Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 26 (2023), No. 5, pp. 1367–1379
DOI : 10.47974/JDMSC-1758

Transfer learning based novel intelligent classification for Alzheimer's Dementia using duplex convolutional neural network

Bali Devi [†]
School of Computer Science and Engineering
Manipal University Jaipur
Jaipur, Rajasthan
India

Sumit Srivastava [*]
School of Information Technology
Manipal University Jaipur
Jaipur, Rajasthan
India

Vivek Kumar Verma [§]
School of Information Technology
Manipal University Jaipur
Jaipur, Rajasthan
India

Gaurav Aggarwal [‡]
Department of Information Technology and Engineering
Amity University in Tashkent
Uzbekistan

Abstract

Goal: This work focuses on neuroimaging studies, including MRI analysis via machine learning and deep learning, which has led to an increase in computer vision research and the early detection of neural disorders. Methods: An adaptive implementation of transfer learning (TL) with a top-level VGG16 architecture is set up with pretrained weights for large MRI images dataset. Convolutional neural networks (CNN) are a bespoke version of a multi-

[†] *E-mail:* baligupta03@gmail.com
[*] *E-mail:* sumit.310879@gmail.com (Corresponding Author)
[§] *E-mail:* vermavivek123@gmail.com
[‡] *E-mail:* gaggarwal@amity.uz

layer view. Through experimentation on the ADNI dataset, the algorithm was trained and tested in binary and multiclass classification using the MRI scanning of individuals. Results: The machine was trained on the CNN model of binary classification and the CNN model for multiclass classification and was trained using TL (using the VGG16 model) with CNN; 25 epochs of batch size of 32 were considered. To validate the contributions of this study, we demonstrated that the proposed model used for binary classification gave an accuracy of 96.89% and an F1 score of 96.90%, and for multiclass classification, we obtained an accuracy of 99.89% and an F1 score of 94.82%. Conclusion: The suggested approach is highly generic, as it is simple and parameter invariant, and therefore applicable to any MRI dataset.

Subject Classification: *92C55, 68U10, 54H30, 14J10.*

Keywords: *Alzheimer's disorder, Binary classification, Convolutional neural network, Multiclass classification, Transfer learning.*

Abbreviations

Transfer Learning: TL
Support vector machines: SVM
Alzheimer's Disease: AD
Oral Contraceptives: OCs
Linear Discriminant Analysis: LDA
Convolutional Neural Networks: CNN
Computer-aided diagnosis: CAD

1. Introduction

AD is a neurodegenerative condition that leads to the permanent death of nerve cells in the brain [21]. These nerve cells are crucial for cognition and memory. Eventually, this disease leads to a loss in cognitive abilities that precedes the onset of dementia [7]. There is presently no treatment for Alzheimer's disease, and the findings implies that the underlying brain alterations responsible for the disease may really begin more than 20 years before symptoms manifest [9]. It is estimated that almost 5.7 million Americans take oral contraceptives, making AD the sixth leading cause of mortality in the country according to data and figures (OCs). Even though there is currently no medication for Alzheimer's disease (AD), researchers have agreed that doing so early on can have significant advantages. Various neurologists and medical researchers spend much time researching how to detect Alzheimer's disease early, and it continues to achieve promising results. Several studies use MRI data for classifying AD [1], and MRI plays a critical part in that classification. One in every 85 people will develop Alzheimer's disease by 2050. Automatic analysis of MRI scans may help detect AD in its early stages [17-19]. Computer diagnostics have recently been shown to be more accurate than clinicians in predicting Alzheimer's

disease, getting it a vital area of research for the field of computer diagnostics. Support vector machines (SVMs) are a significant new area of study for computer diagnostics [3, 21]. Early success with new deep learning techniques such as HR autosensing, CNN, and TL has an excellent statistical method, but what exists is a deep learning method that trains deep bottom-up architectures with some limitations, especially in medical imaging where medical authentication data can be exclusive and protected [8]. The use of cross-border agencies for ethics and confidentiality and training a deep network with many images requires a vast computational network. Deep network learning requires many parameters to be set carefully and tediously [1], and suboptimal tuning can lead to overfitting or underfitting and consequently poor performance. CNN training on these images increases fidelity and improves the classification of medical images. However, the trained CNN has some limitations from the start, one of which is the requirement for large amounts of data. Because of this, a different alternative approach known as TL may be utilized to tackle this problem. This method takes significantly less time and only needs a small data set.

