

The Art of Seeing: A Computer Vision Journey into Object Detection

Mohammad Salman Khan* and Ayesha Imran

Konya Technical University: Konya, Turkey

*Corresponding Author

Mohammad Salman Khan, Konya Technical University: Konya, Turkey.

Submitted: 2024, May 16; Accepted: 2024, Jun 20; Published: 2024, Jun 25

Citation: Khan, M. S., Imran, A. (2024). The Art of Seeing: A Computer Vision Journey into Object Detection. *Adv Mach Lear Art Inte*, 5(2), 01-07.

Abstract

Object detection is a basic task in computer vision with numerous applications ranging from surveillance and autonomous driving to medical imaging and augmented reality. Recently, Machine and Deep learning approaches have significantly advanced the State-of-the-Art in object detection, enabling remarkable progress in accuracy, robustness, and efficiency. This paper presents a detailed review of recent researches and developments in Computer Vision, Object Detection and Sensing Techniques. We discuss key concepts, methodologies, and challenges in object detection, focusing on deep learning-based approaches. Additionally, we explore emerging trends such as instance segmentation, few-shot learning, and privacy-preserving techniques in object detection. Furthermore, we discuss benchmark datasets, evaluation metrics, and open research challenges in the field. Keeping in view the current researches and Research Techniques, this research claims to guide researchers and enthusiasts towards understanding the latest advancements and future directions in this exciting area of computer vision. We discuss key topics such as image classification, object detection, image segmentation, and scene understanding. The rapid progress in deep learning has revolutionized computer vision, enabling models to learn hierarchical representations directly from data. We review prominent deep learning architectures such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and generative adversarial networks (GANs), and their applications in various computer vision tasks. Furthermore, we explore recent developments in multi-modal and cross-modal learning, domain adaptation, and interpretability in computer vision models. Additionally, we discuss challenges such as data bias, ethical considerations, and scalability issues faced by the field. By providing a comprehensive overview, this paper aims to inspire further research and innovation in computer vision, advancing its capabilities and broadening its impact on society.

Keywords: Deep Learning, GANs, CNNs, RNNs, Vision-Testing, Object-Detection, Few-Shot Learning, Privacy Preserving

1. Introduction

In recent years, the field of artificial intelligence (AI) has witnessed remarkable advancements, particularly in the domain of computer vision. Object detection and image identification, two fundamental tasks within computer vision, have garnered significant attention due to their wide-ranging applications in various sectors including healthcare, transportation, security, and manufacturing. The ability of AI systems to accurately detect objects within images and identify their contents has revolutionized industries and paved the way for innovative solutions to complex problems. Object detection involves locating and classifying objects within images, while image identification focuses on recognizing the content of entire images or specific regions within them. These activities are crucial for empowering AI systems to apprehend visual data and make up-to-date choices in real-life situations. The development of robust algorithms and architectures for object detection and image identification has been a primary focus of research in both academia and industry. Convolutional neural networks (CNNs)

have become essential in ultra-modern computer vision systems because they can independently acquire hierarchical features from unprocessed pixel information. Over the years, CNN-based approaches have continuously been optimized in object detection and image identification tasks, achieving unprecedented levels of accuracy and efficiency. Despite the significant progress made in the field, challenges persist in achieving robust and reliable object detection and image identification in complex and dynamic environments. Factors such as occlusions, variations in lighting conditions, and the presence of clutter poses significant challenges to existing AI systems. Moreover, the need for scalable and interpretable solutions remains a pressing concern, particularly in safety critical applications such as individualistic driving and medical diagnosis. In this paper, we present a detailed review of the recent advancements in AI techniques for object detection and image identification. We delve into the underlying methodologies, including CNN architectures, and discuss the evolution of key algorithms and techniques. Furthermore, we explore the challenges

and opportunities in vision testing methodologies, emphasizing the importance of accurate evaluation metrics and benchmark datasets. Through a critical analysis of existing approaches and emerging trends, we aim to provide researchers and practitioners with valuable insights into the current state-of-the-art and future directions in this rapidly evolving field.

