Journal of Scientific & Industrial Research
Vol. 81, December 2022, pp. 1276-1284
DOI: 10.56042/jsir.v81i12.69309

Cascade Network Model to Detect Cognitive Impairment using Clock Drawing Test

Sri Lakshmi Talasila* & Vijaya Kumari R
Computer Science, Krishna University, Machilipatnam, 521 004, Andhra Pradesh, India

Received 03 August 2022; revised 27 September 2022; accepted 06 October 2022

The Clock-Drawing Test (CDT) is commonly used to screen people for assessing cognitive impairment. Diagnoses are based on analyzing the specific features of clock drawing with pen and paper. The manual interpretations and understanding of the features are time-consuming, and test results highly depend on clinical experts' knowledge. Due to the impact of smart devices and advancements in deep learning algorithms, the necessity of a consistent and automatic screening system for cognitive impairment has amplified. This work proposed a simple, fast, low-cost, automated CDT screening technique. Initially, transferred deep convolution neural networks (ResNet152, EfficientNetB4, and DenseNet201) are used as feature extractors. The transfer learning technique makes it possible to experiment with existing models and build models much more quickly. Further, the extracted features are cascaded into a feature fusion layer to improve the quality of learning features, and the obtained feature vector become input for the classifier for classification. The performance of the model is experimentally evaluated and compared with the existing state-of-art models on a real dataset. Obtained results demonstrated that the Cascaded Network Model achieves high performance with an accuracy of 97.76%.

Keywords: Automated CDT screening**,** Convolution neural network, Deep learning, Dementia, Feature fusion

Introduction

The prevalence of cognitive impairment is rising as the world's population ages. Parkinson's disease, Alzheimer's Disease (AD) and Vascular Dementia (VaD) are all neurodegenerative disorders that cause memory loss. Dementia is when an adult experiences memory loss and cognitive impairment. Lack of remembrance, attention, language, executive function (managing day to day work and life), perceptual-motor abilities (interacting with the environment), and social cognition (interacting with other people) are all causes of cognitive impairment.

Detecting the cause of cognitive impairment at a primary stage has several benefits. In western countries, over 5 million people have AD, and deaths have risen by 16% during the COVID-19 pandemic. By 2050, the expense of Alzheimer's Disease and Related Dementia (ADRD) is expected to raise approximately $1.1 trillion.[1] In India, more than 5 million individuals have Dementia, which is expected to rise to 7 million by 2030.[(2)] Throughout the world, over 50 million people are assumed to be affected by dementia.[3] The traditional method of determining an individual's degree of cognitive decline is to give neuropsychological tests, commonly done via personal conversations, to check memory, reasoning, language, and motor function.

Clock Drawing Test (CDT) is the generally used test for screening for dementia.[4] During a CDT, a paper is given to each person to draw a clock. Then each participant must draw a circle, clock hands, and numerals from 1 to 12 corresponding to a specific time set, as '11:10', represented by the numbers 11 and 2. In actuality, the instruction to draw '10 past 11' might confuse people, drawing one clock hand on 11 and the other on 10 instead of 2. This entails the person placing their hands on the left and right sides of the circle's upper quadrants, imposing additional demands on the frontal lobe's executive functions. Furthermore, the practice has revealed that this environment is the most susceptible to neurocognitive dysfunctions and can accurately understand the characteristics of cognitive impairment cases by analyzing the CDT images. Detection of cognitive impairment by capturing the images of clocks drawn by geriatrics is advantageous as it detects at the very early stage itself, i.e., when symptoms appear in the adults and whether there is a need to consult the concerned doctors or not. Based on the deviation of the drawing from the correct answer, predictions can be made on the extent of disorderliness to normal brain function. As the test is cost-effective and requires minimal training, countries like India with low resources for healthcare can use this test.

*Author for Correspondence
E-mail: yarlagadda.srilakshmi@gmail.com

Many studies have recognized the value of CDT as a screening test for global cognitive deficiencies due to its ease of use, absence of language and cultural biases, and brief administration time for patients with cognitive function and various types of Dementia. In most experiments, many features are available through neuropsychological tests, blood biomarkers, and medical images. All of these tests demand time, cost, and other in-person appointments.

Technology is to test and detect cognitive impairments in adults at early stages without human involvement. Digital systems powered by Artificial Intelligence (AI) can help with patient diagnosis and treatment. More importantly, AI-driven tools can predict possible health issues that an individual may face in the future based on their historical data. Researchers utilize Machine Learning (ML) and Deep Learning (DL) methods to predict cognitive dysfunction and dementias early. DL differs from conventional ML methods in that it does not require meticulous feature extraction. The Convolutional Neural Network (CNN) deep learning model includes many hidden layers and more parameters and can be used for the early prediction of ADRD with more accuracy.[5] Determining whether a person has cognitive impairment can be modelled as a classification problem to arrive at a binary decision.

