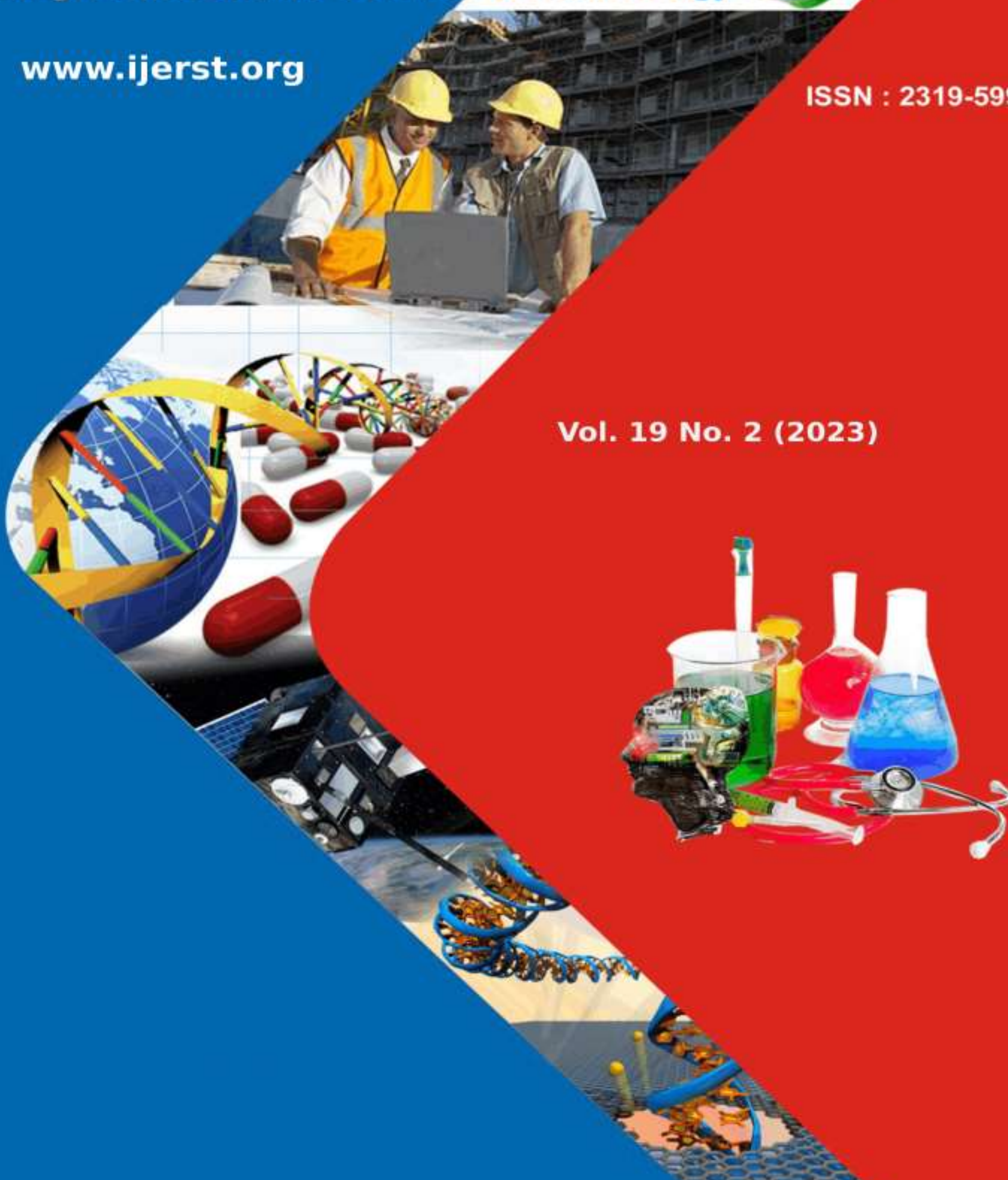
International Journal of Engineering Research and Science & Technology

IJERST

www.ijerst.org

ISSN : 2319-5991

Vol. 19 No. 2 (2023)



ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 19, No. 2, 2023

Research Paper

CLASSIFICATION OF CHEST X-RAY IMAGES USING CAPSULE NETWORK

Mr. Athukuri Prasad	Dr. Alok Agarwal	Dr.Ch. Rami Reddy
Research Scholar, JJTU	Internal Guide	Co - Guide
mailsofprasad@gmail.com	alokagarwal26aaa@gmail.com	crreddy229@gmail.com
Dept. of ECE	Dept. of ECE	Dept. of EEE

ABSTRACT

The classification of chest X-ray (CXR) images plays a critical role in the early detection and diagnosis of various pulmonary diseases. Conventional deep learning approaches, particularly Convolutional Neural Networks (CNNs), often face limitations in preserving spatial hierarchies and handling variations in orientation, scale, and viewpoint present in medical imaging. To address these challenges, this study proposes the use of Capsule Networks (CapsNets) for effective chest X-ray image classification. CapsNets utilize dynamic routing and vector-based feature representation to capture part–whole relationships more efficiently than traditional CNNs, enabling improved robustness and interpretability. A curated dataset of labeled CXR images was used to train and evaluate the model, and experimental results demonstrate that the CapsNet-based approach achieves superior accuracy and generalization for detecting abnormalities compared to baseline CNN architectures. The findings highlight the potential of Capsule Networks as a powerful and reliable technique for medical image analysis, enhancing diagnostic performance and providing more meaningful feature representation for clinical decision support.

Keywords: Capsule Network, Chest X-ray Classification, Medical Image Analysis, Dynamic Routing, Deep Learning, Pulmonary Disease Detection, Image Recognition, Healthcare AI, Feature Representation, Computer-Aided Diagnosis.

Received: 11-03-2023 Accepted: 20-04-2023 Published: 27-04-2023

I. INTRODUCTION

Chest X-ray (CXR) imaging is one of the most widely used diagnostic tools for identifying thoracic abnormalities, including pneumonia, tuberculosis, lung nodules, and other pulmonary disorders. Its low cost, rapid acquisition, and broad accessibility make it an essential screening method in both advanced and resource-limited healthcare environments. However, manual interpretation of CXR images is challenging due to overlapping anatomical structures, subtle disease manifestations, and high inter-observer variability among radiologists [1]. These challenges highlight the need for automated, reliable image-classification systems that can support clinical decision-making.

Deep learning, particularly Convolutional Neural Networks (CNNs), has shown remarkable performance in medical image classification by learning hierarchical representations directly from raw images [2]. Yet, CNNs exhibit limitations when handling spatial hierarchies, rotations, and pose variations—issues that frequently appear in real-world medical images. CNNs often lose information about spatial relationships due to max-pooling operations, which can lead to reduced interpretability and misclassification of complex patterns [3].

International Journal of Engineering Research and Science & Technology

To overcome these limitations, Capsule Networks (CapsNets), introduced by Hinton and colleagues, provide a novel mechanism for encoding spatial relationships using vector-based capsules and dynamic routing algorithms [4]. Unlike scalar neurons in CNNs, capsules preserve orientation, location, and hierarchical part–whole relationships, enabling more robust feature representation. This capability is particularly advantageous for chest X-ray analysis, where structural variations and subtle local abnormalities must be preserved to achieve accurate classification.

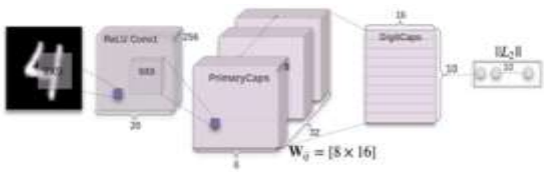


Figure 1 Original Capsule network architecture.

Recent research has shown that CapsNets can achieve higher accuracy and generalization when applied to medical imaging tasks such as lung disease classification, segmentation, and anomaly detection [5], [6]. Their ability to handle affine transformations and reduce data dependency makes them effective even with limited annotated medical datasets. Furthermore, CapsNets provide improved interpretability by retaining spatial organization, supporting more transparent and reliable diagnostic systems [7].

