Babylonian Journal of Machine Learning
Vol.2024, **pp.** 157–165
DOI: https://doi.org/10.58496/BJML/2024/016; ISSN: 3006–5429
https://mesopotamian.press/journals/index.php/BJML

Review Article

Object Detection Using Capsule Neural Network: An Overview

Eman Turki Mahdi[1,*], Shokhan M. Al-Barzinji [1], Waleed Kareem Awad [2]

[1] *Department of computer networks system, College of computer science and information technology, University of Anbar, Ramadi, Iraq.*

[2] *Department of computer science, College of computer science and information technology, University of Anbar, Ramadi, Iraq.*

ARTICLE INFO

Article History
Received 08 Jul 2024
Revised: 02 Aug 2024
Accepted 10 Sep 2024
Published 02 Oct 2024

Keywords

Capsule neural network
conventional neural network
CNN
object detection

ABSTRACT

Appropriate remedies for tasks like language translation, object identification, object segmentation, picture recognition, and natural language processing are needed for today's computer vision tasks. This paper explores the use of Capsule neural Networks (Caps Nets) in object identification and offers a thorough analysis of their developments and uses. It presents an analysis of the transformative impact of Caps Nets on object identification tasks. This distinctive the architecture of neural network is presented as an alternative to convolutional neural networks (CNNs). The paper highlights that how Caps Nets are unique in capturing spatial correlations and hierarchical patterns in visualization data by analyzing the fundamental ideas of the technology. Additionally, it demonstrates how Caps net outperform CNNs in terms of generalization, interpretability, and last in the resistance to spatial distortions, confirming their goodness at object detection. By integrating the Caps networks with the latest scientific findings and advances, the paper shows the current status and potential future paths for object detection methods that use this leading neural network architecture.

1. INTRODUCTION

1.1 Deep learning overview

Over the past decade, computer vision has significantly advanced, primarily due to Deep Learning (DL) [1]. The span military and security sectors of computer vision applications, where it is essential to analyses thousands of video footages, to real-time product flaw identification in manufacturing. These applications, computer vision applications, critical and demand real-time solutions, presenting challenges beyond human capability. The superiority of deep Convolutional Neural Networks (CNNs) and Neural Networks (NNs) in the field of Computer Vision has been made possible by the abundance of available data [2]. CNNs have been a major factor in the incredible achievement that has been observed in this field [1]. Plant disease identification [3], facial (expression) recognition [4], picture processing [5] and gait recognition [6] are only a few of the many applications for which they have been effectively applied. CNNs are good at extracting features, though they are not good at capturing the spatial relationships between features, and they frequently need a large amount of data to train [7]. Hinton [8] suggested a new neural network architecture in response to the shortcomings of convolutional neural networks, pointing out that the pooling operation used to minimize the size and computational load of CNNs was the main cause of their limited functionality. To solve this problem, Capsule neural Networks and dynamic routing algorithms were introduced. [7].

1.2 Capsule neural network overview

Translational equivariance, as opposed to translational invariance as shown in CNNs, is a characteristic of capsule neural networks. Capsule neural Networks can achieve higher degrees of generalization thanks to this special feature. An object or an object part's instantiation parameters are represented by the activity vector of a capsule neural, which is a collection of neurons [6]. Capsule neural networks characterize features as vectors as opposed to scalars, as is the case with standard networks. One-level active capsule use transformation matrices to anticipate the instantiation parameters of higher-level capsule neural network. A higher-level capsule neural activates when several predictions line up [8] Current research on Caps Nets aims to propose more efficient routing algorithms in order to comprehend the contributions of these algorithms [9]. Caps Nets have gained significant attention recently due to their ability to learn more human-aligned visual

**Corresponding author*. maymoonat@uoanbar.edu.iq.

representations. The disentangled illustrations captured by Caps Nets often match the human-understandable appearances of input objects, such as rotations and translations.

However, a little research has been done on explaining the various Caps Net classifications. Recent saliency techniques, as Grad CA[10], are mainly developed for CNNs and combine values of activation and received gradients in certain layers, usually deep convolutional layers. Applying these saliency algorithms becomes difficult with Caps Nets, where deep convolutional layers are replaced by an iterative routing mechanism. Furthermore, the routing technique makes it more difficult to identify interpretable input features that are pertinent to a classification module, which can work in an end-to-end manner with a variety of neural network topologies. Simplifying techniques like top-K pooling [11] and self-attention pooling[12] have been suggested as solutions .

