Title: Facial expression recognition based on deep learning

Article preview

Computer Methods and Programs in Biomedicine
Volume 215, March 2022, 106621
Facial expression recognition based on deep learning
Author links open overlay panelHuilin Ge, Zhiyu Zhu, Yuewei Dai, Biao Wang, Xuedong Wu

https://doi.org/10.1016/j.cmpb.2022.106621
Highlights
 * ?
Autonomous driving, virtual reality and all kinds of robots integrated into our life rely on facial expression recognition technology.
 * ?
Facial expression recognition and computer vision is based on deep learning technology and convolutional neural network.
 * ?
Whether it is two-stage target detection or single-stage target detection, performance of algorithm is measured by detection speed and accuracy.
 * ?
Large variety of training data with accurate expression tags can fundamentally improve expression recognition rate.
Abstract
Background and objective
Facial expression recognition technology will play an increasingly important role in our daily life. Autonomous driving, virtual reality and all kinds of robots integrated into our life depend on the development of facial expression recognition technology. Many tasks in the field of computer vision are based on deep learning technology and convolutional neural network. The paper proposes an occluded expression recognition model based on the generated

countermeasure network. The model is divided into two modules, namely, occluded face image restoration and face recognition.
Methods
Firstly, this paper summarizes the research status of deep facial expression recognition methods in recent ten years and the development of related facial expression database. Then, the current facial expression recognition methods based on deep learning are divided into two categories: Static facial expression recognition and dynamic facial expression recognition. The two methodswill be introduced and summarized respectively. Aiming at the advanced deep expression recognition algorithms in the field, the performance of these algorithms on common expression databases is compared, and the strengths and weaknesses of these algorithms are analyzed in detail.
Discussion and results
As the task of facial expression recognition is gradually transferred from the controlled laboratory environment to the challenging real-world environment, with the rapid development of deep learning technology, deep neural network can learn discriminative features, and is gradually applied to automatic facial expression recognition task. The current deep facial expression recognition system is committed to solve the following two problems: (1) Overfitting due to lack of sufficient training data; (2) In the real world environment, other variables that have nothing to do with expression bring interference problems.
Conclusion
From the perspective of algorithm, combining other expression models, such as facial action unit model and pleasure arousal dimension model, as well as other multimodal models, such as audio mode, 3D face depth information and human physiological information, can make expression recognition more practical.
Introduction
Image recognition is a technology that uses computer to process, analyze and understand images to identify different patterns of targets and objects. It is a main research direction in the field of computer vision and plays an important role in intelligent data acquisition and processing based on image. Image recognition technology can efficiently complete the detection and recognition of specific target objects (such as handwritten characters, products or faces), image classification and marking, and subjective image quality evaluation. At present, image recognition technology has a broad commercial market and optimistic application prospects in Internet application products such as image retrieval, commodity recommendation, user behavior analysis and face recognition. Moreover, it has long-term development potential in high-tech industries such as UAV, autonomous driving and intelligent robot, as well as many fields such as geology, medicine and biology. Early image recognition systems mainly used directional gradient histogram [1] and scale invariant feature transformation [2], and then input the extracted features into the classifier for classification and recognition. These functions are basically manually designed. For different recognition problems, the extracted features will directly affect the performance of the system. Therefore, researchers need to further study the unsolved problem areas in order to design better adaptive features to improve the performance of the system. The image recognition system in this period is often for a specific recognition task, and the data scale is small, the generalization ability is poor, so it is not easy to achieve the ideal recognition effect in daily application.
Deep learning is a branch of representation learning based on data of

artificial neural network. Deep learning includes supervised learning, semi-supervised learning and unsupervised learning. In deep learning, deep neural network, deep belief network and recurrent neural network have been widely used in speech recognition, computer vision, audio recognition, unmanned driving and natural language processing. Rina dechter first introduced deep learning in 1986. In addition, Igor aizenberg introduced the concept of artificial neural network in 2000. In fact, Alexey ivakhnenko and Lapa proposed supervised feedforward learning network as early as 1965. In 1986, Geoffrey Hinton proposed the back-propagation algorithm of multilayer perceptron (MLS), and used the SIGMOD activation function for nonlinear transformation, so as to solve the problems of nonlinear learning and classification. In 2006, Geoffrey Hinton et al. Pointed out that the pre-training weight should be used to initialize the model and fine tune the model according to the supervised training. With the further development of deep learning, the proposal of lenet [3] in 1998 marked the emergence of convolutional neural network (CNN).

