STUDY ON THE KEY RESEARCH ON THE FUNDAMENTAL TECHNOLOGIES OF PEDESTRIAN DETECTION APPLYING DEEP LEARNING.

Chen Huihong [1], Vivekanandam Balasubramaniam [1]

[1] Lincoln University College, Petaling Jaya, Malaysia.

*Corresponding author: Chen Huihong, Lincoln University College, Petaling Jaya, Malaysia.

ABSTRACT

Numerous forms of intelligent technology rely on visual perception for optimal functionality. Vehicles that drive themselves, robots, and surveillance systems are all included in this category. The field of deep learning has already made significant progress despite the fact that it is still relatively new. The performance of more recent models in computer vision, which make use of deep architectures such as CNNs, is superior to that of its predecessors. With the assistance of deep learning, pedestrian detection is made a great deal less difficult. An examination of the current state of the art in research and accessible methods is presented in this article with the purpose of assisting in the identification of pedestrians. This article takes a look at a few new models that are considered to be pioneers in their respective fields. Among the models that fall under this category are CNNs that are built on regions. Some examples of these CNNs are R-CNN, Fast R-CNN, and Faster R-CNN. Transformer-based versions and single-shot detectors (SSD, YOLO) are two other kinds of detectors. Methods for preparing data for processing, methods for adding new data, and methods for adjusting processed data to include non-maximum suppression are all covered in this research. Moreover, this study also includes methods for modifying processed data. During the presentation, each of these subjects will be discussed in further depth. It is possible to get significant insights into the operation of diverse systems by comparing the performance of these systems on various criteria such as recall, average miss rate, and accuracy. In conclusion, the purpose of this study was to investigate the current status of deep learning pedestrian detection systems in all of its exquisiteness.

Keywords: Technologies, pedestrian, learning, learning pedestrian, fundamental technologies.

INTRODUCTION

The identification of pedestrians is a rapidly expanding field of research in the field of computer vision, and there are various potential applications for deep learning in this area. Intelligent robotics, traffic safety, intelligent driving, security and monitoring systems, drone monitoring and rescue, and a great deal more are some of the technologies that fall under this category. In the course of the history of study and development pertaining to object detection, there have been two significant epochs. There has been a period of time during which deep learning-based object identification algorithms have been rapidly developing, and there has also been a period during which more traditional methods of object recognition have been in the process of being investigated. The following are the three primary stages that were involved in the implementation of this algorithm: From the very beginning, researchers use sliding windows and several other methods in order to conduct a comprehensive search of the various regions of the image. After that, the properties of the object region are acquired by using feature extraction techniques such as the Scale-Invariant Feature Transform (SIFT) and Histograms of Orientated Gradients (HOG). In the end, the retrieved object area characteristics are input into classification methods such as Adaboost and Support Vector Machine (SVM) in order to provide the results. As a result of this study's evaluation of current advancements and identification of continuing barriers in the development of trustworthy and intelligent pedestrian detection systems, it is likely that the consequences will include improved mobility, more effective surveillance, and smarter cities. Ongoing challenges include effective person recognition in partly occluded environments, pedestrian management in busy places, and correct identification in tough environmental situations (such as when the weather is poor or when it is midnight). Complex processes, such as generative data augmentation, transfer learning, multi-scale feature aggregation, and fascinating modules, have lately been the subject of a significant amount of attention from the technical community (Iftikhar et al., 2022).