2. CAPSULE NEURAL NETWORKS FUNDAMENTAL AND PRINCIPLES

The principles of Caps Nets revolve around their unique structure and function, which differ from traditional convolutional neural networks (CNNs). The basic principles of capsule neural networks are as follows:

1. **Hierarchical Structure:** Information is transferred between layers of capsule neural network in a hierarchical structure known as a capsule neural network. A cluster of neurons that are in charge of identifying a certain entity or feature in the incoming data is represented by each capsule neural.[12]
2. **Capsule:** A capsule neural network's fundamental units are capsule. Unlike CNN neurons, capsules are designed to record not just the existence of a feature but also its posture, the orientation, and size. Each capsule neural produces a vector, sometimes called the "activity vector," which represents the specifications of the recognized object's production [12], [13].
3. **Dynamic Routing:** Dynamic routing techniques are used by capsule neural networks to allow interaction between capsule neural network at different levels. Capsule neural network in upper tiers can agree on the instantiation parameters using dynamic routing, which uses the predictions of capsule in lower layers. Capsule neural Networks, as opposed to CNNs, can handle variations in the spatial connections between features because of this mechanism. [13]
4. **Translational Equivariance:** CNNs exhibit translational invariance, whereas capsule neural networks exhibit translational equivariance. Because Capsule neural Networks are better adapted to handle input data changes like rotations and translations, feature recognition is therefore more accurate [14].
5. **Routing by Agreement:** A routing-by-agreement method was used by capsule neural networks, in which the agreement of predictions made by lower-level capsule neural networks is a prerequisite for the activation of higher-level capsule neural networks. This technique, routing by agreement, improves the network's efficacy and interpretability by ensuring that, the only pertinent attributes are spread it's throughout [14].
6. **Disentangled Representations:** Obtaining disentangled portrayals of incoming data considered as objective of capsule neural networks. It represents a particular object. This characteristic improves interpretability and generalization by helping Caps Net depict hierarchical connections between features better [15].

By adopting these ideas, capsule neural networks may be able to better understand spatial connections between features and provide more human-friendly visual representations, increasing their effectiveness in a range of tasks, including object recognition.

2.1 Comparing Capsule Neural Networks and Conventional Neural Networks

Key distinctions between Caps Nets and CNNs are as follows:

1. **In Feature Representation:** CNNs Features are expressed as scalar values, which are usually acquired by pooling after convolutional operations.[4] , while Capsule neural Networks use vectors to represent features, encoding not just the features' existence but also their attributes like size, orientation, and pose. Richer feature information may be captured by Caps Nets thanks to this vector representation.[16]
2. **In Hierarchy Structure:** Normally , CNNs had an expanding number of convolutional and pooling layers that process the input data to extract ever more abstract characteristics.[8] Insipid of Caps Nets which are hierarchical structures in which features at different levels of abstraction are represented by capsule neural network at different levels. Caps Nets can more accurately capture the spatial relationships between features owing to their hierarchical structure. [14]

3. **In Routing Mechanism:** The aggregate characteristics of CNNs are throughout local receptive fields using a fixed pooling procedure, which may result in the loss of spatial information [3]. When capsule neural networks use dynamic routing algorithms. Information is routed by higher-level capsule neural network according to agreements made by lower-level capsule neural network. Caps Nets can manage changes in the spatial relationships across features and preserve spatial hierarchies with the use of this routing technique [9].
4. **In Translational Parity :** The susceptible of CNNs are less to translations in the input data since they display translational invariance [3]. Although, capsule neural networks are more generalizable to transformations like rotations and translations because they exhibit translational equivariance. This property makes Caps Nets more robust to changes in the orientation or placement of objects inside photos [17].
5. **In Generalization and Interpretability:** The effective of CNNs in extracting features from data, they may be challenging to interpret because to their scalar activation representation of features. But it differ in Caps Nets , which These networks learn disentangled representations of input data by using capsule neural networks, each of which represents a unique object or attribute. This feature improves Caps Nets' ability to generalize to new inputs and makes them simpler to understand [18].

Figure 1 and Figure 2 below illustrate the architectural differences between capsule neural networks and traditional neural networks.