2. Proposed Method

2.1 Machine Learning (ML)

Machine Learning (ML) is a diverse field that draws from various disciplines like probability theory, statistics, and algorithm complexity theory. Its essence lies in enabling computers to mimic human learning processes, acquiring new knowledge and refining existing ones to enhance performance. ML is integral to artificial intelligence, offering a fundamental avenue to imbue computers with intelligence across various domains. In today's data-driven era, where "data is king", companies scramble to a vast amount of user data, hoping to decipher user preferences and behaviors. Yet, simply having data isn't enough. The vast amount of data surpasses the direct computational capacity, requiring specialized learning algorithms to efficiently extract valuable insights. Machine learning serves this purpose, facilitating the extraction of meaningful information from unstructured data. The development of machine learning can be mapped out across three distinct stages: connectionism in the 1980s, statistical learning methods in the 1990s, and the rise of deep neural networks in the 21st century, fueled by advancements in data availability and computing power. Deep learning, particularly has revolutionized AI applications, enabling significant progress in various fields. Machine learning algorithms endeavor to discern implicit rules from historical data for prediction or classification. They operate by seeking complex functions that map input data in to desired outcomes. Importantly, the goal is not merely to perform well on training data but to generalize effectively to new samples, reflecting the model's true efficacy. Machine learning is a versatile field that tackles uncovering patterns in data. It combines techniques from statistics, data mining, and collaborates with other disciplines like computer vision, speech recognition, and natural language processing to unlock the secrets hidden within information. Its applications span diverse domains, from speech recognition systems like Apple's Siri to computer vision technologies like image recognition and object detection in Baidu Maps. The combination of machine learning with other processing techniques has spawned interdisciplinary fields with promising applications. Deep learning's advancements have significantly enhanced computer image recognition, paving the way for unprecedented growth in the computer vision industry.

2.2 Artificial Intelligence (AI)

In the realm of computer vision, researchers have proposed various models for image recognition, with template matching being one of the most widely used algorithms. This method determines the consistency of target image features in a picture database by comparing them with predicted images. By combining the principles of Image Recognition Technology in artificial intelligence with computational data processing, simple extraction and inspection of image data can be achieved. However, when dealing with fuzzy

image data or large volumes of information, recognition efficiency may decrease. Therefore, there's a continuous quest for improved and more efficient image recognition technologies that simplify principles while enhancing functionality and image processing. The principle of Image Recognition Technology in artificial intelligence involves utilizing computers to process images and extract data from them. Through the analysis and experimentation conducted by experts, the technical principle of image processing technology has been elucidated. Essentially, this principle mirrors human visual perception, where acquired information is analyzed in the brain based on impressions. Despite its simplicity, when faced with large amounts of data, the recognition rate of image recognition technology may decline. Hence, researchers are exploring optimized methods to enhance the quality and efficiency of image processing. Artificial intelligence-based image recognition technology offers the advantages of convenience and intelligence, which directly impact its application efficiency in scientific and technological advancements. The intelligence aspect enables smart selection and recognition features, akin to facial unlocking functions in mobile devices. Additionally, it simplifies tasks like facial recognition and authentication, bringing convenience to users' lives. With the growing popularity and convenience of image recognition technology, its adoption has become widespread in various aspects of society, facilitating ease of use and accessibility. Because image recognition technology is based on artificial intelligence, the process closely resembles human brain image recognition processes. However, the implementation is manifested through technological means. The specific process involves several stages: acquiring data, preprocessing it to highlight essential features, extracting and selecting features crucial for recognition, and designing classifiers and making classification decisions to improve processing efficiency.

2.3 Image Preprocessing

In the process of image accession, it's constantly subject to various foreign conditions and arbitrary interference. Directly acquired analogous images constantly contain daedal unworkable grounds or redundant data, which interferes with the further operation of images. Therefore, some necessary pre-processing ways need to be performed on the initial data image. Typically, image processing tasks involve various techniques such as converting color images to grayscale, enhancing image quality, restoring degraded images, segmenting images into different regions, and applying smoothing or sharpening filters, among others. These techniques are commonly employed to manipulate and improve images for various purposes, ranging from enhancing visual quality to extracting useful information for analysis. To streamline computer processing, minimize resource consumption, and accelerate operations, color images are typically converted to grayscale before undergoing digital image processing. In grayscale images, each pixel's intensity is represented by a single value ranging from 0 to 255, where 0 represents the darkest shade (black) and 255 represents the lightest shade (white). This simplification reduces the computational load by eliminating the need to process separate color channels, thereby enhancing the speed and efficiency of image processing tasks. Additionally, the grayscale representation