BACKGROUND OF THE STUDY

Pedestrian detection has risen to prominence as a research issue in computer vision because to its usefulness in several applications such as intelligent transportation systems, human-computer interaction, surveillance security, and autonomous cars. This is due to the fact that both uses benefit greatly from pedestrian detection. As a result, it has grown into a strategically vital field of study. Preventing accidents, improving safety, and making automated decision-making easier all need the capacity to accurately identify and localise pedestrians in different environments. Some of the many reasons why accidents may be prevented include the following. Being able to execute it in different environments is also crucial. Classifiers such as Support Vector Machines (SVM) and AdaBoost and feature descriptors such as Haar-like features and Histogram of Orientated Gradients (HOG) were once used in pedestrian recognition systems. People walking on foot were located using these techniques. Various techniques were used to identify pedestrians. Because of difficulties like occlusions, cluttered backdrops, and changes in light, and size fluctuations, these technologies are not very practical in real-world scenarios. The intricacy makes it hard for these technologies to do what they set out to do. Regardless matter how useful these technologies have been in contained environments, the truth remains that this is still the case. The game of pedestrian identification has changed drastically as stable hierarchical representations can be built and automated characteristic extraction is now a reality. As a result of advancements in deep learning, and CNNs in particular, this capability is now within reach. A few models that have shown significant potential for real-time applications include YOLO, SSD, and Faster R-CNN. Several applications have made use of these models, including autonomous driving and urban surveillance (Mirama et al., 2021). This is because these models have become much better at what they do, both quickly and accurately. People who are slender, partially covered, or crowded need these models to perform harder. Critical technologies to resolve these difficulties have been the focus of extensive R&D efforts in recent years. In order to identify people of varying sizes, multi-scale feature fusion is crucial. Attention processes aid networks in zeroing in on critical visual information. The tiny form factor of embedded and mobile electronics allows them to function in real time. It is possible to enhance the performance and generalisability of models via data augmentation and transfer learning. Under the present conditions, the study represents a component of a continuing endeavour to enhance pedestrian identification. As a result of researcher's studies, researchers believe that future technology will be safer and more responsibly developed for end users. Researching these vital technologies and creating enhancements to detecting systems for use in the actual world will increase researcher's chances of achieving this objective (Zhu et al., 2022).

PURPOSE OF THE STUDY

The major objective of this investigation is to investigate and evaluate pedestrian detection systems that make use of deep learning. This inquiry is covered under the scope of this investigation. There are a broad variety of applications that are growing more dependent on pedestrian detection. Some examples of these applications include autonomous automobiles, intelligent surveillance systems, and improved driver support systems. Because of this, there is a growing need for algorithms that are able to efficiently and precisely locate pedestrians. This demand is expected to continue to grow. The purpose of this study is to investigate the history of deep learning techniques in relation to this topic and to identify the significant contributions that have influenced the approaches that are currently in development. To be more specific, what is the objective of working on this paper? For the aim of arriving at a conclusion on the benefits and drawbacks of the various deep learning models, datasets, and performance metrics that are presently being utilised in the industry, the objective of this study is to give a complete analysis in order to arrive at a conclusion. The secondary objective of this project is to give insight on various techniques to strengthen these technologies in order to overcome challenges such as occlusion, changeable walking postures, and dynamic surrounds of the outside world. This will be accomplished during the course of the research. The researchers in the academic world and the software developers who are interested in enhancing pedestrian detection systems via the use of deep learning would find this collection of papers to be of great value.

LITERATURE REVIEW

It has been a significant aspect of computer vision research for a considerable amount of time to locate humans walking. This is because locating humans is essential for technologically advanced surveillance systems, autonomous vehicles, and monitoring traffic flow in urban areas. A great deal of positive change has come to these locations as a result of it. Individuals fabricated the components of early detection systems by hand. Ultimately, these elements proved to be quite significant. These characteristics performed admirably in a controlled test that made use of classifiers such as AdaBoost and SVM. Perhaps the failure of these techniques might be attributed to the fact that they were not realistic. Various factors hinder their performance, such as their transparency, the potential for size fluctuations, and the need to adapt to inconsistent light conditions. Through deep learning, models are able to learn hierarchical and unique qualities from the data they are given. Increasingly precise CNNs were required to make progress, which was dependent upon their utilisation. R-CNNs, which were one of these CNNs, have been updated to include Fast and Faster, which are the most current versions. CNN contributed to the strengthening of the base. These two-stage detectors require a significant amount of computational resources. They are not suitable for real-time applications. On the other hand, they were quite adept at finding particular locations. When it came to resolving issues with efficiency, simplified detectors such as YOLO and SSD proved to be of tremendous assistance (Sukkar et al., 2024). It is possible that solid-state drives (SSDs) will be able to achieve the optimal balance between speed and accuracy if they are able to anticipate the need for limit boxes of varying sizes. In terms of comprehensive real-time detection, YOLO is the finest solution available. In spite of the fact that there may not be sufficient datasets including tagged pedestrians, data augmentation and transfer learning have been shown to be effective in resolving this issue. Researchers are now conducting significant research on the development of quick and compact neural networks, such as Mobile Net and EfficientDet. It is advantageous for researchers to employ these designs since they may be implemented on low-power devices without compromising the quality of detection. The Internet of Things (IoT), edge networks, and mobile phone applications are all examples of circumstances in which the identification of pedestrians is of utmost importance. This might be beneficial in such instances. In spite of the fact that a solution may seem to be appealing, researchers are required to continually search for novel approaches to address issues such as occlusion, poor lighting, and sluggish processing speed. These technological problems could be able to be resolved by the use of attention structures, lightweight designs, and multi-scale amalgamation. Simultaneously, researchers have the freedom to employ any or all of these strategies (Chen et al., 2021).