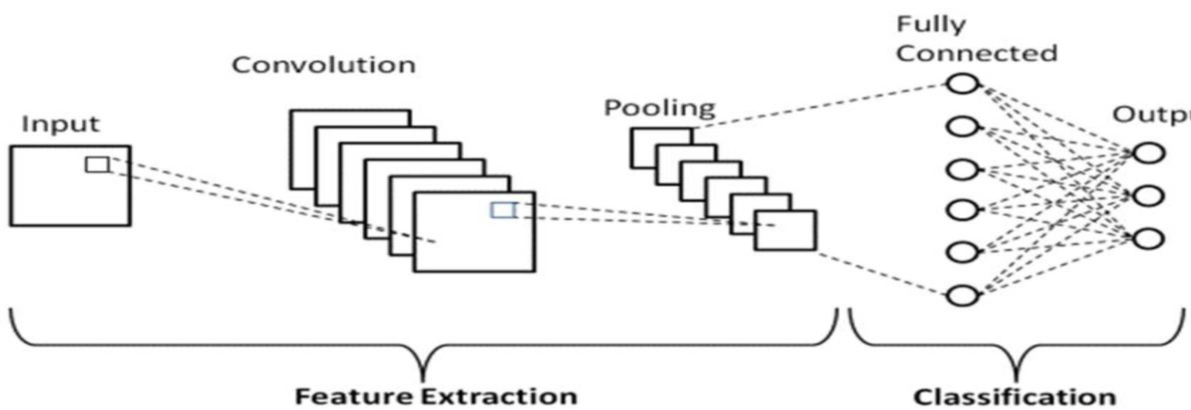

Fig .1. CNN architecture diagram [2].

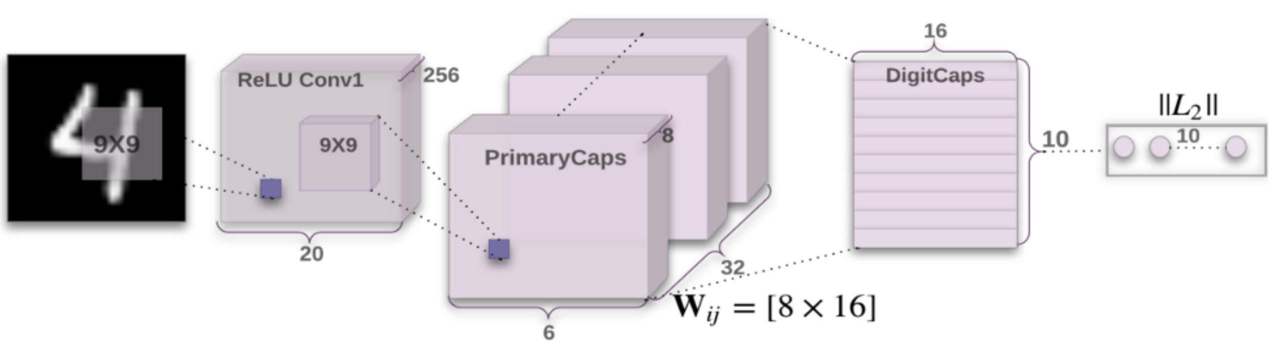

Fig .2. Typical Caps Net architecture [1].

3. ADVANCES AND INNOVATIONS IN THE TECHNOLOGY OF CAPSULE NEURAL NETWORKS

Capsule neural networks have become a popular and exciting model for deep learning in recent years. Unlike traditional CNNs, caps nets offer a new approach to expressing hierarchical patterns and spatial connections within data. The basic units of caps nets, capsules, in contrast to CNNs, which process and abstract features implement standard activations and pooling methods. These capsules are designed to collect and store comprehensive spatial data with complex relationships between features [19].

This paper shows several developments in capsule neural networks, with the various strategies that have been created and their revolutionary influence on deep learning techniques. It is examining the factors that enable capsules to outperform conventional CNNs in maintaining spatial hierarchies and linkages. By emphasizing these advancements, our research intends to demonstrate how contribution, capsule networks contribute, to the progress of deep learning techniques and provide novel answers to the issues of representing and processing complicated data [20].

