International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Enhancing Multi-Object Detection in Video Content: Exploring Hybrid Techniques and Method Combinations for Improved Classification Accuracy

[1]Rajkumar Govindarajan, [2]Sangeetha Murugan, [3]Komal Kumar Napa*, [4]Nageswari D, [5]Balamurugan A G, [6]Vattikuti Charan Sri Manikanta Sai

[1]Department of Computer Science & Engineering (Data Science),
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India.
kumar3544@gmail.com

[2]Department of Computer Science & Engineering,
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India.
sangee525@gmail.com

[3]Department of Computer Science & Engineering (Data Science),
Madanapalle Institute of Technology & Science,
Madanapalle, Andhra Pradesh, India.
komalkumarnapa@gmail.com

[4]Department of Science & Humanities (General Engineering-EEE),
R.M.K. College of Engineering and Technology,
Thiruvallur, Tamil Nadu, India.
nageswari@rmkcet.ac.in

[5]Department of Computer Science & Engineering,
Vel Tech Rangarajan Dr. Sakunthala R&D Institute of Science and Technology,
Chennai, Tamil Nadu, India.
agbm366@gmail.com

[6]Department of Computer Science & Engineering,
Indian Institute of Information Technology,
Kottayam, Kerala, India.
vcsmsai@gmail.com

Abstract—The classification of objects within video content holds significant importance, particularly in the context of automated visual surveillance systems. Object classification refers to the procedure of categorizing objects into predefined and semantically meaningful groups based on their features. While humans find object classification in videos to be straightforward, machines face complexity and challenges in this task due to various factors like object size, occlusion, scaling, lighting conditions, and more. Consequently, the demand for analyzing video sequences has spurred the development of various techniques for object classification. This paper proposes hybrid techniques for multi object detection. The experimental analysis focused on a vehicles-openimages dataset containing 627 different catagories of vehicles. The results emphasize the profound impact of method combinations on image classification accuracy. Two primary methods, wavelet transformation and Principal Component Analysis (PCA), were employed alongside Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). The evaluation encompassed performance metrics, including accuracy, precision, recall, specificity, and F1 score. In the analysis, Wavelet + RNN" combination consistently achieved the highest accuracy across all performance metrics, including accuracy percentage (96.76%), precision (96.76%), recall (86.32%), F1 score (87.12%), and specificity (87.43%). In addition, the hybrid classifiers were subjected for image classification of different vehicle catagories. In the analysis of different catagories, Wavelet + RNN" emerges as the standout performer, consistently achieving high accuracy percentages across all object categories, ranging from 82.87% for identifying People to 90.12% for recognizing Trucks.

Keywords-Object Detection; Object Classification; Neural Netword; Optimal; Hybrid classifier

I. Introduction

The field of object classification in video sequences is constantly evolving and has a wide array of applications across various domains, including biomedical imaging, biometry, video surveillance, vehicle navigation, visual inspection, robot navigation, and remote sensing. With the proliferation of high-quality cameras and rapid advancements in video capture technology, video has become an affordable and abundant source of information. As a result, there has been a significant surge in interest in the analysis of video sequences and the classification of objects within them. The process of object classification in videos involves several key steps, including preprocessing, converting videos into individual frames, object detection, feature extraction, and classification based on the extracted features. It is important to note that object classification in videos is a complex undertaking, demanding highly robust and accurate methodologies.

