*Editorial*

DOI: 10.5582/irdr.2023.01015

# Recent deep learning models for dementia as point-of-care testing: Potential for early detection

**Kenji Karako[1], Peipei Song[2,3,*], Yu Chen[1,*]**

[1] Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo, Chiba, Japan;
[2] Center for Clinical Sciences, National Center for Global Health and Medicine, Tokyo, Japan;
[3] National College of Nursing, Japan.

**SUMMARY** Deep learning has been intensively researched over the last decade, yielding several new models for natural language processing, images, speech and time series processing that have dramatically improved performance. This wave of technological developments in deep learning is also spreading to medicine. The effective use of deep learning in medicine is concentrated in diagnostic imaging-related applications, but deep learning has the potential to lead to early detection and prevention of diseases. Physical aspects of disease that went unnoticed can now be used in diagnosis with deep learning. In particular, deep learning models for the early detection of dementia have been proposed to predict cognitive function based on various information such as blood test results, speech, and the appearance of the face, where the effects of dementia can be seen. Deep learning is a useful diagnostic tool, as it has the potential to detect diseases early based on trivial aspects before clear signs of disease appear. The ability to easily make a simple diagnosis based on information such as blood test results, voice, pictures of the body, and lifestyle is a method suited to point-of-cate testing, which requires immediate testing at the desired time and place. Over the past few years, the process of predicting disease can now be visualized using deep learning, providing insights into new methods of diagnosis.

***Keywords*** deep learning, dementia, prediction, point-of-cate testing

Deep learning has continued to attract attention since its performance exceeded that of humans in image recognition tasks at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC2012), an image processing contest held in 2012 (*1*). Deep learning has been intensively researched over the last decade, yielding several new models for natural language processing, images, speech and time series processing that have dramatically improved performance. In the field of image recognition, ResNets capable of recognizing objects in basic image recognition tasks have emerged (*2*). Yolo has been proposed for object detection tasks (*3,4*), and it is capable of detecting object regions in an image in real time. As well as recognizing objects in images, deep learning is also capable of generating new images. Models for image generation have been proposed, first called GANs, that produce images similar to the training data (*5*), and recently images that might be mistaken for the real thing can now be produced. Stable Diffusion produces images that represent input text (*6*). Natural language processing has also made great strides with the emergence of bert (*7*) and transformer (*8*), which are generic models for languages. Chat GPT (*9*) has emerged in the past few years, and it can generate responses to questions as if they were answers from a person with actual expertise.

**Recent developments in deep learning for diagnostic imaging technology**

This wave of technological developments in deep learning is also spreading to medicine. The impact of the dramatic development of image-oriented models in the early stages of the development of deep learning technology has led to widespread research on the use of deep learning for image-based medical examinations. In particular, models have been proposed to detect various diseases by combining basic imaging diagnostics such as X-rays (*10*), CT scans (*11*), and MRI scans (*12,13*). Diseases are diagnosed with deep learning using images. As well as researching models using still images, detection models have also been studied for endoscopy (*14*) and ultrasonography (*15*), where video is taken. These imaging models help to prevent doctors from missing anything and reduce the time to diagnosis.

## The potential for deep learning in early disease detection

Although the effective use of deep learning in medicine is concentrated in diagnostic imaging-related applications, deep learning has the potential to lead to early detection and prevention of diseases. The detection of diseases requires an examination to analyze and detect physical aspects of disease, and methods of testing and diagnosis have been studied. Physical aspects of disease that went unnoticed can now be used in diagnosis with deep learning. Deep learning analyzes large amounts of data and it learns and identifies factors that are important to diagnosis so that subtle information and aspects of disease that went unnoticed can be identified. Deep learning therefore has the potential to detect aspects of disease in its early stage.

The current work looks at the potential for new methods of detecting dementia early using deep learning and possible applications of those methods.

## Early models for detection of dementia using deep learning

One area where deep learning is being used effectively is in the early detection of dementia. There is no cure for dementia, and the condition needs to be addressing in its early stages to mitigate symptoms and delay progression. Deep learning identifies and analyzes the effects of dementia on physical and cognitive functions, facilitating its diagnosis. As shown in Figure 1, early detection is possible based on blood test results that reflects lifestyle factors that contribute to dementia and speech reflecting cognitive function. In addition, deep learning can also detect the effects of dementia in the face and voice, which are difficult to analyze.

