A Novel CapsNet based Image Reconstruction and Regression Analysis

Dr. Akey Sungheetha
Data Science SIG member,
Image Processing SIG member,
Computer Science and Engineering, School of Electrical Engineering and Computing,
Adama Science and Technology University, Adama, Nazret, Ethiopia.

Dr. Rajesh Sharma R
Image Processing SIG member,
Computer Science and Engineering, School of Electrical Engineering and Computing,
Adama Science and Technology University, Adama, Nazret, Ethiopia.

Abstract: In the field of image processing, all types of computation models are almost evolved to solve the issues through encoded neurons. However, compared with decoding orientation and regression analysis, still the doors are open due to its complexity. At present technologies uses two steps such as, decoding the intermediate terms and reconstruction using decoded information. The performance in terms of regression analysis is lagging due to the decoded intermediate terms. Conventional neural network models perform better in feature classification and representation, though the performance is reduced while handling high level features. Considering these issues in image classification and regression, the proposed model is designed with capsule network as an innovative method which is suitable to handle high level features. The experimental results of the proposed model are compared with conventional neural network models such as BPNN and CNN to validate the superior performance. The proposed model achieves better retrieval efficiency of 95.4% which is much better than other neural network models.

Keywords: Convolution Neural Networks (CNN), Back Propagation Neural Network (BPNN), Capsule Network, Image Classification, Regression

1. Introduction

Conventional deep learning architecture is a computation model which provides solution based on multiple layers to learn the data representation. Multiple abstraction levels are required to learn the data through the layers. Some of the familiar models in deep learning are Convolutional Neural Network (CNN), Deep Neural Network (DNN), Recurrent Neural Network (RNN). Each model attains its significance reach in various fields like image processing, signal processing, etc., Earlier, the neural networks models are used to compute the given problem through single layer architecture, but the performance of the system is not worthy for non linear classifications, later multilayer perceptions are evolved, but it has complexity in terms of data training and testing. Also, the computation time is unacceptable due to its long training process due to the optimal points. To overcome these issues in conventional neural networks deep belief networks are introduced which has more hidden layers to process the input. So that the input transformation is processed better to solve the linear and non linear problems and it consolidates the solution through its output layer. However, a greater number of layers needs more time to process complex classification task, so a pretraining process is introduced to reduce the issues over optimal points which accelerates the training process and the computation time is also reduced in the deep neural networks.

Convolution neural network is one of the familiar models in deep learning architecture due to its data processing ability. The multiple arrays used in CNN represents the features and produces an array of output to analyse the given issue. It is successfully applied in image processing and signal processing applications. The operation of CNN is simple and it obtains output based on convolution between the input array and array of weights which is mentioned as convolution kernel. The two-dimensional array in CNN is calculated based on below equation

$$C(x,y) = (h,k)(x,y) = \sum_i \sum_j h(i,j)k(x-i,y-j) \tag{1}$$

where h is the input array and k is the function of kernel and the convolution output is represented as C which describes the convolution process between the arrays. Figure 1 gives the pictorial representation of convolution process.
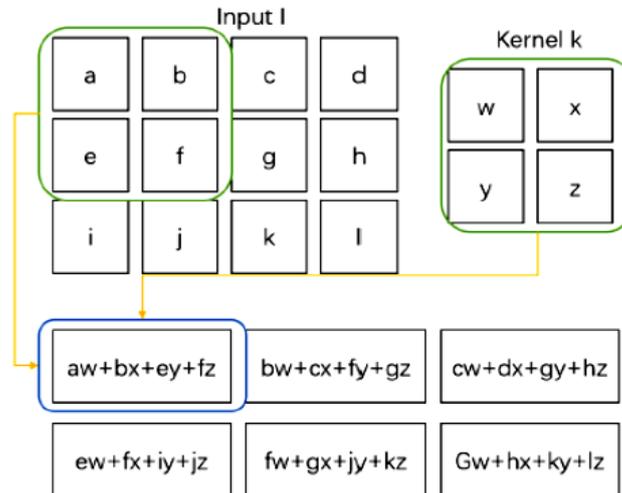

