Signal, Image and Video Processing (2022) 16:2269–2277
https://doi.org/10.1007/s11760-022-02192-5

ORIGINAL PAPER

Improved CAPSNET model with modified loss function for medical image classification

J. Deepika[1] · C. Rajan[2] · T. Senthil[3]

Received: 8 November 2021 / Revised: 22 February 2022 / Accepted: 25 February 2022 / Published online: 29 March 2022

Abstract
A capsule network (CapsNet) is a new neural network model that is recently evolving in the field of image classification. Some of the shortcomings of traditional convolutional neural networks (CNNs) are compensated by the characteristics of CapsNet. It has proven to be effective at a variety of tasks, predominantly in medical image recognition with activation capsules. In this paper, image classification using the special designs in CapsNet is examined in depth. An additional reconstruction loss is used in the proposed work to empower the steering capsules and encode the input's instantiation parameters. The active vectors of higher-level capsules are used for the classification mechanism. The calculation at that point remakes the input picture thus utilizing these active vectors. The directing capsule's yield is sent into a decoder with three completely associated layers, which limits the whole of squared disparities between the calculated unit yields and the pixel power. In comparison to a typical CapsNet, the improved CapsNet method incorporates the extra parameters such as the number of measurements in each capsule sort (essential or directing capsules), the number of essential and directing capsules, and the number of channels within the capsule layer that are used for image classification. The experimental results show promising results in image recognition when compared to other CNN model-based algorithms.

Keywords Accuracy · Capsules · Efficiency · Layered approach · Prediction rate

1 Introduction

Cancer has gotten to be a colossal danger to people's lives and wellbeing in recent years. Either in the case of children or adults, tumors in brain are the deadliest and are highly prevalent among the types [12]. The rightness of determination utilizing neural networks is getting improved by various deep learning techniques [7]. For early disease identification, various software is currently deployed. Of which convolutional neural networks are proved to be effective. In regards, of late, there has been an interest in creating CNNs to categorize such tumor images [19]. CNNs have exploded in popularity in recent years [14]. CNNs have been employed for cancer early detection, but in some circumstances, such as images with affine modification, it fails to produce the required results [17]. The convolutional neural systems overlook numerous spatial relationships between fewer complex objects [8]. CNNs with max-pooling layers have resulted in the rapid development of the deep learning area [21]. As a result, it was the need of the hour to use yet another improved model before a capsule network with CNN capabilities and none of its drawbacks were built. Hence, a modification to the CapsNet is proposed to address the difficulties. It has been utilized to layout the inputs to a midway highlight space rather than performing dynamic routing to deliver class-specific yield capsules. Instead of class-specific capsules, the capsule network provides feature-specific capsules. Also, with the reduced features, the network's generalization potential is strengthened, resulting in better performance [20]. The bonding between capsules is measured in this study based on the existence of a few indicated characteristics, from which the

✉ J. Deepika
deepi.realmail@gmail.com

1 Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamilnadu, India

2 Department of Information Technology, K S Rangasamy College of Technology, Tiruchengode, Namakkal, Tamilnadu, India

3 Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamilnadu, India

yield classes are induced employing a completely connected layer [23].

Recently, CapsNets have been developed to address the drawbacks of CNNs, and they are expected to improve deep learning solutions [24]. CapsNets have been presented to be the best technique for CNN problems lately. By tackling pose and deformation encoding issues, CapsNet has achieved noteworthy results in image identification. They are a new type of deep neural network design that is excelled in a variety of domains, including image identification and natural language processing [18]. Capsule systems are vigorous to turn and may give relative interpretation. This requires essentially fewer training data, as is the case for preparing therapeutic picture datasets such as brain attractive reverberation imaging (MRI) pictures. Both these training datasets gain a special importance toward this research.

Capsule networks are not scalable for more complex tasks on their own. The number of classes is directly dependent on the dynamic routing algorithm in its natural form [11]. As a result, traditional capsule networks tend to overfit the data. Instead of constructing class-specific capsules, we propose using capsule networks for feature extraction in this paper. Furthermore, the number of highlights produced by energetic directing can be essentially littler than the number of classes, avoiding the arrange from over-fitting whereas at the same time learning more non-specific highlights.

2 Background

A brain tumor is one of the most lethal tumors, and its effective treatment would be predicated in part on a correct diagnosis of the tumor type. [1] addressed to this by employing a boosted capsule arrange, named BoostCaps, which takes advantage of boosting approaches' capacity to handle poor learners by gradually boosting the models. The method eliminates the need to experiment with various topologies [10].

CNNs do not completely exploit spatial relations, which is particularly tricky for tumor classification, as the relationship between the tumor and its encompassing tissue may be a key pointer of the tumor's kind [25]. To address this problem, [5] incorporated freshly created CapsNets that are extremely touchy to the picture foundation.

