Application of Convolutional Neural Networks in Autonomous Driving Scene Understanding

Haodi Zhang

Hubei University of Education, Wuhan, China

ABSTRACT

With the rapid development of autonomous driving technology, environmental perception and scene understanding have become key factors to ensure safe driving. As a powerful deep learning algorithm, convolutional neural network (CNN) has demonstrated outstanding capabilities in image processing and visual perception. This paper explores the application of CNN in scene understanding of autonomous driving, and analyzes its advantages in image classification, object detection, semantic segmentation, etc., especially its performance in real-time environmental perception. Through multi-level feature extraction, CNN can identify and understand important information such as road signs, pedestrians, and vehicles from complex traffic scenes, providing accurate decision support for autonomous driving systems. The article also explores the combination of CNN with other technologies such as RNN, reinforcement learning, and multimodal data fusion, looks forward to the development trend of autonomous driving technology in the future, and discusses the challenges faced by the technology and corresponding solutions. Through the research of this paper, it is hoped that a theoretical basis and practical guidance will be provided to further improve the intelligence level and safety of autonomous driving systems.

Keywords: Autonomous driving; CNN; Scene understanding; Image processing; Deep learning

1 INTRODUCTION

As an important breakthrough in the field of intelligent transportation, autonomous driving technology has made significant progress in recent years. With the improvement of computing power, the advancement of sensor technology, and the continuous innovation of big data and artificial intelligence, autonomous driving is no longer just a science fiction concept, but has gradually entered the stage of real application. From the early automated assisted driving system to today's fully autonomous driving, the technological leap has enabled autonomous driving to cope with more complex traffic scenarios and gradually possess the ability to drive safely in different environments such as cities and highways [1]. However, although autonomous driving technology has achieved initial results, how to ensure that the system makes accurate judgments and decisions in complex and dynamic environments is still the key to its widespread application.

Among the many technologies of autonomous driving, scene understanding plays a vital role. Scene understanding is not only about perceiving the surrounding environment, but also about quickly and accurately analyzing and making decisions under rapidly changing traffic conditions. In order to enable the autonomous driving system to have this capability, visual perception has become one of the core technologies [2]. Vehicles need to obtain image or video data of the surrounding environment through sensors such as cameras and convert them into information that can be understood by computers. How to extract key information from a large amount of visual data and quickly identify road signs, pedestrians, vehicles, etc. is one of the main challenges facing autonomous driving technology. The core is image processing and analysis, which makes the application of CNN in autonomous driving particularly important.

As an important algorithm in deep learning, CNNs have powerful image processing capabilities. CNN can automatically extract hierarchical features from the original image by simulating the feature extraction process of human vision [3]. It has excellent performance in image classification, target detection, semantic segmentation, etc. The advantage of CNN is that it can process high-dimensional data, automatically extract spatial features in images, and perform deep learning through multi-level structures to identify potential key objects in complex scenes. Autonomous vehicles can rely on the powerful functions of CNN to process data from sensors in real time, understand various objects, traffic signs and potential dangers in the environment, and provide necessary information support for decision-making systems.

The purpose of this study is to explore the application and advantages of CNNs in autonomous driving scene understanding, focusing on how CNN can improve environmental perception, strengthen decision support, and improve the safety of the overall system during autonomous driving [4]. The article first reviews the development history and challenges of autonomous driving technology, and then introduces the basic concepts of scene understanding and its role in autonomous driving. Subsequently, the application of CNN in image processing is emphasized, especially in specific scenarios in autonomous driving. Finally, this article will look forward to the future combination of CNN and other technologies, propose ideas and methods to further optimize autonomous driving technology, and discuss the challenges and solutions faced.

2 SCENE UNDERSTANDING IN AUTONOMOUS DRIVING

In the research and application of autonomous driving technology, scene understanding is a crucial link. Scene understanding can be simply understood as analyzing the perceived environmental information so that the autonomous driving system can understand and judge the current state of the road and traffic environment, and make reasonable driving decisions. This process involves a large amount of perception data, including images, sounds, radar information, and other sensor data, in order to enable the autonomous driving system to understand the composition of the surrounding environment, predict potential dangers, and respond appropriately, just like human drivers [5].

