Available online at www.HighTechJournal.org

HighTech and Innovation Journal

ISSN: 2723-9535

Vol. 4, No. 3, September, 2023

Deep Learning: A Study of Pattern Recognition for Personalized Clothing

JIng Zhao [1], Hongdi Zhu [1], Bing Liu [1*]

[1] School of Fine Arts & Design, Tianjin Normal University College, Tianjin 300387, China.

Received 04 June 2023; Revised 05 August 2023; Accepted 17 August 2023; Published 01 September 2023

Abstract

Objectives: This article aims to enhance the efficiency of clothing recognition and retrieval by implementing deep learning algorithms for personalized clothing pattern recognition. *Methods*: Based on the you only look once version 4 (YOLOv4) algorithm in deep learning, the CSPDarknet-53 in the original algorithm was replaced by GhostNet, and the original Leaky ReLU activation function was replaced by FMish. Then, an improved YOLOv4 algorithm was obtained. Experiments were carried out on the personalized clothing pattern set, the Fashion Mnist dataset, and the DeepFashion dataset to compare and analyze different algorithms. *Findings*: When replacing CSPDarknet-53 with GhostNet and the Leaky ReLU activation function with FMish, the optimized YOLOv4 algorithm performed significantly better, verifying the reliability of the YOLOv4 improvement. The optimized algorithm achieved an F1 value of 94.22% and a mAP of 95.41% on different datasets, and 39.51% and 49.56% on the DeepFashion dataset, respectively, outperforming other deep learning methods such as the faster-recurrent convolutional neural network. Furthermore, the floating-point operations per second of the optimized YOLOv4 algorithm were 8.72 G, showing a reduction of 49.71% compared to the traditional algorithm. This suggested that it had low complexity and calculation amounts. *Novelty*: The optimized YOLOv4 algorithm performs excellently in recognizing personalized clothing patterns, which can provide a new and reliable approach for recognition and retrieval in the field of clothing.

Keywords: Deep Learning; Personalized Clothing; Pattern Recognition; Activation Function; Recognition Effect.

1. Introduction

With the changes and developments in society, clothing materials, styles, and patterns have also evolved. Clothing design has shifted from being purely practical to becoming more personalized. Unique textures and patterns are increasingly used in fashion design, making personalized clothing a current trend and an important driver of consumption. Under the influence of technological advances and changing perspectives, clothing sales have gradually shifted from traditional offline channels to online channels. For clothing companies, in addition to personalized clothing design, the sales process also plays a crucial role in their development. When shopping for clothing online, consumers typically first determine the style they want or try to search for similar styles after finding a particular pattern. This makes clothing pattern recognition particularly important. As personalized clothing continues to develop, clothing pattern recognition becomes more challenging. Personalized clothing pattern recognition is an image recognition problem, and with the advancement of deep learning technology [1], an increasing number of methods have been applied to image recognition.

* Corresponding author: lbliubing@tjnu.edu.cn

http://dx.doi.org/10.28991/HIJ-2023-04-03-03

Wang et al. [2] studied the recognition of pest images, compared three different deep learning models, and obtained recognition rates of over 80%. This research provides a reference for further studies in the field of agricultural pest recognition. In a convolutional neural network (CNN), Yang et al. [3] merged transfer learning with ensemble learning to detect damage on wind turbine blades. Through experiments, they found that this method outperformed support vector machines. Yang et al. [4] utilized deep learning edge algorithms for the recognition of digestive endoscopy images. The results showed that the accuracy of the approach was 68% higher than that of the simple you only look once (YOLO) algorithm, and both the accuracy and speed were 21% and 85% higher respectively compared to the recurrent convolutional neural network (RCNN).

Anubha Pearline et al. [5] conducted research on plant recognition in images and compared different methods. Through experiments, they found that using logistic regression as a classifier, the VGG19 CNN structure achieved accuracies of 96.53%, 96.25%, and 99.41% on different datasets. Zhang et al. [6] utilized a dual-stream heterogeneous backbone network based on VGG-16 and Res2Net-50 to extract image features, achieving detection of color images. Experimental results on four publicly available datasets demonstrated the outstanding performance of this approach. Tsuiki et al. [7] developed a deep CNN for recognizing lateral head shadow measurement film images and found its high accuracy through experiments. Vinolin et al. [8] designed a deep convolutional neural network called Taylor-ROA-based DeepCNN using the Taylor-rider optimization algorithm to detect forged and original images. The experimental results demonstrated that this approach significantly improved accuracy compared to existing methods. Zhang et al. [9] proposed a method for detecting internal cracks in corn seeds by combining deep learning algorithms with edge detection threshold processing. The experiments show that this method achieves recognition accuracies of 95.08% and 95.75% for cracked and uncracked seeds, respectively.