TL is a machine learning technique that reuses a prebuilt network as a starting point for a model when the second task is performed. CNN has been demonstrated to be a very powerful DL model appropriate for network data such as RGB images and MR images. Due to AlexNet's great success in solving the natural image distribution problem, CNN's 18 applications quickly spread to a wide area. TL is useful for CNN-based research. TL involves using a model trained on a single problem as a starting point for solving related problems. TL is flexible, so pretrained models can be used directly, such as extracting preprocessing functions and integrating them into completely new models [4]. Keras has easy access to many of the most efficient models in ImageNet image detection projects such as VGG, Inception, and ResNet. The TL schedule for MCI classification is provided, which conveys the pressure experienced during AD versus NC as pMCI versus sMCI. We used gradients after sharing our imaging technique in the studied model and identified important biomarkers for much purer imaging results [11]. To the finest of our experience, this is the first successful CNN description of lifelong learning pMCI distribution. In this paper, we will examine how to use transfer learning to improve AD diagnosis. The most important motivation for using TL is to reduce dependency on large training sets. Reduce training sets using an intelligent classification approach to achieve the highest level of performance with smaller training sets [22].

The main results of our method can be reviewed as follows: We apply layer-by-layer TL transmission to modern CNN architectures where CNN upper layer groups are formed gradually, and lower layers remain fixed [5]. Using TL in this way is supposed to give distinct results, and the more levels you train, the further left from the finished network you will be. We found that retraining was only needed at a few top levels to achieve optimal results, which was very encouraging to reduce the training time required. We want to evaluate the robustness of the TL approach on a tiny training set. We will only use randomized training data. It instead utilizes the training data that offers the greatest information since images tend to be random and disordered.

2. Material and Methodology

A. *Material*

The method was developed and evaluated in binary classification leveraging the MRI scans of 450 Alzheimer's patients and 450 healthy controls using the ADNI dataset [2]. We used a small, irregular ADNI data set to multi classify Alzheimer's stages (500 nondemented patients, 448 very mild dementia, 179 mild dementia, and 32 moderate dementia).

B. *Methodology*

The overall methodology for binary and multiclassification is the initiation of the pretrained CNN architecture AlexNet and visual geometry group 16 with transfer learning. The model employs AlexNet, a multilayered neural network, to detect patterns in picture pixels with

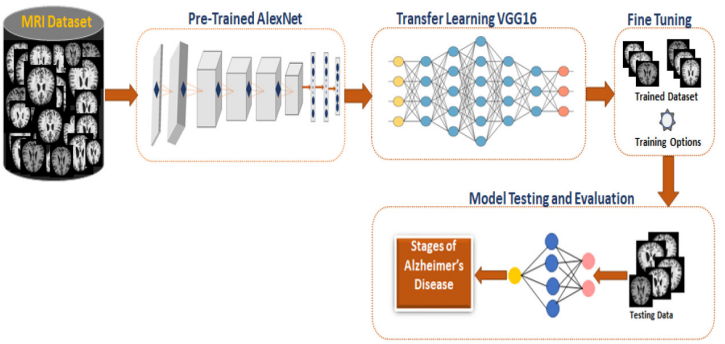


Figure 1

Proposed framework of our work

little to no preprocessing. By contrast, the 16th group of visual geometry transfer learning is utilized to remember the details of how a problem was solved and how that knowledge might be applied to a new problem. The overall proposed novel framework has been shown in Figure 1.