Alhassan and Zainon [6] proposed a modern learning-based strategy to process the automatic segmentation in multimodal MRI images to identify brain tumors, thus the Bat algorithm with clustering. Hence, the techniques of Bat algorithm with fuzzy c-ordered means (BAFCOM), in addition to enhanced capsule networks (ECN) are significantly applied to perform the segmentation and class of mind tumors from MRI snapshots. Afshar et al. [2] demonstrated the feasibility of constructing CapsNet architecture for the classification of brain tumor types. CapsNets, like other deep learning models, do not incorporate forecast uncertainty by returning the uncertain samples.

3 Materials and methods

Capsules are clusters of neurons whose activity vectors represent various pose parameters, and the lengths of these vectors indicate the possibility that a certain entity exists. The feature decomposition module and multi-scale feature extraction module are added to the basic capsule network for the proposed capsule network. The feature decomposition module is used to extract rich features, reducing calculation time and speeding up network convergence. In addition, unlike standard CNNs, the activations in the output layer are also represented as vectors or output capsules. The weights are iteratively changed during the forward proportional to the similarity between the individual activations and the combined activation in dynamic routing, which uses a weighted summation of primary capsules.

The primary capsule layer and the output capsule layer are the two layers that make up the system. The primary capsule layer combines many convolutional outputs into a single capsule unit. The output of the primary layer is sent to the output capsule layer using dynamic capsule routing. Each of the output capsule layer's vectors represents a single output class. The chance of the corresponding class is proportional to the length of the vectors. Table 1 shows the various significant recent works carried out by researchers with the CapsNet model.

Figure 1 illustrates the overall architecture of the proposed CapsNet for feature extraction. The initial layer consists of convolution kernels with ReLU activation. The next layer is multi-scale capsule encoding which performs feature extraction by three stages. The multi-scale encoding unit comprises of two stages: feature extraction and capsule encoding. The last layer is the digit capsules that show all the instances in the classes.

3.1 Primary capsule layer

Convolution layers transforms visual samples groups of activation tensors in capsule networks. In the primary capsule layer, this tensor will be transformed into primary capsules. With single-channel inputs, the 256-dimensional features for images are obtained by convolving 256 numbers of 99 kernels with a stride of 1. The primary capsules are 8-dimensional vectors created by convoluting eight 9×9 kernels with the step of 2, unlike the scalar activation of the standard convolutional layers. Also, 32 chunks of initial capsules are formed by having 32 such groupings of kernels. All the eight-dimensional initial capsules will be padded together

Table 1 Comparison of various significant research works carried out using CAPSNET model

Title	Year	Methodology	Conclusion
A Boosted Capsule Network for Brain Tumor Classification [3]	2020	Considers both brain pictures and tumor harsh border boxes as inputs, permitting get to both the surrounding tissue	Capable of extracting features from the entire brain tissue while simultaneously
Capsule Systems for Brain Tumor Classification Based on MRI Pictures and Coarse Tumor Boundaries [4]	2019	The modified CapsNet engineering employments the tumor coarse boundaries as additional inputs inside its pipeline to extend the CapsNet's center	First, the need for exact tumors explanation is killed; and Capsule Networks' center is extended across the tumor coarse boundaries
Capsule Networks' Interpretability for brain tumor classification via radiomics analyses. [5]	2019	Radiomics features extracted using Capsule networks show a strong association with hand-crafted features	Features learned using a Capsule network successfully distinguishes between different types of cancer
BAT Algorithm With fuzzy C-Ordered Means (BAFCOM) Clustering Segmentation and Enhanced Capsule Networks (ECN) for Brain Cancer MRI Images Classification [6]	2020	The initial centroids and distance inside pixels are computed, using Regions of Interest (RoI)	Performs the segmentation and class of mind tumors from MRI snapshots
BayesCap: A Bayesian Approach to Brain Tumor Classification Using Capsule Networks [1]	2020	Provides mean predictions and entropy as a measure of prediction uncertainty	Reduces overfitting while also allowing for the return of ambiguous predictions to human specialists for additional research
Brain Tumor Type Classification Via Capsule Networks [2]	2018	Proposed a better architecture that maximizes the accuracy of the classification problem at hand	Handles a limited number of training data and their units are equivariant. Improved accuracy by adjusting the number of feature maps in the Capsule network's convolution layers
Automatic Modulation Recognition: A Few-Shot Learning Method [13]	2021	AMR-CapsNet, a new network topology for improving modulation signal classification accuracy with fewer samples	The classification accuracy improves by 20% when compared to CNN
Robotized Classification of Apoptosis in Stage Differentiate Microscopy Utilizing Capsule Organize [15]	2020	A CapsNet, as well as a bi-directional long short-term recurrent structure for consistent predictions	Accuracy is very high when compared to other existing techniques
Capsule Networks – A survey [16]	2019	A deeper discussion about Capsule Networks	Comparative analysis of various Capsnet based algorithms is done
CapsNet-SSP: Multilane capsule network for predicting human saliva-secretory proteins [9]	2020	An end-to-end model based on CapsNet is proposed to identify saliva-secretory proteins from the sequence information	Used for biomedical analysts for analysing unhealthy tissues close or distal to salivary organs utilizing transcriptome or proteomics

for producing $8 \times N_{pc}$ a tensor. Here, N_{pc} denotes the total number of primary capsules.