In this study, we propose a cascade network model that focuses on extracting features from three pre-trained models and cascading them into a feature fusion layer to complete the final classification. Several CNN models were analyzed as part of our work.ResNet152 and DenseNet201 are the deeper networks with more convolution layers that can capture richer and more complex features. ResNet152 is a deep CNN with skip connections and identity blocks that addresses the vanishing gradient problem without compromising the model's generalization power. DenseNet201 is a deep CNN that connects each layer to every other layer in a feed-forward fashion. The core of this network is that it decreases the vanishing-gradient problem, braces feature propagation, boosts feature reuse, and specifically reduces the number of parameters. EficientNetB4 is a CNN that uses a compound scaling method. The method reduces parameter size and FLOPS by order of magnitude and achieves higher accuracy and better efficiency. Considering the characteristics of each of the above models in extracting the features, ResNet152, DenseNet201, and EfficientNetB4 models were selected. A feature fusion was carried out on the features extracted by these models to fine-tune the model. Pre-training the above models using a transfer learning approach is considered to learn generic image features without training from scratch. The model provides a low-cost screening method that accurately detects cognitive impairment using CDT images.

Literature Survey

CDT is the prevalent and widely used neuropsychological test for detecting cognitive impairment. For the past few years, patients with Dementia and Mild Cognitive Impairment have been examined for different forms of performance on the CDT to expose the benefits of the test as a screening tool.

With the evaluation score of the Mini-Mental State Examination and the scale of Hasegawa dementia, the CDT measures each degree of global cognitive dysfunction, making it useful for cognitive impairment screening.[6] A low CDT score has also been associated with the development of dementia[7,8], and the connection is independent of the MMSE score. The CDT detects cognitive dysfunction in numerous psychological syndromes and how effective it is as a supplement to the MMSE in detecting executive dysfunction in a busy outpatient department. A weakened clock drawing provides a gesture of cognitive impairment that leads the analysts to understand better the patient's current level of symptoms.[9,10] The digital CDT (dCDT) technology uses a digital pen to trace the drawing. It supports real-time analysis of complex neuropsychological behaviour, which is hard to achieve with a standard pen and paper.[11]

The ML algorithms analyze and classify the group of dementia people and healthy people, and AD with VaD that used time-based, dynamic, and visuospatial features extracted from the dCDT, which provides the neurocognitive features.[12] The author explored how well ML data analysis algorithms classify patients into groups based on statistically determined variables.[13]A supervised learning model was built for predicting the state of cognitive impairment by collecting the trajectory points of dCDT using electromagnetic tablets.[14]

ML techniques can be used for representing differences in several key graphomotor errors in the CDT inside a single differentiable non-linear 2D manifold to quickly identify features seen in dementia patients.[15] ML algorithm's ability to distinguish across patient groups may increase the CDT's application in primary care and other healthcare settings for cognitive impairment screening. DL models have shown higher accuracies for typical ML applications. The features are trained to find and recognize the most appropriate patterns. In modern healthcare, DL is a well-known research subject, with specific applications attaining the level of a product.

MRI and PET scans are used to diagnose Alzheimer's using a robust deep learning framework[16] by incorporating a stability selection technique, an adaptive learning factor, and a multi-task learning strategy into the framework. An ensemble model is developed to screen cognitive impairment with CDT images and the age of the person.[17] A novel automated and qualitative scoring approach based on mobile sensor data and deep learning algorithms differentiates Dementia and subtypes with the CDT mobile application.[18] CDT and Rey–Osterrieth Complex Figure Test are screening tests to predict cognitive impairment using CNN Algorithms.[19] The cognitive tests using digital pen features are independent of the task. This set of features was shown to outperform traditional task-dependent classification methods on CDT and others.[20] The deep learning architecture employs a CNN and applies gradient-weighted class activation mapping to visualize the features within CDT images to predict participants' cognitive status.[21]

A novel DL model, CNNs, and ensemble learning are applied to acquire anatomical MRI of the brain with the grouping of AD vs Healthy and Mild Cognitive vs not Mild Cognitive by identifying the complex changes with AD.[22]

The literature shows that different CNN models were used to extract features from the CDTs. In the proposed model, distinct features extracted from the three pre-trained networks were concatenated, inspired by the success of exploiting multiple features for classification.[23,24]

Methodology

Detecting cognitive impairment requires several phases: dataset collection, pre-processing, model training (Transfer Learning, Cascade Network Model), and model evaluation. The proposed Cascade Network Model architecture is depicted in Fig. 1.