RESEARCH QUESTION

What is the effect of key technologies in pedestrian detection?

RESEARCH METHODOLOGY

Research Design: The quantitative analysis of the data used the latest version of SPSS, 25. The odds ratio and 95% confidence interval were used to assess the magnitude and direction of the statistical link. The researchers determined a statistically significant criterion of $p < 0.05$. An analytical evaluation was performed to identify the primary elements of the data. Quantitative methods are often used to assess data acquired via surveys, polls, and questionnaires, in addition to data analysed by computational statistical techniques.

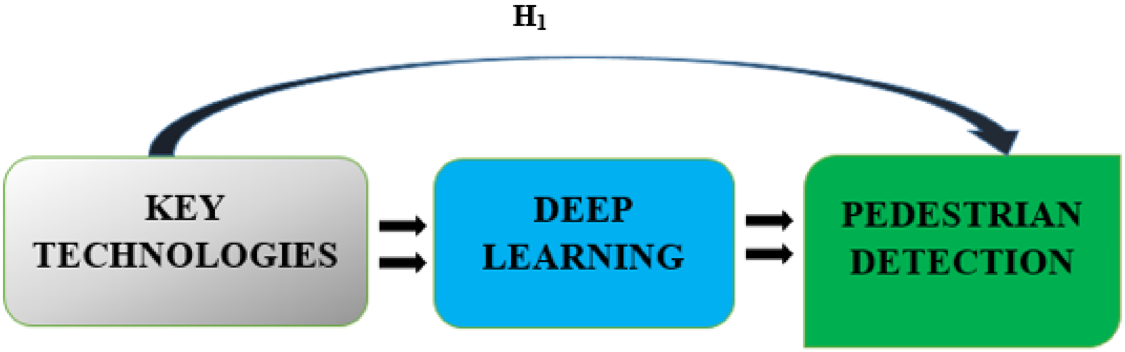
Sampling: A sample of 1455 people was obtained with the Rao-soft method. A total of 1600 questionnaires were sent to a meticulously stratified population, including students, instructors, guardians, and administrators throughout metropolitan areas of China. A total of 1,539 questionnaires were returned, resulting in a remarkable raw response rate. Thirty-nine surveys were discarded due to incompleteness or absence of relevant information. The final sample considered appropriate for utilisation consisted of 1500 genuine respondents.

Data and Measurement: The research primarily used data obtained from a questionnaire survey. The participant's fundamental demographic information was solicited first. Participants were then provided with a 5-point Likert scale to assess the online and offline channels. The researchers meticulously examined several resources, particularly online databases, for this secondary data collection.

Statistical Software: The statistical study was performed with SPSS 25 and Microsoft Excel.

Statistical Tools: The fundamental attributes of the data were comprehended by descriptive analysis. The researcher must analyse the data using ANOVA.