preserves essential visual information for many applications while reducing the complexity of the data. At present-day, the ripest technology operation is the RGB color mode. The digital image described by the RGB mode has three image procurators, and the RGB valuations of three pixels of each pixel singly reflect the brilliance valuations of the three colors at the pixel. The factual color described by the pixel is the result of the color superposition of three nonidentical splendor. Since there are 256 stripes of valuations for each color, there are more than 16 million ($256 \times 256 \times 256$) color variations per pixel. Still, after transformation to a grayscale image, there are only 256 variations of each pixel, so the amount of computation of the computer can be greatly downgraded. The converted grayscale image, like the definition of the initial color image, still contains the correlation characteristics of the initial image's value and brilliance. The objective of improving the appearance of an image through fashioning is to enhance its visual impact, making it better suited for specific purposes. This enhancement involves deliberately emphasizing certain features of the image, thereby accentuating differences between distinct images to cater to particular situations or conditions. By refining the visual presentation, such as adjusting contrast, brightness, or sharpness, the image becomes more effective in conveying intended messages or serving specific needs. This process ensures that the image is optimized for its intended use, whether it's for aesthetic appeal, communication, analysis, or any other specialized application. In a broad sense, as long as the structural relationship between the corridor of the initial image is changed, the purpose is to outdo the operation sequel and the judgment result to meet the special operation conditions. This processing fashion can be called image enhancement processing technology. The image enhancement technology can be roughly codified into two orders, a spatial sphere system, and a frequency sphere system, tallying to non-identical positions of objects reused by the enhancement fashion. The spatial domain-based algorithm involves directly reusing the gray value of the initial pixel when predicting the value of a pixel in the image's plane. The frequency sphere system is to enhance the image on another transfigure sphere of the image. Histogram equalization is a processing system that enhances the operation of digital images predicated on liability proposition. The histogram, also known as the mass division map, is a statistical graph predicated on the report. The histogram of a digital image is a division of the grand number of pixels of nonidentical gray valuations in an image. By examining the histogram of an image, we can understand the brightness levels within the grayscale range of its pixels. In an overly dark image, the histogram's grayscale values are concentrated towards the lower end of the scale, indicating predominantly darker tones. Conversely, in an excessively bright image, the histogram's distribution shifts towards higher grayscale values, suggesting an abundance of lighter tones. Histogram equalization is a technique used to adjust the distribution of pixel intensities in an image, enhancing its contrast and improving visual quality. This process involves transforming the histogram of the original image through a gradient transformation and adjusting the distribution according to a specific rule. The goal is to achieve a more uniform distribution of grayscale values across the image, resulting in a new histogram image with a balanced

and stable division of grayscale values. This equalization helps to reveal more details in both dark and bright areas of the image, making it visually more appealing and easier to interpret. Tallying to the proposition of information proposition, when the division of gray valuations of an image is relatively moderate, the amount of information contained in the image is also voluminous, and the image has a clearer sequel from the visual point of the mortal eye. Median filtering is a nonlinear signal processing technique used in image processing to reduce noise and blur by replacing each pixel's value with the median value of its surrounding pixels. It effectively suppresses noise, including impulse noise, while preserving image details and sharp edges. The median filter is a method used in image processing to reduce noise, particularly effective for types of noise like salt-and-pepper noise. This noise appears as isolated pixels with significantly different values from surrounding pixels. The median filter works by replacing each noisy pixel with the median value of its neighboring pixels, effectively smoothing out isolated noise points. One advantage of the median filter is its effectiveness in filtering out both light and dark noise, making it useful for a variety of noisy image types. However, it may not perform well when there are numerous intricate details in the image, such as points, lines, or edges, as it can blur or distort these features. To address limitations, enhanced algorithms have been developed. These include adaptive median filtering, which adjusts the filter size based on the local characteristics of the image, and switching median filtering, which combines median filtering with other filtering methods based on a predefined threshold. These advanced techniques offer improved noise reduction while preserving important image details in complex scenes.