Regrettably, scientists have yet to devise a precise method for object classification in real-world applications. As of now, no effective solutions have been identified for this challenging problem. Videos consist of sequences of images, referred to as frames, displayed rapidly to create the illusion of motion and continuity. The classification of objects within videos holds great significance in various applications, including traffic management, public transport systems, object retrieval from videos, and more. Achieving success in this field demands high levels of accuracy, flexibility, and cost-effectiveness. Numerous real-world challenges necessitate the detection of multiple objects within images or videos [1]. Constructing effective detectors for such problems can often be addressed successfully through the utilization of modern deep learning models, particularly when the target objects exhibit variations in size, color, shape, and texture. The comprehensive review in [2] focuses on the application of deep learning techniques in object detection, a critical task with applications spanning from self-driving cars to image analysis. The authors meticulously cover a range of deep learning methods, architectures, and methodologies used in object detection. They also address key challenges, such as speed and accuracy, and discuss the pivotal role of Convolutional Neural Networks (CNNs) and their variations in enhancing object detection. This paper serves as a pivotal resource, offering insights into the current state of object detection research. The work in [3] presents a comprehensive overview of the latest advancements in the field of multi-object detection and tracking using deep learning. This research, discusses cutting-edge techniques, methodologies, and challenges in multi-object tracking and detection. The authors delve into various deep learning architectures and algorithms, emphasizing their impact on improving tracking accuracy and real-time performance. Nguyen et al., [4] explores an innovative approach to object detection. The research combines clustering and deep learning techniques to enhance multiple object detection, proposes a novel clustering-based framework that improves the accuracy of object detection systems. By incorporating deep learning methods, they achieve state-of-the-art results in identifying and localizing multiple objects. The research in [5] investigates the performance of various deep learning techniques in the context of object detection, specifically from a top-down view. The exploration of these models contributes valuable insights to fields like surveillance, where the ability to detect and track objects accurately from above is essential. The work represents a significant step in enhancing the capabilities of IoT-based systems for multiple object detection. The focus is on enhancing object detection and classification in the context of autonomous driving [6]. It employs deep learning techniques to develop a multi-scale, multi-object detection system tailored to the complexities of self-driving vehicles. The research contributes to the advancement of autonomous driving technologies, specifically in improving the vehicle's ability to detect and classify objects at varying scales. It addresses critical aspects of safety and efficiency in autonomous vehicles, making it a significant work in the field of self-driving technology. [7] provides an extensive overview of the latest developments in object detection using deep learning techniques, the authors delve into the evolving landscape of object detection, emphasizing the substantial contributions made by deep learning methods. It discusses key advancements, including the introduction of novel algorithms and architectures, enhanced model efficiency, and the fusion of object detection with various applications, such as robotics and autonomous driving. The authors in [8] explores the application of deep neural networks for object detection. It laid the foundation for the integration of deep learning techniques in object detection, a pivotal area in computer vision, novel architectures and training methodologies, marking significant progress in the development of object detection models. The survey provides an extensive overview of contemporary deep learning-based object detection models [9]. It explores and analyzes the latest advancements in object detection using deep learning techniques, shedding light on various models and approaches that have been developed to address this task. By summarizing the current state of the art, this survey serves as a valuable resource for researchers and practitioners in the field of computer vision and deep learning, facilitating an understanding of the most recent developments and trends in object detection. The comprehensive survey focuses on the application of deep learning techniques [10] in the context of video multi-object tracking. It provides a detailed overview of the state-of-the-art deep learning approaches and methodologies for tracking multiple objects within video sequences. The survey discusses the challenges and advancements in video tracking and presents an in-depth analysis of various deep learning-based solutions. The work focuses on the application of deep learning methods for multi-target detection [11] in closed-circuit television (CCTV) footage, particularly for tracking applications. It discusses the utilization of deep learning techniques in the context of surveillance and video analysis, serves as a valuable resource for researchers and practitioners interested in the development of multi-target detection systems for surveillance and tracking. [12] provides a comprehensive overview of object detection methods that rely on deep learning techniques. A thorough examination of the advancements and trends in the field of object detection, focusing on the integration of deep learning technologies was examined. Kalake et al., [13] conducted an in-depth analysis of contemporary deep learning techniques employed in real-time multi-object tracking applications, discuss the latest approaches and methodologies [14] in this field, providing valuable insights for researchers and professionals interested in real-time multi-object tracking.