C. *Experimental Setup and Results*

A domain h is composed of two modules: the feature space denoted as D and the marginal probabilities associated with it, p (D), where D = {d1, d2..., dm} and m is the number of MRI images. Domain D, stated mathematically, is given by the formula: h = {D, P (D)}. The two feature spaces, as well as the marginal probability, would also be different for the two domains. In domain h, there are two components: the label space t, which is the task representation, and the objective predictive function f (), which is a representation of task n in that domain [3].

The formula in mathematical terms is n = {t, ()}. Features d are labeled t, and their relationship to the testing data is learned as part of the training phase using the predictive function f (). We encountered an instance in which just one source, h_s, and one destination, h_t, existed. The source MRI information is designated as follows in Eq. 1

$$h_s = \{(d_{s1}, t_{s1}), (d_{s2}, t_{s2}) \ldots\ldots (d_{si}, t_{si}) \ldots (d_{sm}, t_{sm})\} \quad (1)$$

This example shows an equation where the value after is called d_{si} and the value comparing it to is called t_{si}. In our proposed system, the two feature spaces and their corresponding labels are distinct. In a transfer learning model, a function from the source (f ())) is transferred to the target (f (d)). f (d) is computed based on the information assimilated from the source area and the source tasks and defined as f (d) = p (t | d). Traditionally, $h_s = h_{t,}$ and $n_s = n_t$ in machine learning methods. The source and target areas are different in the proposed framework, leading the modules to be different as well, suggesting either ds ≠ dt or p(ds) ≠ p(dt). The terms t_t and t_s are not identical when it concerns tasks. AlexNet is a CNN that has been pretrained on the AD images that make up the source domain h_s. AlexNet's CNN architecture has more than 60 million parameters [6]. It is risky to take in such a large range of features from training images in a simple way.

The gradient values of each parameter are then analyzed using the neural network chaining technique. The gradient values are used to adjust the parameters. The weight and bias of the updated networks are shown

in Eq. 2 and Eq. 3. Overall, convolutional layers are customized based on duplex architecture design, and they are presented in Eq. 4.

$$\Delta W_p(s+1) = -\frac{Y\beta}{n} W_p - \frac{Y\partial D}{m\partial W_p} + r\Delta W_p(s) \tag{2}$$

$$\Delta K_p(s+1) = -\frac{Y\partial D}{m\partial K_p} + r\Delta K_p(s) \tag{3}$$

Where W, K, p, β, Y, m, r, s and D are the weight, bias, layers, parameters, learning rate, training samples (AD and non-AD), buildup coefficient, step update, and cost analysis function, respectively.

$$Z = \sum_{Q=0}^{Q-1} \sum_{t=0}^{t-1} W_{Qt}\, Y_{r+1,\ m+Q} + b \ \ (0 \le r \le R, 0 \le m \le M) \tag{4}$$

Where W is the input layer of the Q * t dimension convolutional kernel layer, b is the biasness with the kernel, and Z is the feature space or

Algorithm 1: Multiclass Classification using duplex CNN

Input: R(f): Set of restored and enhanced images i.e. R(f)←F(m), I← individual MRI

E(f): Extracted features as per trained CNN network

Output: ND or VMD or MD or MD

```
PROCEDURE CNN (VGG16, AlexN, R(f), E(f), I)
START
    ND = Ø, VMD = Ø, MD = Ø AND MD = Ø ;
    REPEAT
        Addition = false;
        For all MRI Images R(f) do
            random select from training sample Ts;
            VGG16←{ Conv14, MaxPool5, FC- Dense-2);
            E(f) ϵ R(f);
            random select from test sample Tt ;
            VGG16←{ Conv14, MaxPool5, FC- Dense-2);
             AlexN←{ Conv6, MaxPool3, FC-2);
             If Ts & Tt = = R(f) then
                Addition =true;
             End If
        End For
    Until NOT (Additions);
    Count for ND or VMD or MD or MD
End CNN
```

map to produce the feature with the SoftMax function (Eq. 5) to create a probability distribution with V_0 outputs.