Data Pre-processing

The quality of the input image is crucial for training the model to achieve better accuracy. Before considering the training data as inputs, the dataset is applied to normalization and pre-processing. The original CDT images are pre-processed using widely-known pre-processing techniques like position translation, rotation, cross-cutting transformation, scaling, size adjustment, and flip processing.

Data Augmentation

Data augmentation is a method of adding new data points to existing data to enhance the amount of data. To amplify the dataset, this may involve making minor adjustments to the data or applying deep learning models to create extra data points in the latent space of the original data. It lowers the operating costs connected with data collection. It increases the generalizability of an over-fitted data model by producing more training data and exposing the model to different versions of the data. Data augmentation helps to solve the class imbalance

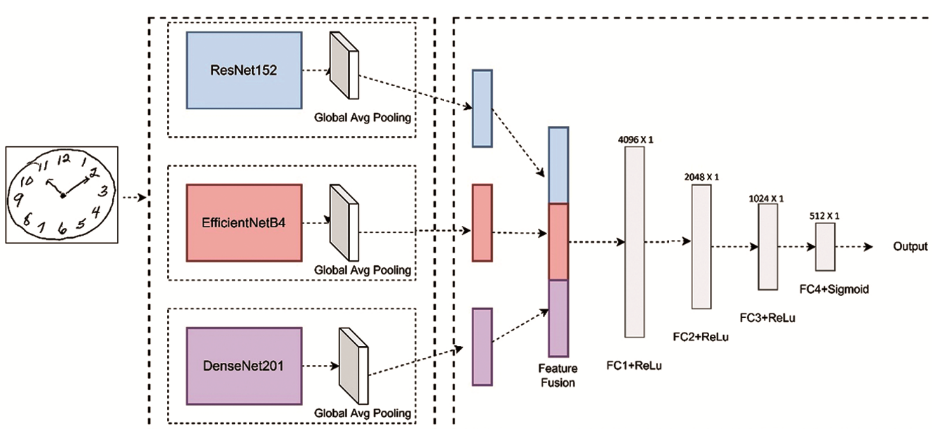

Fig. 1 — Cascade network model architecture

problem in classification, prevents data scarcity, and adds variety and flexibility to data models.

Transfer Learning

It is a common and preferred method to create new ML and DL models. DL models have a sophisticated data-dependent architecture, requiring more data to train. Transfer learning requires information from related tasks with a set of labelled data and applies to a target classification problem to decrease the quantity of training data required. Models are pre-trained before being adapted to specific applications. Pre-trained bottom layers are often used, and the top layers with labelled information are deleted. This approach allows getting a model much faster and experimenting with the existing models differently.[25] For experimentation we used three pre-trained models such as ResNet152[(26)], EfficientNetB4[(27)], and DenseNet201.[(28)] In their original forms, the models were designed to identify 1000 classes of images in the Image Net data set and reach state-of-the-art performance. We adapt these three models to our data set.

Feature Extraction

Feature extraction enhances the accuracy of learned models by extracting corresponding features from the input data. Redundant data is eliminated in this phase of the general framework, and the dimensionality of the data is lowered. Of course, it increases training and inference speed. The feature extraction methods generate new features by combining and transforming the original feature set. The features from the images are extracted using many CNN models.[29] The state-of-the-art models ResNet152, EfficientNetB4, and DenseNet201 were selected for feature extraction.

The ResNet unit was used to train ResNet, a 152-layer neural network, successfully. ResNet significantly improved the performance of neural networks with additional layers than neural networks with simple layers. Although it has fewer parameters than VGGNet, it gives good results. It introduces residual learning, which successfully addresses network degeneration. The rate of error is 3.57%. DenseNet establishes a link between layers, maximizes the use of features, and reduces the gradient. A DenseNet connects all layers directly to one another using Dense Blocks and establishes dense connections between layers. Using a bottleneck layer, translation layer, and lower growth rate shortens calculation time, narrows the network, decreases the parameters and fully controls the overfitting. EfficientNet is a convolutional neural network that uses a design and scaling method using a compound coefficient to scale all depth/width/resolution dimensions evenly. When the input image is larger, the compound scaling method is used , as the network requires more layers to expand the receptive field and more channels to catch more fine-grained patterns,

Global Average Pooling

Compared with the traditional CNN model, the proposed Cascade Network Model used Global Average Pooling (GAP) layer rather than fully connected layers before the classification layer. The GAP layer acts as a link between the convolution and fully connected layers without requiring more parameters to be trained. Each feature map extracted from the preceding convolutional layer is input to the GAP layer to obtain a single input by averaging it. The resultant feature maps are inputs to flatten layer, i.e., the multi-dimensional input into a single-dimensional feature vector as output. For example, if given an input of 299, 7*7 feature maps, a GAP layer forms an output of size 299. This method introduces generalizability and allows for a more progressive reduction in network size. So, it regularizes the network model to prevent overfitting.

