www.cad-journal.net Computer-Aided Design AND APPLICATIONS

CAD Fabric Model Defect Detection Based on Improved Yolov5 Based on Self-Attention Mechanism

Jiahe Shen[1], Guoping Li[2], Raman Kumar[3] and Ranvijay Singh[4]

[1]Shanghai University of Engineering and Technology, School of Electronic and Electrical Engineering, Shanghai, 201620, China, jiaheshen3@163.com
[2]Shanghai University of Engineering and Technology, School of Electronic and Electrical Engineering, Shanghai, 201620, China, guopingli6@126.com
[3]University Centre for Research and Development, Department of Mechanical Engineering, Chandigarh University, Gharuan, Mohali, Punjab, India, raman.me@cumail.in
University Centre for Research and Development, Department of mechanical engineering, Chandigarh University, Gharuan, Mohali Punjab, India, ranvijay.e9210@cumail.in

Corresponding author: Guoping Li, guopingli6@126.com

Abstract. In CAD fabric, there is severe problem of low speed and poor generalization performance in the defect detection algorithms. To solve this problem, improved YOLOv5 based on self-attention mechanism is proposed to detect CAD fabric defects. In the proposed method, concepts of Yolov5 have been used such as extraction of the key information from the feature map and improved target detection network. Aiming at the conflict caused by the unevenness of the special scale in the network feature fusion stage, an adaptive difference fusion model is formulated to propose the algorithm. In the proposed model, the transfer learning has been used to speed up the training process. The experimental results show that the proposed detection scheme can improve the network accuracy by 98.8% and the improve detection rate by 83 frames/s when compared with the existing non-adaptive Yolov5 algorithm. The results show that the proposed detection method can perform well with the required parameters.

Keywords: Defect detection; Deep learning; Target identification; Convolution attention mechanism, adaptive spatial feature fusion.
DOI: https://doi.org/10.14733/cadaps.2024.S6.63-71

1 INTRODUCTION

The quality inspection of products is extremely important and is an indispensable link in the production process of modern manufacturing industry, with the improvement of the intelligent degree of manufacturing industry, the intelligent inspection method based on computer vision gradually replaces the manual inspection and is applied in various fields such as machinery, electrical, textile, etc. The textile industry is one of the largest industries in the world, while China is the largest textile market and exporter in the world.

In the process of fabric production, a large number of fabric defects will be produced. Therefore, the completion of automatic detection of fabric defects can bring huge benefits to enterprises and has a good market prospect [1]. At present, the semi-automatic method is widely used for cloth quality inspection. In the process of cloth rolling, a large number of experienced cloth inspectors are required to manually detect and mark the cloth, and the disadvantages of using manual cloth quality inspection are also very obvious. Visual inspection has certain vision requirements for workers, and workers need to be trained for a long time before entering the post, so the labor cost is high. Because of the high work intensity, workers are prone to make mistakes such as missing inspection and misjudgment due to fatigue [2]. The inspection speed of artificial cloth is slow, about 20-30 m/min, and the workers have subjective consciousness and cannot objectively judge, so it is difficult to unify the inspection standards [3].

In short, the manual inspection method for fabric defects has the disadvantages of slow speed, high cost and low accuracy, resulting in the current situation of high investment and low efficiency of fabric defect detection, therefore. Fabric defect detection equipment mainly consists of visual detection hardware system and fabric defect detection algorithm, when the hardware system is relatively mature, the defect detection algorithm determines the speed and accuracy of defect detection, so how to reduce the hardware cost, optimizing the defect detection algorithm and improving the speed and accuracy of cloth defect detection become the top priority [4].

At present, the algorithms used for defect detection mainly include image processing algorithm and deep learning target detection algorithm, image processing algorithm is used for defect feature extraction of simple background, which has low requirements for sample size and computer performance, but is not suitable for extracting features of complex background. The target detection algorithm based on convolution neural network has many advantages, such as strong robustness, high detection accuracy, and not easily affected by image background, however, it needs a high-performance computer and a large number of sample sets to train the model, and the defect detection accuracy of the cloth image with high resolution and complex texture is insufficient. Therefore, putting forward high-precision defect detection algorithms for high-resolution and complex texture fabric images will have a huge impact on domestic and foreign textile fabric manufacturers, and will provide great help to improve fabric quality [5].

