ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

Overview of Target Detection based on Deep Learning

Zhantao Zhao

CHILDREN'S HOSPITAL OF HEBEI PROVINCE, Shijiazhuang, China

*65835059@qq.com

Abstract

Nowadays,one of the research focuses in the field of computer vision is target detection, which aims at identifying and detecting the designated targets in images. At present , deep learning is a technology with rapid development and wide application range, and its learning ability can perform image recognition, feature extraction, classification recognition and other operations in target detection. The research progress of deep learning was mainly introduced. The characteristics of various network models were analyzed. Finally, the application development of deep learning in the field of target detection was expounded, and the future research development is combined with the current problems and challenges was analyzed.

Keywords

Deep Learning; Computer Vision; Target Detection; Feature Extraction.

1. Introduction

With the rapid development of computer technology, computer vision has become the focus of people's research in order to provide new methods and approaches to the field of target detection. Target detection is an important part of computer vision system, and it is a technology that relies on computer to analyze and process images. Its function is to remove the unwanted background information, and finally keep only the desired target. In this process, we can take a series of algorithms to deal with it, then extract the required target from the image, and finally test it.

At present, target detection is mainly divided into traditional methods and detection methods based on deep learning. Traditionally, sliding windows with different sizes are set for the image and some parts of the image are listed as candidate areas. Then, the features of candidate areas are extracted according to manually set feature operators such as SIFT [1], Harr [2] and HOG [3], and finally the features are classified. Although the traditional target detection method has made some achievements, it also has some drawbacks. First of all, sliding window will lead to time complexity; secondly, the diversity of appearance and background and the change of illumination will lead to poor robustness of artificially designed feature methods, and complicated steps will lead to slow detection speed and low accuracy [4]. Therefore, the traditional target detection methods can no longer meet the needs of modern society. In this case, a variety of deep learning models are proposed. The purpose of deep learning is to build a multi-layer network, which enables computers to learn independently and obtain the internal relations contained in data, so as to extract more data and finally make computer learning more expressive.

2. Deep Learning Development

Hinton et al [5] proposed a network structure with multiple hidden layers in 2006, which can describe the deeper abstract features of an object in detail. At that time, the computer performance was poor, and it could not meet the transportation environment of deep learning, so deep learning encountered

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

a certain bottleneck. Nowadays, with the continuous enhancement of computer hardware level, it greatly improves the computer running ability and computing speed, which makes deep learning have a good development environment. And various models for it have been put forward by people and widely used in all walks of life.

Deep learning is based on the flow graph to describe the whole calculation process from input to output. And depth is a special attribute of this flow graph, which refers to the longest length from one input to another output [6]. Deep learning is a self-learning method of human brain data information by constructing multilayer neural network, which can be used to train a large amount of data information at the same time to realize the simulation and analysis of human brain. Deep learning meets the service requirements of different fields by training and learning data information and extracting features [7]. The deep learning model has a high hierarchical structure and strong autonomous learning ability. It can perform operations such as identification, detection and classification well, and can be applied to deal with various complex data analysis problems. Generally, unsupervised learning includes restricted boltz- mann machine, deep Boltzmann machine, deep belief networks and so on. And supervised learning includes convulsive neural networks, recurrent neural networks, etc. Among them, the level of image classification based on convolutional neural network has far exceeded the level of people's recognition, and its ability based on computer translation has reached the normal level of people.

3. Target Detection Method based on Deep Learning.

3.1 Deep Belief Network

Deep belief networks(DBN) is an unsupervised learning method. It is a probabilistic neural network model based on RBM, which has the ability to express the characteristics of the target. DBN is suitable for nonlinear time series prediction with small number of samples.