Feature Fusion Layer

Feature fusion is a dynamic exploration area that offers the best accuracy compared to distinct manually selected feature sets. It learns image features entirelyto describe their rich internal information.After dimensionality reduction, we can acquire a compact representation of integrated features, which reduces computational complexity and improves recognition performance in an unconstrained environment. The flattened output features of each model from the GAP layer were concatenated into a flattened vector in this layer.

Classifier

In the classifier, the first layer contains hidden units of 4096, followed by the ReLU activation function. The second layer has hidden units of 2048, continued with the ReLU, the third layer with 1024 units, succeeded with ReLU, and the last layer comprises 512 units, with the output layer as one unit with a sigmoid activation function.

Experimentation

Experiments of the model were carried out on the server kaggle, with Tesla P100-PCIE-16 GB GPU graphics card and the python language with PyTorch framework.

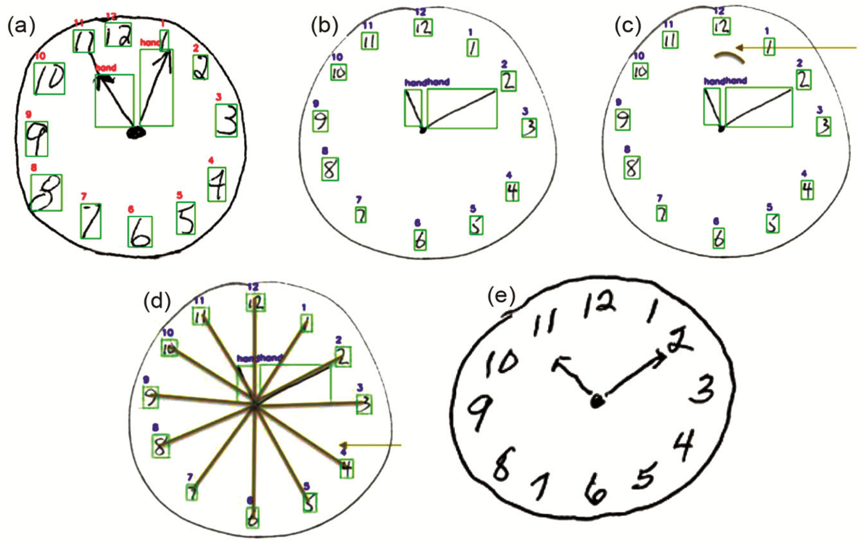

Fig. 2 — (a) Center Dot (b) Clock face (c) Clock hands (d) Digits (e) perseverance

Dataset

The CDT image has five key features: a centre dot, clock face, clock-hands, digits, and perseverance which is shown in Fig. 2.

- Center Dot: The clock image is thoroughly checked to determine whether a dot is present at the intersection point of the two hands and located at the clock's centre or not, as shown in Fig. 2(a).
- Clock face: Lengths of vertical and horizontal axes, areas of the left half, right half, the upper half and lower half of the clock face are carefully examined to explain how disproportionate the clock face is, i.e., how far it deviates from a regular circle, is shown in Fig. 2(b).
- Clock hands: The presence of hour and minute hands, their location, length, and proximity to digits 11 and 2 are examined, as shown in Fig. 2(c).
- Digits: The presence of 12 numbers, their orientation, and location, clockwise or anti-clockwise sequence are examined and are shown in Fig. 2(d).
- Perseverance: Recognition of perseverance while drawing clock-face, hands, centre dots, or digits, measured as an increase in thickness around the clock element (such as hand/digit) relative to normal, shown in Fig. 2(e).

These five features significantly value an individual's ability to diagnose the problem and its progression, making evaluation forecasts and treatment plans. The dataset is collected from the National Health & Aging Trends Study, USA[30], and the class labels are generated and verified by a professionally trained clinical expert. The sample CDT images in the dataset are shown in Fig. 3.