Fig 1. Convolution process in CNN

In general, the convolution neural network uses few layers to extract the input features from the data and the classifier layers are used to map the features. Based on the extracted feature and mapped features the output of the neural network is obtained. The classifier layers consists of several layers which are fully connected and the convolution operation is applied in each layer is a product of convolutional kernel and extracted features. The feature map consists of extracted features from a large pool of correlated values from the entire array structure used. Once the features are selected, the pooling layer merges the data semantically based on similarity which increases the reliability in CNN feature detection. Max pooling and average pooling are the common methods used in CNN which enables the output without changing the input. Even the slight change in the input which has variation in the output array. Figure 2 represents the layer by layer feature extraction process in CNN. Max pooling is used the process, in this once the features are generated from the layers, fully connected features are given into the system to generate the output.
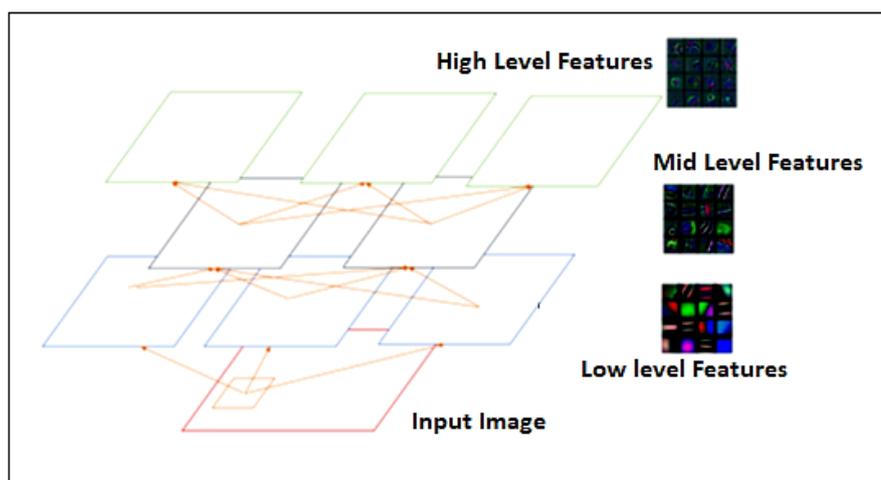

Fig 2. Feature extraction in CNN

The performance of convolution neural network is much better than other neural networks, though there are few limitations in CNN model. CNN provides improved results in image classification tasks but it provides inadequate results while handing high dimension data. On the other hand, CNN based data augmentation is used to capture the possible data and its transformations by adjusting the scalability of the local translations. Similarly, various applications are evolved based on convolution neural network though the research is still going on to obtain better model for obtaining optimal results.

2. Related works

The analyse the issues in conventional neural networks and advantages on capsule network which is used in the proposed model an intense literature survey has been made. Based on the issues observed, the proposed model is designed which reduces the limitations in the conventional model. Bashar *et al* [1] presented CNN based applications which are used in image processing and signal processing applications as a survey work. The key concept used in convolution neural networks and its pooling methods are discussed in the article. Also, the necessary implementations to obtain the kernel function and optimal points are analysed. Similarly, Vijayakumar *et al* [2], presented a survey which provides details about various deep learning models and its techniques. The proposed model summarizes the limitations of each model to obtain a clear idea about various deep learning methods used for image estimation and analysis.

Shaohui *et al* [3], discussed about LeNet which is one the popular model in CNN. The performance of LeNet model is evaluated in the research article for efficient face detection. The image processing applications based on CNN performs much better, though the proposed model in this architecture discussed about the same using LeNet which yields better performance compared to conventional CNN. Mostafa Elsaadouny *et al* [4] reported another feature extraction application using LeNet which uses two convolutional layers to obtain the feature information. The features are extracted from the existing features along with kernel function and it is connected to obtain the classified feature details in the proposed model. however, the performance is better, but the proposed model implementation cost increases while handling large feature data.