Tao et al. [8] first tried to apply DBN network to rolling bearing detection, and proposed a vibration signal detection method based on DBN. It uses DBN to construct encoder to minimize the input and output of vibration signals, and at the same time, it classifies the vibration signals by detecting the energy between signal transmission. This fully demonstrates the ability of DBN network to extract features, but it still needs to be improved because of its many parameters, large amount of calculation and low calculation efficiency. Tao et al. [9] and Chen et al. [10] respectively put forward DBN methods on sensing information and feature fusion, which can obtain information and improve accuracy through fusion. However, feature extraction is complicated and feature fusion has limitations, so further research is needed. Gan et al. [11] put forward a method of hierarchical diagnosis based on DBN network. By taking wavelet packet energy as the input characteristic quantity and using DBN network to classify the fault information, the accurate classification identification of bearings can be realized, and the fault information can be accurately diagnosed. However, the adjustment of parameters in the network is complicated and the training process takes a long time, which is not conducive to fault diagnosis. Huang Shouxi et al. [12] obtained a new DBN network by improving RBM, which learned and extracted facial features through training and classified and recognized the features by stacking automatic encoder. This method has good robustness and high detection accuracy, but because some expressions are fuzzy, the recognition rate is low. Zhong et al. [13] proposed a new DBN network, which trained diversified hidden layers by using recursive greedy learning algorithm and diversified weight parameters, and adjusted the parameters by using supervised training method, so that the classification errors were defined on the marked samples and spread to the whole network. This method has good results in image classification accuracy, time and various performances, but due to the limitation of DBN, there will be over-fitting phenomenon to some extent.

3.2 Convolutional Neural Network

Convolution neural network (CNN) uses convolution, pooling and function mapping to extract data information step by step. Finally, the convolution neural network turns the objective into a function,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

obtains the error value through calculation, and the error value is fed back to each layer by BP algorithm, thus updating the parameters in real time to achieve the purpose of training. CNN is suitable for nonlinear time series prediction with large number of samples and high accuracy requirements. Since the appearance of CNN, researchers have done a lot of work in optimizing the structure of convolutional neural networks and improving the complexity of models, aiming at completing complex tasks with simple models with fewer parameters. Nowadays, the common network models mainly include LeNet, AlexNet, ZF-Net, VGG-Nets, etc [14].

LeNet model is mainly used to solve the problem of handwritten numeral recognition. Now, all the models of deep learning that are widely used by people are improved by LeNet model. There are subtle differences between LeNet-5 model and the original LeNet model in design. LeNet-5 adopts Relu activation function and Softamax regression. It is characterized by simple structure, shallow model depth and general feature extraction ability, which is prone to over-fitting. AlexNet model is a network model that has a deeper understanding and application of CNN. Its characteristic is that it can avoid over-fitting of the model, thus improving the training speed and stabilizing the convergence speed of the model. In addition, the problem of gradient disappearance can be solved by Relu. This model has deeper network structure and more parameters, but it increases the amount of calculation. At this point, deep learning and CNN have been remembered by people, and subsequent related research has appeared frequently [15].

Compared with AlexNet model, ZFNet model has no special improvement in neural network structure, but its performance is much higher than AlexNet model. Its activation function and technology are basically the same as those of AlexNet, only a smaller filter is used. Its feature is that it retains more original pixel information, and its performance is better than that of AlexNet. Xu Laixiang et al. [16] proposed an improved ZFNet network model, which adjusted the number of layers of the model by introducing STN and Dropout layers of the spatial transformation network, which improved the recognition rate. However, due to the lack of infrared image data, it is necessary to further improve the ZFNet structure.

The VGG-Nets model is a neural network structure developed by Oxford University and Google DeepMind. By stacking the convolution layer and the pool layer, a Deep Deep convolutional neural network with more than ten layers is constructed, and some conclusions about the depth and performance of CNN can be obtained from the practical tests of deep neural networks. Compared with AlexNet and ZFNet models, VGG-Nets model has the advantage of greatly reducing the error rate. The VGG-Nets model adopts small convolution kernels of 1×1 and 3×3, its network structure is simple, and the classifier is Softamax logistic regression. Because of the small convolution kernel, the expression ability of the network is improved and the structure of the network is deepened, but at the same time, the amount of computation is increased. Li Xiaolin et al. [17] fused the features of local binary pattern with the features extracted from convolution layer and combined with the connection layer of improved VGG model, thus improving the accuracy of expression recognition, which has strong robustness. This method only uses the initial Softmax function, and does not consider the influence of loss function on accuracy, so it has limitations.