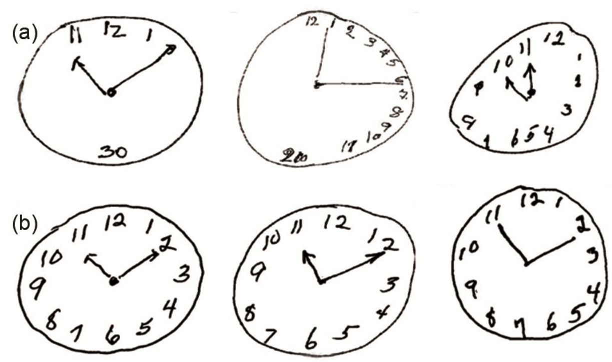

Fig. 3 — Sample CDT images (a) A CDT image of a Cognitive impaired person (b) A CDT image of a normal person

Table 1 — The CDT dataset distribution

Category	Train set	Test set	Total
Normal	1833	1221	3054
Cognitive Impairment	2389	1370	3653
Total	4116	2591	6707

The distribution of the CDT dataset into 4222 train set images, and 2591 test set images are shown in Table 1. The percentage of the training dataset is 61.96%, and the test dataset is 38.03%. The class labels of a normal person and cognitive impaired are set to 0 and 1 in the experiment.

Classification Metrics

The classification metrics are used to assess the model performance. The model's overall performance can be improved using various metrics for performance assessment. Before using the model for production on unseen data, these indicators are crucial for assessing its performance. When the model is deployed on unseen data without proper assessment of the model using various evaluation metrics and only based on accuracy, it can lead to poor predictions. True Positive (TP) is a correctly predicted sample of cognitive impairment cases, False Positive (FP) is a sample of misclassified cognitive impairment cases, True Negative (TN) is a correctly classified normal person, and False Negative (FN) is a sample of a misclassified normal person. F1-score, recall, precision, and accuracy are used to estimate the proposed model performance. The following are the mathematical expression for the evaluation metrics:

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}}$$

$$\text{F1} - \text{Score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TN} + \text{FP} + \text{TP} + \text{FN}}$$

Hyper Parameters

Many Deep Learning models fine-tune the hyperparameters as they are crucial to describe the training details and directly impact the model's output. Fine-tuning the weights in some of the pre-trained model's layers and training the output classifier can improve the performance of the model. The parameters used for each model are shown in Table 2. To assess learning rates, the adapted models used the optimization algorithm. The learning rates range from 0.01 to 0.0001, and batch sizes range from 8 to 32 with 10 to 50 epochs. The model's learning capacity is best in this experiment, with a 0.001 learning rate. The number of epochs and batch size were determined after extensive experimentations.

Results and Discussion

This work employs ResNet152, EfficientNetB4, DenseNet201, and Cascade Network Model to distinguish between cognitive impaired and normal persons. The same NHATS dataset is used to train and test the models, parameter settings, and Fully Connected Layer (FC). The models were compared in accordance with accuracy, F1-score, recall, and precision of the test set. The values for the cascade network model are 97.76%, 97.75%, 97.72%, and 97.85%, respectively. However, for the better classification of the images, the cascade network model uses three distinct models in parallel and extracts the features of the CDT image that can successfully hold the broad feature information of the image. The model extracted feature maps from the convolutional layers of ResNet152, EfficientNetB4, and DenseNet201 as the first layer. The CDT images of cognitive impaired and normal persons were compared with the feature map of each model.

The classification test report of the test set is shown in Table 3. It has shown that the F1-score, recall, precision, and accuracy of ResNet152 on the test set are 94.13%, 94.22%, 94.35%, and 94.13%, respectively. The Comparison with ResNet152 shows that the Cascade Network Model's accuracy, F1-score, recall, and precision increased by 3.63%, 3.62, 3.5%, and 3.5%, respectively. The F1-score, recall, precision, and accuracy of EfficientNetB4 are 94.94%, 95.12%, 95.20%, and 94.94%, respectively, on the test set. Compared with EfficientNetB4, the Cascade Network Model accuracy, F1-score recall, and precision are increased by 2.82%, 2.81%, 2.60%, and 2.65%, respectively. The F1-score, recall, precision, and accuracy of DenseNet201 on the test set are 96.87%, 96.87%, 97.03%, and 96.87%, respectively. The comparison of DenseNet201's accuracy, F1-score, recall, and precision with those of the Cascade Network Model shows an increase of 0.89%, 0.88%, 0.85%, and 0.82%, respectively. It can be seen that all the metrics of the cascade network model have been better.

Similar to test accuracy (Table 3) the values for training accuracy have been calculated. Considering 1 as 100 in Fig. 4, the equivalent percentage values for training and test accuracies are mentioned below, whereas loss is mentioned as the actual value. From Table 3 and Fig. 4 the following observations could

Table 2 — Pre-processing and training phase parameters

Hyper Parameters	ResNet152	EfficientNetB4	DenseNet201	Cascade Network Model
Batch Size	16	16	16	16
Learning rate	1e-3	1e-3	1e-3	1e-3
Loss	Cross Entropy	Cross Entropy	Cross Entropy	Binary Cross Entropy
Optimizer	SGD	SGD	SGD	SGD
Epochs	25	25	25	25

