scientific reports

Check for updates

OPEN

Deep learning-based defects detection of certain aero-engine blades and vanes with DDSC-YOLOv5s

Xubo Li[1], Wenqing Wang[1]✉, Lihua Sun[1], Bin Hu[1], Liang Zhu[1] & Jincheng Zhang[2]

When performed by a person, aero-engine borescope inspection is easily influenced by individual experience and human factors that can lead to incorrect maintenance decisions, potentially resulting in serious disasters, as well as low efficiency. To address the absolute requirements of flight safety and improve efficiency to decrease maintenance costs, it is imperative to realize the intelligent detection of common aero-engine defects. YOLOv5 enables real-time detection of aero-engine defects with a high degree of accuracy. However, the performance of YOLOv5 is not optimal when detecting the same defects with multiple shapes. In this work, we introduce a deformable convolutional network into the structure of YOLOv5s to optimize its performance, overcome the disadvantage of the poor geometric transformability of convolutional neural networks, and enhance the adaptability of feature maps with large differences in the shape features. We also use a depth-wise separable convolution to improve the efficiency of multichannel convolution in extracting feature information from each channel at the same spatial position while reducing the increased computational effort due to the introduction of deformable convolution networks and use k-means clustering to optimize the size of anchor boxes. In the test results, mAP50 reached 83.8%. The detection accuracy of YOLOv5s for common aero-engine defects was effectively improved with only a 7.9% increase in calculation volume. Compared with the metrics of the original YOLOv5s, mAP@50 was improved by 1.9%, and mAP@50:95 was improved by 1.2%. This study highlights the wide application potential of depth science methods in achieving intelligent detection of aero-engine defects. In addition, this study emphasizes the integration of DDSC-YOLOv5s into borescope platforms for scaled-up engine defect detection, which should also be enhanced in the future.

As the main power source of airplanes, normal operation of an aero-engine is the primary prerequisite for flight safety. Because an aero-engine works in an extreme environment with high temperature, high stress, and high speed and under the joint action of a working load and vibration load, severe environmental and operating conditions may result in increasing component defects, the engine is prone to failure. Meanwhile, foreign object damage may further harm the engine and the aircraft[1–3]. Due to the high cost of aero-engine maintenance[4] and the strict requirements of workers and work sites for dismantling and assembling aero-engines, borescope inspection is widely used in condition-based maintenance. Borescope inspection is performed by professionals to detect the internal condition of the engine through borescope instruments. Borescope inspection of defects inside an aero-engine is usually performed manually via video, which not only requires the inspector's expertise but also makes the result susceptible to subjective experience of the operator. Meanwhile, because of the complex and changeable background and the lack of light inside the engine, the inspector needs to manipulate the probe lens angle during the inspection to obtain good quality images. This slows the inspection process, and then tiny defects are easily missed by experienced inspectors, which causes intensive work with low efficiency[5,6]. Therefore, we couple deep learning with computer vision, apply an object detection method to aero-engine borescope inspection to detect defects, and use an improved YOLOv5s[7] algorithm to locate and identify common aero-engine defects.

During the borescope process, the internal condition of the aero-engine is collected through an endoscope and output in the form of images without disassembling the engine[8]. At present, the relevant research on aero-engine

[1]School of Automation, Xi'an University of Posts and Telecommunications, Xi'an 710121, China. [2]College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211156, China. ✉email: wwq@xupt.edu.cn

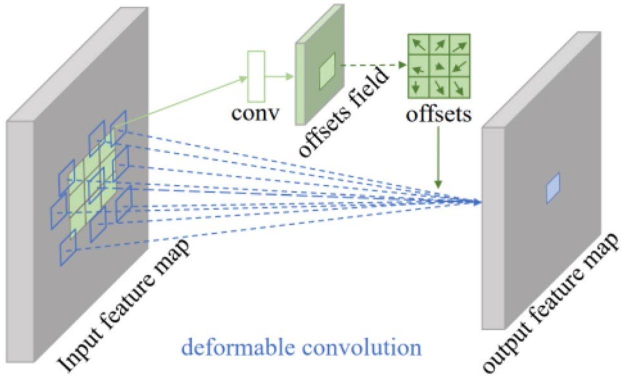

Figure 3. 3×3 deformable convolution process.

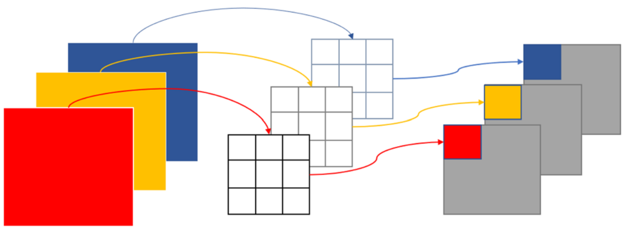

Figure 4. Depth-wise convolution process.

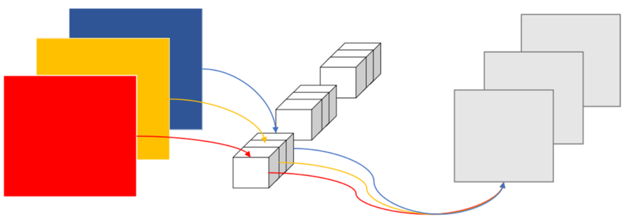

Figure 5. Pointwise convolution process.

feature information of the different channels at the same spatial location cannot be used effectively. Therefore, pointwise convolution is needed to combine these feature maps to generate a new feature map.