3.3 Circulatory Neural Network

Recurrent neural network (RNN), also known as time recurrent neural network, can carry out experimental verification and explain specific problems in specific natural language processing tasks, but its explanation ability is poor, so it is not universal. Common models are: long short term memory network [18], gate recurrent unit [19], Bi-directional Long Short Term Memory Network [20], etc.

Ma et al. [21] combines bidirectional LSTM, CNN and conditional random field, so that the model can learn from characters or words, which does not need to be preprocessed by features or data information, and is suitable for the task of sequential labeling. Mou et al [22] put forward an improved model based on RNN, which inputs hyperspectral pixels into the network model in the form of initial data information for analysis and processing. Zhang et al. [23] put forward a lattice LSTM model based on Chinese named entity recognition. This model can add the word itself to the word vector,

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

thus eliminating the influence of word segmentation errors. However, because Chinese characters have polysemy, this method still has some limitations. Johnston et al. [24] improved the structure of recurrent network to improve the compression performance of the model framework, and also used the weighted pixel loss training based on structural similarity, which can perceive the image more clearly.

Because of the continuous improvement of recurrent neural network, it has made great achievements in image classification and recognition, and provided a new solution for the follow-up research. Although the improved recurrent neural network has some shortcomings, it does not affect the researchers' research on RNN. Researchers will definitely put forward a better network model in the future.

3.4 Generate a Confrontation Network

Generative network (GAN) is mainly composed of two neural network models: generative and discriminant. The main task of model generation is to collect data from random uniform distribution to synthesize output data, and to discriminate the model with formal data or synthesized data as input and output the sample with the true probability. In the training process, the confrontation network is used to generate samples, and the training network is a common method of generating models. GAN is suitable for the task of image generation, which can generate the information in the image for better understanding, but it can't explain the relationship between layers in the network structure, and can't judge whether the network is good or bad.

GAN model has achieved good results in the field of image generation, so researchers began to improve it on the basis of GAN model and developed many new network models. Tang Xianlun et al. [25] put forward a model of conditional depth convolution generation confrontation network by using CNN feature extraction and conditional assistance. This method not only improves the operation speed, but also improves the image recognition rate. However, it is slow in the process of confrontation training, and there is no optimal standard for generating and distinguishing networks, so further research is needed. Shang Xianzhen et al. [26] combined GAN with naive Bayes and put forward a multi-classification diagnosis method. This method can obviously improve the class imbalance and multi-classification problems in data sets, and the recognition accuracy is also improved.

GAN model has many improved models, and Mehdi [27] solves the uncontrollable problem of the network itself by adding additional information to the generated confrontation network. However, this method only uses each tag individually, and does not combine multiple tags, so the search space is limited. Denton et al. [28] generated a higher pixel image by using Gaussian and Lapras pyramid in convolution layer. Arjovsky et al. [29] wanted to solve the problem of unstable network training, so they proposed a Wasserstein GAN model, which removed the last layer of discriminant model and changed some parameters in the generated model and discriminant model. However, this did not solve the problem of unstable network training in the end, but only solved the problem of model collapse. Gulrajani et al. [30] improved the condition of continuity restriction by adopting a new lipschitz continuity restriction technique, which improved the phenomenon of gradient disappearance.

4. Application of Deep Learning

Deep learning has strong feature extraction ability, so it is widely used in various fields of target detection, especially in face, medical images, remote sensing images, pedestrian detection and so on. In the past, because of the shortcomings of traditional methods, target detection has not achieved the desired requirements. However, in recent years, the deep learning technology has been improved, and many researchers have combined the deep learning technology with target detection, finally making the detection effect much better than the traditional methods.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