3.1.1 Output capsule layer

These layers will take the initial capsules from the previous layers to create output capsules by means of dynamic routing. The capsule is a vector that describes one class in each yield capsule layer. A typical CNN yields an output of $N_{\text{class}} \times 1$, where N_{class} is the number of classes present the given dataset. While in CapsNet, the yield capsule layer has a dimension $N_{\text{class}} \times 16$ and each class will be signified as 16-dimensional output capsules. Hence, the data will be comprised in the 16-dimensional vector instead of offering the scalar representation of each class's likelihood. The size of the yield vector indicates the probability of the associated class being present.

3.1.2 Dynamic routing

Dynamic routing is used for obtaining the output capsules from the initial capsules. For determining the individual concepts of each capsule, the trainable weight (W) of dynamic routing is utilized. Here, the dimension of W_{ij} is taken as 8×16, $j \in [1, N_{class}]$ is taken as the index of 16-dimensional output capsule and $i \in [1, N_{PC}]$ is taken as the index of the eight-dimensional initial capsule of dimension. Equation (1) gives the individual concept of the initial capsules i concerning the yield capsules j.

$$\widehat{u}_{j|i} = u_i \times W_{ij}^{DC} \quad (1)$$

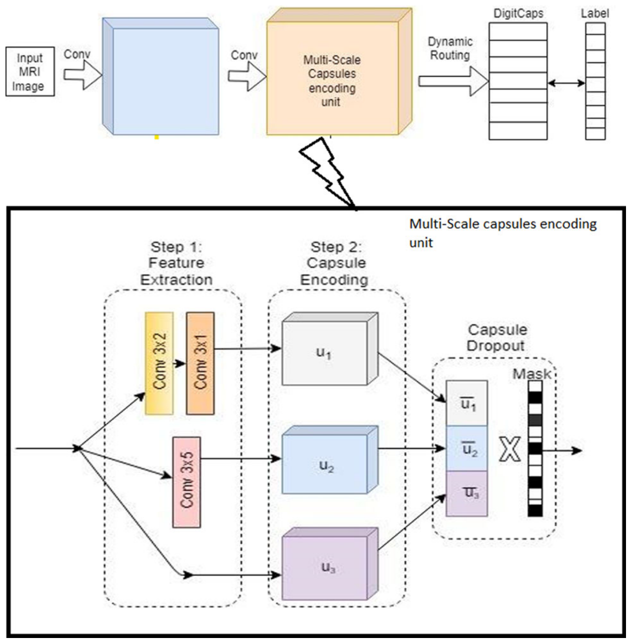

Fig. 1 Overall Architecture of the proposed improved CapsNet for feature extraction, capsule encoding and dropout

where the i^{th} initial capsule is represented as u_i. A yield block of structure $N_{class} \times 16$ is provided for each primary capsule i. Another form of weight of dimension $N_{PC} \times N_{class}$ class is considered for the functioning of dynamic routing, called routing weights, b. Individual thoughts are combined to build the result capsules using them. These weights, unlike W, learns by dynamical routing through further iterations, based on the bonding of concepts to that of the total outputs. At the start of each forward pass, these weights are set to zero. The coupling coefficients c_{ij} are given by Eq. (2).

$$c_{ij} = -\frac{\text{exponent}(b_{ij})}{\sum_k \text{exponent}(b_{ik})} \tag{2}$$

This coupling coefficient will be utilized to combine the individual concepts $\widehat{u}_{j|i}$ to develop combined output capsules. The j^{th} combined output capsules will be acquired by squashing s_j as illustrated in Eq. (3).

$$s_j = \sum_i c_{ij}, \widehat{u}_{j|i} = W_{ij} u_i \tag{3}$$

For ensuring that the longer vectors are compressed to the length just below 1 and shorter vectors are compressed to virtually zero-length, a non-linear squashing function will be utilized. The output capsule v_j is illustrated in Eq. (4).

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \cdot \frac{s_j}{\|s_j\|} \tag{4}$$

The agreement between individual yield capsules $\widehat{u}_{j|i}$ and squashed combined output capsules v_j is calculated using a simple dot product. Here, higher priority will be given for individual capsules that agree with the aggregate outputs. This is accomplished by altering the b_{ij} as in Eq. (5).

$$b_{ij} = b_{ij} + \widehat{u}_{j|i} . v_j \tag{5}$$

3.1.3 Loss function

This will split into major divisions such as the mean square losses and the object presence margin losses for an image generated from output capsule. For the object k, the marginal losses will be calculated as represented in Eq. (6).

$$L_k = T_k(0, m^+ - \|v_k\|)^2 + \lambda(1 - T_k)0, \|v_k\| - m^{-)^2} \tag{6}$$

Here, λ will be set as 0.5E, the upper and lower bounds m + and m will be set as 0.9 and 0.1 and $T_k = 1$ if the object of class k is present.