2 LITERATURE REVIEW

Because of the fabric production in China, there are many deficiencies in the book detection of fabric defects, such as slow speed, high accuracy, low efficiency, etc. Therefore, many research institutes have conducted a long-term-research on fabric defect detection algorithms and achieved some results. With the development of knowledge economy, enterprises and scholars at home and abroad have applied the vision technology to the fabric production in recent years, trying to replace the manual inspection with the automatic fabric inspection technology to improve the production efficiency. The automatic fabric inspection equipment can automatically identify the fabric defects, and then classify and grade the fabric according to the size of the defect area, the harm to the fabric and other indicators, at the same time, it can alarm, cut, stop, mark and other operations for serious defects [6].

With the rapid development of deep learning technology in recent years, a lot of researchers have explored the application of deep learning method in detection of fabric defects. Fabric defect detection based on deep learning technology mainly includes two types: two-stage defect detection based on candidate region and one-stage defect detection based on replication. This method adds a new branch algorithm on the basis of the original residual model, dynamically adjusts the size of the acceptor space with the number of network layers, replaces the residual in the lower level, and proposes several new branches of the backbone.

The number of anchors in FasterRCNN is reduced by using the features of the flaw image, and then the foreground anchors and boundary box regression are generated in the candidate region stage, and finally the feedback is sent to the region of interest for classification. Cascaded regional convolution neural network structure is used to complete defect detection, the multi-scale training is used to adapt to the frame distribution of different scales, and then the density clustering method is used to cluster the width and height of the flaw data to make the network model easier to learn, finally, cascade three detection repressors, and use softening non-maximum suppression instead of non-maximum suppression to effectively improve the accuracy and positioning accuracy of plain cloth. However, for the candidate region-based detection algorithm, in the RPN (Region Proposal Network) stage, the aspect ratio of the anchor is usually fixed, which cannot adapt to extreme situations (targets with a wide aspect ratio), and most of the generated anchors are negative samples, affecting the final detection accuracy. The author designs fabric defect detection method based on Yolov5 development, which can meet the requirement of factory detection and real-time detection, and realizes the technology of textile industry detection [7].

3 RESEARCH METHODS

Target detection algorithm is divided into two frameworks: One is one-stage detection algorithm, which directly designs an end-to-end framework for feature extraction and target detection, such as Yolo series and SSD network; The other is the two-stage detection algorithm, in the process of target recognition, the region frame needs to be detected first, such as R-CNN series network. Yolo series network is one of the most mainstream and effective target detection algorithms based on anchor frame. The latest Yolov5 project has a variety of applications, which also proves the effectiveness of Yolov5 algorithm.

In order to improve the performance of Yolov5 network, the research team introduced Convolutional Block Attention Module (CBAM) in the upper level of the network to limit the invariable features for detection by measuring pixels from semantic and spatial data. The adaptive differential fusion algorithm (ASFF) is introduced based on pyramid listening network (PANET) to prevent the mismatch of the back -propagation technique and improve the network's detection accuracy.

3.1 Yolov5 Infrastructure

Yolov5 is SOTA (state of the art) in the YOLO series of current algorithms. Like other target detection algorithms, it has input, backbone, feature fusion and prediction models. Jolov5 includes four models with different sizes and depths, which are Jolov5s, Jolov5m, Jolov5l and Jolov5x. Referring to the method of EfficientNet, the width and depth of every length of network is equal to that of constant. The task of this research is to identify fabric defects. In order to meet the requirement of market accuracy and achieve higher accuracy and faster detection with the network structure as small as possible, the research group uses Yolov5 as the research object.

Yolov5 enhances the robustness and generality of the network structure: the mosaic data enhancement, the anchor array calculation, the adaptive image compression and the multi-dimensional image compression are introduced at the input end. A new analysis model is introduced in the trunk stage. The Focus model slices the image and splices it according to the channel direction, effectively reducing data loss. The neck part adopts feature pyramid network (FPN) PANet. FPN combines high-level data with low-level characteristics through sampling, and transmits strong features, which improves the ability of network structure to study images, but may lose some data space. PAN transmits the dynamic positioning data to the images obtained by FPN from bottom to top, and uses two to realize the results successfully, which improves the feature fusion ability of the model and improves the robustness of the model [8].

