Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/17/20230952

Application of neural network in object motion tracking in the context of big data

Ruiheng Shao

Knox Grammar School, Wahroonga, Sydney, Australia

ruihengshao@gmail.com

Abstract. In the context of big data, the advancements in computer analytics and computational capability have promoted transformative opportunities. At the forefront of this progress, neural networks have emerged as an essential instrument across diverse domains, catalyzing enhancements in operational efficiency. This paper embarks on an exploration of the applications of neural networks within the realm of object motion prediction and tracking, with the primary objective of improving the efficacy and precision of tracking mechanisms. Amidst the escalating complexities of data proliferation, the imperative for streamlined tracking methodologies has intensified. In response, this study investigates the amalgamation of neural networks with object motion prediction and tracking processes. By using the powerful abilities of neural networks, the research aims to bring a significant change to the field. This change will be seen in improved accuracy and faster tracking efficiency. Combining neural networks with motion prediction and tracking creates a cooperative partnership that develops a new system based on data-driven insights. By merging computational power and advanced algorithms, the paper anticipates a fresh direction for tracking technology.

Keywords: big data, neural network, motion tracking.

1. Introduction

With the rapid development of technology in recent years, there has been an exponential increase in the amount of data available on the Internet. A study conducted in 2018 revealed that an astounding 2.5 trillion bytes of data were created daily, with approximately 90% of all data on the internet generated within the last two years [1]. This growth has ushered in the era of big data, where the analysis of vast quantities of information holds immense potential across various fields. One area that has greatly benefited from this wealth of data is the field of Artificial Intelligence (AI), particularly in terms of AI training. The availability of massive datasets and the relatively low cost of training have positioned AI as a promising technology in numerous domains. The utilization of big data for AI training allows for the extraction of valuable insights and patterns, empowering AI systems to make informed decisions and perform complex tasks with a high degree of accuracy. Moreover, the analysis of substantial amounts of Internet data can enhance the training of other models and algorithms. Traditionally, motion tracking relied on analysing the position of objects in two frames captured by a camera, calculating the necessary camera angle adjustments, and then implementing those adjustments through a steering device [2]. However, with the advent of big data and the application of neural networks trained on extensive datasets, it is anticipated that motion-tracking capabilities can be significantly improved. Integrating big

DOI: 10.54254/2755-2721/17/20230952

data analytics into existing algorithms is expected to enhance tracking behaviours and unlock new possibilities in terms of precision, speed, and efficiency.

2. Overview of the big data era, neural networks, and object tracking

2.1. Big data era

In the context of the global information age, the exponential growth of data production across various facets of human life has given rise to what is commonly referred to as the Big Data Era. Big data encompasses vast and intricate data sets that exceed the processing capabilities of traditional methods within a reasonable timeframe [3]. This explosion of data poses significant challenges to conventional data processing approaches, necessitating the development of new tools and techniques to effectively analyse and derive meaningful insights from such massive quantities of information. However, the Big Data Era also presents a plethora of opportunities for emerging technologies, particularly in the realm of artificial intelligence (AI). The analysis of big data catalyses the active integration and application of AI in diverse fields within contemporary society [4]. The availability of copious amounts of data provides AI systems with the necessary raw materials to learn, adapt, and make accurate predictions or decisions. The fusion of big data and AI enables organizations and individuals to gain valuable insights, identify patterns, and extract actionable knowledge from complex data sets that were previously unattainable. By leveraging advanced algorithms, machine learning techniques, and powerful computing resources, AI can detect hidden correlations, recognize trends, and unveil relationships within the vast sea of information. This capacity to derive meaningful insights from big data not only empowers decision-making processes but also fuels innovation and drives advancements in various industries. The impact of the Big Data Era extends beyond AI and permeates multiple sectors, including healthcare, finance, marketing, transportation, and manufacturing. In healthcare, for instance, the analysis of large-scale patient data can lead to more accurate diagnoses, personalized treatments, and improved patient outcomes. In finance, big data analytics can facilitate more precise risk assessments, fraud detection, and investment strategies. Moreover, in the marketing realm, the analysis of extensive customer data enables targeted advertising, enhanced customer segmentation, and improved marketing campaigns.