3.1.4 Decoder network for regularization

The decoder networks are made up of a series of fully connected layers that recreate the unique inputs. In the masked-out output capsule layers, all capsules save the one matching to the label which will be set to zero. The decoder reconstructs the output capsule after masking them out. During the test, the masking facilitates the generation of class-specific reconstruction. The reconstruction loss will be lowered together with the margin losses to prevent the model from over fitting the training datasets. The reconstruction losses are scaled-down by 0.0005 to guarantee that it does not outnumber the margin loss.

3.2 Improved capsule network

CapsNets have been proven to function best with fewer classes in the past. Higher-class dataset classification, on the other hand, becomes tremendously time-consuming to the point of being practically impossible. It has also been discovered that when the number of classes in a capsule network grows the performance of the network drops substantially. As the number of classes grows, analyzing capsule agreements gets more challenging. The network tends to over fit the data, lowering overall performance.

3.2.1 Capsule connections as feature extractor

For addressing the limitations, an enhancement in the CapsNets is proposed. Here, dynamic routing is utilized for mapping the input to an intermediate feature space rather than using it to build a class-specific output capsule. As a result, the dynamic routing stage is no longer affected by the number of classes, but by the number of features extracted. Instead of class-specific capsules, the capsule network in our proposed model produces feature-specific capsules. Dataset's samples can be represented by features that are lesser than the number of classes in the dataset. Hence, we can target higher-class datasets with fewer characteristics. Because most of the time spent executing the network is spent computing features using the capsule network, increasing the number of output classes will have less effect on the time it takes for network training. Moreover, by restricting the number of features, the network's generalization potential is strengthened, resulting in better performance. Equivariance among capsules is examined in the original capsule networks to analyze the agreement between them concerning the existence of any output classes. The agreement between capsules is calculated in this study based on the existence of some specified traits, where the output classes is assumed using fully connected layers.

3.2.2 Feature capsule layer

Here, intermediate feature capsule layers are utilized instead of final output capsule layers for creating feature capsules from the initial unique capsule. For inputs of size $N_{pc} \times 8$, feature capsule layers will create a feature capsule with $N_{features} \times 16$. Here, $N_{features}$ must be preferably lesser than N_{class} to increase the speed and to reduce the memory footprint compared to standard CapsNets. If a single capsule feature is denoted as v_f for $f \in \left[1, N_{features}\right]$, the yield likelihood for class j for $j \in [1, N_{classes}]$ will be calculated as in Eq. (7).

$$v_j^{'} = \frac{\text{exponential} FC_j\left(v_f\right)}{\sum_k \text{exponential} FC_k\left(v_f\right)} \tag{7}$$

where FC_j is the jth output of fully connected layers.

3.2.3 Regularization using feature capsules

The decoder, unlike traditional CapsNets, attempts to rebuild the input using feature capsules instead of output capsules. By stopping the network from over-fitting and limiting the reconstruction loss, this reconstructed image is employed as a regularization method. To ensure the margin loss is not dominant, the reconstruction losses are scaled down by the factor of 0.0005, similar to a typical capsule network.

3.2.4 The modified loss function

The margin losses for the classes are calculated similar to the initial capsule network. The boundary losses for class k are illustrated as represented in Eq. (8).

$$M_k = T_k(0, m^+ - v_k^{'})^2 + \lambda(1 - T_k)(0, v_k^{'} - m^-)^2 \tag{8}$$

Here, λ will be set as 0.5, the upper and lower bounds m^+ and m^- will be set as 0.9 and 0.1 and $T_k = 1$ if the object of class k is present, else $T_k = 0$.

From intermediate feature capsules, the decoder networks attempt to reassemble the original sample. The decoder, unlike capsule networks, receives the whole collection of feature capsules without any masking. The reconstruction losses are calculated as the simple mean square error among reconstructed images and input. Equation (9) gives the reconstruction loss.

$$R = MSELoss\left(I, I'\right) \tag{9}$$

where I' represents the reconstructed image and I signifies the input images. Conclusively, the net losses will be estimated as in Eq. (9).

$$L_k = M_k + \beta R \tag{10}$$

Here, β denotes the parameter for scaling downwards. It keeps the reform losses from outnumbering the border losses. The β value , is used as 0.0005 in our experiments, which the same as the values are taken in the original capsule network. An additional reconstruction loss is used in the modified capsule network technique to boost the routing capsules for encoding the input's instantiation parameters. During training, they suppress everything except the activity vector of the proper route capsule. The algorithm then uses these activity vectors to recreate the input images. The outputs of the routing capsules are given to three-layer decoders, which reduces the sum of squared discrepancies among logistic unit output and pixel intensities. The CapsNet approach includes additional parameters, like the number of channels in the capsule layers, the number of the primary and routing capsule, and the number of dimensions in each capsule type, compared to a traditional neural network. They mask all but the activity vectors of the correct route capsules during training. The algorithm then reconstructs the input image using the activity vector. Table 2 shows the parameter settings of the proposed approach used for experimentation.

