IJITCE

ISSN 2347- 3657

International Journal of
Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, April 2024

OBJECT PRICE DETECTION USINGR-CNN THROUGH COMPUTER VISION

[1]K. VENKATESWARA RAO,[2]VULLI LIKITHA,[3]CHANDANA BHARATH KUMAR,[4]DUDDETI LAKSHMI DURGA,[5]VISWANADHAM REVANTH KUMAR

[1]Assistant Professor,[2345]Students

Department of CSE, Sri Vasavi Institute of Engineering & Technology (Autonomous), Nandamuru

ABSTRACT

This paper deals with the field of computer vision, mainly for the application of learning in object detection task. On the one hand, there is a simple summary of the datasets and deep learning algorithms commonly used in computer vision. On the other hand, a new dataset is built according to those commonly used datasets, and choose one of the networks called faster R-CNN to work on this new dataset. Through the experiment to strengthen the understanding of these networks, and through the analysis of the results learn the importance of deep learning technology, and the importance of the dataset for deep learning."

Keywords: Computer vision, object detection, deep learning algorithms, datasets, Faster R-CNN, experimentation, dataset analysis.

INTRODUCTION

The advent of computer vision has revolutionized various industries by enabling machines to interpret and understand visual information, akin to human perception. Within the realm of computer vision, object detection tasks stand as fundamental challenges, crucial for applications ranging from autonomous vehicles to surveillance systems and retail analytics. This paper delves into the realm of object detection through the lens of computer vision, exploring the application of learning methodologies in this critical task [1].At the heart of computer vision lies the utilization of deep learning algorithms, which have emerged as the cornerstone for tackling complex visual recognition tasks. Deep learning architectures, characterized by their hierarchical layers of abstraction, have demonstrated remarkable success in object detection, surpassing traditional methods with their ability to automatically learn discriminative features from raw data [2]. The introduction provides a comprehensive overview of the datasets commonly employed in computer vision research, serving as the foundational building blocks for training and evaluating object detection models [3].

In parallel, deep learning algorithms play a pivotal role in driving advancements in object detection capabilities. Through the utilization of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and other deep learning architectures, researchers have made significant strides in achieving state-of-the-art performance in object detection tasks [4]. The introduction offers insights into the various deep learning algorithms commonly employed in computer vision, including Faster R-CNN, YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and Mask R-CNN, among others [5].Furthermore, the introduction highlights the critical role of datasets in facilitating the development and evaluation of object detection models. Datasets such as COCO (Common Objects in Context), PASCAL VOC (Visual Object Classes), and ImageNet provide researchers with standardized benchmarks for assessing the performance of object detection algorithms [6]. Moreover, the introduction emphasizes the importance of data augmentation techniques in expanding the diversity and richness of training datasets, thereby enhancing the robustness and generalization capabilities of object detection models [7].

In addition to leveraging existing datasets, this paper embarks on the creation of a new dataset tailored to the specific requirements of the object detection task. Drawing inspiration from commonly used datasets in computer vision research, the creation of this new dataset aims to address unique challenges and scenarios encountered in real-world applications [8]. The introduction elucidates the methodology employed in curating and annotating the dataset, ensuring its suitability for training and evaluating object detection models [9].Central to the proposed approach is the adoption of the Faster R-CNN (Region-based Convolutional Neural Network) architecture, renowned for its accuracy and efficiency in object detection tasks. By selecting Faster R-CNN as the network of choice for experimentation, this paper aims to explore its efficacy and performance in detecting object prices within the context of the newly created dataset [10]. The introduction provides an overview of the Faster R-CNN architecture, elucidating its underlying principles and mechanisms for object detection [11].

Through a series of experiments conducted on the newly curated dataset, this paper endeavours to deepen our understanding of the intricacies involved in training and evaluating object detection models. By analyzing the results obtained from the experiments, valuable insights are gleaned regarding the efficacy of deep learning technology in object detection, as well as the critical role played by datasets in facilitating model training and evaluation [12]. Furthermore, the introduction sets the stage for the subsequent sections of the paper, wherein detailed methodologies, experimental setups, results, and discussions are presented to elucidate the findings and implications of the study [13].In summary, the introduction provides a comprehensive overview of the objectives, methodologies, and significance of the study within the broader context of object detection using computer vision. By synthesizing insights from existing literature, outlining the chosen approach, and delineating the rationale behind dataset creation and model selection, the introduction sets a solid foundation for the subsequent sections of the paper [14]. Through rigorous experimentation and analysis, this study aims to contribute to the advancement of object detection techniques, shedding light on the importance of deep learning technology and high-quality datasets in tackling real-world challenges in computer vision [15].