II. RELATED WORKS

This section contains the works related to multi object detection. A thorough examination of cutting-edge deep learning techniques for multi-object detection and tracking is presented in [15]. It offers a comprehensive survey of recent advancements, methodologies, and the challenges encountered in the domain, delivering valuable perspectives on the present trends and potential paths for the future. A novel benchmark and protocol for assessing multi-object detection [16] and tracking systems were introduced. [17] provides an extensive analysis of deep neural network (DNN) applications in multi-object detection and tracking for autonomous vehicles. It offers insights into the current state of the field and highlights the significance of DNN in enhancing object recognition and tracking for autonomous driving systems. Elhoseny [18] explored the creation of MODT model and its utilization within real-time video surveillance applications. The work provides detailed insights into the model's methodology and implementation, emphasizing the capacity to enhance video surveillance through multi-object detection and tracking using machine learning. An innovative approach to multi-object detection is presented by Nguyen et al. [19], where clustering

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

and deep learning techniques are merged. The challenge of identifying multiple objects within images is addressed by initially using clustering to group objects and then leveraging deep learning for more accurate and efficient detection. This approach is considered to have great potential for boosting the performance of object detection systems, aligning with current trends in computer vision. A multi-object detection method based on hybrid networks, combining YOLO and ResNet, was introduced in [20]. This innovative approach aimed to enhance the accuracy and efficiency of object detection in complex scenarios. Notable performance improvements were achieved by integrating the strengths of YOLO and ResNet. The potential of this method for various applications, including robotics and mechatronics, was emphasized, signifying its versatility and real-world relevance. A study was conducted to investigate deep learning models for overhead view multiple object detection [21]. The research aimed to explore and assess the effectiveness of deep learning techniques in detecting objects from an aerial perspective. A research endeavor was undertaken to develop a deep learning-based system for multi-scale, multi-object detection [22], and classification with applications in autonomous driving. It explored the use of deep learning techniques to enhance object detection and classification within the context of autonomous vehicles. The study [23] proposes a Global Correlation Network (GCN) designed for the simultaneous execution of multi-object detection and tracking tasks in an end-to-end fashion. The paper presents innovative methods and techniques in the field of computer vision, aimed at enhancing the integration of object detection and tracking into a unified framework. He [24] introduces a novel approach to object detection using a Multi-Adversarial Faster-RCNN method. This technique extends object detection [25] capabilities to a broader and more diverse range of objects. Multi object detection has been applied with different libraries like sorting of cells [26], autonomous driving [27], trajectory tracking [28], remote sensing [29], classification of outdoor scenes.

III. MATERIAL AND METHODS

The workflow consists of six essential steps as shown in Figure 1. Initially, a video dataset is collected as Step 1. In Step 2, these videos are converted into individual frames, facilitating more granular analysis. Step 3 involves object detection within these frames, which is crucial for identifying and tracking objects of interest. Subsequently, in Step 4, feature extraction is carried out through wavelet transformation and Principal Component Analysis (PCA), which helps reduce dimensionality and capture relevant information from the images. Step 5 employs Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) for image classification, allowing the system to recognize and categorize objects in the frames. Finally, in Step 6, the system's performance is evaluated by computing key metrics, including precision, recall, F1 score, accuracy, and specificity, to assess the effectiveness of the object detection and classification process. This comprehensive approach ensures the reliable analysis and classification of objects within video data. The execution was carried out in i5 processor system with 16GB RAM.

