WEP Warwick Evans Publishing

Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 7
TPCEE 2024

Research on Image Recognition Applications Driven by Artificial Intelligence

Yuyun Xiang*

School of Optoelectronic Information Science and Engineering, Hangzhou Dianzi University, Hangzhou, Zhejiang 310000, China

*Corresponding Author: 22320322@hdu.edu.cn

Abstract. With the rapid development of artificial intelligence and machine learning technologies, image recognition technology has become a key force driving progress in many fields. Deep learning and machine learning algorithms have made remarkable progress in image recognition. Therefore, this article deeply explores the role of image recognition technology in three aspects: autonomous driving, medical imaging and face recognition. This article concludes that autonomous driving technology mainly relies on artificial intelligence and machine learning. It perceives the environment in real time through sensors and perception modules, and formulates driving strategies through decision-making and control modules to ensure the safe driving of vehicles. In medical image analysis, image recognition technology is used for feature extraction, image segmentation, and classification to assist doctors in disease diagnosis and treatment planning. In face recognition technology, through the three stages of face detection, face alignment and final recognition, individual identities are confirmed to improve the accuracy and efficiency of recognition. In conclusion, image recognition technology plays a vital role in the above fields, improving the intelligence and reliability of the system.

Keywords: image recognition technology; autonomous driving; medical imaging; and face recognition.

1. Introduction

Artificial intelligence is a technology and system that simulates human intelligent thinking and behavior. It studies how computers imitate and complete various tasks that human intelligence can accomplish. The development of the AI field has been promoted through different branches such as computer vision, natural language processing, machine learning, and deep learning. Artificial intelligence plays a key role in modern society, promoting innovation, improving efficiency and significantly enhancing the quality of life. It has profoundly impacted social and economic development and human life, and has become an important driving force for promoting scientific and technological progress and economic growth. Among them, image recognition is an important application in the field of computer vision, integrating multiple advanced technologies such as big data, cloud computing and blockchain. The images are analyzed and processed through computer algorithms. Its main goal is to enable computers to understand and interpret the content of images, thereby achieving more diverse and intelligent applications and services. Image recognition technology has been widely applied in numerous fields. For example, in autonomous driving, image recognition is one of its key technologies, and it mainly includes three parts: perception technology, planning technology, and control technology [1].

Perception technology refers to the ability of the autonomous driving system to acquire, perceive and understand the surrounding environment through various devices. Compared with traditional cars, autonomous vehicles are equipped with various sensing devices, which can make full use of image recognition technology to accurately identify the driving environment. These sensing devices enable autonomous vehicles to independently analyze and select reasonable driving routes, thereby achieving safe and efficient autonomous driving, thus greatly reducing potential safety hazards and traffic accidents [2]. Image recognition technology is also applied in the medical field. Nuclear medical imaging technology plays an important role in diagnosing multiple myeloma and monitoring disease progression. It can assess the extent and distribution of bone lesions, and detect osteolytic lesions, osteoporosis and bone marrow infiltration [3].

Under the above research background, this article aims to explore the application of image recognition technology in autonomous driving, medical image analysis, and face recognition in security monitoring. First of all, this article introduces the development history of image recognition technology as well as related algorithms and research progress. Then, this article conducts a detailed analysis of the role of image recognition in the three application above fields. Finally, it summarizes the advantages and limitations of image recognition technology and looks forward to its future development direction and potential improvement space.

2. Image Recognition Overview

2.1. The Development History of Image Recognition

Image recognition is an important artificial intelligence technology,and its goal is to enable computers to understand and analyze the content of images. Through deep learning and machine learning algorithms, image recognition systems can automatically extract the features of images and, through multi-level feature extraction and combination, gradually build an understanding of the image content. This technology enables image recognition systems to recognize and classify various kinds of information such as objects, scenes, and texts in images, achieving a comprehensive understanding and processing of image content. Image recognition technology has undergone three main stages: character recognition, data image processing and recognition, and object recognition. In the 1950s, the character recognition stage mainly dealt with text-type data, such as numbers and characters. From 1965 to the early 20th century, the stage of data image processing mainly focused on the recognition and processing of digital images. Digital images have many advantages over analog images, providing impetus for technological development. At present, the object recognition stage involves perceiving objects and environments in the three-dimensional world and belongs to the advanced field of computer vision. This stage is based on digital image processing and recognition, combined with disciplines such as artificial intelligence and systematics, and is widely applied in fields such as industry and exploration robots.