3.2 Improved Yolov5 Structure

3.2.1 Backbone network improvement

In order to improve the special isolation ability of the spinal cord, the research team introduced a solution block maintenance module (CBAM) into the soft wares system at the back stage, this module can infer the weight of attention in space and channel dimensions, so that the network can focus on the key information about the target in the image. The new backbone network introduces a CBAM module after each cross-stage connection network (CSP) structure [9-11].

CBAM is divided into proximity mechanism and channel listening mechanism, which improves the degree of subjectivity to the target from two dimensions of source and channel. The channel listener module aggregates spatial data of images by global merging and global merging of input maps, establishes the correlation model by double layer perceptron model, and finally gets the weight of each channel by sigmoid function. The channel listening mechanism is shown in Figure 1, and the expression is as follows (3.1):

$$M_c(F) = \sigma\left[W_1\left(W_0(F_{avg}^c)\right) + W_1\left(W_0(F_{max}^c)\right)\right]_0 \tag{3.1}$$

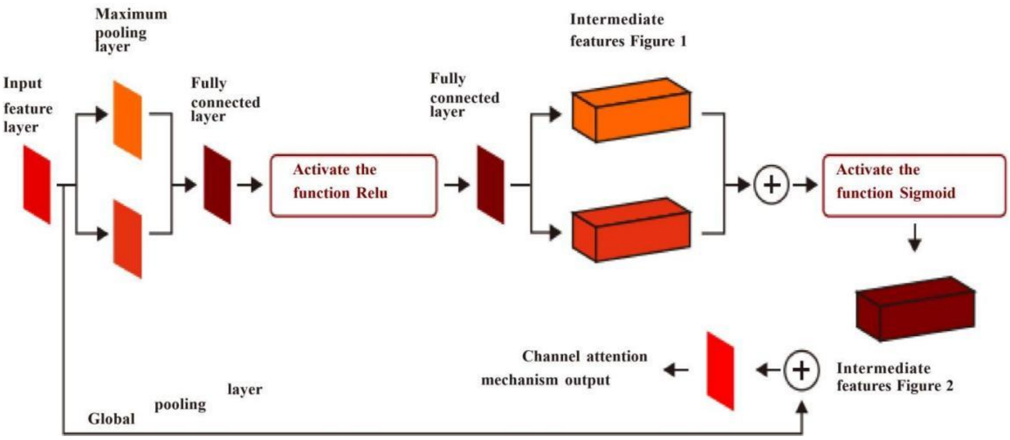

Figure 1: Channel attention mechanism module.

Where: σ means sigmoid activation for all data; F_{max}^c means global average pooling of input characteristic matrix; F_{max}^c represents global maximum pooling of input matrix; W_0 represents the weight of the first fully connected layer; W_1 represents the weight of the second fully connected layer; $M_c(F)$ represents the output characteristic diagram of the channel attention module [12].

The general monitoring module takes the output of the channel listening module as the input characteristics of the module. First, international integration and international integration as a channel have made strategic characteristics. Then the two single-channel feature maps are spliced between channels, and then a resolution of 7 is passed into 7's convolution reduces the dimension of the feature map to obtain a single channel feature map, which is convenient for the final feature fusion. Finally, each feature weight of image spatial position is obtained through sigmoid activation function. The spatial attention mechanism is shown in Figure 2, which is expressed in the following formula (3.2):

$$M_s(F) = \sigma\left(f^{7\times7}([F_{avg}^s; F_{max}^s])\right) \tag{3.2}$$

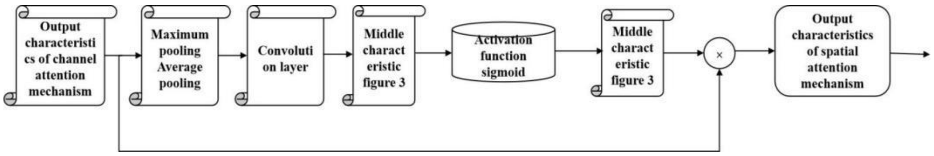

Figure 2: Spatial attention mechanism.