2.2. Neural networks

Neural networks, a fundamental component of machine learning, are mathematical models inspired by the structure and functioning of biological neural networks. In this model, nodes, referred to as neurons, are interconnected to form a network. Each neuron receives input signals, processes them, and produces an output signal. The strength of the connections between neurons, known as weights, determines the significance of each input in the overall computation [5]. What sets neural networks apart is their ability to capture complex patterns and relationships in data through a process called training. During training, the network learns from labelled examples, adjusting its internal parameters, including the weights, to minimize the difference between the predicted outputs and the actual outputs. This process of optimization enables neural networks to generalize their knowledge and make accurate predictions or classifications on unseen data. Neural networks have been trained for several purposes. Computer vision (CV) technology utilises neural networks to recognise the elements in a frame; natural language processing involves neural networks to convert the input texts to logical and grammatical responses; self-driving vehicles use neural networks to make decisions to ensure the safety and comfort of the passengers and driver onboard.

2.3. Object tracking

Object tracking is a critical branch of computer vision with wide-ranging applications in sports events, unmanned vehicles, transportation, and various other fields. Its primary objective is to locate and follow a specific object of interest within a sequence of video frames.[6] In general, object tracking algorithms typically consist of two main strategies. In the first approach, the computer identifies the object to be tracked in the initial frame. This can be achieved through techniques such as object detection, where the

DOI: 10.54254/2755-2721/17/20230952

computer locates and recognizes the object based on predefined features or patterns. Alternatively, the algorithm may rely on background subtraction methods, comparing the background and foreground pixels to distinguish the object of interest. Once the initial identification is complete, the algorithm proceeds to track the object across subsequent frames. The other approach involves extracting distinct features of the object, constructing a model of the target, and then estimating the object's position in each frame by matching the features with those in the model. By analysing video frames either frame by frame or at regular intervals, object-tracking algorithms can continuously identify and track the target object. However, several challenges can significantly impact the accuracy and robustness of the tracking process. Frame noise, caused by various factors such as sensor noise or compression artifacts, can introduce spurious details or distortions, making it difficult to accurately locate and track the object. Motion blur, resulting from the rapid movement of the object or the camera itself, can blur the object's appearance, further complicating the tracking process. Moreover, changes in lighting conditions, such as variations in illumination or shadows, can lead to inconsistencies in the object's appearance, making it challenging to maintain accurate tracking.

3. Big data and neural networks for object tracking

3.1. Target identification

At the initial stage of the tracking algorithm, the computer program focuses on identifying the target object to be recognized. This process involves analysing three key parameters: hue, saturation, and luminance (HSL) within the target region and comparing them with the background. By examining the HSL contrast between adjacent pixel points, the algorithm can discern variations and quantify the differences. These differences are then processed using a neural network model that assigns weights to each parameter. The neural network considers the relative importance of hue, saturation, and luminance in determining the target's characteristics. Through this analysis, the algorithm approximates the size and location of the target based on the obtained results. Once the approximate location of the target is determined, the algorithm proceeds to match the selected pixel points with other targets in a large database. By comparing the characteristics of the selected pixel points with those of known targets, the algorithm can find similar targets within the database. This step helps to classify the identified target into specific categories, such as a ball, car, animal, and so on. To enhance the accuracy of recognition, this process is repeated for the initial frames of the video sequence. By performing the identification and classification procedure on multiple frames, the algorithm aims to reduce the impact of random interference, such as noise or occlusion, on the accuracy of target recognition. This approach increases the robustness of the tracking algorithm, ensuring reliable and consistent results throughout the video sequence. By leveraging the HSL analysis, neural network modelling, and comparison with a comprehensive database, the tracking algorithm can effectively identify and classify target objects. This methodology provides a foundation for subsequent tracking processes, enabling the algorithm to track the target's movement, maintain its position, and predict its future trajectory. It is important to note that the effectiveness of this approach relies on the quality of the initial identification and classification, as well as the accuracy and diversity of the database used for comparison. Continuous advancements in computer vision techniques, machine learning algorithms, and dataset collection contribute to improving the precision and robustness of target recognition in various applications, including surveillance, object tracking, and augmented reality.