4.1 Face Detection

Nowadays, face detection has been applied to people's daily life, and it is a common technical means. Usually, face detection is applied to many aspects of daily life, such as e-commerce, video surveillance, payment methods, access control, etc. With the rapid rise of deep learning technology, face detection has gradually become one of the important research and development directions of deep learning technology. With the continuous development of deep learning, many new algorithms have been proposed, some of which still have some technical defects, and some of them have been relatively successful. Among them, the DeepFace model proposed by Taigman et al. [31] is relatively effective in face detection. This method uses shorter features and reduces the computation time of the network, which has potential in other visual fields. Zhang zhixing et al [32] combined MTCNN with DSC and proposed DSC-MTCNN network. Adding DSC and BN layers to the P-Net layer, R-Net layer and O-Net layer in MTCNN, respectively, can reduce the parameter quantity and prevent the gradient from disappearing. This method reduces the amount of calculation, memory space and complexity of calculation, and the detection accuracy is not greatly affected. Li Nan et al. [33] deepened and widened the network by connecting multiple Inception modules in series and simplifying the convolution kernel, then combining Softmax and Trip- letLoss. This method can still maintain a high recognition rate with fewer parameters, and reduce the complexity of network calculation.

Although the face detection technology has developed well in recent years, it still has obvious shortcomings. For example, how to distinguish twins, how to judge the face changes of different ages and how to eliminate the interference of external environment are all problems that need to be solved. Generally speaking, the face detection technology needs to be further improved.

4.2 Medical Image Detection

There are many kinds of medical images, and the images are greatly influenced by the equipment and environment. These problems affect doctors' diagnosis of patients to some extent. Therefore, medical image recognition technology is one of the most advanced medical diagnosis methods at present. Medical image examination is mainly to quickly and accurately find out the pathological information of patients from a large number of medical images, so as to contribute to medical research and treatment.

It is the main algorithm of CNN medical image detection. CNN model can extract information from images and synthesize high-order feature information, so as to realize medical image detection. Zhao et al [34] combined FCN with CRFs and proposed an integrated model to segment brain tumor images. The original image is input into FCN to get the segmentation probability map, and then the original image and the segmentation probability map are transmitted to CRFs and the segmentation result is optimized, which can ensure the consistency between the result and the position information of the original image. This method improves the calculation speed, but when the image is segmented, the network performance will deteriorate due to different pixels. Li Wei et al. [35] improved CNN and put forward a network model of multi-level second-order feature fusion. CNN was used to extract features from multi-layer sectional images of pulmonary nodules, and feature vectors were obtained through sub-modules and fusion modules, and then the feature vectors were classified, and finally the evaluation results of pulmonary nodules were obtained. Experiments show that this method can effectively improve the classification accuracy. Zhang et al [36] put forward a hybrid image monitoring model based on VGG-16. This model can extract more feature information and improve its robustness, but it will produce errors for single image and mixed image, and it needs further research.

At present, there are still many diseases that can't be effectively diagnosed by this technology because of the lack of a large number of data sets related to pathology. Therefore, the establishment of large data sets about medical images and the improvement of network model are the research focuses of medical image detection at present.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

4.3 Remote Sensing Image Detection

With the improvement of computer level, the remote sensing image data has also increased greatly, and the traditional image recognition methods can't effectively extract valuable information. Deep learning technology has strong learning ability, which can synthesize low-order features into high-order features, and can use deep structure model to identify and classify remote sensing images. Shi Wenxu et al [37] increased the detection performance by adding feature fusion module and feature enhancement module to SSD network model. This method can enhance the adaptability of the network and improve the detection accuracy, but further experiments are needed for small-scale targets. Zhang et al. [38] put forward an unsupervised super-resolution GAN model for remote sensing images. This model uses average pooling method to degrade remote sensing images. Although the generalization ability of the model is improved, there are too many factors affecting image degradation. Zhang et al. [39] used transfer learning method to model HR and LR of remote sensing in different situations, and used channel attention mechanism to fuse features of different depths. This method has obvious advantages for the modeled scenes, but it has a large number of model parameters and poor practical application effect. Generally speaking, it is suitable for remote sensing images with complex structures. Dong et al. [40] designed a dense sampling mechanism, which widened the channel of features and transferred features of different depths to the upsampler for reconstruction. This method can effectively use features of different depths and enhance the utilization rate of features, but it has poor reconstruction ability for tiny targets in remote sensing images, so it is suitable for reconstruction of larger targets.