LITERATURE SURVEY

The literature surrounding object detection using R-CNN through computer vision spans a broad spectrum of research, encompassing diverse methodologies, datasets, and deep learning algorithms. Object detection, a fundamental task in computer vision, involves identifying and localizing objects within an image or video frame. To achieve this goal, researchers have explored various approaches, ranging from traditional methods based on handcrafted features to modern deep learning techniques that automatically learn hierarchical representations from data.A key aspect of object detection research is the availability and utilization of datasets for training and evaluation purposes. Commonly used datasets such as COCO (Common Objects in Context), PASCAL VOC (Visual Object Classes), and ImageNet provide standardized benchmarks for assessing the performance of object detection algorithms. These datasets contain annotated images with bounding boxes delineating object locations, enabling researchers to train models to recognize and localize objects across different categories. Additionally, the creation of specialized datasets tailored to specific domains or tasks further enhances the applicability and effectiveness of object detection algorithms in real-world scenarios.

Deep learning algorithms have emerged as a dominant force in advancing the state-of-the-art in object detection. Convolutional neural networks (CNNs) serve as the backbone of many modern object detection architectures, leveraging their ability to automatically extract hierarchical features from raw image data. Models such as Faster R-CNN, YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), and Mask R-CNN have demonstrated remarkable performance in accurately detecting and localizing objects within complex scenes. These architectures employ innovative techniques such as region proposal networks, anchor boxes, and feature pyramid networks to achieve high precision and recall rates in object detection tasks.The choice of deep learning algorithm often depends on factors such as computational efficiency, accuracy, and real-time performance requirements. While some architectures prioritize speed and efficiency, making them suitable for deployment in resource-constrained environments or real-time applications, others prioritize accuracy and robustness, catering to tasks where precision is paramount. Researchers continually explore novel architectures and optimization techniques to strike a balance between performance and efficiency, pushing the boundaries of what is achievable in object detection.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, April 2024

In addition to algorithmic advancements, the literature also emphasizes the importance of data preprocessing, augmentation, and post-processing techniques in improving the performance of object detection systems. Data preprocessing involves tasks such as resizing, normalization, and augmentation, which help standardize input images and increase the diversity of training data. Augmentation techniques such as random cropping, rotation, and color jittering introduce variations into the training data, enhancing the model's ability to generalize to unseen scenarios. Post-processing steps such as non-maximum suppression (NMS) further refine object localization by eliminating redundant detections and improving localization accuracy.

Moreover, the literature underscores the significance of empirical experimentation and analysis in gaining insights into the strengths and limitations of different object detection algorithms. Through systematic evaluation on benchmark datasets and real-world scenarios, researchers assess the performance of models under varying conditions and identify areas for improvement. Comparative studies, ablation experiments, and qualitative analysis shed light on the factors influencing model performance, facilitating informed decision-making in algorithm selection and parameter tuning.Overall, the literature survey provides a comprehensive overview of the state-of-the-art in object detection using R-CNN through computer vision. By synthesizing insights from diverse research endeavours, this paper aims to contribute to the advancement of object detection methodologies and deepen our understanding of the role of deep learning technology and dataset quality in achieving accurate and reliable object detection performance. Through empirical experimentation and analysis, this study endeavours to strengthen the foundations of object detection research and pave the way for future innovations in the field.

PROPOSED SYSTEM

The proposed system for object price detection using R-CNN through computer vision represents a significant advancement in the field of computer vision, particularly in the domain of object detection tasks. This paper aims to leverage the capabilities of deep learning algorithms to accurately detect and localize objects within images, with a specific focus on identifying and extracting price information from product labels or tags. Building upon existing datasets and deep learning methodologies commonly used in computer vision research, the proposed system introduces a novel dataset tailored to the task of object price detection. This dataset is meticulously crafted to encompass a diverse range of products, pricing formats, and environmental conditions, ensuring comprehensive coverage of real-world scenarios.At the heart of the proposed system lies the utilization of the Faster R-CNN (Region-based Convolutional Neural Network) architecture, a state-of-the-art deep learning model renowned for its effectiveness in object detection tasks. Faster R-CNN consists of two main components: a region proposal network (RPN) responsible for generating region proposals, and a region-based convolutional neural network responsible for classifying and refining these proposals. By leveraging the rich hierarchical features learned by the convolutional layers, Faster R-CNN can accurately localize objects within images while simultaneously predicting their corresponding class labels, including price information in the context of this study.

The proposed system begins with the construction and annotation of the new dataset specifically tailored for object price detection. This dataset is curated to include a diverse array of product images captured under varying lighting conditions, camera angles, and environmental contexts. Each image is meticulously annotated with bounding boxes delineating the location of price tags or labels, providing ground truth annotations for training and evaluation purposes. The dataset is further augmented with variations such as rotation, scaling, and occlusion to enhance model robustness and generalization capabilities.Subsequently, the Faster R-CNN architecture is trained on the newly created dataset using a supervised learning approach. During the training process, the model learns to identify and localize price tags within images by iteratively optimizing its parameters based on a predefined objective function, typically the minimization of a loss function such as cross-entropy loss or mean squared error. The training process involves feeding batches of annotated images into the network, computing the loss between the predicted outputs and ground truth annotations, and adjusting the network parameters using backpropagation to minimize the loss.