Wang et al. [10] introduced an enhanced YOLOv3 algorithm to detect illegal opium poppy cultivation in low-altitude drone inspections. Testing on a self-created dataset revealed that this approach reduced parameters and enhanced recognition accuracy, offering technological support for low-altitude opium poppy detection. While deep learning methods have already achieved mature applications in various fields such as engineering, agriculture, medicine, etc., research regarding clothing pattern recognition remains relatively scarce. Currently, the most commonly used approach in clothing pattern recognition relies on extracting features such as edges and contours for classification and identification, which leads to poor accuracy. However, with the rapid development of e-commerce, there is an increasing demand for clothing pattern recognition, classification, and retrieval. The existing methods fail to meet the needs of consumers searching for interesting clothing items in e-commerce. Therefore, a new method is urgently needed. This paper mainly focuses on the research of deep learning methods. It designed an optimized you only look once version 4 (YOLOv4) method using GhostNet and FMish activation function based on YOLO series algorithms in deep learning. Through experiments, the reliability of this method in recognizing personalized clothing patterns was demonstrated. The research flowchart is shown in Figure 1. The research in this article presents a novel approach for recognizing and retrieving clothing, while also providing theoretical support for further studies on YOLO series algorithms. This not only benefits the future application of YOLO series algorithms in the field of pattern recognition but also offers guidance for enhancing the YOLOv4 algorithm.

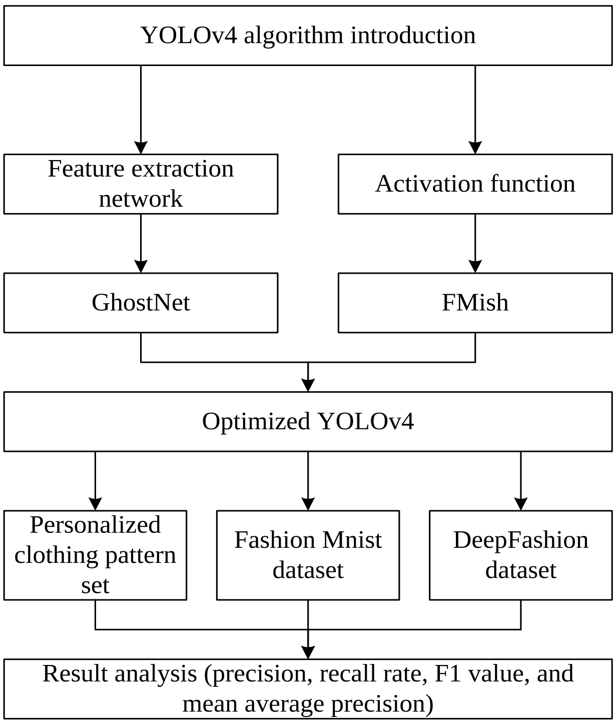

Figure 1. Research flowchart

2. Personalized Clothing Pattern Recognition Method

2.1. Deep Learning and YOLOv4 Algorithm

The pattern of personalized clothing contains a lot of complex feature information, and the traditional recognition method achieves clothing classification by extracting edges, contours, etc. However, the accuracy rate is not high. Deep learning is developed on the basis of artificial neural networks [11], which can extract deep feature information from data. Commonly used methods include CNN, generative adversarial network (GAN), etc. [12], which are widely used in speech recognition, image processing, etc. [13]. This paper chooses the YOLOv4 algorithm from the YOLO series algorithms.

Personalized clothing pattern recognition requires fast recognition speed and high accuracy to meet the needs of consumers. The YOLOv4 algorithm is an enhancement of the YOLOv3 algorithm that significantly optimizes both the accuracy and speed of recognition.

The principle of the YOLOv4 algorithm [14] is described as follows. For an input of 461 × 416 × 3, three different sizes of feature maps are obtained in the backbone feature extraction network (CSPDarknet-53), the spatial pyramid pooling (SSP) module enlarges the perceptual field, and the three feature maps are fused by the path aggregation network (PANet) to further strengthen the learning capacity of the network for features. The training of the YOLOv4 algorithm is achieved by error backpropagation, and the loss function used includes the following three components.

(1) Category loss:

The cross-entropy loss function is used here, and the formula can be written as:

$$Loss_{class} = \sum_{i=0}^{S\times S} I_{ij}^{obj} \sum_{\mathrm{C}\in \mathrm{classes}} \left[\widehat{\mathrm{P}}_{\mathrm{i}}(\mathrm{c}) \log\left(\mathrm{P}_{\mathrm{i}}(\mathrm{c})\right) + \left(1 - \widehat{\mathrm{P}}_{\mathrm{i}}(\mathrm{c})\right) \log\left(\mathrm{P}_{\mathrm{i}}(\mathrm{c})\right)\right], \tag{1}$$

where $S \times S$ represents the size of the output grid, I stands for detection bounding box generated by a priori box, $P_i(c)$ stands for true value of the category, and $\hat{P}_i(c)$ stands for predicted value of the category.

(2) Loss of confidence:

The same cross-entropy loss function is applied in this part, and its equation is:

$$Loss_{confi} = \sum_{i=0}^{S\times S} \sum_{j=0}^{B} I_{ij}^{obj} \left[\hat{C}_i \log(C_i) + \left(1 - \hat{C}_i\right) \log(1 - C_i)\right] - \lambda_{noobj} \sum_{i=0}^{S\times S} \sum_{j=0}^{B} I_{ij}^{noobj} \left[\hat{C}_i \log(C_i) + \left(1 - \hat{C}_i\right) \log(1 - C_i)\right], \tag{2}$$

where B is the number of boxes in each grid, λ_{noobj} represents loss factor, C_i represents the true value of the confidence level, and $\hat{C}_i$ represents the predicted value of the confidence level.