4 Results and discussions

4.1 Experimental scenario

The classification of tumor stages is the proposed system which categorizes the original image into normal and abnormal. Here, the experimentations are completed on

Table 2 Parameter settings

Parameter	Value
Number of epochs	50
Number of iterations	500
Optimizer	Adam
Learning rate	0.0001
Batch size	16
Kernel size	9 × 9
Number of layers	5

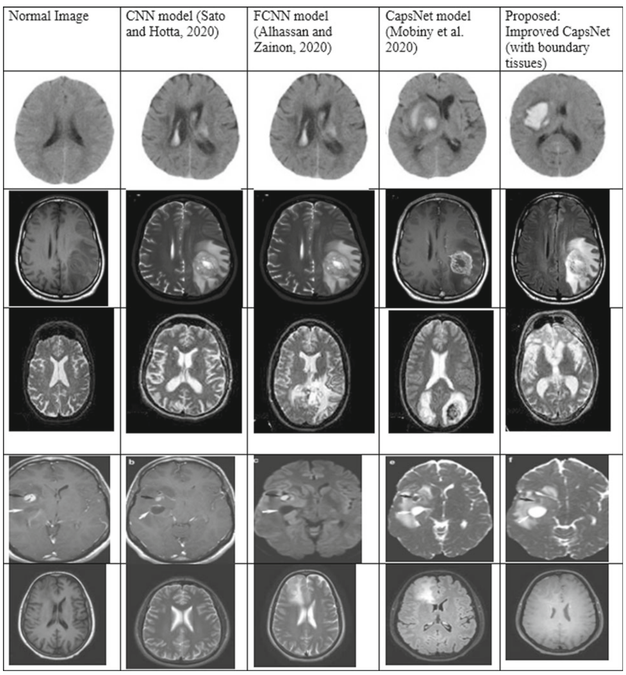

Fig. 2 Sample classification outcomes of the brain MRI images with NN, FCNN, CapNet and improved CapsNet

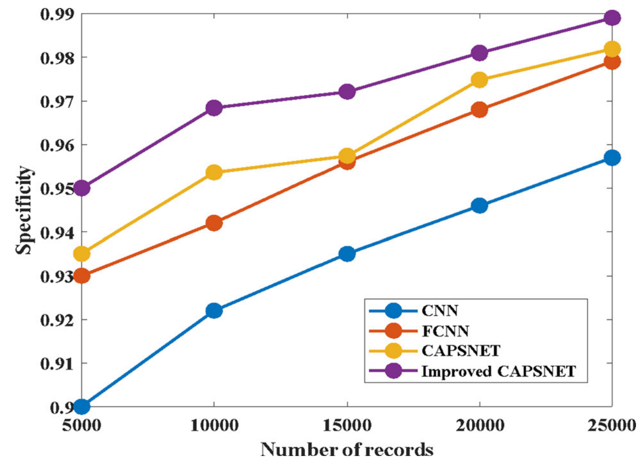

Fig. 3 Specificity analysis

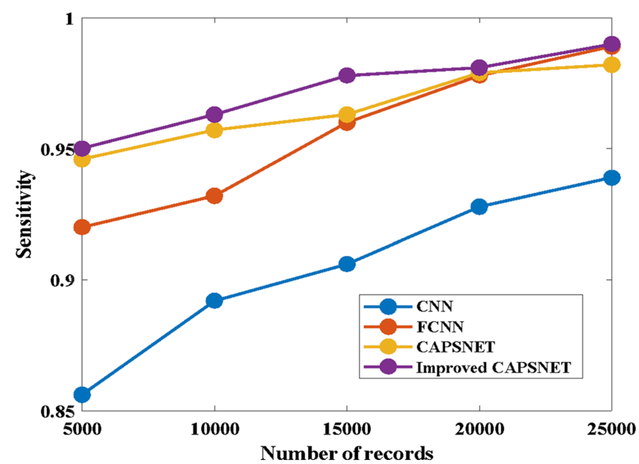

Fig. 4 Sensitivity analysis

medical datasets for evaluating the proposed improved CapsNet algorithm using parameters like specificity, sensitivity, and accuracy. Besides, the other existing algorithms like CNNs and Fuzzy CNNs (FCNN) are also compared with the proposed algorithm for analyzing the efficiency of the proposed algorithm. The MRI images are utilized for training, testing, and validation of the proposed algorithm from the BRATS data source (a benchmark dataset) with 25,000 images. Here, 75% of the images are utilized for training purposes and the remaining 25% of the images are utilized for testing.