$$P_Q = \frac{e^{Y_Q}}{\sum_1^{V_0} e^{Y_{V_0}}} \forall Q = 1,2,3,\ldots..V_0 - 1, V_0 \tag{5}$$

Manual data preparation is performed by separating the data set into training and test set folders, each of which has four folders containing nondemented, very mild demented, mild demented, and moderately demented images (Algorithm 1). Algorithm 1 explains the multiclass classification with the R(f)←F(m) images in all MRI samples and processes it to CNN (VGG16, AlexNet, R(f), E(f), I) with the initially null ND, VMD, MD, and MD individuals. We have performed the customized duplex model of CNN with VGG16 ←{Conv14, MaxPool5, FC-Dense-2) and AlexN←{ Conv6, MaxPool3, FC-2) and as an outcome of our novel model in the form of number of ND, VMD, MD, and MD individuals.

3. Result and Discussion

In several attempts to diagnose AD, machine learning algorithms trained on MRI data have been tried and found wanting. Support vector machines are widely utilized because they formulate a sparsity-based, representation-based classification method. Other common approaches include logistic regression coefficients (like Lasso and Elastic Net), sparse recognition categorization (SRC), and random forest classifiers. Kloppel et al. [12] and others may use T1-weighted MRI images to detect AD patients. In order to analyze structural MRI data, Siddiqui. M, [13] employed dimensional reduction and variance approaches. They utilized a binary SVM classifier in addition to a multiclass classifier to identify AD in MRI scans. Using MRI, demographic, and genetic data, Vemuri et al. [14] built three classifiers to distinguish between patients with AD and healthy persons. Gray constructed a multi-modal classification model for AD diagnosis using MRI and PET data using a random forest classifier. AD was classified using a GLCM by Er et al. [15]. Morra et al. [16] examined the efficacy of several models for AD detection using hierarchical AdaBoost, manual SVM features, and automated SVM features.

A hierarchical structure is used in deep learning architectures, with basic, data-driven low-level features [10] serving as the foundation for more advanced, high-level features. As a result, deep learning models are being used to diagnose Alzheimer's disease and other brain disorders. In our approach, we used the novel concept of a duplex convolutional neural

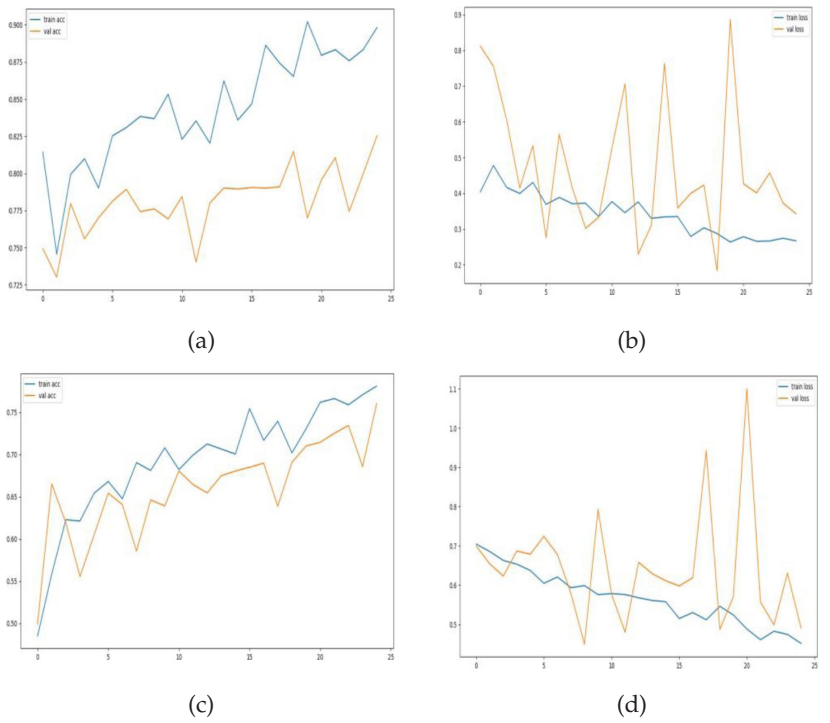
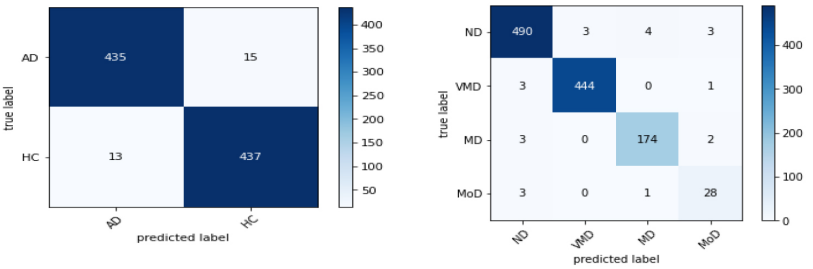