However, due to the backward hardware at that time, convolutional neural network is not in an advantage compared with other machine learning methods (such as SVM [4]). With the further development of computing devices, the Alex net that is proposed by Hinton et al. has made significant achievements in the computer vision competition ILSVRC 2012. Convolutional neural networks have made rapid progress in recent decades.

Section snippets

Review of face recognition technology

Face recognition has the characteristics of easy access, easy operation and diversified features. In the last century, many scholars studied face recognition. However, due to the underdeveloped network, limited resources of face images and poor quality of photos, many scholars mostly studied it from the perspective of algorithm, but the recognition accuracy is low, far from the human eye recognition effect. With the gradual maturity of machine learning technology, there are many powerful

Depth target detection algorithm

In the realm of computer vision, object recognition technology is an algorithm that can detect sample objects in videos and photos. Recent target recognition algorithms mostly rely on high-performance GPU chips based on multilayer neural network and deep learning software framework and built-in thousands of stream processors. Therefore, it is also known as deep object detection (Deep OD). Object detection has always been a vital issue in the realm of computer vision. Before deep learning is

Deep facial expression recognition based on static image

Due to the convenience and availability of network static data processing, a large number of researches are based on static images without considering time information for expression recognition. Direct training of deep network on relatively small facial expression databases will inevitably result in overfitting problem. In order to solve this problem, many related researches use additional auxiliary data to pre-train and build their own network, or directly fine tune based on an effective

Discussion

This paper firstly recommends the background knowledge of facial expression recognition, and summarizes the evolution and development of database and algorithm in the field of facial expression recognition. It points out that deep learning has become the mainstream framework in this field. Then, the expression recognition algorithms based on deep learning are divided into two categories (static expression recognition network and dynamic expression

recognition network). By comparing the
Conclusions
In the field of computer vision, based on the research status at home and abroad, facial expression recognition technology has made great progress and development. However, there are still many challenges and difficulties waiting for researchers to solve. For example, the research on facial expression recognition in real scenes, and there is a certain confusion between different expressions. The expression of facial emotion may vary with region, culture, and environment. There are differences
Declaration of Competing Interest
The authors declare that there is no conflict of interest in this paper.
Acknowledgment
This work is supported by the National Natural Science Foundation of China (No. 62006102).
Special issue articlesRecommended articles
References (38)
* Z. Tang _et al._
Two-phase deep learning model for short-term wind direction forecasting
Renew. Energy
(2021)
* C. He
Overview of face recognition technology
Intell. Comput. Appl.
(2016)
* M.H. Zhu _et al._
Facial expression recognition method based on sparse representation
Pattern Recognit. Artif. Intell.
(2014)
* Y.L. Xue _et al._
Robust facial expression recognition method under occlusion conditions
J. Beijing Univ. Aeronaut. Astronaut.
(2010)
* Y. Chenglin _et al._
Face recognition framework based on effective computing and adversarial neural network and its implementation in machine vision for social robots
Comput. Electr. Eng.
(2021)
* Y. Sun _et al._
Deep learning face representation from predicting 10,000 classes
* Y. Sun _et al._
Deep learning face representation by joint identification verification
* Y. Sun _et al._
Deeply learned face representations are sparse, selective, and robust
* F. Schroff _et al._
Facenet: a unified embedding for face recognition and clustering
* Y. Taigman _et al._
Deepface: closing the gap to human-level performance in face verification
* Y. Wen _et al._
A discriminative feature learning approach for deep face recognition
European Conference on Computer Vision
(2016)
* W. Liu _et al._
Large-margin softmax loss for convolutional neural networks
* F. Wang _et al._