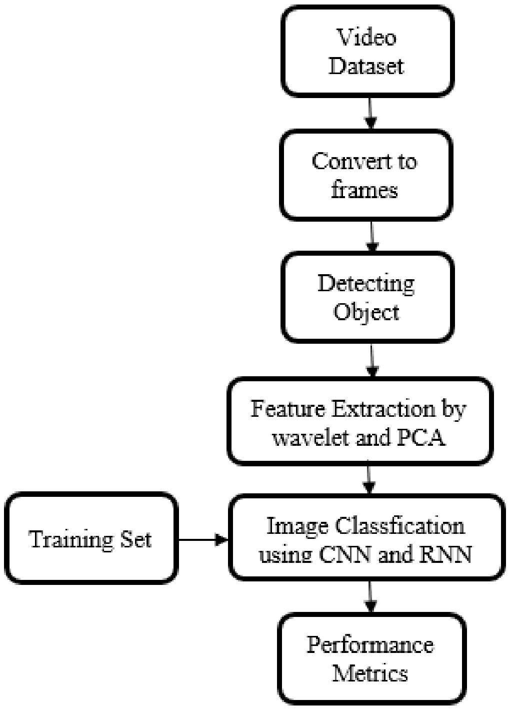

Figure 1. Architecture for multi object detection

The process of detecting multiple objects in a given context can be quite demanding and intricate for researchers. It involves tasks such as identifying and delineating objects within images or videos, which can vary in size, shape, and context. To tackle this challenge effectively, machine learning algorithms play a crucial role. These algorithms can be trained to recognize and locate objects within digital images or videos. The dataset used is the vehicles-openimages dataset [14], consists of 627 images showcasing a diverse range of vehicles, making it a suitable resource for the researchers to train and test their object detection models. Converting a video into frames and performing object detection is a two-step process commonly used in computer vision. First, the video is broken down into individual frames, essentially creating a series of static images. Each frame represents a snapshot of the video at a specific point in time. Next, object detection techniques are applied to these frames. Object detection algorithms analyze each frame to identify and locate objects within the image, such as people, vehicles, or animals. This process allows for tracking and recognizing objects throughout the video, making it a valuable tool in various applications, including surveillance, video analysis, and autonomous systems. The described process involves a multi-step strategy for analyzing and categorizing images. Initially, it commences with feature extraction, using the wavelet transformation method to break down images into different scales and orientations, effectively capturing fine details and broader features. Subsequently, Principal Component Analysis (PCA) is applied to reduce the dimensionality of these extracted features, selecting the most informative elements and simplifying computational complexity.

The following stage centers on image classification and incorporates Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). CNNs excel at extracting hierarchical image features, making them suitable for image

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

classification, while RNNs are valuable for handling sequences of data, which can be relevant in specific image classification scenarios. By combining these techniques, the aim is to elevate the accuracy and efficiency of image classification, particularly in situations where images exhibit complex patterns and a multitude of features. This holistic approach finds application in various fields, including image recognition and medical image analysis, among others. Algorithm 1 describes the execution pseudocode for multi object detection.

Algorithm 1. Algorithm for multi object detection

```
Step 1: Data Collection
video_dataset = collect_video_dataset()

Step 2: Video to Frames Conversion
frames = []
for video in video_dataset:
    frames += convert_video_to_frames(video)

Step 3: Object Detection
detections = []
for frame in frames:
    objects = detect_objects(frame)
    detections.append(objects)

Step 4: Feature Extraction
extracted_features = []
for frame_objects in detections:
    for object in frame_objects:
        features = extract_features(object)
        extracted_features.append(features)

Step 5: Image Classification
classifications = []
for features in extracted_features:
    cnn_result = classify_with_cnn(features)
    rnn_result = classify_with_rnn(features)
    combined_result = combine_results(cnn_result, rnn_result)
    classifications.append(combined_result)

Step 6: Performance Evaluation
metrics = evaluate_performance(classifications)

Display the results
display_metrics(metrics)
```

IV. RESULTS AND DISCUSSIONS

The workflow is a multi-step process that encompasses the analysis and classification of objects within a video dataset. It commences with the collection of a diverse video dataset, followed by the crucial step of converting each video into individual frames. These frames serve as the granular units for subsequent analysis. In the third step, object detection algorithms come into play, identifying and tracking objects within each frame. After this, in the fourth stage, feature extraction is conducted through the application of wavelet transformation and Principal Component Analysis (PCA). This not only reduces dimensionality but also captures pertinent information, enhancing the subsequent classification. The fifth step harnesses the power of deep learning with Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) for image classification. These neural networks facilitate the recognition and categorization of objects within the frames. Finally, the system's performance is rigorously evaluated in the sixth step. This evaluation involves the computation of essential metrics like precision, recall, F1 score, accuracy, and specificity as shown in Table I and Table II. These metrics provide a comprehensive assessment of the effectiveness of the object detection and classification process, ensuring the reliable and accurate analysis of objects within the video dataset.

