren Y, Zhu C and Xiao S. Object detection based on fast/faster RCNN employing fully convolutional architectures. *Math Probl Eng* 2018; 2018: 1–7.
6. Yang T, Wang L and Gu L. Processing defect analysis and nondestructive evaluation technology for Si$_3$N$_4$ bearing ball. *Acta Armamentarii* 2007: 28: 353–357.
7. Yang T, Wang L, Gu L, et al. Surface defects analysis of Si$_3$N$_4$ bearing balls in lapping process. *Lubr Eng* 2006: 45–47 + 70.
8. Zhang J, Liu J, Wang Z, et al. Tribological behavior and lubricating mechanism of Si$_3$N$_4$ in artificial seawater. *Ceram Int* 2020; 46: 14361–14368.
9. Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation. In: *CVF conference on computer vision and pattern recognition (CVPR)*, 2018, vol. 54, pp.564–569.
10. Wu Y, Sha Y and Li S. Wear mode transition in micro-abrasive of silicon nitride ceramic balls. *Ordnance Mater Sci Eng* 2021; 44: 49–57.
11. Kingma D and Ba J. Adam: a method for stochastic optimization. *Comput Sci* 2014; 6980: 775–783.
12. Rao WD, Liu Y, Cheng LJ, et al. Densification mechanism of stereolithographical dense Si$_3$N$_4$ ceramics with CeO$_2$ as sintering additive by field assisted sintering. *J Central South Univ* 2021; 28: 1233–1243.
13. Zhao Z. Review of non-destructive testing methods for defect detection of ceramics. *Ceram Int* 2021; 47: 4389–4397.
14. Ming W, Jia H, Zhang H, et al. A comprehensive review of electric discharge machining of advanced ceramics. *Ceram Int* 2020; 46: 21813–21838.
15. Hu G, Wang N, He W, et al. Unsupervised surface defect detection method based on image inpainting. *J South China Univ Technol (Nat Sci Ed)* 2021; 49: 76–85 + 124.
16. Wang J, Yin X, Wang L, et al. Hashing-based non-maximum suppression for crowded object detection. *J Comput Vision Pattern Recogni* 2020; 200: 84–90.
17. Deng J, Li W and Wu S. Detection of defects in glass ceramics based on grayscale and geometric features. *Opt Tech* 2021; 47: 428–431.
18. Jeong EY, Kim J, Jang WH, et al. A more reliable defect detection and performance improvement method for panel inspection based on artificial intelligence. *J Inf Disp* 2021; 22: 127–136.
19. Li X, Ma H and Lin Y. Structural damage identification based on convolution neural network. *J Vib Shock* 2019; 38: 167–175.
20. Malta A, Mendes M and Farinha T. Augmented reality maintenance assistant using YOLOv5. *Appl Sci* 2021; 11: 254–260.
21. Neuhauser FM, Bachmann G and Hora P. Surface defect classification and detection on extruded aluminum profiles using convolutional neural networks. *Int J Mater Forming* 2020; 13: 591–603.
22. Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection. In: *2016 IEEE conference on computer vision and pattern recognition (CVPR)*, 2016, vol.16, pp.141–147.
23. Zhang A, Huang J and Ji S. Bearing fault pattern recognition based on image classification with CNN. *J Vib Shock* 2020; 39: 165–171.
24. Cappugi L, Castorrini A, Bonfiglioli A, et al. Machine learning-enabled prediction of wind turbine energy yield losses due to general blade leading edge erosion. *Energy Convers Manag* 2021; 245: 256–273.
25. Zhang Z, Shao M, Wang L, et al. A novel domain adaptation-based intelligent fault diagnosis model to handle sample class imbalanced problem. *Sensors* 2021; 21: 3382–3400.
26. El Yousfi B, Soualhi A, Medjaher K, et al. New approach for gear mesh stiffness evaluation of spur gears with surface defects. *Eng Fail Anal* 2020; 116: 104740–104764.
27. Yang P, Su Y and Zhang Z. A study on rolling bearing fault diagnosis based on convolution capsule network. *J Vib Shock* 2020; 39: 55–62 + 68.
28. Geng P, Tang Y and Fan H. Research on influence of data enhancement on deepfake detection model. *Comput Eng Appl* 2021; 57: 10–16.
29. Zhang C, Guo Y and Li M. Review of development and application of artificial neural network models. *Comput Eng Appl* 2021; 57: 57–69.
30. Wang W-T, Zhu S-A and Tang W. Adaptive non-maxima suppression edge detection algorithm and FPGA design to achieve. *J Chin Comput Syst* 2014; 35: 1926–1930.