2.2. Algorithms included in image recognition technology

Image recognition technology mainly uses deep learning and machine learning techniques. Deep learning is a type of machine learning method based on artificial neural network models. It learns and extracts data features through multi-level neural network structures, thereby recognizing and classifying data such as images, texts, and voices. At present, the mainstream deep learning algorithms include Deep Belief Network (DBN), Recurrent Neural Network (RNN), and Convolutional Neural Network (CNN), etc [4].

DBN is a probabilistic generative model proposed by Hinton [5], which is used for unsupervised learning. It usually comprises stacked Restricted Boltzmann Machine (RBM) layers and a logistic regression layer towards the output. The training of DBN is divided into two parts: pre-training and fine-tuning. In the pre-training stage, each RBM layer is trained separately to learn the distribution characteristics of the data. In the fine-tuning stage, the entire network is trained end-to-end as a multilayer perceptron (MLP) to optimize network performance further.

RNN is a dynamic artificial neural network. It performs well when dealing with problems with temporal and spatial correlations and can better capture useful information in time series [6]. It can handle variable-length sequence inputs and introduce temporal information into the model, enabling the model to understand better and utilize the structural information in the sequence data.

CNN is a kind of feedforward neural network, which is established inspired by the receptive mechanism of cells in the animal visual cortex[7]. CNN was first widely used in large-scale image recognition tasks. Its main structure includes the input layer, alternating convolutional layers and pooling layers, and the final fully connected layer (Figure 1). At present, CNN has been widely applied in the fields of computer vision and image processing.

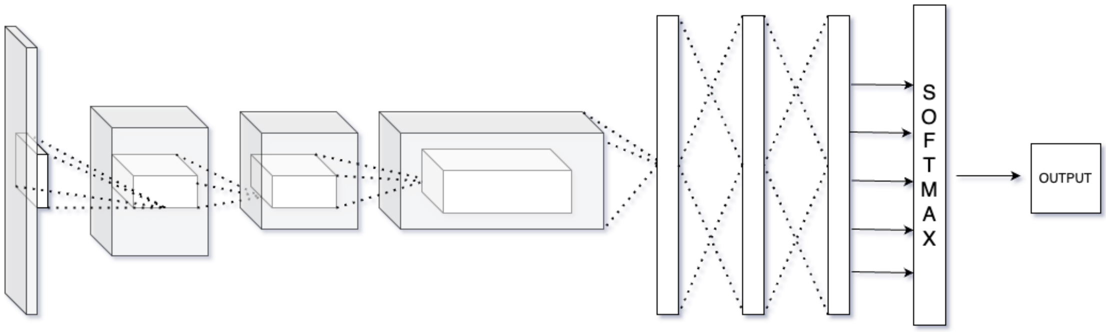

Figure 1. CNN for Image Recognition

3. Image Recognition Application Fields:

Image recognition technology has demonstrated strong potential for application in medical image analysis, autonomous driving, security monitoring, and industrial quality inspection. In medical imaging, it assists doctors in disease diagnosis, improving accuracy and efficiency; in autonomous driving, it is used for environmental perception and path planning to ensure safety; in security monitoring, it is applied to face recognition and behavior analysis to enhance public safety; in industrial quality inspection, it is used to detect product defects and improve production quality and efficiency. This article analyzes these fields mainly because of their wide application and significance in image recognition technology. These fields represent the cutting-edge directions of technology and have significant economic and social values in real life and industrial production.

3.1. Autonomous Driving:

Autonomous driving technology is an advanced technology based on artificial intelligence and machine learning, aiming to achieve autonomous driving of vehicles. Image recognition plays a key role in it. It is used to identify elements such as roads, vehicles, and pedestrians, help vehicles make intelligent decisions and plan paths, and improve driving safety and efficiency. The autonomous driving system consists of sensors, perception, and decision control modules. Intelligent decision-making and vehicle control are achieved through real-time perception and environmental analysis. This article will detail two key modules of the autonomous driving system.

First, there are the sensor and perception module. Environmental perception is divided into two parts: visual and non-visual.

Visual perception mainly relies on image data, including obstacles, road conditions, and vehicle information, which is used to make autonomous driving decisions and optimize the system. Visual perception tasks include traffic target detection and driving environment monitoring, applying reinforcement learning algorithms and deep learning techniques. In terms of object detection, Alam, Ahmed, et al. proposed an enhanced framework based on the Faster R-CNN network to identify and detect fast-moving cars [8]. Song Huajie, Zhou Lei et al. proposed a vehicle detection and recognition algorithm based on the function-improved YOLOv3 to improve the detection accuracy and speed [9]. In terms of the driving environment, Ding Zeliang, Hu Yuhui et al. proposed an adaptive scene road surface extraction method based on deep learning to enhance the adaptability of autonomous driving vehicles to the environment [10].