1. **Dynamic Routing:** The routing of dynamic is a fundamental concept of capsule neural networks, as outlined by Sabour et al. (2017). This technique, dynamic routing, enables communication among capsules at different network layers by iteratively adjusting coupling coefficients based on the congruence between lower and higher-level structures. This process, dynamic routing, allows capsule networks to preserve spatial hierarchy and effectively capture complex relationships among object components.
2. **Routing-by-Agreement:** This technique improves, means routing-by-agreement**,** the routing process, in capsule neural networks, throw building on dynamic routing principles. In Hinton et al. (2018), the expectation-maximization (EM) technique iteratively updates coupling coefficients based on capsule consensus, improving the robustness and convergence of capsule networks, particularly with ambiguous input data [22].
3. **Dynamic Routing with Attention Mechanism:** This technique improved by adding attention mechanisms to dynamic routing offers another means of improving mechanism of capsule neural networks. Its introduced by Zhang et al. (2018). capsule of neural networks use attention weights to selectively focus on relevant data. By improving the network's ability to recognize long-range associations and contextual information, this attention mechanism promotes more efficient representation learning [23].
4. **Sparse Capsule neural Networks:** This technique proposed by Hinton et al. (2018) as an approach to computational inefficiency. These networks reduce unnecessary computation by routing only active capsule neural networks, especially in circumstances with sparse input data. In employing this approach, capsule neural networks become more efficient and flexible, improving their utility [24].
5. **Capsule neural Network with Attention Capsule neural:** A structure of distinctive is provided by adding a focus capsule to capsule neural networks. Attention of capsules assume context and long-term relationships, allowing the system to prioritize appropriate components during routing. Interpretability and performance have been improved by a combination process that combines the advantages of capsule neural networks and attention procedures [5].

Capsule neural networks are a main development in the field of deep learning, providing a lot of structured and understandable approach to feature extraction and representation learning. The numerous techniques covered in this paper show how efficient and flexible capsule neural networks may be when applied to complicated data model situations. As this field of study grows, capsule neural networks will continue to a major part in improving deep learning approaches [25][5].

4. APPLICATIONS OF CAPSULE NEURAL NETWORK

Caps networks received attention for their diverse applications in computing fields and technology. These networks significantly advance the design of deep learning models, then it offered the innovative solutions to the challenges faced by traditional neural networks. These applications separated into two distinct categories: those made particularly for tasks related to recognizing things, like determining faces and objects in photos and videos, as well as more general use instances such as natural language processing (NLP), medical data analysis, and even intelligent gaming [26].

Caps networks offer a good solution that can be used across a wide range of domains, including autonomous vehicles, e-commerce, and robotics to address the needs of different types of subject credits according to their adaptability and versatility. For example, in the healthcare sector, Caps networks used to analyze images for medical purposes with high accuracy, helping to identify diseases faster and more accurately. In the e-commerce domain, these networks can improve user experience with an accurately identifying products and providing personalized recommendations [25]. We have divided their applications into the following two main categories: specialized applications for object recognition and general Various applications that can be used in various domains to improve performance and efficiency.

4.1 Various Domains

Applications of Capsule Networks are classified into two types: generic and domain-specific. These applications are used in many different domains, including medical imaging, machine learning, and natural language processing. Because they can understand intricate spatial correlations between different components in the input data, capsule neural networks perform exceptionally well in this category. Their ability to capture complex patterns as well as hierarchies gives them a flexible tool that may be used in a variety of sectors and enterprises. [27]

1. **Natural Language Processing (NLP):** This is a topic of great interest for many investigations [16], some of them used dynamic routing between capsule neural algorithms. This algorithm has shown promise in capturing hierarchical relationships between textual elements, leading to improved performance in NLP tasks such as sentiment analysis and text classification, which is point of effectiveness in this algorithm. The Studies have demonstrated improved classification accuracy and robustness compared to traditional neural network architectures. [10]
2. **Robotics and Reinforcement Learning(RRL):** In this field, many investigations are being carried out. Several of them made use of Auto-Encoding Robot Control Policies Algorithm. Capsule neural Networks have shown effectiveness in learning compact and interpretable representations of robot control policies, enabling more efficient and adaptive robotic behaviour. Studies have reported improved learning efficiency and generalization capabilities in robotic control tasks demonstrated the effectiveness of Capsule neural Networks in learning complex robot control policies with fewer parameters, leading to more efficient control strategies [28].
3. **Time Series Analysis and Forecasting (TSAF):** Several studies have been proposed in this area [21,22]. Some of them have used the temporal capsule neural network algorithm. Temporal capsule neural networks have proven useful in capturing temporal dependencies and patterns in time series data, resulting in increased prediction accuracy and anomaly detection efficiency.

The outcomes illustrated how the time capsule neural networks work well and in an effective manner in time series forecasting tasks, achieving competitive performance compared to traditional recurrent neural network architectures. The paper informed significant improvements in prediction accuracy and robustness, especially in long-term forecasts [22].

4.2 Case Study

This paper confirms that Caps Nets are more e effective or object detection tasks. Our comprehensive review of the existing literature, proven that the novelty of this significant model is increasingly being used in many studies across a variety of fields. Capsule networks, which having unique architecture, give a significant advantage over traditional methods in terms of accuracy and efficiency in object detection.