Additive margin softmax for face verification
(2018)
* H. Wang _et al._
CosFace: large margin cosine loss for deep face recognition
* S. Chen _et al._
MobileFaceNets: efficient CNNs for accurate real-time face verification on mobile devices
Chinese Conference on Biometric Recognition
(2018)
* M. Sandler _et al._
MobileNetV2: inverted residuals and linear bottlenecks
* R. Girshick _et al._
R-CNN for object detection
* Zhu G., Porikli F., Li H. Tracking randomly moving objects on edge box proposals. arXiv preprint arXiv: 1507.08085,...
* J.R.R. Uijlings _et al._
Selective search for object recognition
Int. J. Comput. Vis.
(2013)
View more references
Cited by (86)
* ### Deep neural network technique for automated detection of ADHD and CD using ECG signal
2023, Computer Methods and Programs in Biomedicine
Citation Excerpt :
In this study, the ECG segments were classified into three categories: CD, ADHD + CD, and CD using 1-Dimensional (1D) Convolutional Neural Network (CNN) models. The CNN model is well-known for its ability to classify images, and as a result, it has been used in applications like face and object identification [26,27], satellite forecasting [28,29] and analysis of medical images such as MRI, CT, X-RAY, and PET [30,31]. Besides 2-dimensional images, CNN models have also been applied to 1-dimensional biosignals, such as ECG to identify arrhythmias [32,33].
Show abstract
Attention Deficit Hyperactivity problem (ADHD) is a common neurodevelopment problem in children and adolescents that can lead to long-term challenges in life outcomes if left untreated. Also, ADHD is frequently associated with Conduct Disorder (CD), and multiple research have found similarities in clinical signs and behavioral symptoms between both diseases, making differentiation between ADHD, ADHD comorbid with CD (ADHD+CD), and CD a subjective diagnosis. Therefore, the goal of this pilot study is to create the first explainable deep learning (DL) model for objective ECG-based ADHD/CD diagnosis as having an objective biomarker may improve diagnostic accuracy. The dataset used in this study consist of ECG data collected from 45 ADHD, 62 ADHD+CD, and 16 CD patients at the Child Guidance Clinic in Singapore. The ECG data were segmented into 2 s epochs and directly used to train our 1-dimensional (1D) convolutional neural network (CNN) model.
The proposed model yielded 96.04% classification accuracy, 96.26% precision, 95.99% sensitivity, and 96.11% F1-score. The Gradient-weighted class activation mapping (Grad-CAM) function was also used to highlight the important ECG characteristics at specific time points that most impact the classification score.
In addition to achieving model performance results with our suggested DL method, Grad-CAM's implementation also offers vital temporal data that clinicians and other mental healthcare professionals can use to make wise

medical judgments. We hope that by conducting this pilot study, we will be able to encourage larger-scale research with a larger biosignal dataset. Hence allowing biosignal-based computer-aided diagnostic (CAD) tools to be implemented in healthcare and ambulatory settings, as ECG can be easily obtained via wearable devices such as smartwatches.

* ### Deep learning and computer vision based occupancy CO<inf>2</inf> level prediction for demand-controlled ventilation (DCV)
2022, Journal of Building Engineering
Show abstract
The present study investigated the potential of the application of a live occupancy detection approach to assist the operations of demand-controlled ventilation (DCV) systems to ensure that sufficient interior thermal conditions and air quality were attained while reducing unnecessary building energy loads to improve building energy performance. Faster region-based convolutional neural network (RCNN) models were trained to detect the number of people and occupancy activities respectively, and deployed to an artificial intelligence (AI)-powered camera. Experimental tests were carried out within a case study room to assess the performance of this approach. Due to the less complexity of people counting model, it achieved an average intersection over union (IoU) detection accuracy of about 98.9%, which was higher than activity detection model of about 88.5%. During the detection, the count-based occupancy profiles were produced according to the real-time information about the number of people and their activities. To estimate the effect of this approach on indoor air quality and energy demand, scenario-based modelling of the case study building under four ventilation scenarios was carried out via building energy simulation (BES). Results showed that the proposed approach could provide demand-driven ventilation controls data on the dynamic changes of occupancy to improve the indoor air quality (IAQ) and address the problem of under- or over-estimation of the ventilation demand when using the static or fixed profiles.