The pointwise convolution operation can be regarded as a conventional convolution operation with a kernel size of 1×1. The feature map processed by depth-wise convolution is subjected to pointwise convolution to ensure that the feature map processed by depth-wise convolution is weighted and combined in the depth direction to generate a new feature map. The number of output features is the same as the number of kernels. The pointwise convolution process can be seen in Fig. 5.

Improved YOLOv5s network

Most object detection backbone networks use deep learning networks to extract features. As the most widely used deep learning network, convolutional neural networks (CNNs) can effectively extract image features, but CNN-based object detection is not the best approach. Deformable convolution enhances the CNN's modeling ability for deformed objects[15] and extracts image features better. Adding a DCN to the YOLOv5s network structure overcomes the shortcomings of the poor geometric transformability of CNNs and enhances the ability to adapt feature maps with large differences in shape features. Conventional convolution extracts features in one convolution, but depth-wise separable convolution extracts features by two convolutional layers, one for extracting features and the other for combining features from different channels. This factorization can effectively reduce the computational effort and the size of the trained model. This paper improves YOLOv5s to propose an aero-engine defect detection method. Based on the original YOLOv5s, the conventional convolution module of the network is replaced by the DConv module for the DCN introduced into YOLOv5 and DSConv to introduce DSC.

We performed a series of experiments in the previous period. When the DConv module is placed in front of the structure, because the feature map contains too many features, the increased offset of the DCN for extracting deformed features increases the computation to the extent that the GPU cannot operate. When the DConv

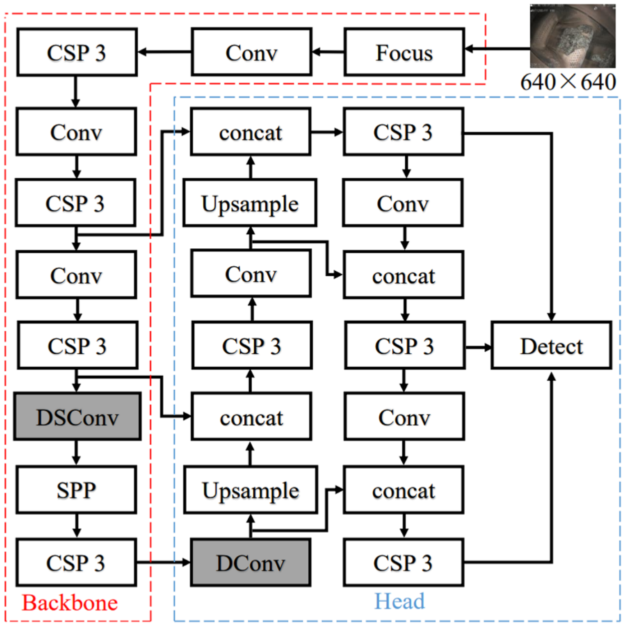

Figure 6. Improved YOLOv5s network structure.

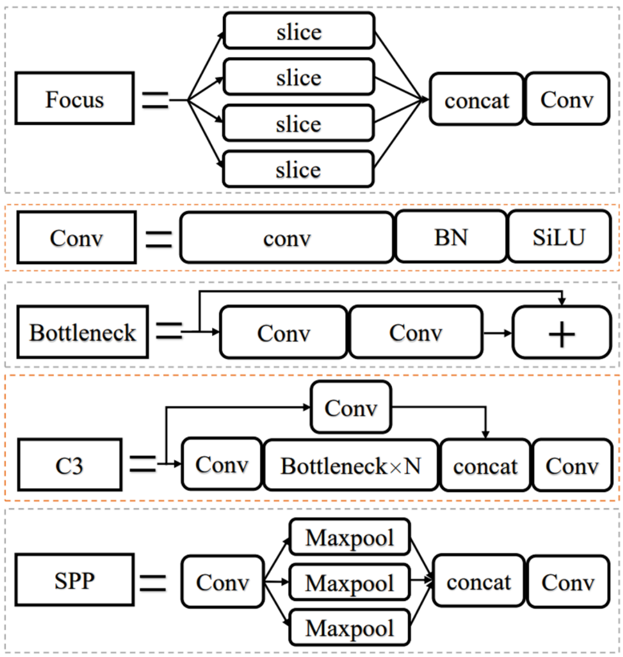

Figure 7. The composition of each module of YOLOv5s.

module is placed further back in the structure, the feature map loses most of the deformation features, so the DCN cannot extract enough deformation features to be useless. Ultimately, we found that placing the DConv module at the intersection of the YOLOv5 backbone and head resulted in the most substantial improvement in the performance of the original network. Similarly, after determining the location of the DConv module, we conducted several experiments by placing the DSC module at different locations of the network and found that the performance of YOLOv5s improved more when placed in front of the DConv module than when placed at back. Eventually, the improved YOLOv5s structure with the best performance was determined, as shown in Fig. 6.

The backbone of the network consists of 1 focus module, 3 conv modules, 4 CSP3 modules and 1 SPP module. The principle of each module can be seen in Fig. 7. The focus module slices the input picture to convert it shape from $640 \times 640 \times 3$ into 4 feature maps of $320 \times 320 \times 3$, which achieves downsampling without losing feature