2.4 Image Recognition

In a broad sense, image technology is a general tenure for colorful image-related technologies. Tallying to the exploration system and the place of abstraction, the imaging technology can be separated into three situations, which are separated into image processing, image dissection, and image understanding. The technology intersects with computer unreality, pattern recognition computer plates, biology, mathematics, drugs, electronics, computer wisdom, and other firmaments get from each other. In extension, with the evolution of computer technology, further exploration of image technology is thick from propositions of neural networks and artificial intelligence. Most of the image processing includes image enhancement, encoding, restoration, and compression. The purpose of recycling the image is to determine whether the image has the required information and sludge out the bruit, and to determine the information. Common or garden styles carry grayscale, binarization, stropping, denoising, etc.; image recognition is to match the reused image to determine the order name. Image recognition can be uprooted on the base of segmentation. The features are screened, and also these features are uprooted and eventually linked tallying to the dimension effects. Image understanding involves the interpretation and definition of an image by analyzing its components and structure through image processing and recognition techniques. Image processing enhances image quality and extracts important information, while image recognition focuses on identifying and classifying objects within

the image. This process includes segmentation to divide the image into meaningful regions, feature extraction to identify significant attributes, and classification to label these segments. By combining these steps, image understanding allows for a comprehensive analysis of the image, enabling applications such as automated surveillance, medical image analysis, and autonomous driving. Thus, image understanding includes image processing, image recognition, and structural dissection. In the image understanding section, the input is an image and the processing stage is the definition of the image. The evolution of image recognition has been marked by a progression from simple rule-based approaches to complex, data-driven techniques enabled by advances in computing power, algorithms, and data availability. Each stage has built upon the successes and limitations of the previous ones, driving the field towards a more robust and versatile image recognition system. Generally, when a sphere has a demand that cannot be answered by the essential technology, the corresponding new technology will be produced. Same is the case with image recognition technology. The invention of this technology is to allow the computer rather than a mortal to process a voluminous number of physical information and break the case of information that cannot be honored or the recognition rate is veritably low. Computer image recognition technology imitates how humans recognize and understand images. It uses algorithms and machine learning to analyze pictures and identify objects or patterns, like faces or animals. By learning from large amounts of data, these systems can accurately recognize and interpret visual information, similar to the way human vision works. In the process of image recognition, pattern recognition is essential. Pattern recognition is an introductory mortal intelligence. Still, with the evolution of computers and the ascent of artificial intelligence, mortal pattern recognition has been unsuitable to meet the requirements of life. Consequently, mortal commodities hope to replace or expand the portion of mortal brain labor with computers. In this expressway, the pattern recognition of the computer is created. Pattern recognition is the process of classifying data based on patterns found within it. It involves identifying and grouping similar items, like recognizing shapes or objects in images, and categorizing them accordingly. For example, it helps in identifying a cat or a tree in a photo by matching the patterns seen in the image. It's a wisdom that's nearly integrated with mathematics. Most of the ideas exercised are liability and statistics. Pattern recognition is substantially separated into three manners: statistical pattern recognition, syntax pattern recognition, and shaggy pattern recognition. Since computer image recognition technology is the same as mortal image recognition, their processes are analogous. Image recognition technology is also separated into the following ways: information accession, preprocessing, point birth and election, classifier project, and bracket resolution. The accession of information refers to the transformation of information similar to light or sound into electrical information through detectors. That's to gain the introductory information of the exploration object and transfigure it into information that the engine can fetch by some means. Preprocessing substantially refers to missions similar to smoothing and altering in image processing, thereby enhancing the important features of the image. The simple understanding is

that the images we study are various. However, we must identify them by the characteristics of these images, if we need to discern them by some system. The process of acquiring these features is point birth. In summary, while features obtained in the early stages of image recognition were valuable for their time, they may not be well-suited for addressing the challenges and requirements of contemporary recognition tasks. The evolution of image recognition has necessitated the development of more advanced feature extraction techniques that capture the complexity, context, and semantics of modern images. Point birth and election is one of the most overcritical ways in the image recognition process, so the understanding of this step is the seat of image recognition. Advancements in image recognition enable the detection and tracking of moving objects in real-time, driven by improved algorithms and deep learning techniques like CNNs and RNNs. This technology is applied in surveillance, autonomous vehicles, and sports analytics, though challenges persist in robustness and ethical considerations. Its main principle is to reuse and make opinions on blurred image information through exceptional modules, and also gain effects with high community, and also confirm image information through webbing. Prescriptive image recognition model LeNet is an earlier CNN model (1994). It has three complication layers (C1, C3, C5), two pooling layers (S2, S4) and one full connection subcaste (F6). The input image is 32 x 32, and the affair image is the liability of 0 to 90 integers. At that time, the inaccuracy rate of the network model was lower than 1. LeNet was arguably the first commercially precious CNN model since it was successfully exercised to identify correspondence canons. AlexNet is a corner in the history of CNN's evolution. Assimilated with the LeNet network, the AlexNet network isn't much bettered in structure but has great advantages in network depth and complication. AlexNet has the following meanings, (1) It reveals the important literacy and suggestive capability of CNN's exploration. (2) GPU is exercised for computation, which shortens the time and cost of training. (3) Training ways similar to the ReLU activation function, data addition, and arbitrary inactivation were acquainted to give slices for posterior CNN.