Table 3 — Classification report on test dataset

Architecture	Classes	Precision (%)	Recall (%)	F1-Score (%)	Support
ResNet152	Normal	98.11	90.27	94.03	1221
	Cognitive Impairment	90.58	98.18	94.22	1370
	Accuracy			94.13	2591
	Macro Avg	94.35	94.22	94.13	2591
EfficientNetB4	Normal	99.75	90.49	94.89	1221
	Cognitive Impairment	90.65	99.75	94.99	1370
	Accuracy			94.94	2591
	Macro Avg	95.20	95.12	94.94	2591
DenseNet201	Normal	99.75	93.98	96.78	1221
	Cognitive Impairment	94.30	99.76	96.96	1370
	Accuracy			96.87	2591
	Macro Avg	97.03	96.87	96.87	2591
Cascade Network Model	Normal	99.50	95.89	97.66	1221
	Cognitive Impairment	96.20	99.54	97.84	1370
	Accuracy			97.76	2591
	Macro Avg	97.85	97.72	97.75	2591

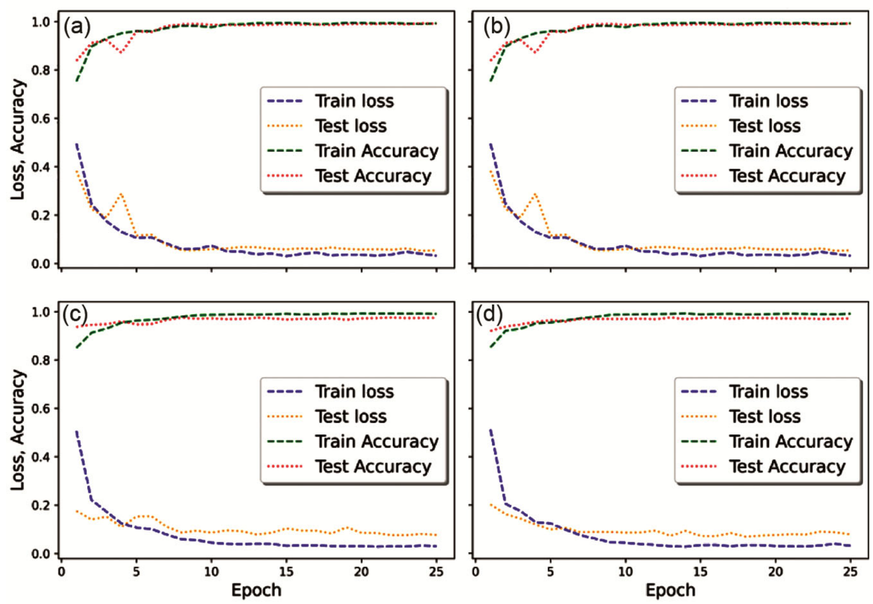

Fig. 4 — CNN models accuracy and loss curves (a) ResNet152, (b) EfficientNetB4, (c) DenseNet201, and (d) Cascade Network

be made. First, the experiment uses ResNet152 to identify cognitive impairment. The accuracy and loss curves of the model is shown in Fig. 4(a). The test accuracy ranges to 94.13%, and the training accuracy ranges to 94.21%. From the figure, it can be observed that the training and testing accuracy curve overlap and reach up to 94.21% (0.9421). The training loss has flattened out at roughly 0.16. Next, EfficientNetB4 model is used to detect cognitive impairment for which the test accuracy is 94.94%, and the training accuracy is 95.25%. The accuracy and loss curves of the EfficientNetB4 model is shown in Fig. 4(b). From analysis, the test and train accuracy of the model is greater than that of ResNet152, and the loss curve result is also good. Further, DenseNet201 is used to detect cognitive impaired persons and the test and training accuracies of the model are 96.87% and 97.02%. The model accuracy and loss curves is shown in Fig. 4(c). The model accuracy and loss are better than ResNet152 and EfficientNetB4. It has a significant training impact, and the loss has been steady. The experiment uses ResNet152, EfficientNetB4, and DenseNet201 and feature fusion from the three networks. The test accuracy reaches 97.76%, and the training accuracy is 98.26%. The model accuracy and loss curves are shown in Fig. 4(d).

The test dataset classification report establishes that the cascade network model has a better classification

Table 4 — Comparative study of the proposed model

Image Type	Architecture	Classes	Accuracy (%)	F1-Score
CDT[31]	DenseNet-121	2	96.65	—
CDT[17]	Ensemble model	2	—	94.4
CDT (Proposed model)	Cascade Network Model (Feature Fusion)	2	97.76	97.75

effect on cognitively impaired and normal persons. The results show that the proposed model has better accuracy, F1 score, recall, and precision than the other architectures. Compared with other models, the Cascade Network Model has enhanced adaptableness to the CDT images and can increase the accuracy of the test dataset. The leading cause is that the ResNet152, EfficientNetB4, and DenseNet201 models miss some in-depth information while extracting features, which will differ from model to model according to their network structure. Cascade Network Model has a cascade architecture that integrates the knowledge of features extracted by the three models, and the result is good. The model is developed by combining ResNet152, EfficientNetB4, and DenseNet201. As a result, the model can retain more features and results in better accuracy.