In the formula: $f^{7\times7}$ means to perform 7×7; F_{avg}^{s} means global average pooling of input characteristic matrix; F_{max}^{s} represents global maximum pooling of input matrix; $M_s(F)$ represents the output characteristic diagram of the spatial attention module [13].

3.2.2 Improvement in feature fusion stage

In order to make full use of the semantic information of high-level and spatial data of low-level network, the research group introduced adaptive spatial structure (ASFF) in the neck stage of Yolov5, and weighted three horizontal spatial maps generated by PANet.

By adding the studied results, the discrepancy in the back -propagation process was restrained, and the characteristics of different scales were fully utilized. The adaptive algorithm for different feature fusion neck is shown in Figure 3 [14-15].

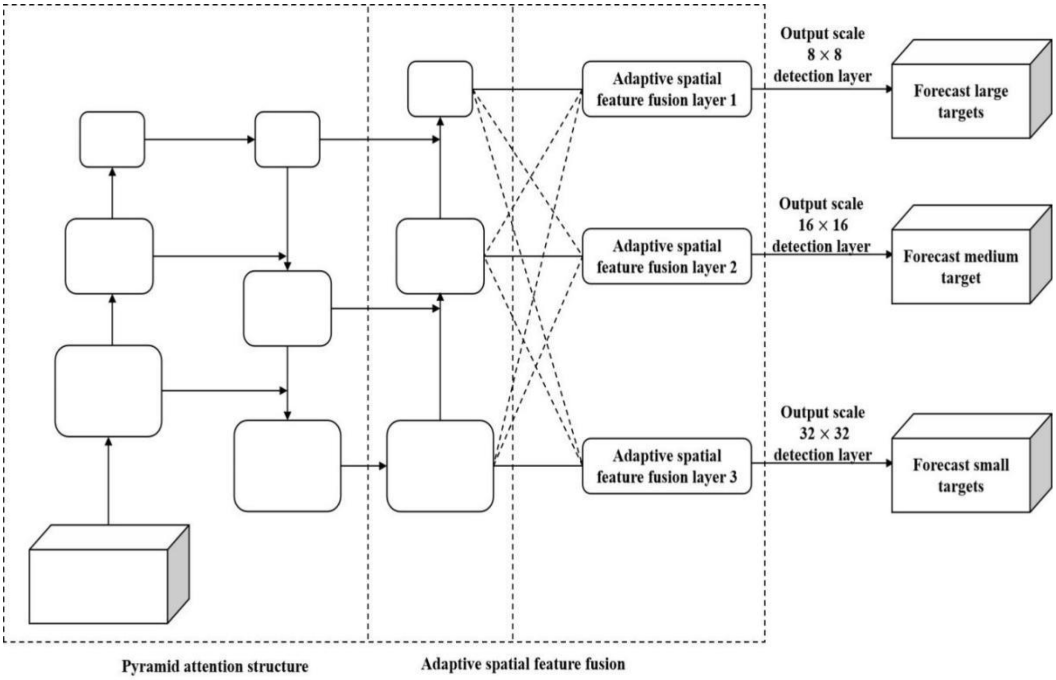

Figure 3: Adaptive spatial feature fusion neck.

The improved Yolov5 model is shown in Figure 4, first, the input image is extracted by CBAM-enhanced backbone, and then the feature map is fused by FPN+PANet+ASFF, making full use of the high-level semantic information and low-level spatial information in the feature map, obtain the target information to the maximum extent, and then send the acquired feature map into the detection head, the detection head outputs three scales 8×8, 16×16 and 32×32 output detection layers are used to detect large, medium and small targets respectively [16].