Figure 2 demonstrates the sample classification outcomes of the MRI images when implemented using the different neural network-based models such as CNN, FCNN, CapsNet and Improved CapsNet. The proposed improved CapsNet showed better results considering coarse tissue boundaries. The high activation points are clearly obtained with the proposed approach.

4.2 Evaluation parameters

The parameters such as specificity, sensitivity, accuracy and F-measure, number of true positive (TP), number of false-positive (FP), number of true negatives (TN), and number of false-negative (FP) are considered for the performance analysis. The analysis is conducted for five different sets of images such as 5000, 10,000, 15,000, 20,000, and 25,000. Figure 3 shows that the specificity is high for the improved CapsNet than the other existing algorithms. The proposed improved CapsNet attains the highest specificity nearly 98.9% for 25,000 brains MRI images.

Sensitivity tests the ability of an algorithm to correctly detect ill patients who do have the condition. Figure 4 illustrates the analysis of sensitivity among the proposed algorithm and other existing algorithms such as CNN, FCNN, CapsNet, and improved CapsNet. The analysis is conducted for five different sets of images such as 5000, 10,000, 15,000, 20,000, and 25,000. The sensitivity for the improved CapsNet is higher than the other existing algorithms like CNN, CapsNet, and FCNN with 99%. Figure 5 illustrates the classi-

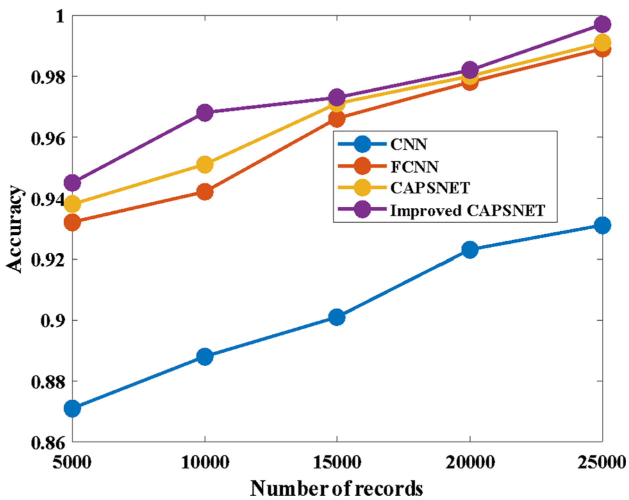

Fig. 5 Classification accuracy

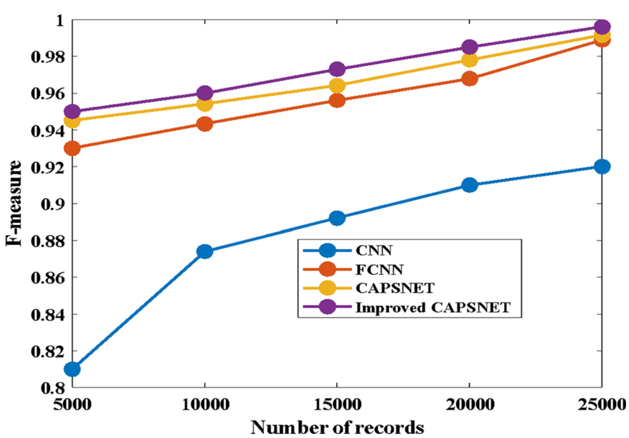

Fig. 6 F-measure analysis

fication accuracy among many traditional algorithms, CNN, FCNN, CapsNet, and improved CapsNet. From the experiments, it is observed that the classification accuracy is high for the improved CapsNet with 99.7%. Figure 6 illustrates the F-measure among the proposed algorithm and other existing algorithms such as CNN, FCNN, CapsNet, and improved CapsNet. It shows that the F-measure value is high for the improved CapsNet than the other existing algorithms like CNN, CapsNet, and FCNN with 99.6%. Computation time is defined as the length of time required to perform a computational process. Figure 7 displays the computation time between the improved CapsNets, CNN, CapsNet, and FCNN. It demonstrates that the computation time is less for improved CapsNets when compared to CapsNet and other considered methods.