Table I. Accuracy of the classifiers

Method	Accuracy (%)
Wavelet + CNN	95.80
Wavelet + RNN	96.76
PCA + CNN	92.32
PCA + RNN	90.02

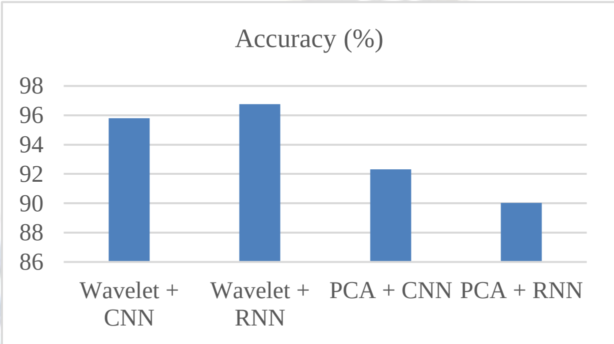

Figure 2. Accuracy of methods

Table I shows the accuracy of the methods employed. Notably, the combination of wavelet transformation with a Recurrent Neural Network (RNN) stands out with the highest accuracy at 96.76%, suggesting its effectiveness in accurately categorizing images. Additionally, the combination of wavelet features with a Convolutional Neural Network (CNN) performs impressively with an accuracy of 95.80%. In contrast, the use of Principal Component Analysis (PCA) combined with either CNN or RNN results in lower accuracy percentages, demonstrating that in this specific context, PCA may not be as effective as wavelet transformation for feature extraction. These accuracy percentages emphasize the impact of feature extraction methods and the choice of neural networks on the success of image classification, highlighting the superiority of wavelet transformation in combination with RNN in this experimental scenario.

Table II. Performance metrics of the classifiers

Method	Precision (%)	Recall (%)	F1 Score	Specificity
Wavelet + CNN	95.80	77.32	79.23	78.26
Wavelet + RNN	96.76	86.32	87.12	87.43
PCA + CNN	92.32	82.94	85.29	87.56