Table 1 below demonstrates presents a detailed summarizing the extensive research on this emerging topic. This table highlights methodologies, applications, and key findings, serving as a valuable resource for researchers and practitioners. It provides insights into the effectiveness of capsule networks in different contexts for object identification, enhancing understanding of their potential and guiding future research and applications in this field [29].

TABLE I. A COMPILATION OF CAPSULE NETWORK APPLICATIONS IN IMAGE OBJECT RECOGNITION, DETAILING THE RESEARCH FIELD, ALGORITHM USED, DATASET EMPLOYED, ADDITIONAL HYPER PARAMETERS, AND PERFORMANCE METRICS ASSESSED.

Research	Felid Algorithm	Dataset	Additional Hyperparameters	Performance Metric	Ref.
Dynamic routing between capsule for image object recognition	Dynamic routing between capsules	Cifar-10, cifar-100, imagenet	Number of routing iterations = 3	Top-1 accuracy, top-5 accuracy	[30]
Matrix capsule with EM routing	Capsule neural routing mechanism	Mnist, fashion-mnist, coco	Learning rate = 0.001, number of routing iterations = 3	Intersection over union (iou), mean average precision (map)	[31]
Interpretable Graph Capsule neural Networks for Object Recognition	Capsule neural graph convolutional networks	Pascal voc, coco, open images	Number of graph convolutions = 3, capsule neural dimension = 16	Mean average precision (map), recall	[32]
Predict traffic flow in complex road Network.	A neural network with capsule that replaces max pooling by dynamic routing	Traffic dataset	Learning rate = 0.0005, exponential decay rate = 0.9999	Mean relative Error (mre), root mean square error (rmse) and mean absolute error (mae)	[33]
Knowledge graph (kg) completion and Search personalization	An embedding model, named CapsE	Wn18rr, fb15k-237, Search17	Adam initial learning rate in {1e-5, 5e-5, 1e-4}, # of filters n in {50, 100, 200, 400}, batch size = 128,	Mean rank (mr), mean reciprocal rank (mrr) and hits@10	[34]

			# of neurons in the Capsule neural in the second capsule neural layer = 10 (d = 10), # of iterations in the routing Algorithm in {1, 3, 5, 7}		
Classification of brain Tumor type	Dynamic routing between capsule neurals	Brain tumor dataset	Non	Prediction accuracy	[35]
3D image generation and quest to find better discriminators for GANs.	Dynamic routing between capsule neurals	Mnist orb	128 batch size, lr:0.002, epochs:20, leaky rel:0.2, Adam optim: beta1:0.5, weight init:normal: mean 0, std:0.02	Training loss and Visualization of outputs.	[36]
Class-imbalance and small dataset size limitations on classification of medical images.	Dynamic routing between capsule neurals	Tupac16 and diaretdb1, mnist and fashion-mnist	Non	F1-score	[29]
Detection of centerline invasion in abnormal driving.	Dynamic routing between capsule neurals	Youtube vehicle accident videos	9 _9 kernel & stride 1 in conv layer. Conv2/caps layer: 9 _9 kernel with stride of 2. Accuracy. Confusion Matrix. Model loss. Validation Train accuracy.		[37]
Classification of 2d hela datasets (protein classification).	Dynamic routing between capsule neurals	2d hela	Input image size: 382 _382 Accuracy. Margin loss. Reconstruction loss. Total loss.		[38]
Emotion recognition	Dynamic routing between capsule neurals	D.E.A.P	Input image size of 18 _18. Accuracy.		[39]
Object Detection by Network Based on CapsNet Architecture and Attention Layers	Model based on Feature Pyramid Network (FPN) with capsule attention layers	Utilizes anchor-based approach with 79.5% MAP in experiments	Research community can access the datasets used in the study are publicly available.		[40]
Cervical image classification based on image segmentation preprocessing and a CapsNet network model	Dynamic routing between capsule neurals	Clinical images of cervical cancer	Stride 1 in conv1 and 2 in conv2. Padding 0. Used cij of 1600 _ 3. Resnet. Inception v2 Confusion Matrix. Accuracy. Margin loss. Reconstruction loss. Total loss		[5]
Design of an Intelligent Approach on Capsule Networks to Detect Forged Images	SE-CapsNet (Attention Capsule Network)	CASIA dataset	Random noise added for evaluation	Accuracy	[41]
Image-Based Scam Detection Method Using an Attention Capsule Network	SE-CapsNet (Attention Capsule Network)	Public Ponzi scheme dataset	Not specified	F1 score	[42]
Exploring Deep Anomaly Detection Methods Based on Capsule Net	AnoCapsNet	Multiple datasets for anomaly detection	Prediction-probability-based (PP-based), Reconstruction-error-based (RE-based)	Various anomaly detection metrics	[33]
Capsule Networks for Object Detection in UAV Imagery	Capsule Networks (CapsNets)	Car and Solar Panel datasets	Cross-dataset transfer learning	Object detection accuracy	[24]
Fast Medicinal Leaf Retrieval Using CapsNet	CapsNet	Not specified, likely medicinal leaf image dataset	Feature vector with control points and discrete Fourier transform	Classification accuracy	[35]
Capsule-Networks Towards Object-Detection Capsule Object-Detector (COD)	Capsule Object-Detector (COD)	Pascal VOC 2007 dataset	Not specified	Mean Average Precision (mAP)	[31]