(3) Regression box position loss:

This part uses the complete intersection over union (CIOU) as the loss function, and the equation can be written as:

$$Loss_{CIOU} = 1 - IOU + \frac{\rho^2(b^2,b)}{c^2} + \upsilon\alpha, \tag{3}$$

$$\upsilon = \frac{4}{\pi^2}\left(arctan\frac{\widehat{w}}{\widehat{h}} - arctan\frac{w}{h}\right) h^2, \tag{4}$$

$$\alpha = \frac{\upsilon}{(1-IOU)+\upsilon}, \tag{5}$$

where IOU stands for intersection over union, $\rho^2(b^2, b)$ is the Euclidean distance between the center points of the two bounding boxes, c stands for the Euclidean distance of the diagonal in the union set region of the two bounding boxes, υ stands for the stability parameter to measure the aspect ratio of the bounding box, w and h are the width and height of the real frame, $\widehat{w}$ and $\widehat{h}$ are the width and height of the prediction frame, and α represents the balance parameter.

Ultimately, the loss function of the YOLOv4 algorithm is:

$$\mathrm{Loss} = \mathrm{Loss}_{\mathrm{class}} + \mathrm{Loss}_{\mathrm{confi}} + \mathrm{Loss}_{\mathrm{CIOU}}. \tag{6}$$

2.2. Optimized YOLOv4 Algorithm

To further enhance the effectiveness of the YOLOv4 algorithm for personalized clothing pattern recognition, the YOLOv4 algorithm is optimized. Firstly, the CSPDarknet-53 network used in the YOLOv4 algorithm is replaced by GhostNet. GhostNet is a staged convolutional computation module [15] that consists of two steps: (1) performing traditional convolutional operations to obtain the feature map; (2) performing deep convolutions based on the obtained feature map. This approach can reduce the computational effort of the convolutional operation, thus strengthening the algorithm performance.

GhostNet consists of a series of Ghost bottlenecks, and Ghost bottleneck contains two Ghost modules, as illustrated in Figure 2.

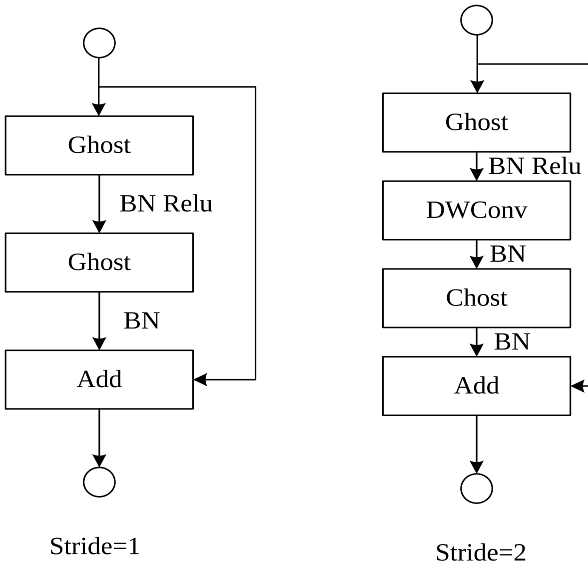

Figure 2. Ghost bottleneck

In the Ghost bottleneck, when stride=1, the first and second Ghost modules are employed to increase and decrease the number of channels respectively, and batch normalization (BN) and Relu nonlinear activation are used in the middle. When stride=2, deep convolution (DWConv) is added to improve the performance. Using GhostNet to replace the CSPDarknet-53 network in the traditional YOLOv4 algorithm, Figure 3 illustrates the network structure of the optimized YOLOv4 algorithm.