Non-visual perception acquires surrounding information through sensors, such as lidar, millimeter-wave radar, etc., providing real-time environmental data and reducing data requirements and hardware requirements. Compared with visual perception, using non-visual perception information as the input of reinforcement learning can reduce data requirements, computing resources and hardware requirements. Meanwhile, the sensors of the autonomous driving system provide a large amount of data for the system. The system can conduct real-time analysis and processing based on the data. Data-driven approaches are beneficial for the system to improve its intelligence and flexibility.

Secondly, the decision-making and control module of autonomous driving is also a key component responsible for formulating driving strategies and controlling the vehicle to perform actions. It mainly includes path planning, behavior decision-making, vehicle control, motion planning and interaction decision-making.

Path planning refers to the image recognition technology that identifies and classify various conditions such as the position, speed, and trajectory of the car through technologies like deep learning, and plans a driving path for the car to avoid all obstacles, comply with traffic rules, minimize the driving distance and other factors. Through image recognition technology, behavior decision-making identifies elements such as surrounding vehicles, pedestrians, traffic signals, etc., combined with the results of path planning and environmental factors, and formulates safe and reasonable driving strategies. Image recognition technology provides crucial environmental perception and information acquisition, helping the autonomous driving system make accurate decisions and ensuring the safe and reliable driving of vehicles. Vehicle control refers to converting behavior decisions into specific vehicle control instructions to control the vehicle's movement actions, such as acceleration, deceleration, braking and other vehicle instructions. Motion planning, is the core part of vehicle control. It converts path planning and behavior decision-making into vehicle motion trajectories and needs to consider factors such as vehicle dynamics constraints, dynamic changes in the environment, and obstacle avoidance to ensure the vehicle travels safely and smoothly.

Interaction decision-making refers to interacting with other participants such as vehicles and pedestrians, and formulating corresponding coping strategies to ensure traffic safety and smoothness.

These modules cooperate to form the core of the autonomous driving system, enabling the vehicle to perceive the environment, make decisions and drive safely intelligently. The design and optimization of these modules are crucial for achieving efficient and safe autonomous driving (Figure 2).

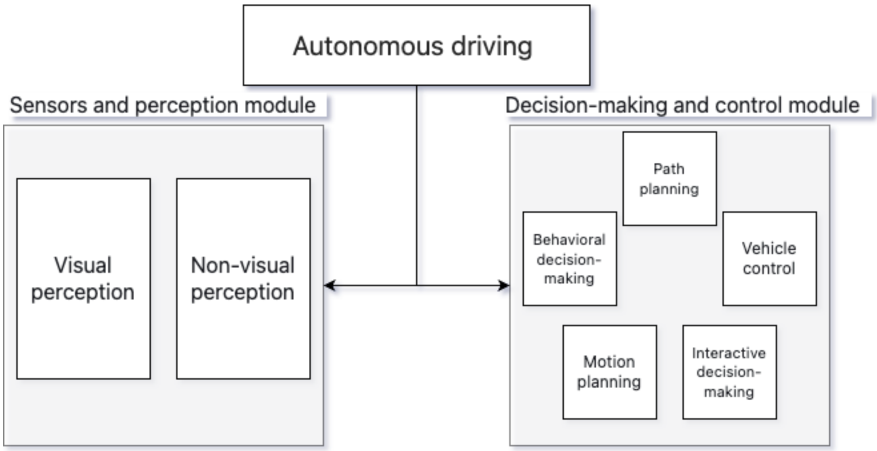

Figure 2. Schematic diagram of autonomous driving

3.2. Medical Image Analysis:

Medical image analysis utilizes medical imaging techniques to obtain data for disease diagnosis, treatment planning, and efficacy evaluation. Technologies such as X-rays, CT scans, MRI, and ultrasound generate images, and image recognition technology reveals the human body's internal structure and functional information. This article will explore feature extraction, image segmentation, image classification and recognition in medical image analysis.

Before processing medical image data, preprocessing is often required, such as denoising, enhancement, and geometric correction, to improve the image quality and accuracy. Feature extraction is crucial in medical image analysis. By capturing features such as shape, texture, and density, it describes the attributes of the lesion area. The research by Miao Yang, Zhang Shuo et al. utilized deep learning to extract features of highly malignant tumors such as hypopharyngeal cancer and HPC lesions from MRI images, providing diagnosis and treatment support for doctors and accelerating the location of lesion areas [11]. These feature information provide key clues, contributing to accurately describing disease states and tissue structures.