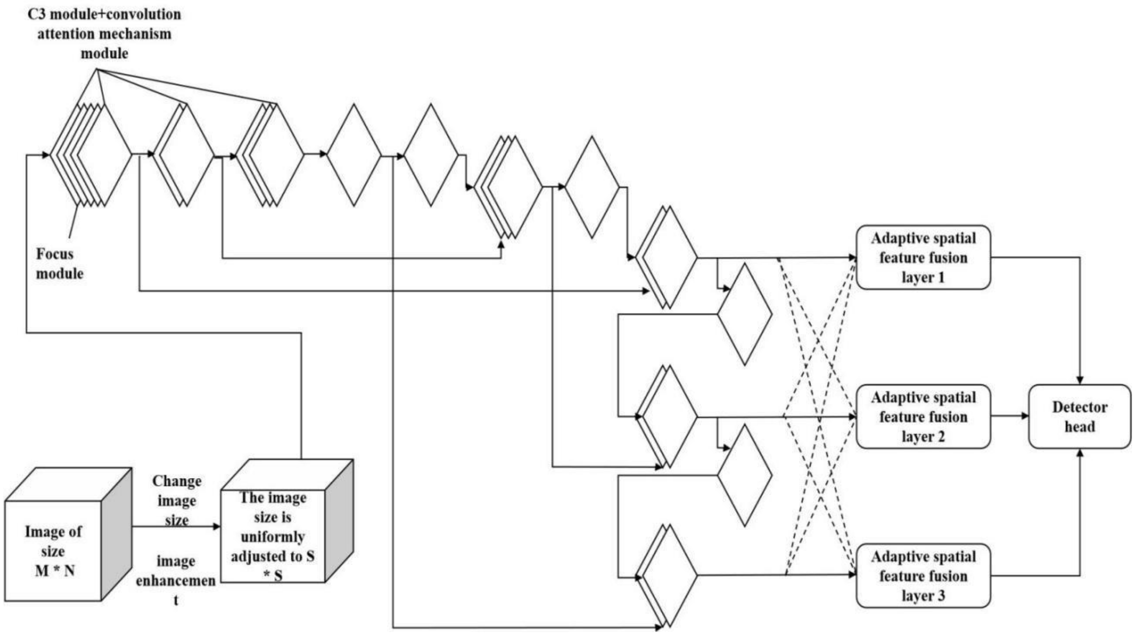

Figure 4: Improved Yolov5 model.

4 RESULT ANALYSIS

4.1 Software and Hardware Environment

The CPU is Intel (R) CoreTM i9-9900K, the main frequency is 3.60GHz, the memory is 32GiB, and the graphics card is NVIDIA; The software configuration is Win10 operating system, the graphics card driver is NVIDIAGEForceRTX2080Ti and CUDA10.1, the in-depth learning framework uses PyTorch-1.7.1, LableMe as the annotation tool, and the programming language is Python 3.7 [17-18].

4.2 Network Training

Compared with the commonly used pedestrian and vehicle data sets, the collection of cloth defect data is more difficult, the main problem is that it is difficult to collect the data of some types of defects, resulting in the insufficient number of samples of some types of defects. The collected fabric defect data is the basic data set, the fabric data set has the characteristics of rich patterns, complex background, multiple types of defects, and high image resolution. The data set used in this topic contains almost all the difficulties in the fabric defect detection process. It is of great significance for textile production to complete this kind of fabric defect detection task [19].

The validation data set of this algorithm is obtained by industrial CCD camera, and 3000 images are manually collected, the images contain complex texture fabrics and simple texture fabrics, and there is at least one defect on the sample surface, which is manually cut to 256 pixels × 256 pixels. Draw a CAD map of the obtained image and input the value to the deep learning network. The information training of the deep training model is based on a large amount of data. In order to improve the robustness and generalization ability of the model, the images are automatically rotated and analyzed, and the brightness, chroma and saturation of the model are changed. The original 3000 images are expanded to 10000 images. 7000 images were selected from training set, 2000 images based on configuration certificates, and 1000 images based on sampling rate. The data set contains 6 defects, including burr, hole (including scratch hole and broken hole), yellow stain, broken warp and weft, ink drop and damage.

After many experiments, the threshold has been set to 0.001%. The threshold is set to 0.001%. After 300 iterations, the program is set to 0.0001, the size is 16, the weight attenuation coefficient is 0.0005, and the power is 0.937. The power is 0.0005. The optimizer uses random gradient descent, and the total number of trainings is 1000.

4.3 Comparison Before and After Algorithm Improvement

In consideration of the applications used in the factory, the research team adopted relevance and average relevance (mAP) as the evaluation index of the model. The accuracy can be accurately evaluated the target location and target detection capability of the algorithm;P_{mA} is used to evaluate the network performance and is applicable to various label defect image classification programs in this study, the improved network is compared with the common SSD network, the formula is as follows (4.1):

$$P = \frac{T_P}{T_P+F_N};$$
$$P_{mA} = \frac{\sum_{n=1}^{N} P(n)\Delta P(n)}{C}. \tag{4.1}$$

Where: T_P means that the sample that should be a positive sample is considered as a positive sample by the algorithm and the prediction is correct.