The perception module in the autonomous driving system is mainly responsible for acquiring and processing data from various sensors. These sensors usually include cameras, radars, laser radars (LiDAR), ultrasonic sensors, etc. Through these devices, the vehicle can obtain visual and distance information of the surrounding environment, and further perform target detection, obstacle recognition, road sign recognition, lane line tracking, etc. The perception module is crucial to the safety and efficiency of the autonomous driving system because it determines the system's understanding accuracy and real-time response capabilities of the environment [6]. In this module, CNNs are widely used in image and video processing to achieve more accurate environmental recognition by automatically extracting meaningful features from images.

However, the implementation of scene understanding is not easy, especially in autonomous driving, where the environment is usually complex and changeable. When processing environmental data, autonomous driving systems often face three major challenges: diversity, dynamics, and complexity. First, diversity refers to the changes in scenes under different driving environments [7]. Different environments such as day and night, sunny and rainy, urban and rural have a great impact on sensor data, so autonomous driving systems need to have sufficient adaptability to maintain efficient perception capabilities under various conditions. Second, dynamics means that changes in the traffic environment are constantly occurring. Whether it is pedestrians, other vehicles, or traffic lights, these targets and signs change in space and time. Therefore, the system needs to have the ability to perceive in real time and respond quickly to adapt to emergencies or rapidly changing situations. Finally, complexity is reflected in the multi-level and multi-target structure of the scene itself. In complex urban traffic, there are multiple targets at the same time, the road conditions are ever-changing, and vehicles need to process a large amount of information while driving at high speeds, which places strict requirements on the accuracy and efficiency of the perception system.

In the face of these challenges, autonomous driving systems need to use efficient algorithms and powerful computing power to not only identify various targets in the environment, but also understand the relationships and interactions between them. Only through accurate scene understanding can reliable information be provided to the decision module to make safe and efficient driving decisions.

3 CNN BASICS

CNN is a deep learning model widely used in image recognition and processing. Its basic principle is to imitate the working mode of biological neural network and gradually extract the features of the image through a multi-level structure. In CNN, each layer automatically learns different levels of features from the original image, so that the model can effectively perform image classification, object detection, etc. Through end-to-end training, CNN can self-optimize with the help of a large amount of labeled data.

The structure of CNN consists of multiple layers, mainly including convolution layer, pooling layer and fully connected layer. The convolution layer is the core part of CNN, and its function is to extract local features in the image through convolution operation. The essence of convolution operation is to slide the filter on the input image and calculate the weighted sum of the local area to obtain the feature map [8]. This method can effectively capture the low-level features of the image such as edges, textures, and colors. The pooling layer is responsible for downsampling the feature map after convolution, reducing the size of the feature map, thereby reducing the amount of calculation and memory consumption. Common pooling methods include maximum pooling and average pooling, which retain important features and discard redundant information by extracting the maximum value or mean of the local area. The fully connected layer is located at the end of CNN and is usually used to map high-dimensional feature information to the final output space for classification or regression.

CNN is widely and successfully used in image recognition. In autonomous driving scenarios, CNN is used to identify road signs, traffic signals, pedestrians and other vehicles. Among these, CNN can extract useful visual information from complex visual data through its multi-level feature extraction capabilities and make accurate recognition based on this. In addition to classification, CNN can also be used for object detection and semantic segmentation.

Compared with traditional image processing methods, deep learning methods, especially CNN, have significant advantages in image recognition. Traditional image processing methods usually rely on manually designed feature extraction algorithms, such as edge detection and texture analysis, which perform well in some simple cases, but often have limited effects in complex image data. The introduction of deep learning, especially CNN, has changed this situation. CNN can handle more complex and diverse image data by automatically learning features in images instead of relying on manually designed features, and can adaptively optimize the feature extraction process.

4 APPLICATION OF CNN IN AUTONOMOUS DRIVING

CNNs are widely used in autonomous driving and have promoted the development of autonomous driving technology in many aspects. First, CNNs are widely used for image classification and are essential for autonomous driving systems. During autonomous driving, vehicles need to judge and identify multiple key elements such as road signs, traffic lights, pedestrians, etc. in real time. Image classification automatically analyzes image data obtained from cameras through CNN and classifies them into different objects or categories [9]. This enables autonomous driving systems to respond quickly, such as stopping, changing lanes, or slowing down, thereby improving driving safety. CNNs can effectively identify the color changes of traffic lights, determine whether to stop, or predict potential collision risks by identifying pedestrians.