The proposed model results are compared with recent works and shown in Table 4. It shows that the accuracy and F1 score of the proposed model are better than those of comparative network models. Chen *et al*.[31] used the DenseNet121 deep learning classifier, and its accuracy is 96.65%. Amini *et al*.[17] used an Ensemble CNN model to extract the depth features of CDT images and classified them with an F1-Score of 94.4%. Our Proposed Cascade Network model achieves Accuracy and F1-score of 97.76% and 97.75%, respectively. The Cascade Network Model can more accurately detect cognitive impaired persons from normal persons and is considered good in-depth architecture. From the above analysis, it can be found that the classification metrics of the comparative research work on the test dataset are lower than that of our work.

Conclusions

CDT is an effective tool for detecting cognitive impairment. The experiment used the three CNN models ResNet152, EfficientNetB4, DenseNet201, and one Cascade Network Model to classify CDT images into two categories: cognitive impaired and normal. Features were extracted by applying these model architectures and classified classes through a Fully Connected Layer. The experimentation results found that the Cascade Network Model best classified cognitive impaired and normal persons under the same conditions. A remarkable improvement is there in the classification, with an accuracy of 97.76%, an F1-score of 97.75% with 97.85% precision, and 97.72% recall. The model can assist psychiatrists in the early detection of cognitive impairment. A limitation of the work is that age of the persons is not considered for classification. The model suggested in this paper can help to classify different cognitive impairment levels by taking more classes. In future, the model can be integrated to develop a simple mobile application that captures the image of the CDT drawn by the person that can be used to test the cognitive function.

References

1 Alzheimer's Association, *Alzheimer's Disease Facts and Figures*, https://www.alz.org/ alzheimers-dementia/facts-figures, 2020 (September 20, 2021).

2 Alzheimer's & Related Disorders Society of India, https://www.alzint.org/member/alzheimers-related-disorders-society-ofindia-ardsi, 2021, (September 20, 2021).

3 DeTure M A & Dickson D W, The neuropathological diagnosis of Alzheimer’s disease, *Mol Neurodegener*, **14(1)** (2019) 1–18.

4 Freedman M, Leach L, Kaplan E, Shulman K & Delis D C, Clock drawing: A neuropsychological analysis, *Oxford University Press*, *USA* (1994).

5 Janakiramaiah B, Kalyani G & Jayalakshmi A, Automatic alert generation in a surveillance systems for smart city environment using deep learning algorithm, *Evol Intell*, **14(2)** (2021) 635–642.

6 Ryu S Y, Lee S B, Kim Y I & Lee K S, The utility of the clock drawing test for cognitive impairment screening, *Alzheimers Dement*, **2** (2006) Poster 2, Article 113.

7 Umegaki H, Suzuki Y, Yamada Y, Komiya H, Watanabe K, Nagae M &Kuzuya M, Association of the qualitative clock drawing test with progression to dementia in non-demented older adults, *J Clin Med,* **9(9)** (2020) 2850.

8 Cacho J, Benito-León J, García-García R, Fernández-Calvo B, Vicente-Villardón J L & Mitchell A J, Does the combination of the MMSE and clock drawing test (mini-clock) improve the detection of mild Alzheimer's disease and mild cognitive impairment?, *J Alzheimers Dis*, **22(3)** (2010) 889–896.

9 Mittal C, Gorthi S P & Rohatgi S, Early cognitive impairment: role of clock drawing test, *Med J Armed Forces India*, **66(1)** (2010) 25–28.

10 Eknoyan D, Hurley R A & Taber K H, The clock drawing task: common errors and functional neuroanatomy, *J Neuropsychiatry Clin Neurosci*, **24(3)** (2012) 260–265.

11 Piers R J, Devlin K N, Ning B, Liu Y, Wasserman B, Massaro J M & Libon D J, Age and graphomotor decision

making assessed with the digital clock drawing test: the Framingham heart study, *J Alzheimers Dis*, **60(4)** (2017) 1611–1620.

12 Cohen J, Penney D L, Davis R, Libon D J, Swenson R A, Ajilore O & Lamar M, Digital clock drawing: differentiating “thinking” versus “doing” in younger and older adults with depression, *J Int Neuropsychol Soc*, **20(9)** (2014) 920–928.