The results before and after the algorithm improvement are shown in Table 1 and Figure 5.

Algorithm	*Accuracy P/%*	*Average precision of mean P_{mA}/%*	*Detection speed/(frame · s − 1)*
Yolov5	0.938	0.955	100
Yolov5+CBAM+ASFF	0.988	0.936	83
SSD	0.784	0.786	48

Table 1: Comparison of results before and after algorithm improvement.

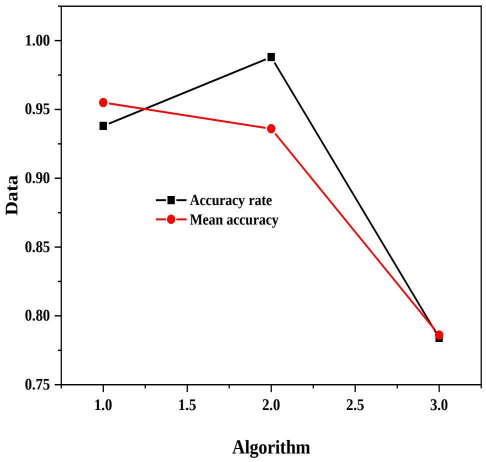

Figure 5: Accuracy and average accuracy of the algorithm before and after improvement.

According to the results of the improved network model, the improved algorithm has a high detection accuracy for six defects, including burr, hole, yellow stain, broken warp and weft, ink drop and damage, which is estimated to reach 98.8%.

5 CONCLUSION

The research group proposed a fabric defect detection algorithm based on the improved Yolov5 algorithm, which solved the problems of low detection rate, poor performance and low real-time of the traditional method for fabric defect detection. The experimental results show that further solving the supervisory mechanism and modifying the deviation of difference fusion for Yolov5 simple network can improve the supervision of the model to the defect target, improve the connectivity of information system and spatial information, and improve the accuracy of defect detection to 98.8% when detection speed is only reduced by 17 frames/s. fabric defect detection algorithm developed by the research group has high real-time and accuracy, and can meet the requirement of the actual factory. The algorithm also relies on a large number of different models, and designs fabric surface defects detection algorithm based on small sample size is the key of the next step.

6 ACKNOWLEDGEMENT

The study was supported by intelligent collaborative computing and cross network collaborative transmission in media convergence network National key R&D plan2019YFB1802700"

Jiahe Shen, https://orcid.org/0009-0002-0802-7008
Guoping Li, https://orcid.org/0009-0002-7477-8600
Raman Kumar, https://orcid.org/0000-0002-5040-7920
Ranvijay Singh, https://orcid.org/0000-0001-8430-6186

REFERENCES

[1] Ren, Z.; Fang, F.; Yan, N.; Wu, Y.: State of the art in defect detection based on machine vision, International Journal of Precision Engineering and Manufacturing-Green Technology, 9(2), 2022, 661-691. https://doi.org/10.1007/s40684-021-00343-6

[2] Tulbure, A.-A.; Tulbure, A.-A.; Dulf, E.-H.: A review on modern defect detection models using DCNNs-Deep convolutional neural networks, Journal of Advanced Research, 35, 2022, 33-48. https://doi.org/10.1016/j.jare.2021.03.015

[3] Zeng, N.; Wu, P.; Wang, Z.; Li, H.; Liu, W.; Liu, X.: A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection, IEEE Transactions on Instrumentation and Measurement, 71, 2022, 1-14. https://doi.org/10.1109/TIM.2022.3153997

[4] Czimmermann, T.; Ciuti, G.; Milazzo, M.; Chiurazzi, M.; Roccella, S.; Oddo, C.-M.; Dario, P.: Visual-based defect detection and classification approaches for industrial applications—A survey, Sensors, 20(5), 2020, 1459. https://doi.org/10.3390/s20051459