Figure 2

Our model performance (a) Plotting Accuracy -binary classification (training & validation/test set) (b) Plotting Loss -binary classification (training & validation/test set) (c) Plotting Accuracy -multiclass Classification (training & validation/test set) (d) Plotting Loss -multiclass classification (training & validation/test set).

network with customized VGG16 and AlexNet and evaluated the model with training and validation accuracy with loss.

The score is good, which was expected, as the dataset size is very small, and training was performed for only 25 epochs. On increasing epochs, the results were similar, as it was observed that the validation loss was not improving after a certain epoch. The model performed very well for binary classification with an accuracy of 82%, but for multiclass classification, the accuracy was 76% on the test set, as the data set was much lower and irregular. However, training these data was difficult and took considerable time to train even after reducing the batch size to 32 (Figure 2).



(a) (b)

Figure 3

Confusion Matrix (a) Binary classification (AD: Alzheimer Dementia, and HC: Healthy Control (b) Multiclass classification (MD: Mild Demented, MoD: Moderate Demented, ND: Non-Demented, VMD: Very Mild Demented)

This research evaluated the assessment scores, such as precision, recall, f1 score and FPR, via a confusion matrix. Various assessment criteria [19] are used to assess the accuracy of classification results produced via the customized VGG16 and AlexNet architecture. Each measure is described in more detail below. We have also presented the scores of all the assessment matrices in Table 1 by using the confusion matrix in Figure 3.

Precision: This is defined by [20] as the proportion of correctly predicted genuinely positive occurrences among all instances classified as positive. A mathematical formula is used to represent it in Eq 6.

$$Precision(Pn) = \frac{TP}{TP+FP} \tag{6}$$

Sensitivity–Recall: According to the definition, the fraction of true positive anticipated occurrences to all of the truly positive predicted cases in the ground data is 1. It provides information on the categorization enactment of occurrences that have been positively tagged. It is denoted by the mathematical formula in Eq. 7.

$$Sensitivity - Recall\ (Rn) = \frac{TP}{TP+FP} \tag{7}$$

Specificity: This is known as the proportion of true negative anticipated occurrences to the total of genuinely negative predicted individuals in the ground data. It provides information on the classification presentation of cases with negative labels. A mathematical formula is used to represent it in Eq 8.