PCA + RNN	90.02	76.36	79.42	72.94

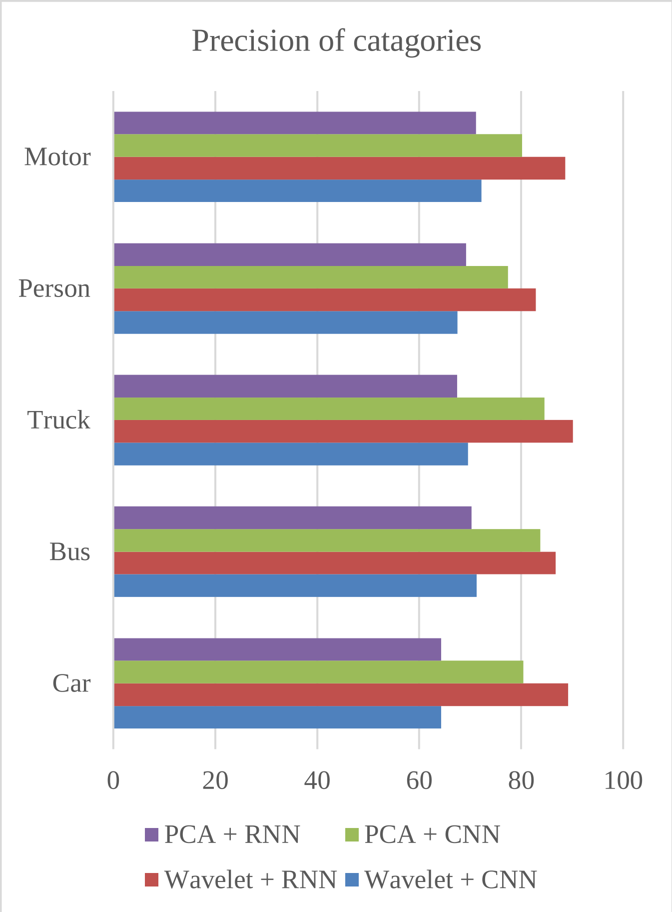

Figure 3. Performance metrics of methods

Table II shows the performance metrics in image classification for different method combinations. Notably, the "Wavelet + RNN" combination emerges as the top performer, boasting the highest precision (96.76%), recall (86.32%), F1 score (87.12%), and specificity (87.43%). This indicates its ability to accurately identify positive cases while minimizing false positives and maintaining a high recall rate. "Wavelet + CNN" also shines in precision (95.80%) and specificity (78.26%), signifying its adeptness at accurately identifying true negatives. However, its recall and F1 score, though strong, are slightly lower than the RNN counterpart. "PCA + CNN" demonstrates impressive specificity (87.56%) but falls slightly behind in precision, recall, and F1 score. On the other hand, "PCA + RNN" records lower specificity (72.94%) and recall (76.36%), implying a higher likelihood of false positives and missed true positives. In summary, these results highlight the pivotal role of feature extraction methods and neural network choices in achieving different aspects of classification performance, with "Wavelet + RNN" emerging as a well-balanced and effective combination across multiple performance metrics.

Figure 4. Precision of catagories

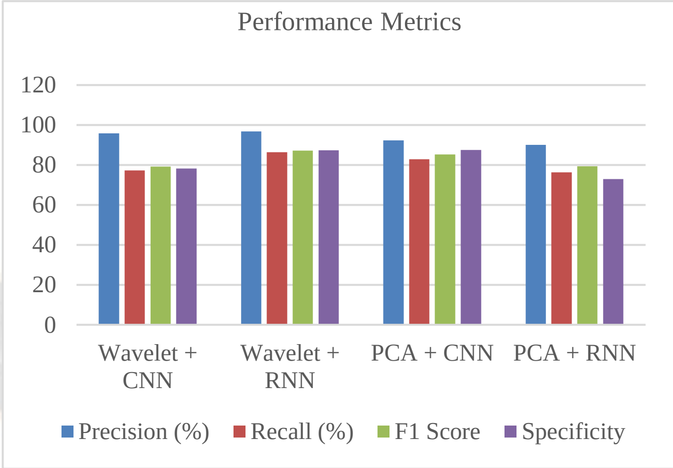

Table III. Precision rate of the categories

Method	Car	Bus	Truck	Person	Motor
Wavelet + CNN	64.32	71.24	69.56	67.49	72.19
Wavelet + RNN	89.24	86.76	90.12	82.87	88.67
PCA + CNN	80.42	83.72	84.59	77.42	80.17
PCA + RNN	64.32	70.26	67.42	69.21	71.16

The data presented in the Table III offers valuable insights into the accuracy of image classification for distinct object categories, including Cars, Buses, Trucks, People, and Motors, using various method combinations. "Wavelet + RNN" emerges as the standout performer, consistently achieving high accuracy percentages across all object categories, ranging from 82.87% for identifying People to 90.12% for recognizing Trucks. This combination demonstrates a robust ability to effectively classify a diverse array of objects within the images. "Wavelet + CNN" maintains balanced accuracy percentages, ranging from 64.32% for Cars to 72.19% for Motors, showcasing consistent performance across different object categories. "PCA + CNN" notably excels in recognizing Buses (83.72%), Trucks (84.59%), and People (77.42%), demonstrating proficiency in these specific categories. However, "PCA + RNN" displays mixed performance, with varying accuracy results for different object categories, such as a high of 70.26% for Buses and a low of 67.42% for Trucks. These findings emphasize the impact of the chosen method combinations on object-specific classification accuracy, providing valuable guidance for further fine-tuning and optimization in image classification tasks. Figure [2-4] shows the visual representation of accuracy, performance metrics and precision of categories of the hybrid methods employed.