5. CHALLENGES

Despite numerous pieces of research presented in literature, several challenges remain for the widespread and effective use of Caps Nets in item detection. Compared to traditional CNNs, Caps Nets necessitate sophisticated architectures and extended training for the reason of their intrinsic complexity in both design and training procedures [10]. Due to their inherent complexity, training becomes more expensive and creates difficult obstacles to implementation and practical deployment. In addition, large amounts of heterogeneous training data are required for effective training of CAPs, and can be difficult to obtain, especially for specialized applications or domains [23]. Insufficient datasets can exacerbate training challenges by causing overfitting or poor performance. Additionally, the complex routing techniques employed by Caps Nets result in extra processing expenses, result in higher processing costs, highlighting the need for efficient resource allocation and routing algorithms, principally for extensive item recognition tasks [12]. The lack of comprehensive grasp of the fundamental mechanics of Caps Nets, notwithstanding their success, has slowed back the development of techniques to enhance their resilience and efficiency in recognizing object applications [23]. Caps Nets have the potential to transform object detection as well as additional computer vision applications, opening the door for their general adoption and useful implementation. However, the computational complexity of Caps Nets' restricts their scalability and application, particularly in situations of resource-limited by presenting infrastructure requirements, training time, and operations expenses [9]. The diversity and representativeness of the training data additionally, impact how effective Caps Nets are, since a lack of variability in the dataset may lead to biases or may problems with generalization [10]. The obstacles must tackles, concerted efforts in data gathering, computational optimization, algorithm development, and expanding our comprehension of the fundamental ideas of Caps Nets are needed [10][12][23].

6. RECOMMENDATIONS

Capsule Neural Networks are the newest processes in image classification, which break the away from conventionally well-recognized techniques of Convolutional Neural Networks. The data which used with Caps Nets becomes much more than binary; it is difficult network of inter-connected capsule neurons that capture the spatial correlations in the image. That's an creative way of changing our understanding of data-that, movies from conventional pixel-based analysis to its holistic explanation. One of the eminence features of Caps Nets includes robustness. They are capable to withstand a set of changes, such as rotation and scale, just like seasoned travellers slogging through choppy waters. With this robustness, the Caps Nets manage to keep their accuracy and consistency across a wide variation of datasets-something particularly important in these modern times when the images differ greatly both in forms and dimensions. However, another incredible capacity of Caps Nets is how to understand the greater context in which individual components of functions. The Caps Nets get an overview of the overall structure of an image, not just the separate aspects of it. It's as if it were having an experienced mentor who escorts us through life and empowers us with visions beyond the ordinary.

Furthermore, Caps Nets are good with interpreting professionals, exposing hidden meanings in images and denotes the complex relationships among various components. They are such as professional storytellers: by capturing emotive and specific information from the visual data, they help to advance knowledge of the latent concepts represented. While Caps Nets have several advantages, there are also some disadvantages in terms of difficulties such as training and processing requirements. Any trade-offs involved in the use of the Caps Net must be considered carefully, as they required, more resources than CNNs do on both time and dimensions of hardware. Concluding, Capsule Neural Networks give us another, very unusual perspective, and enabling to start a new way using approaches other than those common in picture classification. They encourage us to learn more about the foundations of visual data and explore the complex dance of spatial interactions. Let us move with caution into this uncharted territory, knowing what resources we expend with the paths that we take to fully unleash Caps Nets and change the way in image processing and more gets done.