Once trained, the Faster R-CNN model undergoes rigorous evaluation on a separate validation set to assess its performance in object price detection. Performance metrics such as precision, recall, and mean average precision (mAP) are computed to quantify the accuracy and robustness of the model across different product categories and environmental conditions. The evaluation results provide valuable insights into the strengths and limitations of the proposed system, facilitating informed decisions regarding model selection, hyperparameter tuning, and dataset refinement.Furthermore, the proposed system incorporates a comprehensive analysis of the experimental results to gain deeper insights into the importance of deep learning technology and dataset quality in object price detection tasks. By systematically analyzing the model's performance across various experimental settings and scenarios, researchers can elucidate the factors influencing detection accuracy and identify areas for improvement. This analysis sheds light on the efficacy of the Faster R-CNN architecture in handling complex pricing formats, occlusions, and environmental variations, highlighting its potential for real-world applications in retail, inventory management, and pricing optimization. Finally, the proposed system for object price detection using R-CNN through computer vision offers a robust and effective solution for automating the extraction of price information from product images. By leveraging deep learning methodologies and a meticulously curated dataset, the system achieves high accuracy and reliability in detecting and localizing price tags within images, demonstrating the efficacy of deep learning technology in addressing real-world object detection tasks. Through empirical experimentation and analysis, this study contributes to the advancement of computer vision research and underscores the importance of dataset quality and algorithmic sophistication in achieving superior performance in object detection applications.

METHODOLOGY

The methodology employed in this study revolves around the task of object price detection using R-CNN (Region-based Convolutional Neural Network) through computer vision. The overarching objective is to leverage deep learning techniques to accurately detect and extract price information from images of various products. To achieve this goal, a systematic approach is followed, encompassing dataset preparation, model selection, training, experimentation, and result analysis.The first step in the methodology involves a comprehensive review and summary of existing datasets and deep learning algorithms commonly utilized in computer vision research. This includes gathering information on publicly available datasets containing annotated images of products, as well as understanding the underlying principles and architectures of deep learning algorithms suitable for object detection tasks. By analyzing the characteristics and limitations of existing datasets and algorithms, researchers gain valuable insights into the state-of-the-art methodologies and potential areas for improvement.

Subsequently, a new dataset tailored specifically for object price detection is constructed based on the insights gleaned from the literature review. This new dataset is designed to encompass a diverse range of products, pricing formats, and environmental conditions to ensure comprehensive coverage of real-world scenarios. Each image in the dataset is meticulously annotated with bounding boxes delineating the location of price tags or labels, providing ground truth annotations for training and evaluation purposes. Additionally, the dataset is augmented with variations such as rotation, scaling, and occlusion to enhance model robustness and generalization capabilities.Following dataset preparation, the next step involves selecting an appropriate deep learning model for object price detection. In this study, the Faster R-CNN architecture is chosen as the network of choice due to its effectiveness in object detection tasks. Faster R-CNN consists of two main components: a region proposal network (RPN) responsible for generating region proposals, and a region-based convolutional neural network responsible for classifying and refining these proposals. By leveraging the hierarchical features learned by the convolutional layers, Faster R-CNN can accurately localize objects within images while simultaneously predicting their corresponding class labels.

Once the Faster R-CNN architecture is selected, the model is trained on the newly created dataset using a supervised learning approach. During the training process, the model learns to identify and localize price tags within images by iteratively optimizing its parameters based on a predefined objective function, typically the minimization of a loss function such as cross-entropy loss or mean squared error. The training process involves feeding batches of annotated images into the network, computing the loss between the predicted outputs and ground truth annotations, and adjusting the network parameters using backpropagation to minimize the loss.After training, the performance of the trained Faster R-CNN model is evaluated on a separate validation set to assess its effectiveness in object price detection. Performance metrics such as precision, recall, and mean average precision (mAP) are computed to quantify the accuracy and robustness of the model across different product categories and environmental conditions. The evaluation results provide valuable insights into the strengths and limitations of the proposed system, facilitating informed decisions regarding model selection, hyperparameter tuning, and dataset refinement.

Finally, the experimental results are systematically analyzed to gain deeper insights into the importance of deep learning technology and dataset quality in object price detection tasks. By scrutinizing the model's performance across various experimental settings and scenarios, researchers can elucidate the factors influencing detection accuracy and identify areas for improvement. This analysis not only reinforces the significance of deep learning methodologies in computer vision but also underscores the critical role of high-quality datasets in achieving superior performance in real-world applications. Through this comprehensive methodology, researchers aim to advance the state-of-the-art in object price detection and contribute to the broader field of computer vision research.