CONCEPTUAL FRAMEWORK



RESULT

Factor Analysis: Factor Analysis (FA) is often used to identify latent variables within observable data. Utilising regression coefficients for evaluation is a conventional practice when explicit visual or diagnostic cues are lacking. Models are crucial for success in Financial Analysis. Modelling inherently entails mistakes, interferences, and discernible connections. Datasets generated from multiple regression analyses may be evaluated using the Kaiser-Meyer-Olkin (KMO) Test. Researchers assert that the model and the variables in the sample are representative. The data exhibits redundancy. Data is more intelligible when shown in reduced volumes. Any value between 0 and 1 may function as the KMO output. A KMO value ranging from 0.8 to 1 is considered sufficient for sample size. Kaiser contends that these are the acceptable ranges:

Kaiser has delineated additional admission criteria.

A bleak range of 0.050 to 0.059 and a poor range of 0.60 to 0.69

The standard range for middle grades is 0.70 to 0.79.

Demonstrating a quality point score between 0.80 and 0.89.

The range from 0.90 to 1.00 astounds them.

The outcomes of Bartlett's sphericity test are as follows: The chi-square value is around 190, with a significance level of 0.000.

This confirms that the statements made for sampling purposes are genuine. The researchers used Bartlett's Test of Sphericity to evaluate the relevance of the correlation matrices. A Kaiser-Meyer-Olkin measure score of 0.918 indicates a satisfactory sample size. The p-value derived from Bartlett's sphericity test is 0.00. The association matrix does not possess a unique value, hence satisfying Bartlett's circularity test.

Table 1. Evaluation of Sampling Adequacy using KMO and Bartlett's Test; the Kaiser-Meyer-Olkin measure is 0.918.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.918
Bartlett's Test of Sphericity	**Approx. Chi-Square**	3252.968
	df	190
	Sig.	.000

Bartlett's Test of Sphericity further validated the importance of the association criteria. The Kaiser-Meyer-Olkin metric of sampling adequacy is 0.918. Utilising Bartlett's sphericity test, researchers obtained a p-value of 0.00. The results of Bartlett's sphericity test indicated that the correlation matrix is flawed.

INDEPENDENT VARIABLE

Key Technologies: Modern computer vision systems are built on a number of distinct basic technologies. These technologies, which employ deep learning, help the algorithms that are used to find pedestrians. CNNs play a crucial role in this process by automatically learning feature hierarchies from the provided pictures. YOLO (You Only Look Once), SSD (Single Shot Multibox Detector), and Faster R-CNN are three important models that have helped make real-time recognition both precise and accurate. These models are some of the most advanced that have helped researchers reach this aim. The detecting system has worked much better once it started using hybrid models and designs based on transformers. This progress has been especially apparent when things are highly intricate. Another important technique is transfer learning, which is the process of fine-tuning pre-trained models using big datasets (like ImageNet) for the purpose of detecting pedestrians. Another crucial method is called data augmentation. This method makes the model more generalisable by using a wider range of training data. Non-maximum suppression, or NMS for short, is a strategy that helps enhance detection results by cutting down on the number of duplicate bounding boxes. To achieve this specific goal, the method involves removing border boxes. The use of these technologies together may lead to several good results, such as better accuracy, faster inference times, and greater adaptability when faced with real-world situations that are hard to deal with. These outcomes are all things that they may make happen in the future (He et al., 2021).

DEPENDENT VARIABLE

Pedestrian Detection: Pedestrian Detection (PD) is an ADAS that helps drivers avoid hitting pedestrians, motorcycles, or even pets by letting them know when they are in their way. Pedestrian Detection warns drivers of a possible risk as soon as it sees, hears, or feels it. Some pedestrian detection systems do more than just warn vehicles. Pedestrian detection systems are an important part of modern car safety technology that helps save people from being hurt or killed by cars. These systems employ a multitude of sensors to detect pedestrians who are near to a car. When this system sees anything, it tells the driver or, in certain cases, applies the brakes. Better accuracy and speed: unlike traditional detection systems that rely on fixed sensors, AI-driven solutions are more accurate since they adapt to changes in the environment. AI systems can quickly make sense of huge amounts of data, which lets them find pedestrians in both busy and dark areas. Detecting pedestrians is one kind of driving assistance technology. Sensors, cameras, and AI can find people who are near to cars. This technology can look at the area around it and identify human forms and movements. This approach might help save people and cars from getting into accidents or at least lower the number of them. Pedestrian detection systems can identify and track individuals because they employ a mix of different technologies. When used together, these devices create a large buffer zone around the car. High-resolution cameras are essential for pedestrian detection. The strategically positioned cameras enable the automobile to see pedestrians, crosswalks, and sidewalks. The technology can tell the difference between people and other things and see human outlines and movements by looking at the captured photographs in real time (Cao et al., 2021).