AlexNet is another CNN model which is evolved to train the kernel as well as the features. Rui Wang *et al* [5] reported the research work using AlexNet which uses five deep layers and three numbers of fully connected layer. It helps to reduce the vanishing issues in deep layer which is considered as a remarkable change in the image processing. Anju Unnikrishnan *et al* [6], reports the issues in image classification through deep AlexNet. The proposed model is developed considering the issues in training parameters. Selection of training parameters which reduces the system computation time and increases the efficiency. The proposed model is designed to evaluate the satellite images with the proposed reduced training parameters and achieves better classification accuracy. Siyuan *et al* [7], reported the issues in conventional brain image detection through AlexNet. The limitations in conventional models are used to design the proposed model to attain better classification accuracy in image analysis.

In CNN models while using deep layers the features are degraded due to its internal parameters. To reduce the gradient issues in deep layers, ResNet is proposed. Zhenyu *et al* [8], reported the gradient issued in CNN through ResNet. In this a residual block is created along with a bypass layer to reduce the vanishing effects, additionally the proposed mode attains better connectivity between the layers with residual blocks to provide improved performance compared to conventional models. Shuai Peng *et al* [9] reported the advantages of ResNet in image processing through face recognition application. The proposed mode gives details about the limitations about conventional CNN model and the advantages of proposed model through a real time experimentation process. The proposed model achieves better recognition results compared to other models. Shaopeng Liu *et al* [10] discussed about the advantages of ResNet through scene classification application. In the proposed model data augmentation with filter is used along with ResNet to achieve better classification results with minimum parameters. From the survey it is observed that, most of the CNN models are concentrated over the classification and feature extraction. Though the performance is satisfactory but it has some flaws in terms of optimal function, kernel function, gradient feature vanishing, increased training time and parameters. Considering these issues, the proposed model is designed to obtain better image classification accuracy with better reconstruction. It is observed from the survey most of the research works are not worked with regression analysis. So, the proposed model is designed a combined model to perform image reconstruction along with regression analysis.

3. Proposed Model

In convolution neural network, the features are presented as an array which are invariant to its locations. Using a suitable pooling layer, the features are selected and the system doesn't consider the location of the features. This process forces the system to focus over the features to obtain decision based on pooling layer and not on the locations. This leads into improper or incorrect classification results. To reduce the issues in convolutional neural network, capsule networks are introduced. It collects the features from the array and also obtains the information location to improve the classification performance. Capsule is a group of neurons in which the instantiation parameters are represented as activity vectors to specify the entity of an object. While in capsule networks those vectors are used to represent the features. The relative object is described based on a three-dimensional vector which includes the existence probability, relative size and height. Thus, capsule networks provide more information in a vector format so that better information decision could be obtained for the given issue. The information flow in the capsule network is performed using dynamic routing algorithm (DRA) and the information is transferred from low level capsule to high level capsule. DRA decides the flow of collected information from low to high level using proper capsule which provides equal weightage to all the capsules in the network. Figure 3 depicts the process in capsule network in detail.
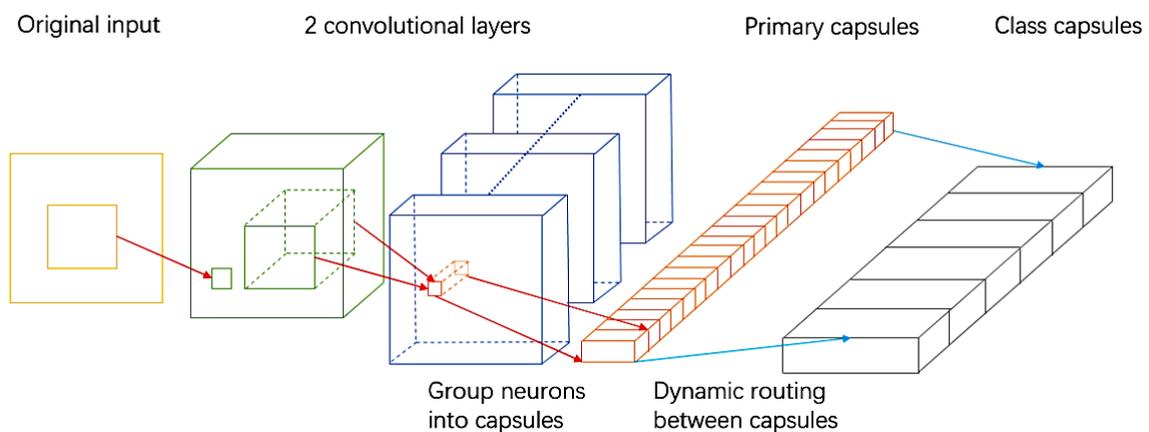