In terms of target detection, the advantages of CNNs have also been widely used. Target detection is not only about identifying objects in an image, but also about identifying the locations of these objects. Target detection in autonomous driving includes identifying surrounding vehicles, pedestrians, and other obstacles. CNN extracts image features through convolutional layers, and then locates and classifies targets in the image through region proposal networks (RPNs) or other detection algorithms. Target detection is the core of the perception module in the autonomous driving system. Only by accurately identifying and locating surrounding objects can the autonomous driving system effectively avoid collisions and drive safely in complex traffic environments.

Semantic segmentation is another important application of CNN in autonomous driving. Unlike object detection, the goal of semantic segmentation is to classify images at the pixel level and label each pixel in the image as a certain category, such as roads, lane lines, buildings or pedestrians. This is very important for the decision-making of autonomous driving systems because it helps the system understand the structure of the entire scene and identify which areas are drivable and which areas are potentially dangerous [10]. Through high-precision semantic segmentation, autonomous driving systems can achieve more refined environmental perception and further enhance the system's ability to understand complex environments.

In addition, CNN is also widely used in depth estimation and three-dimensional reconstruction, especially in autonomous driving, which is crucial for enhancing environmental perception. Depth estimation refers to obtaining image data through devices such as cameras and inferring the depth information of each object in the scene. This not only helps to determine the distance between the object and the vehicle, but also can reconstruct the three-dimensional scene model by generating a depth map. Through depth estimation, autonomous driving systems can better perform three-dimensional modeling and spatial positioning, which plays a fundamental role in coping with complex road environments and performing path planning.

Finally, the realization of path planning and decision-making is inseparable from visual understanding. Through the scene information extracted by CNN, the autonomous driving system can judge multiple factors such as road conditions, vehicle location, pedestrian distribution, etc., so as to generate a reasonable driving route. CNN helps the system understand the structure of the current environment, traffic rules and other dynamic changes, and make driving decisions based on these understandings. When the system recognizes that there is a traffic jam or obstacle ahead, it can decide whether to change lanes or slow down, thereby effectively avoiding potential risks. In path planning, visual understanding is not limited to identifying static traffic signs, but also includes real-time analysis of dynamic scenes, such as the driving trajectories of other vehicles and the movement of pedestrians.

Therefore, CNN not only provides efficient and accurate perception capabilities for autonomous driving, but also provides an important visual basis for vehicle decision-making and path planning, enabling autonomous driving systems to achieve safe and precise autonomous driving in the complex real world.

5 CHALLENGES AND SOLUTIONS OF CNN IN AUTONOMOUS DRIVING

Although the application of CNNs in autonomous driving has made significant progress, it still faces many challenges, mainly in terms of data sets and annotations, computational efficiency, generalization ability and robustness. First, the data set and annotation problem is a major problem in the application of CNN in autonomous driving. In order to train an efficient CNN model, it is necessary to rely on a large amount of labeled data. However, the high-quality labeled data required for autonomous driving is not only expensive to obtain, but also challenging in terms of the accuracy and diversity of annotations in different scenarios. Especially in complex urban roads, complex weather and lighting conditions, the existing labeled data often cannot cover all driving scenarios. Therefore, how to improve the diversity and quality of data has become a key issue in the field of autonomous driving. To address this problem, researchers have proposed a variety of solutions, such as data augmentation technology, synthetic data generation and self-supervised learning, which help expand the training data set and improve the performance of the model.

Secondly, computational efficiency is another major challenge in the application of CNN in autonomous driving. Since the autonomous driving system requires real-time processing of a large amount of data obtained from sensors such as cameras and radars, the CNN model must have extremely high computational efficiency in order to respond quickly during driving. Modern CNN models are usually very large, containing millions of parameters and requiring huge computational effort. Especially for high-resolution image data, the processing speed is far from meeting real-time requirements. To address this challenge, researchers have adopted many optimization strategies, including network structure optimization, pruning technology, quantization technology, and hardware acceleration. These methods can effectively reduce the computational complexity of CNN models and improve their performance in real-time applications.