3. Experiments

Selecting the car license plate detection as our primary experiment, certain tests were done. The software environment used is as follows: Windows 10, VC++, MATLAB. The hardware test environment is as follows: CPU: Intel Core i5 7th Generation, memory 16 GB. Items to be identified (3.2.2) H, U, V, W. (2) Manually taking a screenshot of the imported license plate. Due to the unique nature of license symbols, the number of non-numeric symbols is small and it is impossible to find many test patterns that track real probability. Therefore, some characters are manually added to the training model for training. All test samples in the plate image are divided by 100 from the symbol. In the process of optimizing neural network weights and thresholds, the genetic algorithm first encodes all the values thresholds, then synthesizes a long sequence in a specific order. Once the network model is determined in this way, each chromosome sequence defines all the parameters of the neural network, including thresholds and weights. To calculate each person's security, we randomize (decide)

the individual's strings. The network calculates the input training model and returns the sum of squared error (fit) based on the output model. In genetic algorithms, neural networks play the role of computational work. Through the iterative process, the structure of the neural network is fixed, including the number of hidden layers, the number of nodes, and the connection process between nodes. This paper selects: (1) neural network based on genetic selection algorithm, (2) BP neural network, (3) KNN algorithm for benchmarking. First, according to the following formula:

$$h = r + \sqrt{i + o} \quad (1)$$

$$h = r + \sqrt{i \times (o + I) + I} \quad (2)$$

Equations 1 and 2 are used to calculate the quantity of hidden layers, wherein **R** is a random range between 1 and 10, **I** is the size of enter layer and **O** is the dimension of output layer. Then, the chromosome generated through genetic set of rules is decoded, and the health of the sample inside the check set is tested immediately. The adaptive genetic set of rules can hold the populace range and ensure the convergence of the genetic set of rules. However, the template-primarily based picture matching set of rules calls for that the image to be expected and the template have to be completely consistent before it can be diagnosed efficaciously. In practical situations, identical images can appear differently due to various factors like lighting, distortion, or selection. This complexity requires machine learning systems to not only recognize images that perfectly match predefined templates but also identify images that deviate slightly due to different conditions. In simpler terms, image recognition systems need to be flexible enough to recognize objects even when they look a bit different in different environments. While the group fitness is big, we must select a smaller pass possibility. When the fitness is small, a huge pass opportunity is used to hurry up the search. For example, system 3, health is a vital feature (i.e. speculation) to assess the benefits and drawbacks of chromosomes. The health function constructed in this paper is based totally at the minimum score where elegance mistakes are general. The algorithm that goes with the flow is as follows: Calculate the health of the given speculation, Input: H chromosome. Output: fitness cost D. The fitness calculation as shown in the system is as follows.

$$P_c = \left\{ P_{c1} - \frac{(P_{c1} - P_{c2})(f_m - \bar{f})}{P_{c1}(f_m - \bar{f})} \right. \quad (3)$$

$$f = \frac{I}{I + \delta^2} \quad (4)$$

After starting the system with a baseline chromosome, we assess each hypothesis in the test set based on its accuracy. Then, we rank them based on their strengths and weaknesses. As we continue, we select the best candidates based on their performance and discard the weaker ones during further iterations.

4. Discussion

The results of the tests show that the optimal individual has been

optimized by the genetic algorithm of 5000 generations. When it evolved to the 3000th generation, it basically converges. In Figure1 the fitness curve is shown.