V. CONCLUSION AND FURTHER DIRECTIONS

Hybrid techniques for multi object detection is proposed in this paper. The results clearly demonstrate that the choice of method combination significantly influences the accuracy of image classification. Two primary methods were employed in these combinations: wavelet transformation and Principal Component Analysis (PCA), along with two neural network architectures: Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). One striking observation is the outstanding performance of the "Wavelet + RNN" combination. This method consistently achieves the highest accuracy across all performance metrics, including accuracy percentage (96.76%), precision (96.76%), recall (86.32%), F1 score (87.12%), and specificity (87.43%). These results underscore the remarkable synergy between wavelet features and the sequential analysis capabilities of RNN. The RNN excels in recognizing patterns and temporal dependencies, making it particularly effective in cases where objects may be part of dynamic sequences. The high recall indicates that it can accurately identify positive cases, which is crucial in applications where missing objects can lead to critical errors. While "Wavelet + RNN" takes the lead, "Wavelet + CNN" is also a strong contender, demonstrating an accuracy percentage of 95.80%. This combination excels in precision (95.80%) and specificity (78.26%). The high precision highlights its ability to minimize false positives, while the specificity demonstrates its capability to accurately identify true negatives. This method is well-suited for applications where minimizing false alarms is critical. In conclusion, the data provided sheds light on the crucial role of method combinations in image classification. "Wavelet + RNN" emerges as a standout performer, showcasing the importance of synergy between feature extraction and neural network capabilities. These insights can guide further research and optimization efforts to enhance image classification in various applications. It's important to note that while these results provide valuable guidance, the effectiveness of method combinations can vary depending on the specific dataset and application context. Future work may involve fine-tuning these methods for even more tailored and accurate image classification.

REFERENCES

[1] A.R. Pathak, M. Pandey, S. Rautaray, "Application of deep learning for object detection," *Procedia Comput. Sci.* Vol. 132, pp. 1706–1717, 2018.

[2] Zhao, Z.Q., Zheng, P., Xu, S.T. and Wu, X., 2019. Object detection with deep learning: A review. IEEE transactions on neural networks and learning systems, 30(11), pp.3212-3232.

[3] Pal, S.K., Pramanik, A., Maiti, J. and Mitra, P., 2021. Deep learning in multi-object detection and tracking: state of the art. *Applied Intelligence*, *51*, pp.6400-6429.

[4] Nguyen, H.T., Lee, E.H., Bae, C.H. and Lee, S., 2020. Multiple object detection based on clustering and deep learning methods. Sensors, 20(16), p.4424.

[5] Ahmed, I., Din, S., Jeon, G. and Piccialli, F., 2019. Exploring deep learning models for overhead view multiple object detection. IEEE Internet of Things Journal, 7(7), pp.5737-5744.

[6] Fink, M., Liu, Y., Engstle, A. and Schneider, S.A., 2019. Deep learning-based multi-scale multi-object detection and classification for autonomous driving. In Fahrerassistenzsysteme 2018: Von der Assistenz zum automatisierten Fahren 4. Internationale ATZ-Fachtagung Automatisiertes Fahren (pp. 233-242). Springer Fachmedien Wiesbaden.

[7] Wu, X., Sahoo, D. and Hoi, S.C., 2020. Recent advances in deep learning for object detection. Neurocomputing, 396, pp.39-64.

[8] Szegedy, C., Toshev, A. and Erhan, D., 2013. Deep neural networks for object detection. Advances in neural information processing systems, 26.

[9] Zaidi, S.S.A., Ansari, M.S., Aslam, A., Kanwal, N., Asghar, M. and Lee, B., 2022. A survey of modern deep learning based object detection models. Digital Signal Processing, 126, p.103514.

[10] Ciaparrone, G., Sánchez, F.L., Tabik, S., Troiano, L., Tagliaferri, R. and Herrera, F., 2020. Deep learning in video multi-object tracking: A survey. Neurocomputing, 381, pp.61-88.