After the feature extraction is completed, the next key technology is image segmentation. Image segmentation can accurately separate the lesion area from the image. Extracting features helps doctors locate the lesion, quantify the features, and assist in diagnosis and treatment planning. The segmentation methods include threshold segmentation, edge detection, region growing, etc. Taking HPC lesions as an example, this is a rare and highly malignant tumor, accounting for only 6% of all head and neck cancers. Since there are many lesions in the pharynx and they vary in size, achieving accurate image segmentation is of crucial importance. Through precise image segmentation technology, different-sized lesions can be effectively located and distinguished [11]. However, HPC lesions have multiple receptive field attributes, and it is difficult to sample reasonably with a single receptive field. In the experiment, target detection and semantic segmentation were used to handle image segmentation, but huge challenges were still faced.

The classification and recognition of medical images using machine learning, deep learning and artificial intelligence technologies usually require training classifiers or neural network models. CNN is a commonly used model that extracts image features and performs classification through multiple convolutional layers and pooling layers. The convolutional layer extracts features by sliding the convolution kernel on the input image, while the pooling layer reduces the size of the feature map, lowers the computational complexity, and enhances the robustness of the network to changes. The fully connected layer connects the extracted features into a neural network. The model is trained through a large number of data sets to classify and identify unknown medical image data accurately.

3.3. Face Recognition

Face recognition is a biometric identification technology aiming to confirm an individual's identity by analyzing the facial image's unique features. This process usually includes three stages: face detection, face alignment, and face recognition, completing the entire process from detection to final recognition.

Face detection is the first step of face recognition, and its task is to locate and mark the face area in the image accurately. By applying different algorithms and models, such as CNN and Haar Cascade Classifier, the system can detect and locate the face areas in the image, usually represented in rectangular bounding boxes. Face detection algorithms based on deep learning are mainly divided into two types: One-stage face detection algorithms and two-stage face detection algorithms [12]. Among the Two-stage face detection algorithms, the one with better performance is the face region-based fully convolutional network (Face R-FCN) proposed by Wang et al. [13], which is suitable for small-scale face detection. Meanwhile, the one-stage face detection algorithms include You Only Look Once (YOLO) [14].

Next comes the face alignment stage. Face alignment is to precisely position the face image to ensure that all face images maintain a consistent posture and position during the feature extraction and recognition stages. This step usually includes standardizing the detected faces, such as rotation, scaling and cropping, to make the face images meet the input requirements of the model. Face alignment helps to reduce the variation factors in face recognition, improve the accuracy of subsequent feature extraction and matching, and thus significantly enhance the efficiency and fault tolerance ability of face recognition.

Finally, it comes to the face recognition stage. Face recognition is used to identify or verify an individual's identity by extracting the features of the face image and comparing them with the face features stored in the database. In this process, feature extraction techniques, such as convolutional neural networks in deep learning, are usually utilized to train the model to extract distinguishable features from face images. Subsequently, various matching algorithms (such as Euclidean distance, cosine similarity, etc.) are used to compare the extracted features with the features stored in the database to achieve face recognition or authentication functions.

4. The Advantages, Disadvantages and Prospects of Image Recognition

As one of the important applications of computer vision, image recognition has many advantages. Firstly, image recognition can achieve automatic processing. It can handle a large amount of image data without human intervention, thereby improving efficiency and accuracy. Secondly, image recognition can handle large-scale image data, which is helpful for quickly extracting effective information and processing and recognizing in real time. Finally, image recognition is widely applied in multiple fields, including healthcare, security, retail, agriculture, autonomous driving and other fields, which helps improve work efficiency and the quality of life. With the continuous progress of technologies such as deep learning, the accuracy and performance of image recognition have been continuously improved, making its application in various fields more extensive and reliable.

Although image recognition technology has many advantages, it also faces some disadvantages and challenges. Firstly, image recognition has high requirements for data quality. Blurry or distorted images will affect the accuracy. And in complex scenarios, such as uneven lighting and occlusion, the algorithm's accuracy may be affected. Moreover, constructing high-quality labeled datasets requires a significant investment of human resources. Secondly, image recognition technology may involve user privacy and data security issues. Unauthorized image collection and usage may lead to privacy leaks. To sum up, image recognition has high requirements for the quality of datasets.