3.2. Target prediction

Based on the results obtained from the target identification process described above, different object motion trajectory prediction models are established for different types of target objects. These models aim to estimate and predict the path that the object will follow in subsequent frames. The specific type of model employed depends on the characteristics and behaviour of the target object. For instance, in the case of ball motion, such as in sports like soccer or volleyball, parabolic models can be used. These models approximate the curved trajectory of the ball's motion through the air. By considering factors

DOI: 10.54254/2755-2721/17/20230952

such as initial velocity, launch angle, and gravitational forces, the algorithm can predict the future positions of the ball accurately. Similarly, for objects like cars, linear or circular motion models are often employed. These models assume that the motion of the car follows a straight path or a circular arc, respectively. By considering factors such as speed, direction, and potential changes in movement, the algorithm can estimate future positions and track the car accordingly. Additionally, to enhance the tracking process, the algorithm calculates the approximate distance between the camera and the target object. This calculation is based on the equivalent focal length of the camera lens and the ratio of the area of the object's image projected onto the light-sensitive element in the camera to the actual size of the target object. By considering these factors, the algorithm can estimate the relative distance or depth of the object from the camera's perspective. The calculation of the approximate distance provides valuable contextual information for the tracking algorithm. It helps in refining the tracking process by incorporating depth cues and enables a better understanding of the object's spatial relationship with the camera and the surrounding environment. Furthermore, factors such as occlusion, motion blur, and changes in lighting conditions can introduce uncertainties that may affect the accuracy of the predictions. By utilizing these trajectory prediction models and distance calculations, the tracking algorithm can anticipate the future positions of the target object and maintain continuous and accurate tracking of its motion.

3.3. Target searching

Based on the mathematical model calculated earlier, the tracking algorithm predicts the approximate location of the target object in the next frame. To assess the similarity between pixels in the vicinity of the predicted location and the target object, a comparison is performed. This comparison examines areas of pixels from near to far, and similarity values between 0 and 1 for each area are obtained. In addition to the similarity value, the algorithm calculates the ratio between the distance of the region from the predicted position and the maximum distance possible in the frame. This ratio, also ranging between 0 and 1, provides an indication of the relative position of the region within the frame. To determine the overall likelihood of the region being the actual location of the target, these two values are weighted separately. Each value is subjected to a non-linear conversion, ensuring that they remain within the range of 0 to 1. The two converted values are then multiplied together, generating a resulting product. The tracking algorithm employs a specific threshold value to evaluate the product. If the product exceeds this threshold, the search is considered successful, and the area is identified as the actual location of the target. This threshold serves as a criterion for determining the level of confidence in the match between the predicted location and the region under evaluation. In cases where no region surpasses the threshold in the entire screen, the algorithm resorts to an alternative approach. It identifies the area with the largest product, considering it as the target's location, albeit with a lower level of confidence. This fallback mechanism helps in situations where the target is partially obscured or when tracking conditions are challenging. The camera angle is fine-tuned based on the previously calculated distance. This adjustment allows for precise alignment of the camera's viewpoint with the target object, ensuring that the subsequent frames capture the object consistently and accurately. The process described above demonstrates the decision-making mechanism employed by the tracking algorithm to determine the actual location of the target object. By combining similarity values, distance ratios, non-linear conversions, threshold evaluation, and fine-tuning of the camera angle, the algorithm achieves robust and reliable object tracking. The effectiveness of this approach depends on various factors, including the quality of the initial identification, the accuracy of the distance calculation, and the appropriateness of the threshold value.

3.4. Fine-tuning parameters

The fine-tuning of the model parameters plays a crucial role in optimizing the tracking process. This fine-tuning process involves adjusting the parameters based on the disparity between the actual position of the target object and the position predicted by the model. The objective is to minimize the difference between the predicted position and the actual position for each previous target. If the predicted position