7. CONCLUSION

Capsule Neural Networks (Caps Nets) is a major shift from the standard Convolutional Neural Networks (CNNs) paradigm in the field of image classification and recognition, where it takes an alternate diagonal approach in which information is held in a hierarchical manner. This hierarchical model improves the network's ability to capture and store complex object relationships and spatial hierarchies within images. Caps Nets use small layers of neurons named capsules that, along with each other, encode different properties of the item, such as its pose, orientation, and spatial relations. This innovative approach improves Caps Nets to be more efficacious for any changes in direction and movement, thus making them highly accurate in object detection.

Although Caps Nets have several characteristics, certain challenges do exist with them too. They are inherently much more complex than traditional CNNs, since they implement sophisticated routing algorithms and take huge computational resources for training and development. The basic principles of mechanical engineering of Caps Nets are not fully understood, increasing the complexity of these networks and making optimization and interpretation harder. Likewise, with specializations, generating large-scale, broad, or structured training datasets for the Caps Nets may be tricky, if not impossible. These limitations may be formed as obstacles that the Caps Nets might not be able to address to their full potential, and hence it will not be able to contribute to advancing image recognition and other computer vision applications.

Conflicts Of Interest
The author's paper emphasizes that there are no conflicts of interest that could impact the research integrity.
Funding
None.
Acknowledgment
The author acknowledges the institution for their commitment to fostering a research-oriented culture and providing a platform for knowledge dissemination.

References

[1] D. Kulkarni, "Deep convolution neural networks for image classification" 2022.
[2] A. Iscen, A. Fathi, and C. Schmid, "Improving image recognition by retrieving from web-scale image-text data" in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 19295–19304.
[3] Y. E. García-Vera, A. Polochè-Arango, C. A. Mendivelso-Fajardo, and F. J. Gutiérrez-Bernal, "Hyperspectral image analysis and machine learning techniques for crop disease detection and identification: A review," Sustainability, vol. 16, no. 14, p. 6064, 2024.
[4] W. Sun, H. Zhao, and Z. Jin, "A facial expression recognition method based on ensemble of 3D convolutional neural networks," Neural Computing and Applications, vol. 31, no. 7, 2019, doi: 10.1007/s00521-017-3230-2.
[5] X. Q. Zhang and S. G. Zhao, "Cervical image classification based on image segmentation preprocessing and a CapsNet network model," International Journal of Imaging Systems and Technology, vol. 29, no. 1, 2019, doi: 10.1002/ima.22291.
[6] W. K. Awad, E. T. Mahdi, and M. N. Rashid, "Features extraction of fingerprints based bat algorithms," International Journal on Technical and Physical Problems of Engineering, vol. 14, no. 4, 2022.
[7] J. Hai, Y. Hao, F. Zou, F. Lin, and S. Han, "Advanced retinexnet: A fully convolutional network for low-light image enhancement," Signal Processing: Image Communication, vol. 112, p. 116916, 2023.
[8] E. Juralewicz and U. Markowska-Kaczmar, "Capsule neural network versus convolutional neural network in image classification: Comparative analysis," in Lecture Notes in Computer Science, Springer, 2021, pp. 17–30, doi: 10.1007/978-3-030-77977-1_2.
[9] Y. Liu, D. Cheng, D. Zhang, S. Xu, and J. Han, "Capsule networks with residual pose routing," IEEE Transactions on Neural Networks and Learning Systems, 2024.
[10] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsule neurons," [Online]. Available: http://arxiv.org/abs/1710.09829, 2017.
[11] S. Chen and W. Guo, "Auto-encoders in deep learning—a review with new perspectives," Mathematics, vol. 11, no. 8, p. 1777, 2023.
[12] J. P. Bharadiya, "A tutorial on principal component analysis for dimensionality reduction in machine learning," International Journal of Innovative Science and Research Technology, vol. 8, no. 5, pp. 2028–2032, 2023.
[13] A. R. Kosiorek, S. Sabour, Y. W. Teh, and G. E. Hinton, "Stacked capsule neural autoencoders," in Advances in Neural Information Processing Systems, 2019.
[14] X. Jia and L. Wang, "Attention enhanced capsule neural network for text classification by encoding syntactic dependency trees with graph convolutional neural network," PeerJ Computer Science, vol. 7, 2022, doi: 10.7717/PEERJ-CS.831.
[15] Z. Zhu, G. Peng, Y. Chen, and H. Gao, "A convolutional neural network based on a capsule neural network with strong generalization for bearing fault diagnosis," Neurocomputing, vol. 323, 2019, doi: 10.1016/j.neucom.2018.09.050.
[16] J. Hollósi, Á. Ballagi, and C. R. Pozna, "Simplified routing mechanism for capsule neural networks," Algorithms*, vol. 16, no. 7, 2023, doi: 10.3390/a16070336.
[17] S. Tiwari and A. Jain, "Convolutional capsule neural network for COVID-19 detection using radiography images," International Journal of Imaging Systems and Technology, vol. 31, no. 2, 2021, doi: 10.1002/ima.22566.
[18] M. Puttagunta and S. Ravi, "Medical image analysis based on deep learning approach," Multimedia Tools and Applications, vol. 80, no. 16, 2021, doi: 10.1007/s11042-021-10707-4.
[19] M. K. Patrick, A. F. Adekoya, A. A. Mighty, and B. Y. Edward, "Capsule neural networks – A survey," Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 1, pp. 1295–1310, 2022, doi: 10.1016/j.jksuci.2019.09.014.
[20] N. Zhang, S. Deng, Z. Sun, X. Chen, W. Zhang, and H. Chen, "Attention-based capsule neural networks with dynamic routing for relation extraction," in Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018, 2018, doi: 10.18653/v1/d18-1120.
[21] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, "Scalable object detection using deep neural networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2147–2154.