4.3 Efficiencies of improved capsule network

Each capsule in improved CapsNets resembles the human brain in terms of processing information. They have a huge

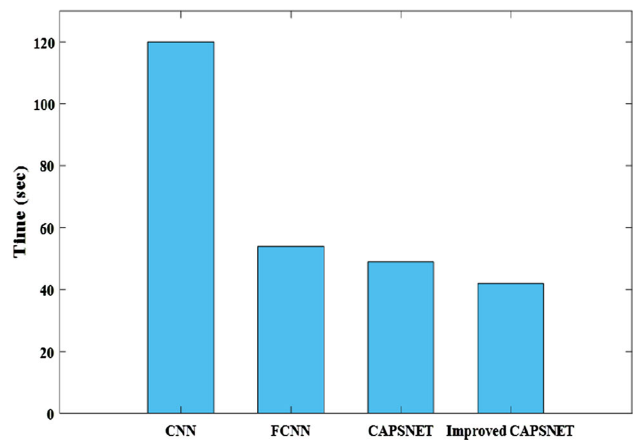

Fig. 7 Computation time

potential for recognizing even complex objects from images of poor quality. The learning ability of CapsNets is better than the CNN which divides the total images into small divisions and hierarchically relates them. Also, they represent the image with more details compared to CNN. When compared to CNN, the improved CapsNet significantly improves the performance of identifying overlapped pictures and sound. Hence, these improved capsule networks outperform the CNN in terms of detecting and quantifying structural damages, as the CNN is incapable of realizing object rotations and the existence of scaling within the object. Unlike CNN, the improved CapsNet will keep the object's location within the object. The capsule network's activation vectors are interpretable and perform better than CNN in classification accuracy using the MRI image dataset. Table 3 illustrates the performance matrices of the proposed improved CapsNet and other existing algorithms like CNN, CapsNet, and FCNN. From these table values, it is evident that the proposed improved CapsNet produces better sensitivity, specificity, F-measure, and accuracy. For 25,000 brain MRI images, the proposed improved CapsNet produces around 99.7% accuracy.

5 Conclusion

With the improvements in deep learning, convolutional networks have attained a lot of success. CNNs detect features in images and use this information to learn how to recognize objects. Layers near the beginning identify very basic elements like edges, whereas layers further down can detect more complex aspects in the images. It then makes a final forecast based on all the qualities it has learned. Even if they are successful, their core architecture has flaws that make them ineffective for some purposes. To address this flaw, capsule networks were developed, which fix the problem by

Table 3 Performance matrices of the proposed improved CapsNet and other existing approaches

No. of Images	5000	10,000	15,000	20,000	25,000
Specificity					
CNN	0.900	0.922	0.935	0.946	0.957
FCNN	0.930	0.942	0.956	0.968	0.979
CapsNet	0.935	0.953	0.957	0.974	0.981
Improved CapsNet	0.950	0.968	0.972	0.981	0.989
Sensitivity					
CNN	0.856	0.892	0.906	0.9278	0.9389
FCNN	0.920	0.932	0.960	0.978	0.989
CapsNet	0.946	0.957	0.963	0.979	0.982
Improved CapsNet	0.950	0.963	0.978	0.981	0.990
Accuracy					
CNN	0.871	0.888	0.901	0.923	0.931
FCNN	0.932	0.942	0.966	0.978	0.989
CapsNet	0.938	0.951	0.971	0.980	0.991
Improved CapsNet	0.945	0.968	0.973	0.982	0.997
F-measure					
CNN	0.810	0.874	0.892	0.910	0.920
FCNN	0.930	0.943	0.956	0.967	0.988
CapsNet	0.945	0.954	0.964	0.978	0.991
Improved CapsNet	0.950	0.960	0.973	0.985	0.996

implementing groupings of neurons that encode spatial information as well as the likelihood of an object being present. The probability of a feature appearing in the image is represented by the size of a capsule vector, and the direction of the vector represents the feature's posture information. It understands the prediction process by recreating the object and comparing it to the tagged examples from the training dataset. The vector representation of a feature in the capsules changes when it is moved around in an image, but not the chance that it exists. Lower-level capsules identify features, which are subsequently transmitted to higher-level capsules that fit them well. The capsule networks have higher classification accuracy than CNNs and FCNNs because they extract the necessary features before classification. As a result, capsule networks outperform convolutional neural networks in image classification.