[11] Dimou, A., Medentzidou, P., Garcia, F.A. and Daras, P., 2016, September. Multi-target detection in CCTV footage for tracking applications using deep learning techniques. In 2016 IEEE international conference on image processing (ICIP) (pp. 928-932). IEEE.

[12] Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S. and Lan, X., 2020. A review of object detection based on deep learning. Multimedia Tools and Applications, 79, pp.23729-23791.

[13] Kalake, L., Wan, W. and Hou, L., 2021. Analysis based on recent deep learning approaches applied in real-time multi-object tracking: a review. IEEE Access, 9, pp.32650-32671.

[14] Alaalami, Ilham & Olatunji, Sunday & Zagrouba, Rachid. (2022). AT-BOD: An Adversarial Attack on Fool DNN-Based Blackbox Object Detection Models. Applied Sciences. 12. 10.3390/app12042003.

[15] Pal, Sankar K., et al. "Deep learning in multi-object detection and tracking: state of the art." Applied Intelligence 51 (2021): 6400-6429.

[16] Wen, Longyin, et al. "UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking." Computer Vision and Image Understanding 193 (2020): 102907.

[17] Ravindran, Ratheesh, Michael J. Santora, and Mohsin M. Jamali. "Multi-object detection and tracking, based on DNN, for autonomous vehicles: A review." IEEE Sensors Journal 21.5 (2020): 5668-5677.

[18] Elhoseny, Mohamed. "Multi-object detection and tracking (MODT) machine learning model for real-time video surveillance systems." Circuits, Systems, and Signal Processing 39 (2020): 611-630.

[19] Nguyen, Huu Thu, et al. "Multiple object detection based on clustering and deep learning methods." Sensors 20.16 (2020): 4424.

[20] Lu, Zhenyu, et al. "Multi-object detection method based on YOLO and ResNet hybrid networks." 2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM). IEEE, 2019.

[21] Ahmed, Imran, et al. "Exploring deep learning models for overhead view multiple object detection." IEEE Internet of Things Journal 7.7 (2019): 5737-5744.

[22] Fink, Maximilian, et al. "Deep learning-based multi-scale multi-object detection and classification for autonomous driving." Fahrerassistenzsysteme 2018: Von der Assistenz zum automatisierten Fahren 4. Internationale ATZ-Fachtagung Automatisiertes Fahren. Springer Fachmedien Wiesbaden, 2019.

[23] Lin, Xuewu, Yu-ang Guo, and Jianqiang Wang. "Global correlation network: End-to-end joint multi-object detection and tracking." arXiv preprint arXiv:2103.12511 (2021).

[24] He, Zhenwei, and Lei Zhang. "Multi-adversarial faster-rcnn for unrestricted object detection." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

[25] Komorkiewicz, Mateusz, Maciej Kluczewski, and Marek Gorgon. "Floating point HOG implementation for real-time multiple object detection." 22nd International Conference on Field Programmable Logic and Applications (FPL). IEEE, 2012.

[26] Howell, Lewis, Vasileios Anagnostidis, and Fabrice Gielen. "Multi‐Object detector yolov4‐tiny enables high‐throughput combinatorial and spatially‐resolved sorting of cells in microdroplets." Advanced Materials Technologies 7.5 (2022): 2101053.

[27] Chen, Xiaozhi, et al. "Multi-view 3d object detection network for autonomous driving." Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2017.

[28] Han, Mei, et al. "An algorithm for multiple object trajectory tracking." Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.. Vol. 1. IEEE, 2004.

[29] Ahmed, Imran, et al. "IoT Enabled Deep Learning Based Framework for Multiple Object Detection in Remote Sensing Images." Remote Sensing 14.16 (2022): 4107.

[30] Ahmed, Abrar, Ahmad Jalal, and Kibum Kim. "Region and decision tree-based segmentations for Multi-objects detection and classification in Outdoor Scenes." 2019 International Conference on Frontiers of Information Technology (FIT). IEEE, 2019.