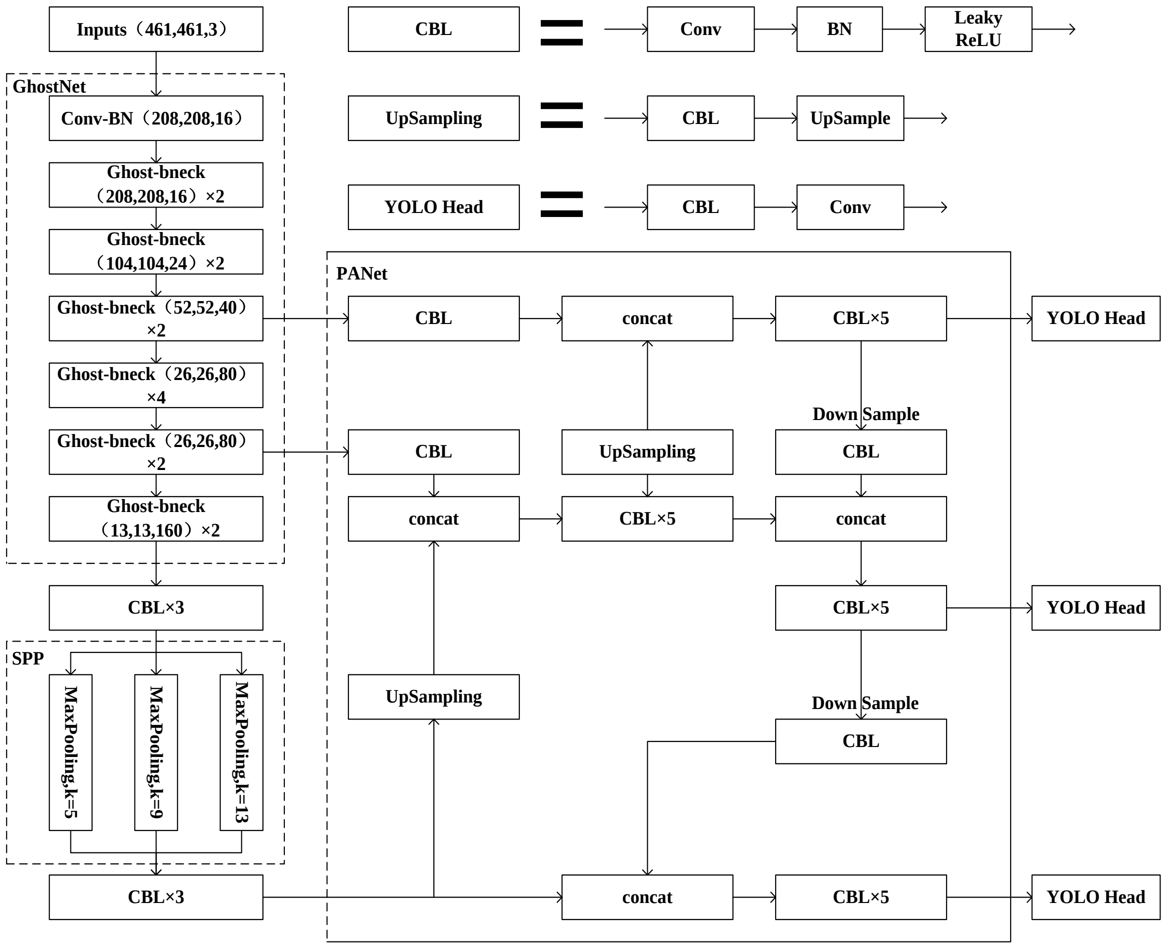

Figure 3. The network structure of the optimized YOLOv4 algorithm

As shown in Figure 3, the YOLOv4 algorithm is applied to the pattern recognition of personalized clothing. For the input clothing pattern, the GhostNet module is applied to extract the pattern features of the personalized clothing, and three feature maps of different sizes are obtained and connected to PANet and SPP, respectively. The SPP layer increases the perceptual field through maximum pooling, while fusion between different feature layers is achieved in PANet. Finally, the YOLO head module discriminates and adjusts the prediction frames obtained from each feature layer to obtain the final prediction frame.

To ensure the stability of the training, the activation function is indispensable in the network. According to Figure 3, it can be found that the Leaky ReLU activation function is involved in the improved YOLOv4 algorithm:

$$y_{LeakyReLU} = \begin{Bmatrix} x, x > 0 \\ \gamma, x \leq 0 \end{Bmatrix} = max(0, x) + \gamma min(0, x), \tag{7}$$

where γ is a constant very small gradient, generally taken as 0.01. The main problem with Leaky ReLU is that the transmission reliability on negative intervals is not high; therefore, to improve this, it is replaced by the FMish function:

$$y_{FMish} = \frac{x \ln(1+e^x)}{\sqrt{1 + \ln^2(1+e^x)}}. \tag{8}$$

FMish can ensure the stability of the training more effectively, as the gradient at the zero point does not change abruptly, while avoiding the problem of oversaturation.

3. 1. Experiment and Analysis

3.1. Experimental Environment and Evaluation Indicators

The experiments were conducted in a Windows 10 environment with 32 GB of memory, using Python 3.8 as the programming language and Pytorch 1.7.1 as the deep learning framework. The following indicators were used to assess the deep learning method designed in this paper:

1) Precision: $\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$,
2) Recall rate: $\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}$,
3) F_1 value: $\mathrm{F}_1 = \frac{\mathrm{2PR}}{\mathrm{P+R}}$,
4) Mean average precision (mAP): $\mathrm{mAP} = \frac{\sum_{i=1}^{n} \mathrm{AP(i)}}{n}$, $\mathrm{AP} = \int_0^1 \mathrm{P(R)dR}$.

In the above equations, TP stands for the number of positive samples forecasted to be positive, FP stands for the number of negative samples forecasted to be positive, FN stands for the number of positive samples forecasted to be negative, n is the number of categories, F_1 value represents the harmonic mean of precision and recall rate, and AP is the integral of the P-R curve.

3.2. Dataset

(1) The personalized clothing pattern set designed in this paper: it contains three different personalized patterns, all of which can be used in clothing pattern design. The patterns contain different shapes and texture variations. One of them is shown in Figure 4, which embroiders the imagery of flying cranes, red sun, and floating clouds through different colored silk threads, supplemented by beadwork such as rice beads and tube beads to emphasize the pattern's outline and enhance its ornamental nature. The use of materials considers both soft and rigid, cold and warm, achieving color coordination and the unity of thickness and softness. In addition, the design of the pattern also contains a good symbolic meaning of rising day by day. It was processed into a 28×28 grayscale image, and the expansion of the dataset was realized by random flipping. The training and test sets were divided in a ratio of 8:2.