Fig 3. Capsule Network

When the information is need to be passed from low level to high level capsule, the high-level capsule need to agree with routing algorithm to allow the data transfer. In the routing process all the low-level capsules are allowed to send data to high level capsule with equal weights. In the high level the received information from all the low-level capsule are summarized and the decision is obtained. In which the high-level capsules are represented as rectangles and the low-level capsules are represented as dots. The cluster predictions of each capsule and other capsules are represented as coloured dots and each one is different from others. Based on the agreement the high-level capsules are used to predict the cluster with minimum iteration by increasing the weight. If the weights are reached into a threshold level the high-level capsules complete the process and further it doesn't accept any information from the low-level capsules and holds the iteration process. figure 4 depicts the information about capsule classes and its reconstructed information array in detail.
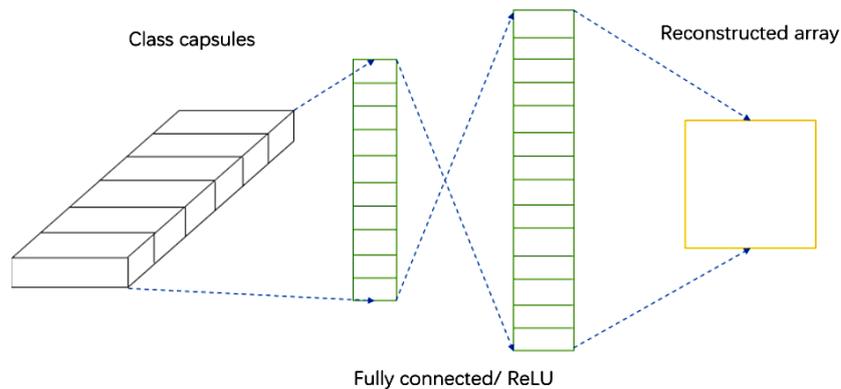
Fig 4. Array construction in capsule networks