Relationship between Key Technologies and Pedestrian Detection: To make intelligent vision systems better, researchers need to understand better how main technologies are connected to identifying pedestrians. Researchers need to do this in order to be able to make those adjustments. On the other side, the importance derives from the fact that these systems can be changed. To find and understand complicated visual patterns, researchers need technologies that use deep learning, including CNNs. So, for these technologies to work well, they need to be able to tell when people are moving about in various places. Some models that show how these approaches may be better at finding persons going on foot without losing accuracy include Faster R-CNN, YOLO, and SSD. Each of these models has its own set of unique features. Transfer learning lets models that have been trained on vast, generic datasets do tasks that only apply to pedestrians with a little amount of data. The models have learnt from the general datasets, which is why this is the case. Transfer learning is a great way to study since it lets researchers utilise what researchers already know to fill in the gaps in researchers' knowledge. That's why transfer learning is becoming increasingly popular. The goal of data augmentation is to improve the detection process by making the training samples more varied and stronger. This will lead to better results for the samples in the end. This is done with the help of data. The models work better in tough situations, including when there isn't much light or when something is in the way. This is due to the characteristic that makes them work better. One example of a strategy that could provide more accurate and complicated findings than other approaches is non-maximum suppression. This strategy cuts down on the number of times researchers have to detect anything, which is something that other techniques need to do. These technological improvements have a big effect on how well the system works in real time, how accurately it can identify people walking, and how well it can handle unexpected situations (Nirmal et al., 2021).

Based on the above discussion, the researcher established the following hypothesis to examine the link between Key Technologies and Pedestrian Detection.

"H_{01}: There is no significant relationship between Key Technologies and Pedestrian Detection"

"H_1: There is a significant relationship between Key Technologies and Pedestrian Detection"

Table 2. H1 ANOVA Test.

ANOVA					
Sum					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	39588.620	536	5456.515	1018.576	.000
Within Groups	492.770	963	5.357		
Total	40081.390	1499			

This investigation will provide substantial outcomes. The F statistic is 1018.576, accompanied by a p-value of .000, which is below the .05 significance threshold. The hypothesis posits that **"H_1: There is a significant relationship between Key Technologies and Pedestrian Detection"** The alternative hypothesis is accepted, whereas the null hypothesis is rejected.

DISCUSSION

According to the conclusions of this research, several AI technologies that are situated on the foundation of deep learning have been shown to have a significant association with the efficiency of pedestrian detection systems. It has been shown that the use of various diverse methodologies has the potential to significantly enhance the accuracy, diversity, and efficiency of detection models. Concurrent artificial neural networks (CNNs), data enhancement, transfer learning, and non-maximal suppression are a few examples of the approaches that fall under this category. SSD, YOLO, and faster R-CNN are three modern ideas that have shown effectiveness in situations requiring real-time settings. Research has proven the success of these concepts. When it comes to the designs, it is feasible to successfully strike a balance that is desirable between speed and accuracy. It is very vital to have these technologies in place to recognise pedestrians in a timely and dependable way. There are many different uses for these technologies, some examples of which include self-driving automobiles, advanced surveillance systems, and smart city infrastructure. Enhancing the efficiency of a model may be accomplished via the use of a method that is referred to as data augmentation. This method is applicable under a wide range of circumstances, such as when one's field of view is restricted, when the lighting conditions are changing, or when the settings are being dynamically altered. Transfer learning, on the other hand, is a method that makes use of information that has been taught in the past to reduce the amount of data that is required for certain tasks. The mobile and integrated pedestrian detection systems are getting more advanced as a result of the development of newer, lighter models and designs that make use of transformers to discern between humans. Transformers are used since they are a crucial component of these systems and the adjustments that are made to them. The findings of the study led the researchers to the opinion that these kinds of technologies are necessary for the creation of intelligent machines that are able to solve problems that are faced in the real world.