(2) Fashion Mnist dataset [16]: it contains about 70,000 clothing images and ten different clothing categories (trouser, T-shirt, coat, sandals, shirt, dress, pullover, bag, sneaker, and ankle boots). The number of samples in the training and test sets was 60,000 and 10,000, respectively.

(3) DeepFashion dataset [17]: it contains about 800,000 clothing images and 50 different clothing categories, and five of them were selected for analysis in this paper, as shown in Table 1.

Figure 4. Personalized clothing pattern design

Table 1. DeepFashion dataset

Clothing category	Training set/n	Test set/n
Blazer	33,591	8,514
Dress	34,072	8,853
Jeans	33,394	8,242
Shorts	35,632	9,411
Sweater	32,021	8,341

3.3. Result Analysis

First, for personalized clothing pattern recognition, the YOLOv4 algorithm was also used to compare the effects of different feature extraction networks. Table 2 represents the results.

Table 2. Effects of feature extraction networks on the recognition effect of personalized clothing patterns

	CSPDarknet-53	GhostNet
Precision/%	91.24	93.27
Recall rate/%	85.77	87.44
F1 value/%	88.42	90.26
mAP/%	91.76	93.58

From Table 2, it was seen that for pattern recognition of personalized clothing, the traditional YOLOv4 algorithm, i.e., when using CSPDarknet-53 as the feature extraction network, had a precision of 91.24%, a recall rate of 85.77%, an F1 value of 88.42%, and a mAP value of 91.76%, while after using GhostNet instead of the original CSPDarknet-53, the algorithm achieved a precision of 93.27% (improved by 2.03%), a recall rate of 87.44% (improved by 1.67%), an F1 value of 90.26% (improved by 1.84%), and a mAP value of 93.58% (improved by 1.82%), i.e., all the indicators suggested significant improvements. These results showed that GhostNet performed better in feature extraction than CSPDarknet-53, i.e., it could extract features more effectively from complex personalized patterns, thus achieving better performance in personalized clothing pattern recognition.

Then, the effect of activation functions Leaky ReLU and FMish on the recognition performance of personalized clothing patterns was compared in the case of using GhostNet, the results are displayed in Table 3.

Table 3. The effect of activation function on the recognition of personalized clothing pattern

	Leaky ReLU	FMish
Precision/%	93.27	94.57
Recall rate/%	87.44	89.61
F1 value/%	90.26	92.02
mAP/%	93.58	95.16

From Table 3, it was seen that when FMish was used as the activation function in the optimized YOLOv4 algorithm, the algorithm showed some improvement in all indicators. In comparison, when FMish was used instead of Leaky ReLU, the algorithm achieved a precision of 94.57% (improved by 1.3%), a recall rate of 89.61% (improved by 2.17%), an F1 value of 92.02% (improved by 1.76%), and a mAP value of 95.16% (improved by 1.58%) in personalized clothing pattern recognition, proving the reliability of FMish as an activation function and the reliability of YOLOv4 improvement.

On the Fashion Mnist dataset, the improved YOLOv4 algorithm was compared with other deep learning methods:

① the Faster-RCNN algorithm [18],

② the single shot multibox detector (SSD) algorithm [19],

③ the YOLOv3 algorithm [20],

④ the YOLOv4 algorithm,

⑤ the attention-YOLOv4 algorithm [21].

Figure 5 illustrates the results.

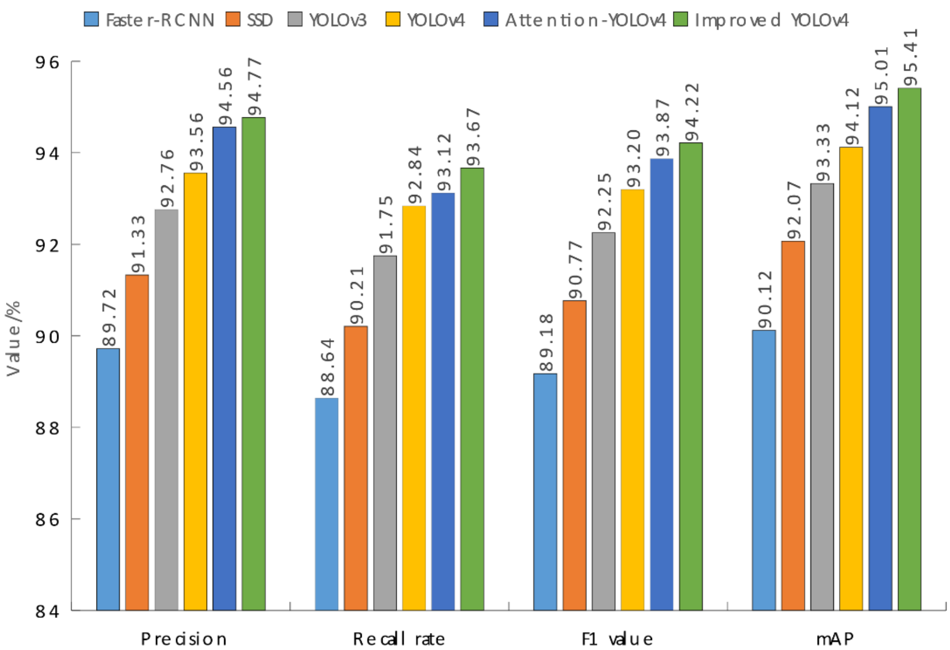

Figure 5. Comparison of the recognition results of different methods on the Fashion Mnist dataset