References

1. Afshar, P., Mohammadi, A., Plataniotis, K.N.: BayesCap: a Bayesian approach to brain tumor classification using capsule networks. IEEE Signal Process. Lett. **27**, 2024–2028 (2020)
2. Afshar, P., Mohammadi, A., Plataniotis, K.N.: Brain Tumor type classification via capsule networks. In: 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 3129–3133 (2018). https://doi.org/10.1109/ICIP.2018.8451379
3. Afshar, P., Plataniotis, K.N., Mohammadi, A.: BoostCaps: A boosted capsule network for brain tumor classification. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1075–1079 (2020). https://doi.org/10.1109/EMBC44109.2020.9175922
4. Afshar, P., Plataniotis, K.N., Mohammadi, A.: Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries. In: ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 1368–1372 (2019). https://doi.org/10.1109/ICASSP.2019.8683759
5. Afshar, P., Plataniotis, K.N., Mohammadi, A. Capsule networks interpretability for brain tumor classification via radiomics analyses. In: 2019 IEEE International Conference on Image Processing (ICIP), pp. 3816–3820 (2019). https://doi.org/10.1109/ICIP.2019.8803615
6. Alhassan, A.M., Zainon, W.M.N.W.: BAT algorithm with fuzzy C-ordered means (BAFCOM) clustering segmentation and enhanced capsule networks (ECN) for brain cancer MRI images classification. IEEE Access **8**, 201741–201751 (2020). https://doi.org/10.1109/ACCESS.2020.3035803
7. Anupama, M.A., Sowmya, V., Soman, K.P.: Breast cancer classification using capsule network with pre-processed histology images. In: 2019 International Conference on Communication and Signal Processing (ICCSP), pp. 0143–0147 (2019). https://doi.org/10.1109/ICCSP.2019.8698043
8. Basu, A., Kaewrak, K., Petropoulakis, L., Caterina, G.D., Soraghan, J.J.: Modified capsule neural network (Mod-CapsNet) for indoor home scene recognition. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1–6 (2020). https://doi.org/10.1109/IJCNN48605.2020.9207084
9. Du, W., Sun, Y., Li, G., Cao, H., Pang, R., Li, Y.: CapsNet-SSP: Multilane capsule network for predicting human saliva-secretory proteins. BMC Bioinf. **21**(1), 237 (2020). https://doi.org/10.1186/s12859-020-03579-2

10. Gumusbas, D., Yildirim, T., Kocakulak, M., Acir, N.: Capsule network for finger-vein-based biometric identification. In: 2019 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 437–441 (2019). https://doi.org/10.1109/SSCI44817.2019.9003019
11. Hollósi, J., Pozna, C.R.: Improve the accuracy of neural networks using capsule layers. In: 2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI), pp. 000015–000018 (2018). https://doi.org/10.1109/CINTI.2018.8928194
12. Katarya, R., Arora, Y.: Study on text classification using capsule networks. In: 2019 5th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 501–505 (2019). https://doi.org/10.1109/ICACCS.2019.8728394.
13. Li, L., Huang, J., Cheng, Q., Meng, H., Han, Z.: Automatic modulation recognition: a few-shot learning method based on the capsule network. IEEE Wirel. Commun. Lett. **10**(3), 474–477 (2021). https://doi.org/10.1109/LWC.2020.3034913
14. Mazzia, V., Salvetti, F., Chiaberge, M.: Efficient-CapsNet: capsule network with self-attention routing. Sci Rep **11**, 14634 (2021). https://doi.org/10.1038/s41598-021-93977-0
15. Mobiny, A., Lu, H., Nguyen, H.V., Roysam, B., Varadarajan, N.: Automated classification of apoptosis in phase contrast microscopy using capsule network. IEEE Trans. Med. Imaging **39**(1), 1–10 (2020). https://doi.org/10.1109/TMI.2019.2918181
16. Patrick, M.K., Adekoya, A.F., Mighty, A.A., Edward, B.Y.: Capsule networks—a survey. J. King Saud Univ. Comput. Inf. Sci. (2021)
17. Poncelet, J., Hamme, H.V.: Multitask learning with capsule networks for speech-to-intent applications. ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8494–8498 (2020). https://doi.org/10.1109/ICASSP40776.2020.9053832
18. Quan, H., Xu, X., Zheng, T., Li, Z., Zhao, M., Cui, X.: DenseCapsNet: Detection of COVID-19 from X-ray images using a capsule neural network. Comput. Biol. Med. **133**, 104399 (2021). https://doi.org/10.1016/j.compbiomed.2021.104399
19. Sabour, S., Frosst, N., Hinton, G.E.: Dynamic routing between capsules. In: Advances in Neural Information Processing Systems 30. In Proceedings of the Annual Conference on Neural Information Processing Systems, NIPS, Long Beach, CA, USA, pp. 3856–3866 (2017)
20. Sato, T., Hotta, K.: CNN to capsule network transformation. In: 2020 Digital Image Computing: Techniques and Applications (DICTA), pp. 1–2 (2020). https://doi.org/10.1109/DICTA51227.2020.9363395
21. Sun, K., Yuan, L., Xu, H., Wen, X.: Deep tensor capsule network. IEEE Access **8**, 96920–96933 (2020). https://doi.org/10.1109/ACCESS.2020.2996282
22. Xi, E., Bing, S., Jin, Y.: Capsule network performance on complex data (2017) ArXiv abs/1712.03480
23. Xiong, Y., Su, G., Ye, S., Sun, Y., Sun, Y.: Deeper Capsule Network for Complex Data. In: 2019 International Joint Conference on Neural Networks (IJCNN), pp. 1–8 (2019). https://doi.org/10.1109/IJCNN.2019.8852020
24. Yao, H., Gao, P., Wang, J., Zhang, P., Jiang, C., Han, Z.: Capsule network assisted IoT traffic classification mechanism for smart cities. IEEE Internet Things J. **6**(5), 7515–7525 (2019). https://doi.org/10.1109/JIOT.2019.2901348