Generalization is also an important issue facing CNN in autonomous driving. Autonomous driving systems need to perform consistently in a variety of different environments. However, CNN models perform well on specific training data sets, but not so well in new, unseen environments. Since the application scenarios of autonomous driving are extremely complex and varied, how to improve the generalization ability of CNN in different environments is an urgent problem to be solved. To solve this problem, researchers have adopted technologies such as transfer learning and generative adversarial networks (GANs) to enable the model to achieve better generalization effects on smaller training sets. At the same time, designing more robust loss functions and more representative training data sets is also an effective way to improve generalization ability.

Finally, strong robustness is a key requirement for the application of CNN in autonomous driving. The robustness of the autonomous driving system is directly related to driving safety, especially under complex weather and lighting changes, the quality and accuracy of sensor data will be greatly affected. In rainy days, haze weather or strong sunlight, the perception effect of cameras and radars will be significantly reduced, resulting in inaccurate scene understanding, which in turn affects decision-making and control. In order to improve the robustness of the CNN model, researchers have proposed many methods, such as multimodal data fusion technology, combining different types of sensor information such as cameras, radars, and LiDAR, to make up for the shortcomings of a single sensor in harsh environments. In addition, data enhancement technology can also be used to simulate different weather and lighting conditions to help CNN models perform more stably and reliably under various complex conditions.

Overall, although CNN faces challenges in data sets, computational efficiency, generalization ability, and robustness in autonomous driving, the application prospects of CNN in the field of autonomous driving are still very broad through continuous optimization of algorithms, enhancement of data set diversity, improvement of computational efficiency, and strengthening of robustness training.

6 COMBINATION OF CNNS AND OTHER DEEP LEARNING MODELS

In the field of autonomous driving, the combination of CNNs and other deep learning models has become an effective way to improve system performance and adapt to complex situations. First, the combination of CNN and recurrent neural networks (RNNs) provides autonomous driving systems with powerful time series data processing capabilities. CNNs are usually used to extract spatial features from images, while RNNs are able to process time series data, especially to model the dynamic environment that changes constantly during vehicle driving. Autonomous driving systems need to process video stream data or continuous image sequences. RNNs can predict the content of the current frame based on the image information of the previous frame, which is crucial for predicting future road conditions, vehicle positions, and the behavior of dynamic obstacles. During vehicle driving, RNNs can help predict the movement trajectory of the vehicle in front or the behavior of pedestrians, thereby providing temporal continuity and rationality for path planning and decision-making. Therefore, the combination of CNN and RNN can not only improve the ability to recognize static images, but also enhance the perception and prediction capabilities of dynamic environments, making autonomous driving systems more intelligent and flexible in complex real-time situations.

In addition to combining with RNNs, CNNs also play an important role in multimodal data fusion. Autonomous driving systems usually rely on multiple sensors to perceive the surrounding environment, including different types of sensors such as cameras, LiDAR, and radars. Each sensor has unique advantages and limitations, and a single sensor may have errors or blind spots in complex environments. Multimodal data fusion aims to effectively integrate data from different sensors to improve the perception accuracy and robustness of the system. CNN can extract spatial features of images when processing data from visual sensors, but visual sensors fail when facing low light or bad weather. In this case, LiDAR can supplement the lack of visual data by providing high-precision distance measurement. Radar can provide stable perception data under complex weather conditions. By fusing these heterogeneous data, the autonomous driving system can obtain more comprehensive environmental perception information, thereby improving the ability to understand complex scenes. This multimodal fusion not only enhances the reliability of the system in various situations, but also ensures the safety of autonomous driving under different environmental conditions.

In addition, self-supervised learning and reinforcement learning in autonomous driving have gradually become important technical means. Self-supervised learning is a method of unsupervised learning that automatically generates labels and learns by training with unlabeled data. In autonomous driving, due to the lack of high-quality labeled data, self-supervised learning is widely used to automatically extract features from a large amount of unlabeled sensor data, thereby improving the learning efficiency of the model. This method can reduce the reliance on manually labeled data, while still effectively improving the performance of the model in the case of large-scale data sets. Reinforcement learning learns through the interaction between the agent and the environment. The autonomous driving system can continuously optimize the driving strategy through reinforcement learning to achieve the best decision-making effect. In the actual driving process, reinforcement learning can help autonomous driving vehicles adjust driving speed, path planning and obstacle avoidance strategies based on real-time environmental feedback. Autonomous driving vehicles learn how to deal with complex traffic conditions and emergencies through repeated interactions with the road environment, thereby continuously improving the intelligence of driving decisions.