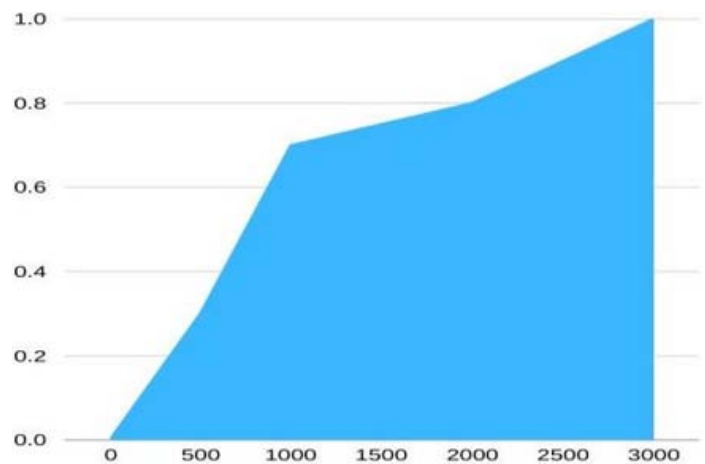


Figure 1: Fitness Curve (Chart)

The BP's parameters of the neural network are: the learning rate is 0.08 and the impulse is 0.1, which is directly trained using the training sample set. The accuracy of the training set finally obtained after training is 0.9756. KNN does not require training and is an inert classification. Under the assumption that the fitness is **Fitness (h1) = 0.9543**, **Fitness (h2) = 0.9456**, and **Fitness (h3) = 0.9325**, the results of the classification are shown in Table 1.

Classification	Noiseless Test	Noisy Test	Training Time	Recognition Time
h 1	92.41%	86.90%	7H20M	11
h 2	94.2%	84.45%	7H20M	10
h 3	93.01%	72.56%	7H20M	11
BP	91.23%	80.96%	10M	11
KNN	91.67%	87.54%	Null Value	40

Table 1: Collected Data

Among the three suppositions named H1, which performs outstanding in training set, doesn't achieve as required in the test set. The effects show off that inheritable algorithm can't fully break the over befitting case of neural network. In the bracket trial, the common delicacy is high. Character neutralization, distortion, disposal, and blur in an image cause textbook features to lose their spatial organization. In the case of bruit hindrance, it's delicate to directly codify with a many simple features. The license plate recognition effects show off some differences with nonidentical stages of adaption 2, 3 and 4. It can be discerned from the recognition affect that the recognition delicacy of Fig. 2 isn't high, and the prey area cannot be directly obtained. The recognition result in Fig. 3 is good, and all the prey areas in the image can be obtained, and the prey area in Fig. 4 identifies the disfigurement.