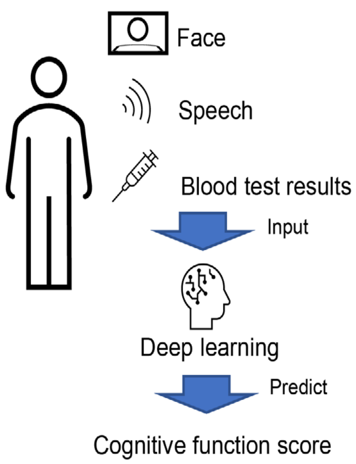


**Figure 1. Deep learning predicts the cognitive status of patients based on the various effects of dementia on the body.** A variety of information is used as input data, such as blood test results, voice signals, spoken text, and facial photographs.

The first approach to predicting cognitive function was predicting the Mini-Mental State Examination (MMSE) Score, which is used to assess cognitive function, based on blood test results reflecting lifestyle in order to detect dementia of lifestyle-related origin in its early stage (*16*). The proposed model uses a three-layer neural network model, which is a basic deep learning technique, to predict the MMSE Score using 23 items on a blood test performed on 202 elderly patients (73.48 ± 13.1) with various underlying diseases. Common items on the test such as the total blood cell count, red blood cell count, and the hemoglobin level are used. Although the training data were limited, the correlation coefficient between the predicted MMSE Score and the actual MMSE Score was 0.66, indicating a close correlation.

A model for predicting cognitive function based on speech has been proposed as a second approach. The risk of dementia is predicted based on speech when answering a few fixed questions or answering for about five minutes. This approach includes models that predict cognitive function based on the information contained in sound and models that convert speech into textual information and that predict the risk of dementia based on aspects of cognitive function evident in the text. In a model that predicts using speech information as it is (*17*), speech is treated as signals and deep learning extracts the phonetic features that emerge due to dementia. In a model that predicts dementia by converting speech into textual information (*18*), deep learning analyzes sentences that reflect the decline in comprehension and judgement occurring with dementia, and it evaluates cognitive functions. Although the information used in both methods of prediction is the same speech, the predictions focus on different aspects of dementia.

Once sound is converted to textual information, signals that should have been present are missing as a result of dementia. There are models that use both speech signals and textual information to predict the risk of dementia (*19*).

A final approach is to predict the risk of dementia based on images of the face. A model (*20*) has been proposed to predict whether there is cognitive decline or not based on photographs of the face taken without facial expressions. A deep learning model was created using data from 121 patients with cognitive decline and 117 normal individuals, and it performed well at discrimination with a sensitivity of 87.31%, a specificity of 94.57%, and an accuracy of 92.56%.

## Effective use of deep learning as point-of-cate testing

Models have been studied to indirectly predict the risk of dementia by examining the facial appearance, voice, or blood test results that exhibit aspects of dementia. Photographs of the face cost almost nothing, voice recordings take 10 minutes at most, and blood tests are performed during physical examinations. Therefore, the risk of dementia can easily be determined with little effort. In addition to the benefit of being able to determine the risk of dementia, models can also ascertain there is a risk of dementia before symptoms become apparent. Thus, deep learning can serve as point-of-cate testing (PoCT) that encourages an individual to undergo a thorough examination, although there are concerns about prediction performance (*21*).

Originally, deep learning had a black box problem, in which the prediction process was not evident from the outside, and what caused the accurate or inaccurate prediction of cognitive function was not known. Over the past few years, however, a method of revealing the prediction process called SHAP (*22*) has been proposed. It can also reveal which blood test results are responsible for the poor prediction of cognitive function. Such information would not only increase the transparency of the deep learning diagnostic process but could also be an important source of information for determining the appropriate course of action for each individual's condition. When predicting cognitive function based on blood test results, for example, knowing which blood test results had the greatest impact when poor cognitive function was predicted would allow the suggestion of a coping strategy tailored to that individual. In addition, revealing the prediction process could provide new insights into the early detection of dementia.

When deep learning models predict dementia based on photographs of the face and speech, analyzing regions of the face, facial expressions, and phonetic features that are particularly affected could help to detect changes due to dementia that were previously overlooked. This could help lead to new diagnostic methods.

Deep learning, which has developed rapidly in recent years, is being used not only as a means to help doctors diagnose diseases, but also as a completely new means with which to detect diseases early. Physical aspects of disease that went unnoticed can now be used in diagnosis with deep learning. Research on the early detection of dementia in particular is being conducted to predict aspects of dementia from multiple perspectives using information such as blood test results, voice, and photographs of the face. The available data are currently limited, but if data are collected in the future, then this could be an effective PoCT tool. Combining all of these predictive models based on blood test results, photographs of the face, and voice should lead to further improvements in performance. Deep learning has the potential to allow early detection of dementia and other diseases based on their physical aspects.

*Funding*: None.