* ### SMD-YOLO: An efficient and lightweight detection method for mask wearing status during the COVID-19 pandemic
2022, Computer Methods and Programs in Biomedicine
Citation Excerpt :
Therefore, it is of great practical significance to realize the detection for mask wearing status in public places (such as hospitals, campuses etc.). In recent years, a large number of studies have used deep learning to complete object detection and are widely used in biomedicine [6?8], lesions detection [9?11], face detection [12?14] and other fields [15?18]. The existing machine learning and deep learning methods have achieved some results in the task of face mask detection [19?22], however, there are still limitations in the complicated environments.
Show abstract
At present, the COVID-19 epidemic is still spreading worldwide and wearing a mask in public areas is an effective way to prevent the spread of the respiratory virus. Although there are many deep learning methods used for detecting the face masks, there are few lightweight detectors having a good effect on small or medium-size face masks detection in the complicated environments.
In this work we propose an efficient and lightweight detection method based on YOLOv4-tiny, and a face mask detection and monitoring system for mask wearing status. Two feasible improvement strategies, network structure optimization and K-means++ clustering algorithm, are utilized for improving the detection accuracy on the premise of ensuring the real-time face masks recognition.

Particularly, the improved residual module and cross fusion module are designed to aim at extracting the features of small or medium-size targets effectively. Moreover, the enhanced dual attention mechanism and the improved spatial pyramid pooling module are employed for merging sufficiently the deep and shallow semantic information and expanding the receptive field. Afterwards, the detection accuracy is compensated through the combination of activation functions. Finally, the depthwise separable convolution module is used to reduce the quantity of parameters and improve the detection efficiency. Our proposed detector is evaluated on a public face mask dataset, and an ablation experiment is also provided to verify the effectiveness of our proposed model, which is compared with the state-of-the-art (SOTA) models as well.

Our proposed detector increases the AP (average precision) values in each category of the public face mask dataset compared with the original YOLOv4-tiny. The mAP (mean average precision) is improved by 4.56% and the speed reaches 92.81 FPS. Meanwhile, the quantity of parameters and the FLOPs (floating-point operations) are reduced by 1/3, 16.48%, respectively.

The proposed detector achieves better overall detection performance compared with other SOTA detectors for real-time mask detection, demonstrated the superiority with both theoretical value and practical significance. The developed system also brings greater flexibility to the application of face mask detection in hospitals, campuses, communities, etc.

Haixia Xiao, Zhengfa Hu
 * ### A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis
Neurocomputing, Volume 323, 2019, pp. 62-75
Zhu Zhiyu, ?, Gao Huijun
Show 3 more articles
Article Metrics
Citations
 * Citation Indexes86
Captures
 * Readers93
Social Media
 * Shares, Likes & Comments1
View details
 * About ScienceDirect
 * Remote access
 * Shopping cart
 * Advertise
 * Contact and support
 * Terms and conditions
 * Privacy policy

services we have added to our pages. If you do not allow these cookies then some or all of these services may not function properly.
Cookie Details List?
Performance Cookies
Performance Cookies
These cookies allow us to count visits and traffic sources so we can measure and improve the performance of our site. They help us to know which pages are the most and least popular and see how visitors move around the site.
Cookie Details List?
Targeting Cookies
Targeting Cookies
These cookies may be set through our site by our advertising partners. They may be used by those companies to build a profile of your interests and show you relevant adverts on other sites. If you do not allow these cookies, you will experience less targeted advertising.
Cookie Details List?
Back Button
Cookie List
Search Icon
Filter Icon
Clear
checkbox label label
Apply Cancel
Consent Leg.Interest
checkbox label label
checkbox label label
checkbox label label
Confirm my choices