Based on the observations in capsule networks the proposed model is designed for efficient image recovery and regression analysis. The process starts from selecting the raw input data and then it is pre-processed to remove the unwanted features and by using Fourier transform the image is converted as time frequency image which is used to compare the results of capsule network. Simultaneously the image is processed in capsule network and features are selected to reconstruct the image. Once the features are reconstructed using capsule network it is compared with actual time frequency graph to obtain correct results. Based on the obtained results the image features are classified and the processed for regression analysis. Figure 5 depicts the detailed process of proposed capsule network-based image classification, recovery and regression analysis.
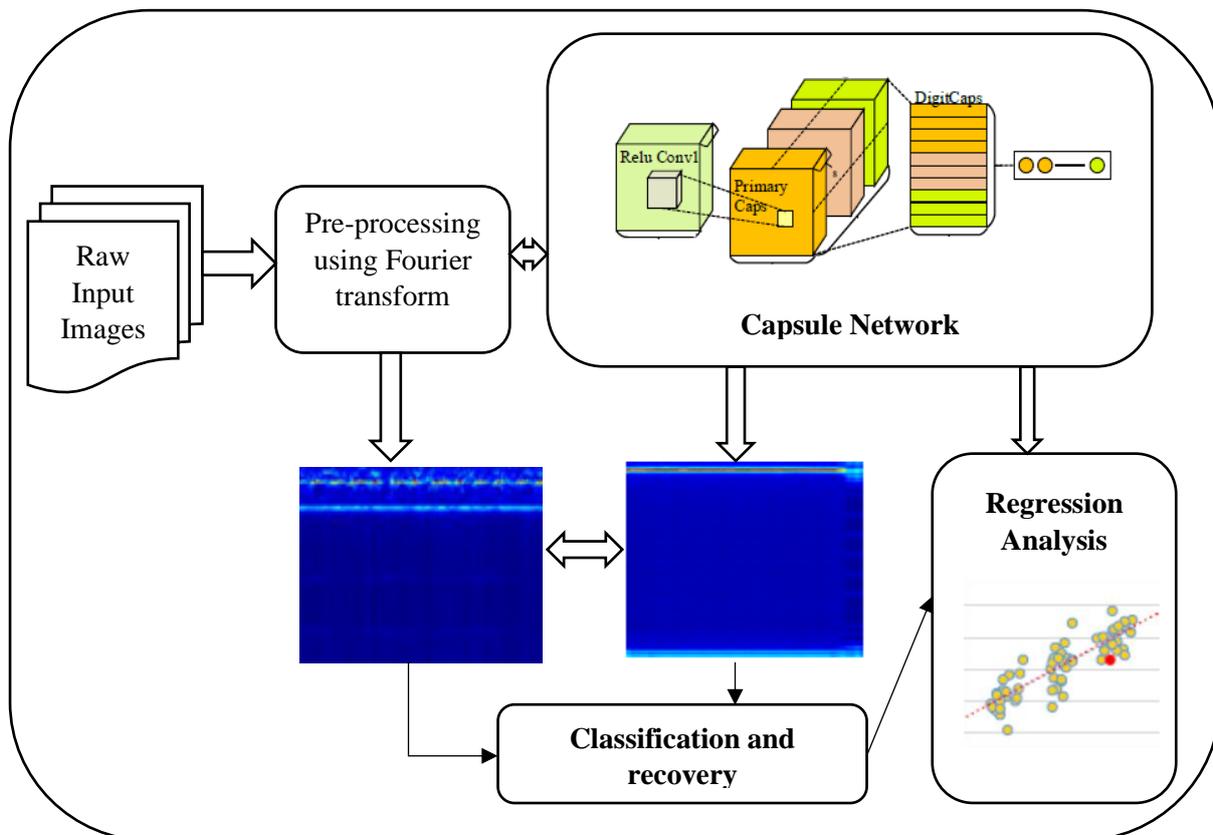

Fig 5. Proposed Recovery and regression model

160

The proposed model designed with Fourier transform function to covert the image into time frequency graph for that two important parameters are selected such as window length and step size. The values are selected to maintain high resolution output and it maintains the input raw data characteristics as well. Also, zero normalization is used to reduce the noise features also to preserve the original data distribution. From the process it is clearly visible that the relationship between the features are present in the image data labels. Conventional neural network like CNN concentrates over the specific features and ignores the relationship between them. This leads into incorrect classification in some cases so that the proposed model is implemented with capsule network to utilize the features effectively to reconstruct the image. The pseudocode of the proposed model routing algorithm is summarized as below

----

Initialize Routing ($\hat{h}_{y|x}\ m, n$)

For all capsule x in layer n and capsule y in layer $(n + 1)$

Set d_{xy} to 0

    for k in iteration do

        for all capsule x in layer n: Set e_{xy} to softmax(d_{xy})

        for all capsule y in layer (n+1) set l_n to $\sum_x d_{xy}\hat{h}_{y|x}$

        for all capsule y in layer (n+1) set l_m to squash(l_n)

    for all capsule x in layer y in layer (n+1): set d_{xy} to $d_{xy} + \hat{h}_{y|x}.l_m$

return l_m

----

4. Result and Discussion

The proposed capsule network is experimented and compared with conventional methods like convolution neural network and back propagation neural network. For intense analysis a dataset is generated with images captured in a smartphone and intentionally some regions are damaged to check the recovery performance of proposed model. Matlab 14.1 is used for experimentation which is installed over intel i7 processer with 8Gb of RAM. Performance of the proposed model is measured and compared in terms of classification accuracy, retrieval efficiency and overall regression analysis ratio. The parameters used in the proposed model is listed in table 1.