Overall, the combination of CNNs and other deep learning models enables autonomous driving systems to handle more complex problems, such as time series data modeling, heterogeneous data fusion, and autonomous decision-making. These advances not only improve the intelligence level of the system, but also lay a solid foundation for the further development of autonomous driving technology.

7 FUTURE OUTLOOK

With the continuous advancement of autonomous driving technology, the application of

CNNs in scene understanding will usher in a new development trend. In the future, CNN will not only be limited to a single function, such as image classification or target detection, but will develop in a more comprehensive and intelligent direction. Through multi-task learning, CNN will be able to process multiple perceptions at the same time and achieve more accurate environmental understanding. In complex urban road environments, autonomous driving systems need to perform road sign recognition, pedestrian detection, vehicle tracking, etc. at the same time. Future CNN models will be able to achieve parallel processing of multiple tasks, improve efficiency and enhance the overall understanding ability of the model. In addition, with the continuous evolution of deep learning models, CNN will be more closely integrated with other algorithms, such as graph neural networks (GNNs), which can better capture the spatial relationship between roads, objects and vehicles, thereby more accurately inferring the layout and future changes of the driving environment. In addition, the application of CNN in autonomous driving will also gradually deepen its integration with other technologies, especially with technologies such as 5G and edge computing. The high speed, low latency and large bandwidth characteristics of 5G networks provide great support for real-time data transmission of autonomous driving technology. Through the 5G network, autonomous vehicles can transmit perception data to the cloud in real time for processing and analysis, or communicate with other vehicles and traffic infrastructure, thereby achieving collaboration between vehicles and between vehicles and roads. This collaboration can effectively improve the performance of autonomous driving systems in complex traffic environments, allowing vehicles to make more accurate and safe decisions based on surrounding traffic conditions. At the same time, edge computing will also become an important part of autonomous driving systems. By processing and calculating data on on-board devices or roadside devices, edge computing can greatly reduce data transmission delays and improve real-time response capabilities. The combination of CNN and edge computing enables autonomous driving systems to quickly process image data locally and respond in a timely manner, thereby improving driving safety and efficiency.

However, the rapid development of autonomous driving technology has also brought about some ethical and legal issues, which will be an important challenge for future development. First of all, how autonomous driving systems make decisions in the face of emergencies involves moral and ethical issues. When an autonomous driving vehicle faces an unavoidable collision, the system needs to decide whether to protect the safety of passengers in the car or choose to avoid hitting pedestrians or other vehicles. This type of decision-making problem is not only a technical issue, but also involves social ethical thinking. Therefore, how to design a reasonable decision-making framework so that the autonomous driving system can make ethical choices in different situations has become an urgent problem to be solved. Secondly, the popularization of autonomous driving technology has also brought challenges at the legal level. Issues such as the attribution of responsibility for autonomous driving vehicles, the handling of traffic accidents, and the protection of data privacy all require the improvement of the legal system. When an autonomous driving vehicle has an accident, how to define whether the responsibility is the responsibility of the owner, developer, or manufacturer, the existing legal system cannot provide clear guidance. In addition, the massive data collected by autonomous driving vehicles involves user privacy and security issues. How to balance the relationship between technological progress and data protection is also an issue that needs to be considered in law and ethics.

In short, with the advancement of technology and changes in social needs, the development of CNN in autonomous driving will continue to move towards a more intelligent, accurate and efficient direction. At the same time, the combination with cutting-edge technologies such as 5G and edge computing will provide stronger support for autonomous driving systems and enhance their real-time and collaborative capabilities. However, the development of technology has also brought new ethical and legal issues, and all sectors of society need to work together to formulate appropriate norms and standards to ensure that autonomous driving technology can meet social ethical and legal requirements while achieving intelligent driving.

8 CONCLUSION

The application of CNNs in autonomous driving has shown its important role in key areas such as environmental perception, target recognition, and path planning. Through CNNs, autonomous driving systems can extract effective features from visual data, identify road signs, traffic lights, pedestrians, other vehicles, etc., and then make accurate decisions and controls. In the actual scenarios of autonomous driving, CNNs not only help the system understand and interpret complex image information, but also improve the real-time perception of the environment. Combined with technologies such as target detection, semantic segmentation, and depth estimation, CNNs enable autonomous driving systems to better handle the recognition of dynamic and static objects and enhance the driving ability of vehicles. In addition, the combination of CNNs with other deep learning technologies such as recurrent neural networks (RNNs) and reinforcement learning has further improved the performance of autonomous driving systems in time series data processing and autonomous learning. Overall, CNNs provide strong perception support for autonomous driving and provide a foundation for higher-level intelligent driving decisions.