The optimized YOLOv4 algorithm achieved better results than the other methods in all indicators, as observed from Figure 5. In terms of the F1 value comparison, the optimized YOLOv4 algorithm was 94.22%, which was increased by 5.04% compared to the Faster-RCNN method, 4.01% compared to the SSD approach, 1.97% compared to the YOLOv3 algorithm, 1.02% compared to the YOLOv4 algorithm, and 0.35% compared to the attention-YOLOv4 algorithm. In terms of the mAP comparison, the optimized YOLOv4 algorithm was 95.41%, which increased by 5.29% compared to the Faster-RCNN algorithm, 3.34% compared to the SSD method, 2.08% compared to the YOLOv3 algorithm, 1.29% compared to the YOLOv4 algorithm, and 0.4% compared to the attention-YOLOv4 algorithm. It can be concluded that the optimized YOLOv4 algorithm had the best performance in clothing recognition on the Fashion Mnist dataset and could achieve good classification of different kinds of clothing.

The results of different methods after recognizing the DeepFashion dataset are illustrated in Figure 6.

According to Figure 6, the optimized YOLOv4 algorithm exhibited the best recognition performance on the DeepFashion dataset. However, its recognition performance on the DeepFashion dataset was obviously lower than that on the Fashion Mnist dataset, which may be because the larger amount of clothing data and more complex clothing categories in the DeepFashion dataset resulted in more cases of recognition errors. Specifically, the mAP of the optimized YOLOv4 algorithm on the DeepFashion dataset was 49.56%, which increased by 13.45% compared to the Faster-RCNN method, 5.29% compared to the SSD method, 4.5% compared to the YOLOv3 algorithm, 1.33% compared to the YOLOv4 algorithm, and 0.53% compared to the attention-YOLOv4 algorithm, proving the effectiveness of the method in recognizing different clothing categories.

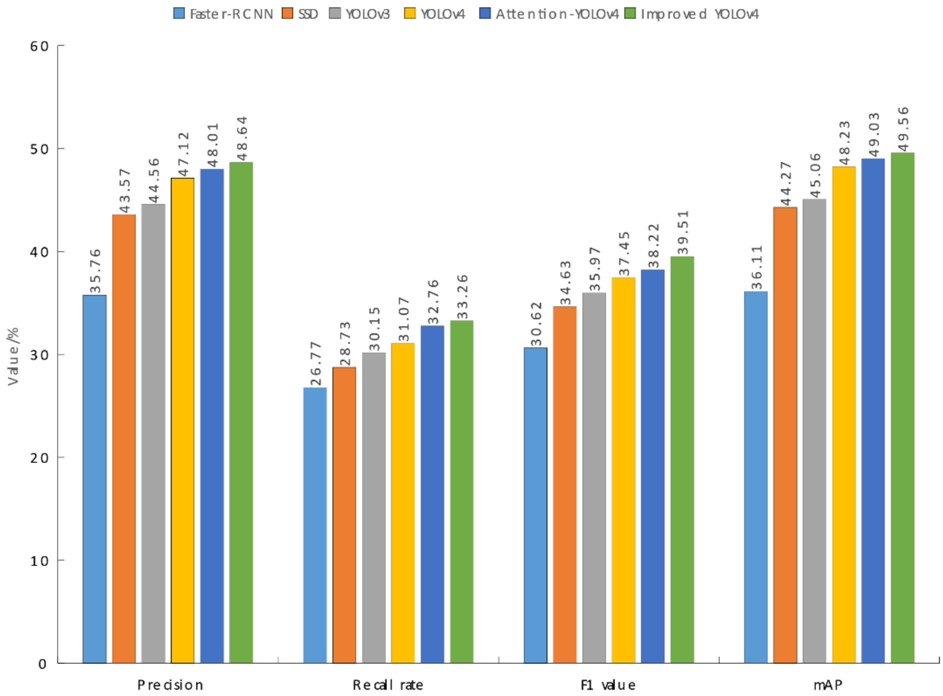

Figure 6. Comparison of recognition results between different methods on the DeepFashion dataset

Finally, the complexity of several YOLO algorithms was compared in terms of floating-point operations per seconds (FLOPs). Using the DeepFashion dataset as an example, the results are presented in Figure 7.