[22] Z. Sun, G. Bebis, and R. Miller, "Object detection using feature subset selection," Pattern Recognition, vol. 37, no. 11, pp. 2165–2176, 2004.
[23] C. P. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," in Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), Jan. 1998, pp. 555–562.
[24] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," Multimedia Tools and Applications, vol. 82, no. 6, pp. 9243–9275, 2023.
[25] F. Jalled and I. Voronkov, "Object detection using image processing," arXiv preprint arXiv:1611.07791, 2016.
[26] R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra, and J. M. Z. Maningo, "Object detection using convolutional neural networks," in TENCON 2018-2018 IEEE Region 10 Conference, Oct. 2018, pp. 2023–2027.
[27] C. Szegedy, S. Reed, D. Erhan, D. Anguelov, and S. Ioffe, "Scalable, high-quality object detection," arXiv preprint arXiv:1412.1441, 2014.
[28] R. Padilla, S. L. Netto, and E. A. Da Silva, "A survey on performance metrics for object-detection algorithms," in 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), Jul. 2020, pp. 237–242.
[29] S. K. Divvala, D. Hoiem, J. H. Hays, A. A. Efros, and M. Hebert, "An empirical study of context in object detection," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2009, pp. 1271–1278.
[30] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," Advances in Neural Information Processing Systems, vol. 26, 2013.
[31] A. R. Pathak, M. Pandey, and S. Rautaray, "Application of deep learning for object detection," Procedia Computer Science, vol. 132, pp. 1706–1717, 2018.
[32] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 10781–10790.
[33] A. Basharat, A. Gritai, and M. Shah, "Learning object motion patterns for anomaly detection and improved object detection," in 2008 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2008, pp. 1–8.
[34] H. S. Ranganath and G. Kuntimad, "Object detection using pulse coupled neural networks," IEEE Transactions on Neural Networks, vol. 10, no. 3, pp. 615–620, 1999.
[35] L. Malagon-Borja and O. Fuentes, "Object detection using image reconstruction with PCA," Image and Vision Computing, vol. 27, no. 1-2, pp. 2–9, 2009.
[36] X. Wang, M. Yang, S. Zhu, and Y. Lin, "Regionlets for generic object detection," in Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 17–24.
[37] P. Zhou, B. Ni, C. Geng, J. Hu, and Y. Xu, "Scale-transferrable object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 528–537.
[38] P. Zhou, B. Ni, C. Geng, J. Hu, and Y. Xu, "Scale-transferrable object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 528–537.
[39] A. Mohan, C. Papageorgiou, and T. Poggio, "Example-based object detection in images by components," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 4, pp. 349–361, 2001.
[40] P. S. Rajpura, H. Bojinov, and R. S. Hegde, "Object detection using deep CNNs trained on synthetic images," arXiv preprint arXiv:1706.06782, 2017.
[41] R. Girshick, P. Felzenszwalb, and D. McAllester, "Object detection with grammar models," Advances in Neural Information Processing Systems, vol. 24, 2011.
[42] F. S. Khan, R. M. Anwer, J. Van De Weijer, A. D. Bagdanov, M. Vanrell, and A. M. Lopez, "Color attributes for object detection," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2012, pp. 3306–3313.