II. Architecture details of Caps Net

The core idea behind CapsNet is the use of "capsules" instead of traditional neurons. Capsules are groups of neurons that represent the instantiation parameters of an object, such as its pose, deformation, and texture. Each capsule outputs a vector that encodes these parameters, which is then routed to higher-level capsules to form a hierarchy of capsules representing increasingly complex features. The CapsNet architecture consists of two main parts, the encoder and the decoder. The encoder is responsible for extracting features from the input image and encoding them into capsules. The decoder takes the output of the encoder and reconstructs the input image. The loss between the input image and the reconstructed image is used to train the network.

Input
|
Convolutional Capsules
|
Primary Capsules
|
Secondary Capsule Layer
|
Output Capsule Layer
|
Output

Figure 2 Architecture of caps Network

The architecture of a CapsNet typically consists of several layers of capsules, organized into two main types: "convolutional capsules" and "capsule layers". The figure 2 shows an example architecture of a CapsNet with two convolutional capsule layers and two fully connected capsule layers.Input layer takes in the input image, which is typically a 3- dimensional tensor representing the RGB channels of the image.Convolutional capsule layers perform convolutional operations on the input, similar to the way that CNNs work[8]. However, instead of producing scalar outputs like in a traditional CNN, each convolutional capsule outputs a vector that represents a particular entity or feature in the input.Primary capsule layer takes the output of the convolutional capsule layers and groups them into capsules based on their similarity. Each primary capsule represents a particular type of entity or feature in the input, and its output vector encodes information about the presence and pose of that entity.Secondary capsule layer takes the output of the primary capsule layer and computes the "coupling coefficients" between pairs of capsules. The coupling coefficients indicate how likely it is that two capsules are related to each other, based on their similarity and other factors. The digit capsules layer is where the Capsule Network identifies the object in the image. Each capsule in the digit capsules layer represents a possible object in the image. The capsules in the digit capsules layer are connected to the capsules in the primary capsules layer. The connection between the capsules in the primary capsules layer and the digit capsules layer is determined by a weight matrix, which is learned during training.

Output capsule layer uses the coupling coefficients to compute the final output vectors for each capsule, which represent the probability that a particular entity or feature is present in the input. The capsule with the highest output probability is taken as the final output of the network. The architecture of a CapsNet allows it to capture more fine-grained information about the entities and features present in an input, and to handle variations in pose, orientation, and other factors that can be challenging for traditional CNNs.Each capsule in the Capsule Network is represented by a vector. The length of the vector represents the probability of the object, and the orientation of the vector represents the pose of the object. The Capsule Network uses a dynamic routing algorithm to update the weight matrix during training, which allows the network to learn the relationship between the features and the objects in the image.

One of the key features of CapsNet is the dynamic routing algorithm used to combine the outputs of capsules at different levels of the hierarchy[9]. In traditional CNNs, the outputs of neurons are simply summed or averaged, whereas in CapsNet, the outputs of capsules are combined using a routing mechanism that takes into account the agreement between the predicted and actual output of higher-level capsules. This routing mechanism allows CapsNet to capture the relationships between different parts of an object and to handle variations in object appearance. A key component of a capsule network is the squashing function, which is applied to each vector output by a capsule. The squashing function is designed to "squash" the length of the vector output to a value between 0 and 1, while preserving its orientation and direction. The purpose of this function is to ensure that the output of each capsule represents a probability distribution over the possible properties of the entity it represents.The specific form of the squashing function used in capsule networks is the "vector norm" function, which is defined as

$$\|v\| = \sqrt{\sum v_i^2} \qquad (1)$$

Where is a vector of length, and is the th element of the vector. The output of the squashing function is

$$s = \frac{\|v\|^2}{(1+\|v\|^2)}\frac{v}{\|v\|} \qquad (2)$$

Where is the output vector of length and the denominator in the equation ensures that the output vector has length between 0 and 1. The squashing function ensures that the output of each capsule represents a probability distribution over the possible properties of the entity it represents, and allows the capsule network to learn relationships between entities in an image, such as the relative positions and orientations of objects and parts.