Table 1
Evaluation assessment metrics results of binary and multiclass classification

Classification Type		Precision	Recall	Accuracy	Specificity	F1-Score
Binary Classification		96.68	97.11	96.89	96.66	96.90
Multiclass Classification	**ND**	98.19	98.00	99.80	97.00	98.10
	VMD	99.32	99.10	99.77	98.20	99.21
	MD	97.20	97.20	100	96.13	97.20
	MoD	82.35	87.20	100	86.50	84.78
	Overall	94.27	95.38	99.89	94.46	94.82

$$Specificity\ (Fn) = \frac{TN}{TN+FP} \tag{8}$$

F1-Score: An F1-Score is determined by taking the harmonic mean of the recall and accuracy scores. It is represented by a mathematical formula in Eq. 9

$$F1-Score(F1) = \frac{2*Pn*Rn}{Pn+Rn} \tag{9}$$

4. Conclusion

In this study, we used the novel intelligent classification for Alzheimer's dementia using a duplex convolutional neural network [23]. We have customized VGG16 and AlexNet as per the convolutional layer requirement and produced a duplex convolutional layer. For the features, we have automatically evaluated 2990 features and trained and tested the model using MRI images. For the classification, we used binary and multiclass classification and implemented our duplex CNN model with extracted features. We have taken AD and HC in binary classification, and in multiclass classification, we have taken ND, VMD, MD, and MoD. Regarding the results, we demonstrated that the proposed model used for binary classification accuracy of 96.89% and an F1-score of 96.90%, and for multiclass classification, we obtained an accuracy of 99.89% and an F1-score of 94.82%. In future, we will perform parameter optimization and augmentation to work with a larger MRI dataset.

Acknowledgments

It is an honor for us to offer our heartfelt gratitude to Manipal University Jaipur, India for supplying us with a research facility and

platform to perform experiments during the implementation of this work. We wish to extend our gratitude to the School of Computing and Information Technology, and CIDCR Lab, Manipal University Jaipur for support and motivation toward our research work.

References

[1] Sarraf, S.; Tofighi, G. DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI. bioRxiv 2016.

[2] Guerrero, R.; Wolz, R.; Rao, A.; Rueckert, D. Manifold population modeling as a neuro- imaging biomarker: Application to ADNI and ADNI-GO. NeuroImage, 94, 275-286 (2014).

[3] Plocharski, M.; stergaard, L.R.; Initiative, A.S.D.N. Extraction of sulcal medial surface and classification of Alzheimer's disease using sulcal features. *Comput. Meth. Biomech. Biomed. Eng*. 133, 35-44 (2016).

[4] Ben Ahmed, O.; Mizotin, M.; Benois-Pineau, J.; Allard, M.; Catheline, G.; Ben Amar, C. Alzheimer's disease diagnosis on structural MR images using circular harmonic functions Fdescriptors on hippocampus and posterior cingulate cortex. *Comput. Med. Imag. Grap.* 2015, 44, 13-25 (2015).

[5] A. Payan and G. Montana, ''Predicting Alzheimer's disease: A neuroimaging study with 3D convolutional neural networks,'' (2015), arXiv:1502.02506. [Online]. Available: https://arxiv.org/abs/1502.02506.

[6] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng, ''Early diagnosis of alzheimer's disease with deep learning,'' in Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI), pp. 1015-1018 (2014).

[7] V.L. Villemagne, S. Burnham, P. Bourgeat and B. Brown, "Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study," The Lancet Neurology, vol. 12, no. 4, pp. 357-367 (2013).

[8] S. Belleville, C. Fouquet, S. Duchesne and D.L. Collins, "Detecting early preclinical Alzheimer's disease via cognition, neuropsychiatry, and neuroimaging: qualitative review and recommendations for testing," *Journal of Alzheimer's Disease*, vol. 42, no. 4, pp. 375-382 (2014).

[9] J. Gaugler, B. James, T. Johnson, A. Marin and J. Weuve, "2019 Alzheimer's disease facts and figures," Alzheimers & Dementia, vol. 15, no. 3, pp. 321-87, Mar. (2019).