13 Binaco R, Calzaretto N, Epifano J, McGuire S, Umer M, Emrani S, ... & Polikar R, Machine learning analysis of digital clock drawing test performance for differential classification of mild cognitive impairment subtypes versus Alzheimer’s disease, *J Int Neuropsychol Soc*, **26(7)** (2020) 690–700.

14 Zheng X, Zhang W, Wang X, Li R, Liu M, Xu F, Li Y, Zheng J & Nie Z, Extended application of digital clock drawing test in the evaluation of Alzheimer's disease based on artificial intelligence and the neural basis, *Curr Alzheimer Res*, **18(14)** (2021) 1127–1139.

15 https://anest.ufl.edu/2022/06/16/research-expands-use-of-clock-drawing-test-for-dementia-screening/

16 Li F, Tran L, Thung K H, Ji S, Shen D & Li J, A robust deep model for improved classification of AD/MCI patients, *IEEE J Biomed Health Inform*, **19(5)** (2015) 1610–1616.

17 Amini S, Zhang L, Hao B, Gupta A, Song M, Karjadi C, & Paschalidis I C, An ai-assisted online tool for cognitive impairment detection using images from the clock drawing test, *medRxiv* (2021), https://doi.org/10.1101/2021.03.06.21253047.

18 Park I & Lee U, Automatic, qualitative scoring of the clock drawing test (CDT) based on u-net, CNN and mobile sensor data, *Sensors*, **21(15)** (2021) 5239.

19 Youn Y C, Pyun J M, Ryu N, Baek M J, Jang J W, Park Y H, ... & Kim S Y, Use of the clock drawing test and the rey–osterrieth complex figure test-copy with convolutional neural networks to predict cognitive impairment, *Alzheimer's Res Ther*, **13(1)** (2021) 1–7.

20 Heimann-Steinert A, Latendorf A, Prange A, Sonntag D & Müller-Werdan U, Digital pen technology for conducting cognitive assessments: a cross-over study with older adults, *Psychol Res*, **85(8)** (2021) 3075–3083.

21 Sato K, Niimi Y, Mano T, Iwata A & Iwatsubo T, Automated evaluation of conventional clock-drawing test using deep neural network: Potential as a mass screening tool to detect individuals with cognitive decline, *Front Neurol*, **13** (2022) 896403–896403.

22 Pan D, Zeng A, Jia L, Huang Y, Frizzell T & Song X, Early detection of Alzheimer’s disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning, *Front Neurosci*, **14** (2020) 259.

23 Zheng L, Zhao Y, Wang S, Wang J & Tian Q, Good practice in CNN feature transfer, *arXiv preprint* (2016) arXiv: 1604.00133, https://doi.org/10.48550/arXiv.1604.00133.

24 Kawahara J & Hamarneh G, Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers, in *Machine Learning in Medical Imaging*, edited by L Wang, E Adeli, Q Wang, Y Shi, H I Suk, MLMI 2016, Lecture Notes in Computer Science (Springer, Cham) vol 10019, (2016) 164–171 https://doi.org/10.1007/978-3-319-47157-0_20.

25 Pan S J & Yang Q, A survey on transfer learning, *IEEE Trans Knowl Data Eng*, **22(10)** (2009) 1345–1359.

26 He K, Zhang X, Ren S & Sun J, Deep residual learning for image recognition, *Proc IEEE Conf on Comput Vision Pattern Recognit*, (2016) (pp 770–778).

27 Tan M & Le Q, Efficientnet: Rethinking model scaling for convolutional neural networks, in *IEEE Int Conf Machine Learn*, (California, USA) 9–15, June 2019, (pp 6105–6114) PMLR.

28 Huang G, Liu Z, Van Der Maaten L & Weinberger K Q, Densely connected convolutional networks, *Proc IEEE Conf on Comput Vision Pattern Recognit*, (2017) (pp 4700–4708).

29 Subramanian M, Narasimha Prasad L V, Janakiramaiah B, Mohan Babu A & Sathishkumar Ve, Hyperparameter optimization for transfer learning of VGG16 for disease identification in corn leaves using Bayesian optimization, *Big Data*, **10(3**) (2022) 215–229, DOI: 10.1089/big.2021.0218

30 NHATS Public Use Data. (Insert Round or Rounds), sponsored by the National Institute on Aging (grant number NIA U01AG032947) through a cooperative agreement with the Johns Hopkins Bloomberg School of Public Health, Available at www.nhats.org.

31 Chen S, Stromer D, Alabdalrahim H A, Schwab S, Weih M & Maier A, Automatic dementia screening and scoring by applying deep learning on clock-drawing tests, *Sci Rep*, **10(1)** (2020) 1–11.