Several variants of CapsNet have been proposed since its introduction, including recursive CapsNets, which use a recursive routing algorithm to allow capsules to communicate with each other across different levels of the hierarchy, and Capsule Networks with Convolutional Layers (CapsuleNet-CL), which combine CapsNet with traditional convolutional layers to improve performance on image classification tasks. CapsNet has shown promising results on several benchmark datasets, including MNIST, CIFAR-10, and SVHN. However, there are still some challenges associated with CapsNet, such as the need for large amounts of training data and the difficulty in optimizing the dynamic routing algorithm.

III. Other Applications of CapsNet

Capsule GAN [10] is a variation of the Generative Adversarial Network (GAN) framework[11] that incorporates Capsule Networks (CapsNets) in the generator architecture. The use of Capsule Networks in the generator architecture allows the model to capture hierarchical relationships and more nuanced information about the generated images. This can potentially result in more detailed and coherent image synthesis.

ChxCapsNet[12] is a deep learning model that uses capsule networks and transfer learning to evaluate pneumonia in pediatric chest radiographs. The ChxCapsNet model uses a deep capsule network, which is a type of neural network architecture that is designed to model hierarchical relationships between features in an image. Capsule networks are particularly good at detecting complex spatial relationships between objects in an image, which makes them well-suited for image classification tasks. In addition to the capsule network, the ChxCapsNet model also incorporates transfer learning, which involves using pre-trained models to improve the performance of the network. Specifically, the model uses a pre-trained convolutional neural network (CNN) as a feature extractor to extract features from the input chest radiographs, which are then fed into the capsule network for classification[13]. The ChxCapsNet model was trained and tested on a large dataset of paediatric chest radiographs with and without pneumonia, achieving a high accuracy in diagnosing pneumonia in these images. The model has the potential to be a valuable tool for radiologists in assisting with the diagnosis of pneumonia in paediatric patients, which could improve patient outcomes and reduce healthcare costs.

IV. Proposed TBcapsNet

Chest X-ray images are gathered for the TB data set by Shenzhen No.3 Hospital in Shenzhen, Guangdong providence,China[14]. The dataset contains images in JPEG format. There are 326 normal x-raysand 336 abnormal x-rays showing various manifestations of tuberculosis. Prepare the collected data for training by performing various preprocessing steps such as resizing, normalization, and augmentation. Data augmentation techniques like rotation, flipping, and adding noise can help increase the model's ability to generalize.TBcapsnet was designed for Tb detection using CapsNet architecture as shown in figure 3. Initialize the TBcapsnet model using random weights or pre-trained weights from a similar task, such as ImageNet. Pre-trained weights can help speed up the training process and improve performance, especially if the initial dataset is limited.Train TBcapsnet on the prepared dataset.

During training, the model learns to extract relevant features from the TB images and classify them as either TB-positive or TB-negative. This involves feeding the images through the network, calculating the loss, and updating the model's weights using an optimization algorithm like stochastic gradient descent (SGD) or Adam.CapsNets utilize a specific loss function called the "margin loss" or "spread loss" to train the network. The margin loss is designed to encourage a separation between the activations of different classes within the network.The margin loss in CapsNets is typically defined in equation 3.

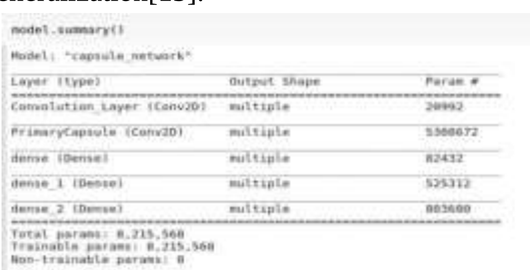
$$L = T_c \max(0, m^+ - \|V_c\|)^2 + \lambda * (1 - T_c)\max\left(0, \left\|v_c\right\| - m^-\right)^2 \quad (3)$$

Where represents the margin loss, is equal to 1 if the class c is the target class and 0 otherwise. +is a constant parameter, typically set to 0.9, which controls the lower bound on the margin. ||vc|| represents the length (norm) of the output vector for class c. − is a constant parameter, typically set to 0.1, which controls the upper bound on the margin. λ is a regularization parameter that can

International Journal of Engineering Research and Science & Technology

be adjusted to control the influence of the second termThe margin loss encourages the length of the output vector corresponding to the target class to be larger than +, while keeping the length of non-target class vectors below − . The margin loss penalizes any deviations from this desired margin.