CONCLUSION

Increasing the effectiveness of pedestrian detection systems requires fundamental deep learning strategies, as shown by the findings of this study. Convolutional neural networks, data augmentation, transfer learning, and maximum suppression are some of the techniques that researchers have used to enhance pedestrian identification in tough and dynamic environments. These techniques have resulted in the detection of pedestrians being faster, more accurate, and more reliable. Real-time detection of humans walking may be accomplished by the use of models such as SSD, YOLO, and Faster R-CNN. As a result of the fact that this skill is necessary for both intelligent surveillance systems and autonomous vehicles, researchers are confident in the claim that researchers have made. Solutions that are based on transformers are getting lighter and more efficient, which is excellent news for devices that have limited resources. The number of individuals who are eager to get further knowledge on these subjects is astounding. In spite of the fact that the organisation has made significant progress, it is essential to bear in mind that there are still issues that need to be resolved. The problem of occlusion, scenes that are overly crowded, and uneven lighting are only some of the many concerns that need to be addressed. It is imperative that research be carried out in these vital areas, as well as the development of innovative ideas, in order to improve the reliability and safety of pedestrian detection systems. With the use of this newly discovered knowledge, researchers are now able to develop algorithms that are able to identify pedestrians and steer clear of these problems. After the researcher concludes their study, they will conduct an exhaustive investigation. This study may be valuable in the development of artificial intelligence vision systems for applications in the real world. A great deal of information is included within the pages of this book.

REFERENCES

1. Cao, J., Pang, Y., Xie, J., Khan, F. S., & Shao, L. (2021). From handcrafted to deep features for pedestrian detection: A survey. IEEE transactions on pattern analysis and machine intelligence, 44(9), 4913-4934.
2. Chen, L., Lin, S., Lu, X., Cao, D., Wu, H., Guo, C., ... & Wang, F. Y. (2021). Deep neural network-based vehicle and pedestrian detection for autonomous driving: A survey. IEEE Transactions on Intelligent Transportation Systems, 22(6), 3234-3246.
3. He, F., Karami Olia, P., Jamili Oskouei, R., Hosseini, M., Peng, Z., & BaniRostam, T. (2021). Applications of deep learning techniques for pedestrian detection in smart environments: a comprehensive study. Journal of advanced transportation, 2021(1), 5549111.
4. Iftikhar, S., Zhang, Z., Asim, M., Muthanna, A., Koucheryavy, A., & Abd El-Latif, A. A. (2022). Deep learning-based pedestrian detection in autonomous vehicles: Substantial issues and challenges. Electronics, 11(21), 3551.
5. Mirama, V. F., Diez, L. E., Bahillo, A., & Quintero, V. (2021). A survey of machine learning in pedestrian localization systems: Applications, open issues and challenges. IEEE Access, 9, 120138-120157.
6. Nirmal, I., Khamis, A., Hassan, M., Hu, W., & Zhu, X. (2021). Deep learning for radio-based human sensing: Recent advances and future directions. IEEE Communications Surveys & Tutorials, 23(2), 995-1019.
7. Sukkar, M., Jadeja, R., Shukla, M., & Mahadeva, R. (2024). A Survey of Deep Learning Approaches for Pedestrian Detection in Autonomous Systems. IEEE Access.
8. Zhou, H., & Yu, G. (2021). Research on pedestrian detection technology based on the SVM classifier trained by HOG and LTP features. Future Generation Computer Systems, 125, 604-615.
9. Zhu, Z., Hu, Z., Dai, W., Chen, H., & Lv, Z. (2022). Deep learning for autonomous vehicle and pedestrian interaction safety. Safety science, 145, 105479.