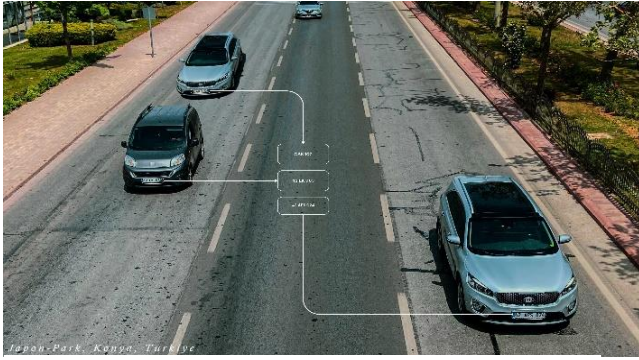


Figure 2: H1 Recognition Results Under Fitness

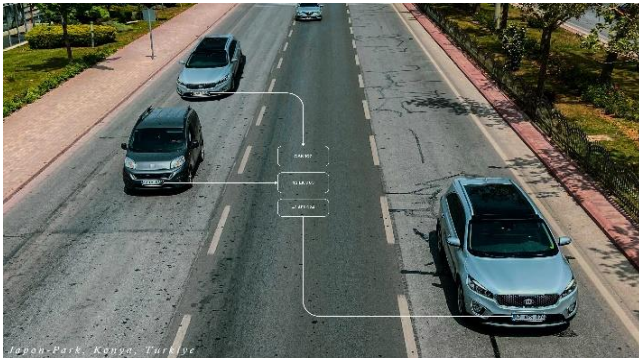


Figure 3: H2 Recognition Results Under Fitness

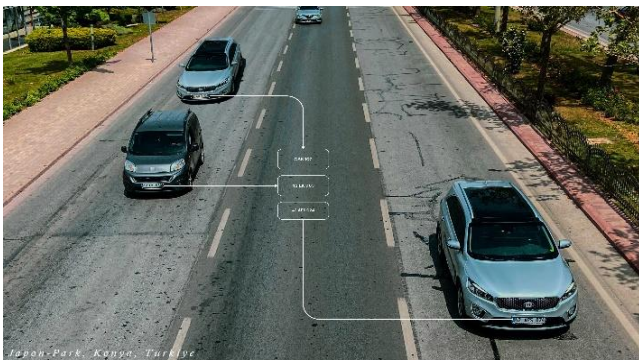


Figure 4: H3 Recognition Results Under Fitness

The common accurate rate in the complete bracket trial, when the fitness in this trial is $\text{Fitness}(h2) = 0.9456$, the interpretation of the recognition image is better. From the compliance of recognition rate, we can detect that the recognition rate of neural network grounded on inheritable algorithm training is better than that grounded on traditional system training. In extension, the traditional KNN bracket system also shows good effects. Although the recognition time increases with the boost of the training slices' number, it has good simplicity and interference. In non-real-time dissection terrain, the cross-use of multitudinous styles can be considered to farther ameliorate the recognition rate.

5. Conclusion

As an important system in the field of artificial intelligence, machine literacy has been extensively used in business identification

exploration in recent times. Because of its intelligence, good conception and high recognition effectiveness, it has gradually come the mainstream of image recognition exploration. This paper studies the operation of image recognition technology grounded on machine literacy in license plate recognition. In order to complete the exploration of this paper, a lot of exploration on the current development of license plate recognition exploration is carried out, and the vertical and perpendicular exploration and exploration are carried out in the field of recognition. Some introductory technologies of license plate recognition are studied, similar as image processing, pattern bracket, machine literacy, artificial intelligence and so on. In order to complete this trial, a large quantum of target data was collected, but in the field of target recognition, it's veritably delicate to gain large-scale effective data. This is also the primary problem that hinders the operation of deep literacy in the field of image recognition. To this end, it's necessary to find a more effective way to carry out homemade data expansion grounded on the original database, so that deep literacy can be effectively applied. Data in life is ubiquitous, but tagged data isn't common. also, it's easier to collect data in the field of image recognition, but manually collecting the collected data is a time-consuming and labor-ferocious task. To this end, unsupervised literacy algorithms are also the focus of exploration in deep literacy, similar as generating combative network models. In the correction process of the license plate, this paper substantially focuses on the direct information handed by the framed license plate. However, also a targeted algorithm should be developed, If the license plate position module provides a license plate without a frame. At the same time, in view of the control of the conception delicacy of the classifier in the license plate character recognition, this paper combines the inheritable algorithm with the optimal result hunt tool which is better than the total system to break the global space of the weight of the neural network. After experimental verification, the three results with the loftiest fitness are attained from the inheritable algorithm. The conception effect after decrypting to the neural network is fairly good.

References

1. Zhu, B., Ye, S., Jiang, M., Wang, P., Wu, Z., Xie, R., ... & Wei, Y. M. (2019). Achieving the carbon intensity target of China: A least squares support vector machine with mixture kernel function approach. *Applied energy*, 233, 196-207.
2. Zhang, Y., Kwong, S., Wang, X., Yuan, H., Pan, Z., & Xu, L. (2015). Machine learning-based coding unit depth decisions for flexible complexity allocation in high efficiency video coding. *IEEE Transactions on Image Processing*, 24(7), 2225-2238.
3. Wang, B., Chen, L. L., & Zhang, Z. Y. (2019). A novel method on the edge detection of infrared image. *Optik*, 180, 610-614.
4. Yuan, Y., Lin, J., & Wang, Q. (2015). Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization. *IEEE transactions on cybernetics*, 46(12), 2966-2977.
5. Wang, X., Li, X., & Leung, V. C. (2015). Artificial intelligence-based techniques for emerging heterogeneous network: State

of the arts, opportunities, and challenges. *IEEE Access*, 3, 1379-1391.

Copyright: ©2024 Mohammad Salman Khan, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.