At present, remote sensing image detection is mainly based on manual recognition, and its main data comes from satellite images. However, satellite images have huge data and low resolution, which makes it very difficult to detect targets. Therefore, the key point of remote sensing image detection is to accurately extract valuable information, and deep learning can achieve the above operations. However, the current technology is still insufficient, and there are still many defects in practical application. Therefore, it is an important research direction to create appropriate models and optimize algorithms.

4.4 Pedestrian Detection

Pedestrian detection is mainly used in video surveillance, vehicle navigation and self-driving Compared with target detection, pedestrian detection is more complicated. Different pedestrians are static and dynamic at the same time, and they are easily affected by position, posture, illumination, background and image blur, which increases the difficulty of pedestrian detection. Zhang et al. [41] found that the effect of Faster RCNN in pedestrian detection is not ideal. After analyzing it, a method of RPN processing small targets and negative samples, and then using random forest to classify them is obtained. Gu Wei et al [42] put forward a multi-channel feature model based on multi-feature fusion, which can reduce the computational burden and improve the efficiency. This has high recognition accuracy for pedestrian detection, and it is suitable for blocked pedestrians. Because this method uses two-stage detection method, there is no way to solve the problem of highly overlapping targets. Xie Yongming et al. [43] proposed an improved Faster R-CNN algorithm, which can effectively remove the interference of background information in the image. The detection accuracy of this method is obviously improved, but the detection speed is slightly decreased, so it is necessary to optimize the network to improve the detection efficiency.

5. Outlook

Deep learning technology has become a research hotspot because of its superior learning ability and superiority in complex environment. Although the technology has made great achievements, it still needs further development. Here are some discussions on future research.

(1) Aiming at the problem that the target in the image is tiny, occluded and shadowed, how to detect the target to meet the application requirements of real life has become a problem that needs to be solved at present. In order to improve the detection of the target, attention module can be used in

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

shallow structure to integrate information, and the real-time monitoring function of the target can be realized by changing the calculation method, and the detector with adaptive scale can be used to locate the detected target.

(2) For small data sets, because they can't carry out deep network training, even if transfer learning is used to adjust the data sets, it will make the effect worse. However, for large data sets, the current data sets lack diversity, and the data sets need to be manually marked, and they are easily influenced by external factors. Therefore, the study of large-scale and diverse data sets is one of the key studies.

(3) The current detection algorithm needs to mark the image data set completely, but in the data set, it takes a long time to manually mark the target, which also imposes a serious burden on the algorithm and is difficult to realize. Realizing weak supervised target detection can effectively improve this problem, and it can detect unlabeled images according to a small number of labeled images, thus greatly reducing the difficulty. Therefore, weak supervised target detection method is a research focus.

(4) How to combine multiple tasks in one network and improve the detection accuracy is a challenge for researchers. The network architecture of accumulating multiple sub-features is an important method to improve the performance of target detection. When multiple computer vision tasks are carried out simultaneously, it can obtain rich information and greatly improve the performance of a single computer vision task.

References

[1] Lowe D G.Distinctive image features from scale-invariant keypoints [J].International Journal of Computer Vision ,2004 ,60 (2) :91-110.

[2] Lienhart R,Maydt J.An extended set of haar-like features for rapid object detection [C] / /International Conference on Image Processing.NewYork : IEEE,2002 : I-900-I-903.

[3] Yuan Xiaoping, Ma Xunqi, Liu Sai . An improved algorithm of pedestrian and vehicle detection based on YOLOv3[J].Science Technology and Engineering,2021,21 (8) : 3192-3198.

[4] Ling Chen , Zhang Xintong , Ma Lei.Remote sensing image processing technology and its application based on mask R-CNN algorithms [J].Compute Science,2020,47 (10) : 151-160.

[5] Hinton G E ,Salakhutdinov R R.Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313 (5786) : 504.