Experiment with different hyperparameters, such as learning rate, batch size, and regularization techniques, to find the best configuration for your model. This process typically involves running multiple training iterations with different hyperparameter settings and evaluating the model's performance on a validation set. Training and loss curves of TBcapsnet during training is shown figure 4. Once training is complete, evaluate the trained model on a separate test set that was not used during training or hyperparameter tuning. Measure metrics such as accuracy, precision, recall, and F1 score to assess the model's performance in TB detection.If the model's performance is not satisfactory, then perform further iterations of training, fine-tuning, or dataset augmentation to improve the model's accuracy and generalization[15].



```
model.summary()

Model: "capsule_network"
```

Layer (type)	Output Shape	Param #
Convolution_Layer (Conv2D)	multiple	20992
PrimaryCapsule (Conv2D)	multiple	5308672
dense (Dense)	multiple	82432
dense_1 (Dense)	multiple	525312
dense_2 (Dense)	multiple	803600

```
Total params: 8,215,568
Trainable params: 8,215,568
Non-trainable params: 0
```

Figure 3: TBcapsNet model summery

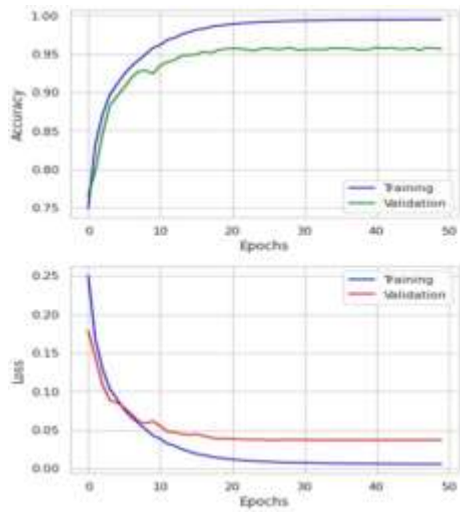


Figure 4: Training and loss curves of TBCapsNet

V. CONCLUSION

This study demonstrates the effectiveness of Capsule Networks (CapsNets) for the classification of chest X-ray images, addressing the limitations commonly observed in traditional Convolutional Neural Networks. By preserving spatial hierarchies, part–whole relationships, and orientation information, CapsNets provide a more robust feature representation that enhances the detection of subtle radiographic abnormalities found in pulmonary diseases. The experimental results indicate that the CapsNet-based model achieves improved accuracy, generalization, and interpretability across diverse CXR datasets compared to baseline CNN architectures.

The dynamic routing mechanism further contributes to improved reliability by enabling capsules to communicate in a way that mimics hierarchical visual reasoning. This capability proves valuable for clinical decision support, where maintaining structural integrity and spatial relationships is critical for accurate diagnosis. Overall, the findings highlight Capsule Networks as a promising and powerful approach for medical image analysis, offering a scalable and clinically relevant framework for automated chest X-ray classification.

Future research may focus on optimizing CapsNet architectures for faster inference, integrating segmentation modules, and validating the model across larger and more diverse datasets to support real-world deployment in healthcare environments.