However, in the face of future technological development, autonomous driving still faces many challenges. First, data diversity and quality, computational efficiency, and system robustness are still key factors restricting technological progress. With the diversification of autonomous driving scenarios, the system needs to have higher generalization capabilities and be able to adapt to changes in different environments. In addition, the real-time processing of large amounts of sensor data requires increasingly higher computing resources, and the model structure needs to be further optimized to reduce computational complexity. Future technical development suggestions include increasing research on multimodal data fusion, and improving the accuracy of environmental perception through the combination of sensor information such as vision, lidar, and radar; at the same time, with the emergence of new technologies such as 5G and edge computing, autonomous driving systems can achieve more accurate real-time responses through more efficient data transmission and computing methods. In addition, how to further solve the ethical and legal issues in autonomous driving technology is also a direction that cannot be ignored in future technological development. In order to ensure the popularization of autonomous driving, it is not only necessary to rely on technological innovation, but also to make corresponding supporting preparations in terms of legal frameworks and ethical norms.

In general, the application of CNNs in the field of autonomous driving has achieved remarkable results, but the development of technology still faces many challenges. In the future, autonomous driving technology will further develop in many aspects such as deep learning, hardware acceleration, data fusion, and legal ethics to promote the full popularization of intelligent driving. Through multidisciplinary collaborative innovation and cross-domain cooperation, autonomous driving technology will gradually be commercialized and popularized on a global scale, and ultimately bring humans a safer, more convenient and intelligent way of travel.

REFERENCES

[1] Liu, X., Yan, W. Q., & Kasabov, N. (2024). Moving vehicle tracking and scene understanding: A hybrid approach. *Multimedia Tools and Applications*, *83*(17), 51541-51558.

[2] Cao, Z., Gao, Y., Bai, J., Qin, Y., Zheng, Y., & Jia, L. (2024). Efficient dual-stream fusion network for real-time railway scene understanding. *IEEE Transactions on Intelligent Transportation Systems*.

[3] Liao, G., Li, J., & Ye, X. (2024, March). VLM2Scene: Self-supervised image-text-LiDAR learning with foundation models for autonomous driving scene understanding. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 4, pp. 3351-3359).

[4] Zhao, T., Chen, G., Gatewongsa, T., & Busababodhin, P. (2024). Forecasting Agricultural Trade Based on TCN-LightGBM Models: A Data-Driven Decision. *Research on World Agricultural Economy*, *6*(1), 207–221. https://doi.org/10.36956/rwae.v6i1.1429

[5] Sahoo, L. K., & Varadarajan, V. (2025). Deep learning for autonomous driving systems: technological innovations, strategic implementations, and business implications-a comprehensive review. *Complex Engineering Systems*, *5*(1), N-A.

[6] Li, H., Chu, H. K., & Sun, Y. (2024). Temporal consistency for RGB-thermal data-based semantic scene understanding. *IEEE Robotics and Automation Letters*.

[7] Tirumalapudi, R., & Sirisha, J. (2024). Onward And Autonomously: Expanding The Horizon Of Image Segmentation For Self-Driving Cars Through Machine Learning. *Scalable Computing: Practice and Experience*, *25*(4), 3163-3171.

[8] Salman, H. A., & Kalakech, A. (2024). Image enhancement using convolution neural networks. *Babylonian Journal of Machine Learning*, *2024*, 30-47.

[9] Zhang, Y., & Tang, Q. (2024). Accelerating autonomy: an integrated perception digital platform for next generation self-driving cars using faster R-CNN and DeepLabV3. *Soft Computing*, *28*(2), 1633-1652.

[10] Abdullahi, M., Oyelade, O. N., Kana, A. F. D., Bagiwa, M. A., Abdullahi, F. B., Junaidu, S. B., ... & Chiroma, H. (2024). A systematic literature review of visual feature learning: deep learning techniques, applications, challenges and future directions. *Multimedia Tools and Applications*, 1-58.