Table 1 Parameters used in the proposed model

S.No	Parameters	Value
1	Number of images in the data set	500
2	Batch size	100
3	learning rate	0.001
4	Number of epochs	4-8
5	Pooling	max
6	size of filer	100

The classification accuracy of the proposed model is depicted in figure 6. From the figure it is observed that the proposed model attains better classification accuracy than conventional models. The reduced performance in CNN is due to selection of features and the proposed model overcomes the limitation of CNN and achieved better classification accuracy.
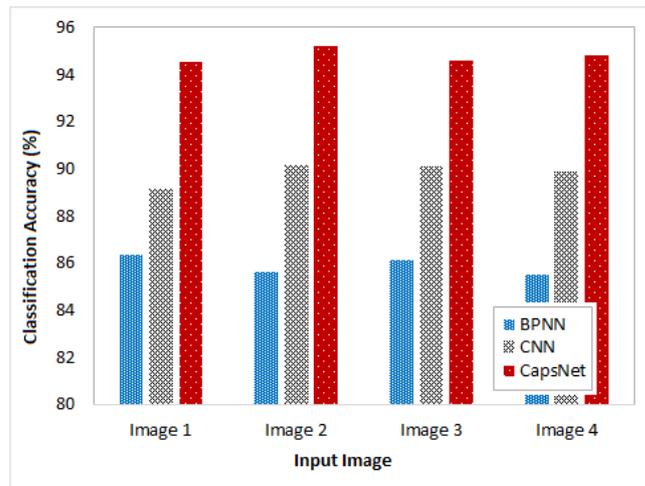
Fig 6. Classification Accuracy Comparison

Figure 7 depicts the regression analysis of the proposed model. The proposed model performs better analysis than other two models. The raw image used in the experimentation is damaged intentionally and the proposed model finds almost all the damaged portions in the image and based on this detail the images are reconstructed.
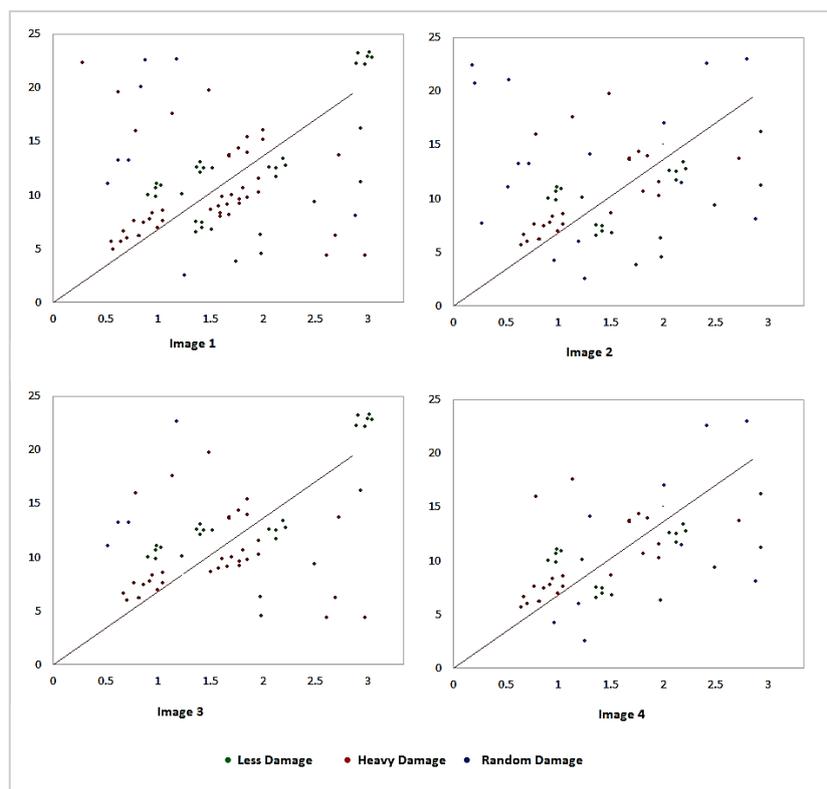