*Conflict of Interest*: The authors have no conflicts of interest to disclose.

## References

1. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Li FF. ImageNet large scale visual recognition challenge. Int J Comp Vision. 2015; 115:211-252.
2. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 IEEE Conf Comp Vision Pattern Recog (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778.
3. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. 2016 IEEE Conf Comp Vision Pattern Recog (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788.
4. Redmon J, Farhadi A. Yolov3: An incremental improvement. 2018; arXiv:1804.02767. https://arxiv.org/abs/1804.02767
5. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial networks. Commun ACM. 2020; 63:139-144.
6. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-resolution image synthesis with latent diffusion models. Proceedings IEEE/CVF Conf on Comp Vision Pattern Recog. 2022; pp. 10684-10695.
7. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. 2018; arXiv:1810.04805. https://doi.org/10.48550/arXiv.1810.04805
8. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need. 2017; arXiv:1706.03762. https://arxiv.org/abs/1706.03762
9. OpenAI, ChatGPT: Optimizing language models for dialogue. *https://openai.com/blog/chatgpt/* (Accessed February 1, 2023).
10. Hemdan EE, Shouman MA, Karar ME. COVIDX-Net: A framework of deep learning classifiers to diagnose covid-19 in X-ray images. 2020; arXiv:2003.11055. https://arxiv.org/abs/2003.11055
11. Lakshmanaprabu SK, Sachi Nandan Mohanty, Shankar K, Arunkumar N, Gustavo Ramirez. Optimal deep learning model for classification of lung cancer on CT images. Future Gen Computer Systems. 2019; 92:374-382.
12. Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ. Deep learning for brain MRI segmentation: State of the art and future directions. J Digit Imaging. 2017; 30:449-459.
13. Ueda D, Yamamoto A, Nishimori M, Shimono T, Doishita S, Shimazaki A, Katayama Y, Fukumoto S, Choppin A, Shimahara Y, Miki Y. Deep learning for MR angiography: Automated detection of cerebral aneurysms. Radiology. 2019; 290:187-194.
14. Min JK, Kwak MS, Cha JM. Overview of deep learning in gastrointestinal endoscopy. Gut Liver. 2019; 13:388-393.
15. Karako K, Mihara Y, Arita J, Ichida A, Bae SK, Kawaguchi Y, Ishizawa T, Akamatsu N, Kaneko J, Hasegawa K, Chen Y. Automated liver tumor detection in abdominal ultrasonography with a modified faster region-based convolutional neural networks (Faster R-CNN) architecture. Hepatobiliary Surg Nutr. 2022; 11:675-683.

16. Sakatani K, Oyama K, Hu L. Deep learning-based screening test for cognitive impairment using basic blood test data for health examination. Front Neurol. 2020; 11:588140.
17. Luz S, Haider F, de la Fuente S, Fromm D, MacWhinney B. Detecting cognitive decline using speech only: The ADReSSo challenge. 2021; arXiv:2104.09356. https://arxiv.org/abs/2104.09356
18. Syed ZS, Syed MSS, Lech M, Elena Pirogova E. Tackling the ADRESSO challenge 2021: The MUET-RMIT system for Alzheimer's dementia recognition from spontaneous speech. Interspeech. 2021; 3815-3819.
19. Balagopalan A, Novikova J. Comparing acoustic-based approaches for Alzheimer's disease detection. 2021; arXiv:2106.01555. https://arxiv.org/abs/2106.01555
20. Umeda-Kameyama Y, Kameyama M, Tanaka T, Son BK, Kojima T, Fukasawa M, Iizuka T, Ogawa S, Iijima K, Akishita M. Screening of Alzheimer's disease by facial complexion using artificial intelligence. Aging (Albany NY). 2021; 13:1765-1772.
21. Karako K, Song P, Chen Y, Tang W. Increasing demand for point-of-care testing and the potential to incorporate the Internet of medical things in an integrated health management system. Biosci Trends. 2022; 16:4-6.
22. Lundberg S, Lee SI. A unified approach to interpreting model predictions. 2017; arXiv:1705.0787. https://arxiv.org/abs/1705.07874

Received February 7, 2023; Accepted February 22, 2023.

*Address correspondence to:
Peipei Song, Center for Clinical Sciences, National Center for Global Health and Medicine, 1-21-1 Toyama, Shinjuku, Tokyo 162-8655, Japan.
E-mail: psong@it.ncgm.go.jp

Yu Chen, Department of Human and Engineered Environmental Studies, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 227-8568, Japan.
E-mail: chen@edu.k.u-tokyo.ac.jp

Released online in J-STAGE as advance publication February 24, 2023.