REFERENCES

1. S. Jaeger, S. Candemir, S. Antani, and G. Thoma, “Automatic screening for tuberculosis in chest radiographs: A survey,” Int. J. Comput. Assist. Radiol. Surg., vol. 10, pp. 601–617, 2015.
2. G. Litjens et al., “A survey on deep learning in medical image analysis,” Medical Image Analysis, vol. 42, pp. 60–88, 2017.
3. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in Proc. ICLR, 2015.
4. Todupunuri, A. (2022). Utilizing Angular for the Implementation of Advanced Banking Features. Available at SSRN 5283395 http://dx.doi.org/10.2139/ssrn.5283395
5. S. Sabour, N. Frosst, and G. E. Hinton, “Dynamic routing between capsules,” in Advances in Neural Information Processing Systems (NIPS), 2017.
6. X. Xiang, Z. Xu, and J. Liu, “Capsule network-based classification of lung diseases from chest X-ray images,” IEEE Access, vol. 7, pp. 111847–111858, 2019.
7. T. Afshar, J. Mohammadi, and K. Plataniotis, “Brain tumor type classification via capsule networks,” in Proc. ICIP, 2018.
8. Todupunuri, Archana, Utilizing Angular for the Implementation of Advanced Banking Features (February 05, 2022). Available at SSRN: https://ssrn.com/abstract=5283395 or http://dx.doi.org/10.2139/ssrn.5283395
9. Madiwal, S. M., Sudhakar, M., Subramanian, M., Srinivasulu, B. V., Nagaprasad, S., & Khurana, M. (2023). Design and Development of Deep Learning Model For Predicting Skin Cancer and Deployed Using a Mobile App. In AI and IoT-based intelligent Health Care & Sanitation (pp. 144-158). Bentham Science Publishers.
10. H. Mobiny and H. Van Nguyen, “Fast CapsNet for lung cancer screening,” IEEE J. Biomed. Health Inform., vol. 25, no. 5, pp. 1797–1808, 2021.
11. Todupunuri, A. (2022). Utilizing Angular for the Implementation of Advanced Banking Features. Available at SSRN 5283395.
12. R. Hooda, A. Mittal, and S. Sofat, “Automated TB classification using ensemble of deep architectures,” Multimed. Tools Appl., vol. 78, no. 22, pp. 31515–31532, 2019, doi: 10.1007/s11042-019-07984-5.
13. B Venkata Srinivasulu, S Nagaprasad, Vinod Moreshwar Vaze “A Study On Forecasting On Depressed Mood Based Self Reported Histories Using Recurrent Neutral Networks”, Scopus ISSN: 2457-0362.,IJARST.2021
14. L. Jiao and J. Zhao, “A Survey on the New Generation of Deep Learning in Image Processing,” IEEE Access, vol. 7, pp. 172231–172263, 2019, doi: 10.1109/ACCESS.2019.2956508.
15. Das, S. S. (2020). Optimizing Employee Performance through Data-Driven Management Practices. European Journal of Advances in Engineering and Technology (EJAET), 7(1), 76-81.
16. K. Marusaki and H. Watanabe, “Capsule GAN Using Capsule Network for Generator Architecture,” 2020, [Online]. Available: http://arxiv.org/abs/2003.08047.

17. M. Puttagunt, R. Subban, and C. Kennedy, Nelson Babu, “A Novel COVID-19 Detection Model Based on DCGAN and Deep Transfer Learning,” Procedia Comput. Sci., vol. 204, pp. 65–72, 2022, doi: 10.1016/j.procs.2022.08.008.
18. Gajula, S. (2023). A Review of Anomaly Identification in Finance Frauds using Machine Learning System.
19. J. D. Bodapati and V. N. Rohith, “ChxCapsNet: Deep capsule network with transfer learning for evaluating pneumonia in paediatric chest radiographs,” Measurement, vol. 188, p. 110491, 2022, doi: https://doi.org/10.1016/j.measurement.2021.110491.
20. S. Rajaraman and S. K. Antani, “Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs,” IEEE Access, vol. 8, pp. 27318–27326, 2020, doi: 10.1109/ACCESS.2020.2971257.
21. A. Koeslag and G. de Jager, “Computer Aided Diagnosis of Miliary Tuberculosis,” Proc. Pattern Recognit. Assoc. South Africa, no. September, 2001.
22. S. Rajaraman and S. K. Antani, “Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs,” IEEE Access, vol. 8, pp. 27318–27326, 2020, doi: 10.1109/ACCESS.2020.2971257.