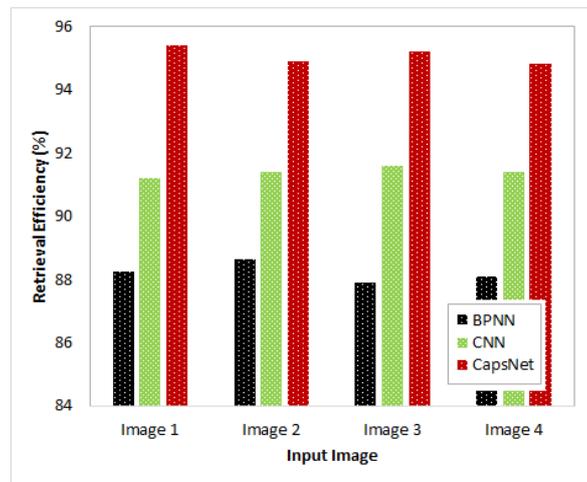**Fig 7. Regression analysis**

Fig 8. Retrieval Efficiency

The retrieval efficiency of the proposed model is depicted in figure 8. It is observed that the proposed model attains better retrieval efficiency. Though the conventional models are also performing well, but the performance of the proposed model is much better than convolution neural network and back propagation neural network.

5. Conclusion

A novel image reconstruction and regression analysis is presented in this research work. Feature selection and its issues in conventional neural network models are analysed to design the proposed model. The proposed model is implemented with capsule network to attain better performance in image reconstruction. Along with reconstruction, regression analysis is performed to improve the reconstruction efficiency of the proposed model. experimental results are compared with conventional neural network models such as convolution neural network and back propagation neural network to highlight the superior performance of proposed model. from the experimental results it is observed that proposed model performs better than conventional models.

Reference

1. Bashar, A. (2019). Survey on Evolving Deep Learning Neural Network Architectures. *Journal of Artificial Intelligence*, 1(02), 73-82.
2. Vijayakumar, T. (2019). Comparative Study of Capsule Neural Network in Various Applications. *Journal of Artificial Intelligence*, 1(01), 19-27.
3. Shaohui Lin, Ling Cai, Xianming Lin, Rongrong Ji (2016). Masked face detection via a modified LeNet. *Neurocomputing*. 218, 197-202
4. Mostafa Elsaadouny, Jan Barowski, Ilona Rolfes (2020). Extracting the Features of the Shallowly Buried Objects using LeNet Convolutional Network. *European Conference on Antennas and Propagation,*1-4
5. Rui Wang, Jingwen Xu, Tony X. Han (2019). Object instance detection with pruned Alexnet and extended training data. *Signal Processing: Image Communication*, 70, 145-156
6. Anju Unnikrishnan, Sowmya V, Soman K P (2018). Deep AlexNet with Reduced Number of Trainable Parameters for Satellite Image Classification. Procedia Computer Science, 143, 931-938
7. Siyuan Lu, Zhihai Lu, Yu-Dong Zhang (2019). Pathological brain detection based on AlexNet and transfer learning. *Journal of Computational Science*. 30,41-47
8. Zhenyu Lu, Yanzhong Bai, Yi Chen, Chunqiu Su, Shuihua Wang (2020). The classification of gliomas based on a Pyramid dilated convolution ResNet model. *Pattern Recognition Letters*. 133, 173-179

9. Shuai Peng, Hongbo Huang, Weijun Chen, Liang Zhang, Weiwei Fang (2020). More trainable inception-ResNet for face recognition. *Neurocomputing*. 411,9-19

10. Shaopeng Liu, Guohui Tian, Yuan Xu (2019). A novel scene classification model combining ResNet based transfer learning and data augmentation with a filter. Neurocomputing. 338, 191-206