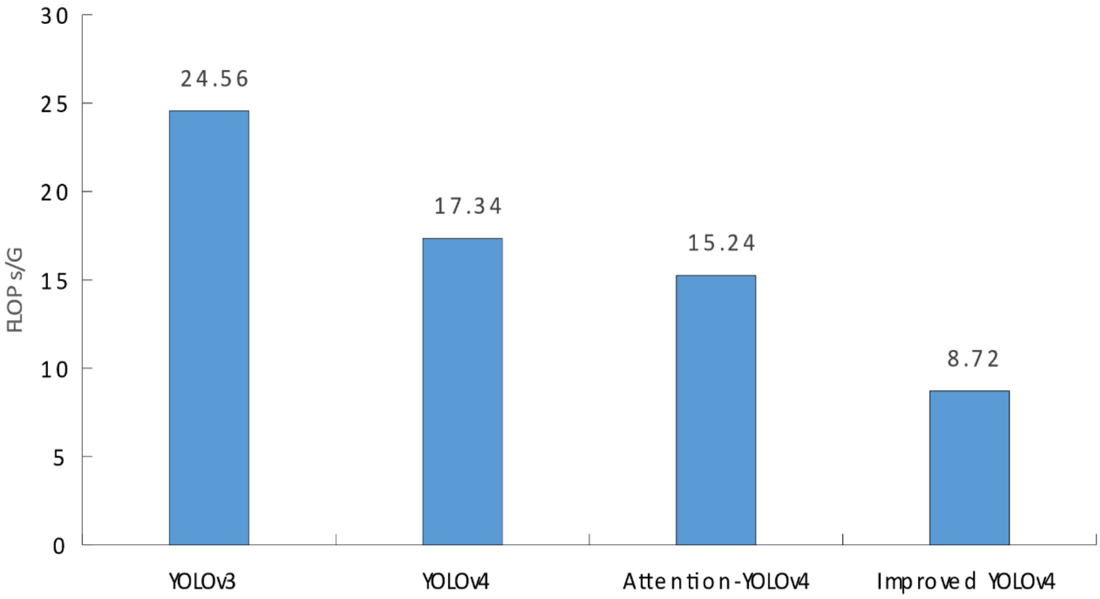

Figure 7. Complexity comparison of YOLO algorithms

From Figure 7, it was observed that the YOLOv3 algorithm had a FLOPs of 24.56 G, indicating its high complexity and heavy computational load. In comparison, the YOLOv4 algorithm had a FLOPs of 17.34G, which was 29.4% lower than the YOLOv3 algorithm. This reduction demonstrated the optimization resulted in a significant reduction in algorithm complexity. The FLOPs of the attention-YOLOv4 algorithm was 15.24 G, which was 12.11% lower than the YOLOv4 algorithm. The improved YOLOv4 algorithm had a FLOPs of 8.72 G, which was 49.71% lower than the YOLOv4 algorithm and 42.78% lower than the attention-YOLOv4 algorithm. By replacing the original CSPDarknet-53 with GhostNet, the optimized YOLOv4 algorithm significantly reduced the computational load and had lower FLOPs than the attention-YOLO4 algorithm, demonstrating its excellent computational efficiency in clothing pattern recognition.

4. Discussion

The problem of recognizing personalized clothing patterns can be considered as an object detection algorithm. With the continuous development of deep learning, the YOLO algorithm has received increasing research attention. In order to achieve a better balance between speed and accuracy, researchers have made various improvements and optimizations to the YOLO series algorithms and applied them in different fields for experimentation. This article focused on the pattern recognition problem of personalized clothing and improved the YOLOv4 algorithm in order to design an enhanced version of YOLOv4, which was then compared with some existing deep learning methods.

From the experimental results, firstly, this study demonstrated the reliability of the improvements made to the YOLOv4 algorithm through experiments on a personalized clothing pattern dataset. By comparing the results in Table 2 and Table 3, it can be observed that both CSPDarknet-53 and Leaky ReLU in the traditional YOLOv4 algorithm performed worse than GhostNet and FMish used in the improved version. The precision, recall rate, and other performance indicators of the improved YOLOv4 algorithm were significantly increased, indicating that the direction of YOLOv4 improvement in this study is correct. By replacing the feature extraction network and activation function, the algorithm's recognition effectiveness has been effectively enhanced.

Furthermore, based on the results of comparisons with other deep learning methods on Fashion Mnist and DeepFashion datasets, the improved YOLOv4 algorithm demonstrated superior performance. It exhibited higher recognition efficiency compared to both other deep learning methods and other modified versions of YOLOv4. The comparison of algorithm complexity (Figure 6) showed both improved YOLOv4 algorithms had lower FLOPs than the YOLOv3 and YOLOv4 algorithms. However, in this comparison, the improved YOLOv4 algorithm had FLOPs below 10 G, which was 49.71% lower than the YOLOv4 algorithm and 42.78% lower than the attention-YOLOv4 algorithm. These results provide sufficient evidence of the computational efficiency advantage of the improved YOLOv4 algorithm proposed in this paper.

The comprehensive experimental results reveal that the improved YOLOv4 algorithm, designed in this paper, significantly enhances the accuracy and efficiency of the algorithm by improving both the feature extraction network and activation function, making it applicable for practical clothing pattern recognition.

5. Conclusion

This paper focused on the recognition of personalized clothing patterns. The traditional YOLOv4 algorithm was optimized using the deep learning method. Through experimentation, it was found that the optimized YOLOv4 algorithm outperformed the original YOLOv4 algorithm on various datasets, showcasing the reliability of the enhancements and its potential application in practical personalized clothing recognition, classification, and retrieval. In future research, we will explore additional optimization possibilities for the YOLOv4 algorithm and conduct experiments on a broader range of datasets.