[5] Lu, H.; Zhu, Y.; Yin, M.; Yin, G.; Xie, L.: Multimodal fusion convolutional neural network with cross-attention mechanism for internal defect detection of magnetic tile, IEEE Access, 10, 2022, 60876-60886. https://doi.org/10.1109/ACCESS.2022.3180725

[6] Liu, Q.; Wang, C.; Li, Y.; Gao, M.; Li, J.: A fabric defect detection method based on deep learning, IEEE Access, 10, 2022, 4284-4296. https://doi.org/10.1109/ACCESS.2021.3140118

[7] Wu, Y.; Qin, Y.; Qian, Y.; Guo, F.; Wang, Z.; Jia, L.: Hybrid deep learning architecture for rail surface segmentation and surface defect detection, Computer-Aided Civil and Infrastructure Engineering, 37(2), 2022, 227-244. https://doi.org/10.1111/mice.12710

[8] Aggarwal, K.; Mijwil, M.-M.; Al-Mistarehi, A.-H.; Alomari, S.; Gök, M.; Alaabdin, A. M.-Z.; Abdulrhman, S.-H.: Has the future started? The current growth of artificial intelligence,

machine learning, and deep learning, Iraqi Journal for Computer Science and Mathematics, 3(1), 2022, 115-123. https://doi.org/10.52866/ijcsm.2022.01.01.013
[9] Fu, H.; Song, G.; Wang, Y.: Improved YOLOv4 marine target detection combined with CBAM, Symmetry, 13(4), 2021, 623. https://doi.org/10.3390/sym13040623
[10] Shwartz-Ziv, R.; Armon, A.: Tabular data: Deep learning is not all you need, Information Fusion, 81, 2022, 84-90. https://doi.org/10.1016/j.inffus.2021.11.011
[11] Pereira, T. D.; Tabris, N.; Matsliah, A.; Turner, D.-M.; Li, J.; Ravindranath, S.; Murthy, M.: SLEAP: A deep learning system for multi-animal pose tracking, Nature Methods, 19(4), 2022, 486-495. https://doi.org/10.1038/s41592-022-01426-1
[12] Lauriola, I.; Lavelli, A.; Aiolli, F.: An introduction to deep learning in natural language processing: Models, techniques, and tools, Neurocomputing, 470, 2022, 443-456. https://doi.org/10.1016/j.neucom.2021.05.103
[13] Zhan, Z.-H.; Li, J.-Y.; Zhang, J.: Evolutionary deep learning: A survey, Neurocomputing, 483, 2022, 42-58. https://doi.org/10.1016/j.neucom.2022.01.099
[14] Hanif, A.; Muaz, M.; Hasan, A.; Adeel, M.: Micro-Doppler based target recognition with radars: A review, IEEE Sensors Journal, 22(4), 2022, 2948-2961. https://doi.org/10.1109/JSEN.2022.3141213
[15] Gao, J.; Sun, H.; Han, J.; Sun, Q.; Zhong, T.: Research on recognition method of electrical components based on FEYOLOv4-tiny, Journal of Electrical Engineering & Technology, 17(6), 2022, 3541-3551. https://doi.org/10.1007/s42835-022-01124-0
[16] Kawakami, K.; Vingilis Jaremko, L.; Friesen, J.-P.; Meyers, C.; Fang, X.: Impact of similarity on recognition of faces of Black and White targets, British Journal of Psychology, 113(4), 2022, 1079-1099. https://doi.org/10.1111/bjop.12589
[17] Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B.: A Review of Yolo algorithm developments, Procedia Computer Science, 199, 2022, 1066-1073. https://doi.org/10.1016/j.procs.2022.01.135
[18] Fang, B.; Long, X.; Sun, F.; Liu, H.; Zhang, S.; Fang, C.: Tactile-based fabric defect detection using convolutional neural network with attention mechanism, IEEE Transactions on Instrumentation and Measurement, 71, 2022, 1-9. https://doi.org/10.1109/TIM.2022.3165254
[19] Chen, K.; Song, X.; Yuan, H.; Ren, X.: Fully convolutional encoder-decoder with an attention mechanism for practical pedestrian trajectory prediction, IEEE Transactions on Intelligent Transportation Systems, 23(11), 2022, 20046-20060. https://doi.org/10.1109/TITS.2022.3170874