DOI: 10.54254/2755-2721/17/20230952

deviates significantly from the actual position, indicating a substantial discrepancy in the model's performance, the algorithm will change the model used. For instance, if the linear motion model fails to accurately predict the target's position, the algorithm may switch to a circular motion model that better aligns with the observed behaviour of the target object. This adaptive approach allows the algorithm to select an appropriate model that captures the target's motion more accurately. Moreover, the algorithm analyses the background and the surrounding environment to identify any external forces or factors that may interfere with the target's motion trajectory. By considering the influence of external forces, such as wind, gravity, or other objects in the scene, the algorithm can make informed decisions on model selection and parameter adjustment. This adaptability ensures that the tracking algorithm can effectively handle dynamic scenarios and adjust its predictions accordingly. Simultaneously, the algorithm continuously updates the distance of the target object from the camera. This update is achieved by comparing the projected area of the target object with its actual size. By leveraging the known properties of the camera, such as the lens focal length and the ratio between the projected area and the actual size, the algorithm can estimate and update the distance of the object from the camera. This distance information is valuable for maintaining accurate spatial relationships and can aid in refining the tracking process. The iterative process of fine-tuning the model parameters, adapting to different motion patterns, considering external forces, and updating the distance information enhances the accuracy and robustness of the tracking algorithm. By continuously improving the model's predictions and accounting for various factors that may affect the target's motion, the algorithm can achieve reliable and consistent tracking performance.

4. Advantages of big data and neural networks in object tracking

4.1. Improve tracking efficiency

By estimating the location of the target object beforehand, the algorithm can start its search around the predicted position. This targeted search approach significantly reduces the computational burden, making the algorithm more efficient and scalable. In the optimal case, where the predicted location is close to the actual location of the target, the algorithm can quickly locate the target object with minimal computation. Even in the worst-case scenario, where the predicted location deviates significantly from the actual location, the complexity of the algorithm remains comparable to traditional approaches that involve traversing the entire frame to locate the target. This streamlined process enables real-time tracking performance, even in scenarios with higher frame rates or higher-resolution video. This ensures that the tracking algorithm can handle challenging situations without compromising its efficiency. The integration of big data and neural networks offers several benefits in the context of target tracking. The algorithm can leverage the vast amount of data available to train and optimize the neural network models, enabling more accurate predictions and reducing the search space. The use of big data also facilitates the identification of complex patterns and relationships in the tracking process, enhancing the algorithm's capability to handle diverse scenarios.

4.2. Strong anti-interference capability

In scenarios where the target object is subject to environmental disturbances or noise, the tracking algorithm incorporates two critical factors to ensure efficient and robust performance. The algorithm considers the high similarity between the target captured by the camera and the actual target object it is searching for. Also, the algorithm considers the proximity of potential target regions to the predicted location. This approach of prioritising to search near the predicted area helps to overcome the impact of environmental factors such as ambient light changes, noise, or blur. By focusing the search efforts in the nearby area, the algorithm increases the likelihood of locating the target object even in the presence of disturbances. By combining these two factors, the tracking algorithm enhances its ability to handle environmental variations and disturbances. The emphasis on similarity and proximity reduces the possibility of losing track of the target object due to factors like varying lighting conditions, noise interference, or blurriness in the frame.

DOI: 10.54254/2755-2721/17/20230952

5. Conclusion

In summary, through neural networks and big data comparison, the new algorithm will have the ability to build different mathematical models for different tracking targets to predict their motion trajectories. While tracking, the mathematical model will be optimized by calculating the deviation and analysing the screen, so that the prediction results will be closer to the actual target position. The prediction results can effectively reduce the costs required for object tracking and improve the anti-interference ability against ambient light changes, noise, blur, and other factors during tracking.

References

[1] Marr, B. (2022) How much data do we create every day? the mind-blowing stats everyone should read, Forbes. Forbes Magazine. Available at: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/

[2] D. Murray and A. Basu, "Motion tracking with an active camera," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 5, pp. 449-459, May 1994, doi: 10.1109/34.291452.

[3] Kusnetzky, Dan. What is "Big Data?"/ ZDNet, pp. 1708, Feb 2010. http://blogs.zdnet.com/virtualization/?p=1708

[4] Pence, H. E. (2014). What is Big Data and Why is it Important? Journal of Educational Technology Systems, 43(2), 159–171. https://doi.org/10.2190/ET.43.2.d

[5] What are neural networks? (no date) IBM. Available at: https://www.ibm.com/topics/neural-networks

[6] Meng, L. and Yang, X. (2019) A survey of Object Tracking Algorithms, CNKI. Available at: https://kns.cnki.net/kcms/detail/detail.aspx?doi=10.16383%2Fj.aas.c180277