RESULTS AND DISCUSSION

The results obtained from the experiments conducted in this study provide valuable insights into the efficacy of using R-CNN (Region-based Convolutional Neural Network) for object price detection in the domain of computer vision. Leveraging a new dataset constructed based on commonly used datasets, the Faster R-CNN architecture was employed to detect and localize price tags within images of various products. Through rigorous experimentation, the performance of the model was evaluated across different product categories and environmental conditions, shedding light on the strengths and limitations of the proposed approach.

The experimental results demonstrate the effectiveness of the Faster R-CNN model in accurately detecting and localizing price tags within images. Across a diverse range of product categories and pricing formats, the model achieved high precision and recall rates, indicating its robustness and generalization capabilities. The analysis of detection performance revealed that the model excelled in identifying price tags of varying sizes, orientations, and locations within images, showcasing its versatility and adaptability to real-world scenarios. Furthermore, the model exhibited resilience to environmental factors such as lighting conditions, occlusions, and background clutter, highlighting its efficacy in challenging operating environments. These results underscore the potential of R-CNN-based approaches for object price detection tasks and reinforce the importance of leveraging deep learning technologies for addressing complex computer vision challenges.

Moreover, the experimentation process provided valuable insights into the importance of dataset quality and diversity in deep learning-based object detection tasks. By constructing a new dataset tailored specifically for object price detection, researchers were able to ensure comprehensive coverage of real-world scenarios and pricing formats. The analysis of dataset characteristics revealed that the diversity of products, pricing formats, and environmental conditions significantly influenced the model's performance. In particular, the presence of diverse pricing formats, including labeled prices, handwritten prices, and digital displays, posed unique challenges for the model, necessitating robust feature representation and learning capabilities. Additionally, augmenting the dataset with variations such as rotation, scaling, and occlusion proved instrumental in enhancing the model's resilience to variations in image composition and scene complexity. These findings underscore the critical role of high-quality datasets in facilitating effective deep learning-based object detection and highlight the importance of dataset curation and augmentation in improving model performance.

International Journal of
Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 2, April 2024

Fig 1. Results screenshot 1

Fig 2. Results screenshot 2

Furthermore, the analysis of experimental results provided insights into the broader implications of deep learning technology in computer vision applications. By systematically evaluating the performance of the Faster R-CNN model across different experimental settings and scenarios, researchers gained a deeper understanding of the capabilities and limitations of state-of-the-art object detection techniques. The analysis revealed that while deep learning-based approaches such as R-CNN offer unprecedented accuracy and efficiency in object detection tasks, they are also subject to certain challenges, including the need for large-scale annotated datasets, computational resources, and model optimization techniques. Moreover, the analysis highlighted the importance of continued research and development efforts in advancing deep learning methodologies to address evolving challenges in computer vision. Overall, the results and discussions underscore the transformative potential of deep learning technology in revolutionizing object detection and recognition tasks and pave the way for future advancements in the field of computer vision.

CONCLUSION

Due to its powerful learning ability and advantages in dealing with occlusion, scale transformation and background switches, deep learning-based object detection has been a research hotspot in recent years. This paper provides a detailed review on deep learning-based object detection frameworks which handle different sub-problems, such as occlusion, clutter and low resolution, with different degrees of modifications on R-CNN. The review starts on generic object detection pipelines which provide base architectures for other related tasks. Then, three other common tasks, namely salient object detection, face detection and pedestrian detection, are also briefly reviewed. Finally, we propose several promising future directions to gain a thorough understanding of the object detection landscape. This review is also meaningful for the developments in neural networks and related learning systems, which provides valuable insights and guidelines for future progress.

REFERENCES

1. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

2. He, K., Gkioxari, G., Dollár, P., &Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

3. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125).

4. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).

5. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., & Reed, S. (2016). SSD: Single shot multibox detector. In European conference on computer vision (pp. 21-37).

ISSN 2347–3657

Volume 12, Issue 2, April 2024

6. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.

7. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255).

8. Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The Pascal visual object classes (VOC) challenge. International journal of computer vision, 88(2), 303-338.

9. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). ImageNet large scale visual recognition challenge. International journal of computer vision, 115(3), 211-252.

10. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... &Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In European conference on computer vision (pp. 740-755).

11. Everingham, M., Eslami, S. M. A., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2015). The Pascal visual object classes challenge: A retrospective. International journal of computer vision, 111(1), 98-136.

12. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 936-944).

13. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

14. Zhang, H., Dana, K., Shi, J., Zhang, Z., Wang, X., Tyagi, A., & Agrawal, A. (2018). Context encoding for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7151-7160).

15. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.