[10] Erhan, P.-A. Manzagol, Y. Bengio, S. Bengio, and P. Vincent, ''The difficulty of training deep architectures and the effect of unsupervised pre- training,'' in *Artificial Intelligence and Statistics*. USA: JMLR, 2009, pp. 153-160 (2009).

[11] K.H. Jin, M.T. McCann, E. Froustey, "Deep Convolutional Neural Network for Inverse Problems in Imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509-4522 (2017).

[12] Kloppel, S., Stonnington, C. M., Chu, C., Draganski, B., Scahill, R. I., Rohrer, J. D., … Frackowiak, R. S. J. Automatic classification of MR scans in Alzheimer's disease. Brain, 131(3), 681-689 (2008). doi:10.1093/brain/awm319.

[13] Siddiqui, M., Mujtaba, G., Reza, A., & Shuib, L. Multi-Class Disease Classification in Brain MRIs Using a Computer-Aided Diagnostic System. Symmetry, 9(3), 37 (2017). doi:10.3390/sym9030037.

[14] Vemuri, P., Gunter, J. L., Senjem, M. L., Whitwell, J. L., Kantarci, K., Knopman, D. S., … Jack, C. R. Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies. NeuroImage, 39(3), 1186-1197 (2008). doi:10.1016/j.neuroimage.2007.09.073.

[15] Amulya E R1, Soumya Varma2, Dr Vince Paul, "Classification of Brain MR Images using Texture Feature Extraction" *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 6, Issue 4, April (2017).

[16] Morra, J. H., Zhuowen Tu, Apostolova, L. G., Green, A. E., Toga, A. W., & Thompson, P. M. Comparison of AdaBoost and Support Vector Machines for Detecting Alzheimer's Disease Through Automated Hippocampal Segmentation. IEEE Transactions on Medical Imaging, 29(1), 30-43 (2010). doi:10.1109/tmi.2009.2021941.

[17] Shankar, V. G., Sisodia, D. S., & Chandrakar, P. An efficient MR images based analysis to predict alzheimer's dementia stage using random forest classifier. *Information Systems and Management Science*, 104-119 (2022). https://doi.org/10.1007/978-3-031-13150-9_9.

[18] Shankar, V. G., Sisodia, D. S., & Chandrakar, P. A novel discriminant feature selection–based mutual information extraction from mr brain

images for alzheimer's stages detection and prediction. *International Journal of Imaging Systems and Technology*, 32(4), 1172-1191 (2021). https://doi.org/10.1002/ima.22685.

[19] Devi, B., Srivastava, S., & Verma, V. K. Predictive analysis of alzheimer's disease based on wrapper approach using SVM and KNN. *Information and Communication Technology for Intelligent Systems*, 707-715 (2020). https://doi.org/10.1007/978-981-15-7062-9_71.

[20] Bhardwaj, S., Arora, D., Devi, B., Shankar, V. G., & Srivastava, S. Machine Learning assisted binary and multiclass parkinson's disease detection. *Intelligent Sustainable Systems*, 191-206 (2022). https://doi.org/10.1007/978-981-19-2894-9_15.

[21] Shankar, V. G., Sisodia, D. S., & Chandrakar, P. DataAutism: An early detection framework of autism in infants using Data Science. Data Management, Analytics and Innovation, 167-178 (2019). https://doi.org/10.1007/978-981-13-9364-8_13.

[22] Rawat, A., Mishra, S., Sharma, Y., & Khetarpal, P. High accuracy multilayer autoencoder trained classification method for diagnosis of Parkinson's disease using vocal signals. *Journal of Information and Optimization Sciences*, 43(1), 93-99 (2022, January 2). https://doi.org/10.1080/02522667.2022.2036355.

[23] Shih, W. F., Lin, C. W., Wang, W. F., & Wu, H. H. Association rule mining of care targets from hospitalized dementia patients from a medical center in Taiwan. *Journal of Statistics and Management Systems*, 21(7), 1299-1310 (2018, October). https://doi.org/10.1080/09720510.2018.1498189.