[6] Zheng Yuanpan ,Li Guangyang ,Li Ye.Survey of application of deep learning in image recognition [J]. Computer Engineering and Applications,2019 (12) : 20-36.

[7] Guo Y M ,Liu Y ,Ard O, et al.Deep learning for visual under-standing : a review [J].Neurocomputing, 2016,187 : 27-48.

[8] Tao J,Liu Y L,Yang D L, et al.Fault diagnosis of rolling bear-ing using deep belief networks [C]//Proceedings of the 2015 In- ternationgal Symposium on Material and Environment Engineering. Dordrecht : Atlantis Press,2015 : 566-569.

[9] Tao J,Liu Y L,Yang D L.Bearing fault diagnosis based on deep belief network and multisensor information fusion [J].Shock and Vibration,2016 (6) : 1-9.

[10] Chen Z Y, Li W H.Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network [J].IEEE Transactions on Instrumentation and Measurement,2017,66 (7) : 1693-1702.

[11] Gan M,Wang C, Zhu C A.Construction of hierarchical diagnosis network based on deep learning and its application in the fault pat- tern recognition of rolling element bearings [J].Mechanical Sys- tems and Signal Processing,2016 (72-73) : 92-104.

[12] Huang Shouxi,Qiu Weigen.Facial expression recognition via im- proved deep belief networks [J].Computer Engineering and De- sign,2017,38 (6) : 1580-1584.

[13] Zhong P,Gong Z Q,Li S T,et al.Learning to diversify deep belief networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing,2017,55 (6) :3516-3530.

[14] LeCun Y, Bottou L, Bengio Y,et al. Gradient-based learning applied todocumentrecognition [J].Proceedings of the IEEE,1998,86 (11) : 2278-2324.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

[15] Lin Kun,Lei Yinjie.Research on palm vein recognition algorithm based on improved AlexNet convolution neural network [J].Mdo- ern Electronics Technique,2020,43 (7) : 52-56.

[16] Xu Laixiang,Liu Gang,Liu Sen,et al.Research on infrared tar- get recognition based on improved convolution neural network [J].Fire Control & Command Control,2020,45 (8) : 136-141.

[17] Li Xiaolin,Niu Haitao.Facial expression recognition using feature fusion based on VGG-NET [J]. Computer Engineering& Science,2020,42 (3) : 500-509.

[18] Cho K,Van Merrienbore B ,Gulcehre C, et al.Learning phrase representations using RNN encoder-decoder for statistical machine translatio [C] // Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing.Doha ,Qatar : Association for Computational Linguistics,2014 : 1724-1734.

[19] Greff K,Srivastava R K,Koutnik J,et al.LSTM : A search space odyssey [J].IEEE Transactions on Neural Networks and Learning Systems,2016,28 (10) : 2222-2232.

[20] Schuster M,Paliwal K K.Bidirectional recurrent neural networks [J].IEEE Transactions on Signal Processing,1997,45 (11):2673-2681.

[21] Ma X Z, Hovy E.End-to-end sequence labeling via bi-directional lstm-cnns-crf [C] / / Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics.Berlin : Association for Computational Linguistics,2016 : 1064-1074.

[22] Mou L C,Ghamisi P,Zhu X X. Deep recurrent neural networks for hyperspectral image classification [J].IEEE Transactions on Geoscience and Remote Sensing,2017,55 (7) : 3639-3655.

[23] Zhang Y,Yang J.Chinese NER using lattice LSTM [J].Association for Computational Linguistics,2018,1 : 1554-1564.

[24] Johnston N,Vincent D ,Minnen D , et al.Improved lossy image compression with priming and spatinally adaptive bit rates for re- current networks[C] / / Proceedings of the IEEE Conference on Computer Visionand Pattern Recognition. Salt Lake : IEEE,2018 : 4385-4393.

[25] Tang Xianlun,Du Yiming,Liu Yuwei,et al.Image Recognition With Conditional Deep Convolutional Generative Adversarial Net- works [J]. Acta Automatica Sinica,2018,44 (5) : 855-864.