6. Declarations

6.1. Author Contributions

Conceptualization, J.Z. and B.L.; methodology, J.Z. and B.L.; software, H.Z.; validation, J.Z., H.Z., and B.L.; formal analysis, H.Z. and B.L.; investigation, H.Z.; resources, J.Z. and B.L.; data curation, H.Z.; writing—original draft preparation, J.Z. and B.L.; writing—review and editing, H.Z. and B.L.; visualization, H.Z.; supervision, B.L.; project administration, B.L.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

7. References

[1] Singh, A., & Chakraborty, S. (2020). Deep Active Transfer Learning for Image Recognition. 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, United Kingdom. doi:10.1109/ijcnn48605.2020.9207391.

[2] Wang, K., Chen, K., Du, H., Liu, S., Xu, J., Zhao, J., Chen, H., Liu, Y., & Liu, Y. (2022). New image dataset and new negative sample judgment method for crop pest recognition based on deep learning models. Ecological Informatics, 69, 101620. doi:10.1016/j.ecoinf.2022.101620.

[3] Yang, X., Zhang, Y., Lv, W., & Wang, D. (2021). Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier. Renewable Energy, 163, 386–397. doi:10.1016/j.renene.2020.08.125.

[4] Yang, L., Li, Z., Ma, S., & Yang, X. (2022). Artificial intelligence image recognition based on 5G deep learning edge algorithm of Digestive endoscopy on medical construction. Alexandria Engineering Journal, 61(3), 1852–1863. doi:10.1016/j.aej.2021.07.007.

[5] Anubha Pearline, S., Sathiesh Kumar, V., & Harini, S. (2019). A study on plant recognition using conventional image processing and deep learning approaches. Journal of Intelligent & Fuzzy Systems, 36(3), 1997–2004. doi:10.3233/JIFS-169911.

[6] Zhang, Y., Zhang, Z., Huo, L., Xie, B., & Wang, X. (2021). Image Saliency Detection via Two-Stream Feature Fusion and Adversarial Learning. Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao/Journal of Computer-Aided Design & Computer Graphics, 33(3), 376–384. doi:10.3724/SP.J.1089.2021.18438.

[7] Tsuiki, S., Nagaoka, T., Fukuda, T., Sakamoto, Y., Almeida, F. R., Nakayama, H., Inoue, Y., & Enno, H. (2021). Machine learning for image-based detection of patients with obstructive sleep apnea: an exploratory study. Sleep & Breathing, 25(4), 2297–2305. doi:10.1007/s11325-021-02301-7.

[8] Vinolin, V., & Sucharitha, M. (2022). Taylor-rider-based deep convolutional neural network for image forgery detection in 3D lighting environment. Data Technologies and Applications, 56(1), 103–131. doi:10.1108/DTA-10-2020-0234.

[9] Zhang, Y., Lv, C., Wang, D., Mao, W., & Li, J. (2022). A novel image detection method for internal cracks in corn seeds in an industrial inspection line. Computers and Electronics in Agriculture, 197, 106930–. doi:10.1016/j.compag.2022.106930.

[10] Wang, C., Wang, Q., Wu, H., Zhao, C., Teng, G., & Li, J. (2021). Low-Altitude Remote Sensing Opium Poppy Image Detection Based on Modified YOLOv3. Remote Sensing, 13(11), 2130. doi:10.3390/rs13112130.

[11] Selvathi, D., & Chandralekha, R. (2021). Fetal biometric based abnormality detection during prenatal development using deep learning techniques. Multidimensional Systems and Signal Processing, 33(1), 1–15. doi:10.1007/s11045-021-00765-0.

[12] Nikolova, D., Vladimirov, I., & Terneva, Z. (2021). Human Action Recognition for Pose-based Attention: Methods on the Framework of Image Processing and Deep Learning. 2021 56th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), Sozopol, Bulgaria. doi:10.1109/icest52640.2021.9483503.

[13] Aizezi, Y., Jiamali, A., Abudurexiti, R., Liu, X., Du, J., & Ding, L. (2018). Research on Image Recognition Method Based on Deep Learning Algorithm. 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). doi:10.1109/avss.2018.8639372.

[14] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. doi:10.48550/arXiv.2004.10934.

[15] Zhang, S., & Zhou, X. (2021). MicroNet: Realizing Micro Neural Network via Binarizing GhostNet. 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP). doi:10.1109/icsp51882.2021.9408972.

[16] Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms. doi:10.48550/arXiv.1708.07747.

[17] Liu, Z., Luo, P., Qiu, S., Wang, X., & Tang, X. (2016). DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/cvpr.2016.124.

[18] Pham, M.-T., & Lefevre, S. (2018). Buried Object Detection from B-Scan Ground Penetrating Radar Data Using Faster-RCNN. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. doi:10.1109/igarss.2018.8517683.

[19] Sogabe, M., Ito, N., Miyazaki, T., Kawase, T., Kanno, T., & Kawashima, K. (2022). Detection of Instruments Inserted into Eye in Cataract Surgery Using Single-shot Multibox Detector. Sensors and Materials, 34(1), 47–54. doi:10.18494/SAM3762.

[20] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint. doi:10.48550/arXiv.1804.02767.

[21] Li, Y., Li, J., & Meng, P. (2023). Attention-YOLOV4: a real-time and high-accurate traffic sign detection algorithm. Multimedia Tools and Applications, 82(5), 7567–7582. doi:10.1007/s11042-022-13251-x.