With the continuous advancement of deep learning and neural network technologies, the accuracy and performance of image recognition technology will continue to improve, and it is expected to solve the challenges of image data in complex environments. In the future, image recognition technology may integrate with other perception technologies (such as speech recognition, natural language processing, etc.) to comprehensively analyze and understand multimodal information. In terms of real-time processing capabilities, image recognition technology will pay more attention to real-time performance and be able to analyze and recognize image data in a shorter time. Overall, image recognition technology will continue to grow and develop in the future, bringing more innovation and application possibilities to various fields and helping human society move towards an intelligent and digital future.

5. Conclusion

This article deeply explores the applications of image recognition technology in the fields of autonomous driving, medical imaging and face recognition. Image recognition technology realizes the understanding and analysis of image content through deep learning and machine learning algorithms, promoting technological progress in these fields. In autonomous driving, image recognition technology realizes the perception and decision-making modules, improving driving safety and reliability. In medical image analysis, image recognition technology is utilized to help doctors conduct disease diagnosis and treatment planning through feature extraction, image segmentation and classification recognition. While the face recognition technology accurately confirms the identity of an individual through three steps: face detection, face alignment and face recognition. Deep learning and feature matching algorithms play key roles in extracting and comparing facial features, significantly improving the accuracy and efficiency of face recognition. Overall, through applying deep learning and machine learning algorithms, image recognition technology has achieved an in-depth understanding and analysis of image content, demonstrating strong application potential and practical value in fields such as autonomous driving, medical imaging, and face recognition.

Through in-depth analysis of these fields, this article demonstrates the advantages and prospects of image recognition technology and provides references for future technological development and application promotion. However, image recognition technology also faces data acquisition, volume, and quality challenges. Transfer learning and enhancement techniques can be utilized to achieve effective image recognition to cope with difficulties in cases of data scarcity or domain transfer. In the future, with the continuous advancement of technology and the continuous expansion of applications, image recognition technology is bound to play an essential role in more fields and bring more innovations and changes.

Reference

[1] J. Wang, H. Huang, et al. Review of the Development and Key Technologies of Autonomous Driving, Application of Electronic Technique, 45(6), 28-36, (2019).

[2] C. Zhou, The Application of Artificial Intelligence Technology in Autonomous Driving of Vehicles, Automobile Test Report, 21(5), 16-18, (2024).

[3] J. Zhang, Exploring the clinical value of nuclear medical imaging in the diagnosis of multiple myeloma. Chinese Journal of Modern Drug Application, 14(04), 63-65, (2020).

[4] G. Haskins, U. Kruger, P. Yan. Deep learning in medical image registration: A survey. Machine Vision and Applications, 31, 8, (2020).

[5] G. Hinton, S. Osindero, Y. Teh. A Fast Learning Algorithm for Deep Belief Nets. Neural Computation, 18(7), 1527-1554, (2006).

[6] C. Zhang, M. Zheng, et al. Recurrent Neural Network Model Based on Weight Activity Evaluation, Computer Applications and Software, 41(6), 296-304+319, (2024).

[7] D. H. Hubel, T. N. Wiesel. Receptive Fields Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. Journal of Physiology, 160(1), 106-154, (1962).

[8] M. K. Alam, A. Ahmed, R. Salih, et al. Faster RCNN based robust vehicle detection algorithm for identifying and classifying vehicles. Journal of Real-Time Image Processing, 20(5), 93-103, (2023).

[9] H. Song, L. Zhou, Research on YOLOv3 vehicle detection and recognition algorithm based on function improvement, Journal of Intelligent Science and Technology, 5(4), 535-542, (2023).

[10] Z. Ding, Y. Hu, W. Jian, et al. Adaptive scene road surface extraction method based on deep learning. Journal of Beijing Institute of Technology (Natural Science Edition), 39(11), 1133-1137, (2019).

[11] M. Yang, S. Zhang, et al. Review of Hypopharyngeal Cancer Medical Image Analysis Based on Convolutional Neural Network, Journal of Beijing University of Technology, 51(7), 883-896, (2024).

[12] M. Xu, H. Li, et al. Face5 Series Face Detection Algorithm Based on Improved YOLOv5s-face, Journal of Chongqing University of Technology (Natural Science), 38(6), 194-202, (2024).

[13] Y. Wang, X. Ji, Z. Zhou, et al. Detecting faces using region-based fully convolutional networks [EB/OL], https://arxiv.org/abs/1709.05256, (2023).

[14] J. Redmon, S. Divvala, R. Girshick, et al. You only look once: unified real-time object detection [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 779-788, (2016).