[26] Shang Xianzhen,Han Meng ,Sun Yuzhong , et al.Skin diseases diagnosis method based on generative adversarial networks and na- ive bayes [J].Journal of Frontiers of Computer Science and Tech- nology, 2019,13 (6) : 1005-1015.

[27] Mirza M, Osindero S.Conditional generative adversarial nets [EB /OL].(2014-11-06) [2020-04-09] https://www.research-gatenet/publication/268079622_Conditional_Generative_Adversarial_Nets.

[28] Denton E ,Chintala S ,Szlam A , et al. Deep generative image models using a laplacian pyramid of adversarial networks [C] / / Proceeding of the IEEE International Conference on Computer Vision. Venice : IEEE,2015 : 1486-1494.

[29] Arjovsky M,Chintala S,Bottou L.Wasserstein GAN [EB /OL].(2017-12-06) [2018-12-25]. https://www. researchgate.net /publication / 313044478_Wasserstein_GAN.

[30] Gulrajani I,Ahmed F ,Arjovsky M , et al. Improved training of Wasserstein GANs[C] / / Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach : Curran Associates,Inc.,2017 : 5767-5777.

[31] Taigman Y,Yang M,Ranzato M A, et al. DeepFace : closing the gap to human-level performance in face verification [C] / /Proceed- ings of IEEE Conference on Computer Vision and Pattern Recognition. Columbus : IEEE,2014 : 1701-1708.

[32] Zhang Zhixing ,Liu Jingyi.Intelligent face recognition guidance system based on deep learning[J]. Information Technology and Informatization,2020 (4) : 210-213,217.

[33] Li Nan ,Cai Jianyong ,Li Ke , et al.Face recogniton algorithms based on convolutional neural network with Multi-Inception struc- ture [J]. Computer Systems& Applications ,2020 ,29 (2) :157-162.

[34] Zhao X M,Wu Y H,Song G D,et al.A deep learning model integrating FCNNs and CRFs for brain tumor segmentation [J].Medical Image Analysis,2018,43 : 98-111.

[35] Li Wei,Zhao Xiaole,Duan Yanlong, et al.Classification of pul- monary nodules based on CNN multi-level second-order feature fu- sion [J].Journal of Frontiers of Computer Science and Technolo- gy,2020,14 (9) : 1590-1601.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202209_8(9).0038

[36] Zhang Guobiao,Li Xin,Lu Wei.Biomedical compound figure de- tection based on deep convolutional neural network [J].Journal of Modern Information,2020,40 (7) : 74-81,103.

[37] Shi Wenxu,Bao Jiahui,Yao Yu.Remote sensing image target de- tection and identification based on deep learning [J].Journal of Computer Applications,2020,40 (12) : 3558-3562.

[38] Zhang N,Wang Y,Zhang X,et al.An unsupervised remote sensing single-image super-resolution method based on generative ad- versarial network [J].IEEE Access,2020,8 : 29027-29039.

[39] Zhang S,Yuan Q,Li J, et al.Scene-adaptive remote sensing image super-resolution using a multiscale attention network [J].IEEE Trans-actions on Geoscience and Remote Sensing,2020,58 (7) : 4764-4779.

[40] Dong X Y,Sun X,Jia X P,et al. Remote sensing image super resolution using novel Dense-sampling networks [J].IEEE Trans- actionson Geoscience and Remote Sensing ,2021 ,59(2) :1618-1633.

[41] Zhang L L,Lin L,Liang X D,et al.Is faster R_CNN doing well for pedestrian detection [C] / / European Conference on Computer Vision. Cham : Springer,2016 : 443-457.

[42] Gu Wei ,Li Feifei ,Chen Qiu.Pedestrian detection algorithm based on multiple feature fusion [J].Electronic Science and Tech- nology,2021,34 (5) : 29-34.

[43] Xie Yongming,Wang Honglei. Improved algorithm for long-distance and small-size pedestrian detection in complex backgrounds [J].Computer Engineeting and Design,2021,42 (5) : 1323-1330.