

Code-free deep learning: a step into the future of ophthalmology

BY ROHAN MISRA, CIARA O'BYRNE AND PEARSE KEANE

A new focus has arisen within the research domain of artificial intelligence (AI) in healthcare called code-free deep learning (CFDL), and recent research demonstrates that ophthalmology is becoming one of the leading specialties in this field.

Artificial intelligence (AI) is the science of developing machines with the ability to mimic the complexities of the human mind [1]. In recent years, its application in healthcare has shown very promising results, leading to increasing interest and investment in the field. In fact, the global AI healthcare market size is predicted to reach around US \$187.95 billion by 2030, with a compound annual growth rate of 37% between 2022 to 2030 [2].

Deep Learning (DL): the future of healthcare?

Deep learning (DL) is a subtype of AI which is causing considerable excitement within the field of healthcare. These algorithms are inspired by the neural networks of the human brain and have been revolutionary in the fields of image classification, object detection and natural language processing [3]. Ophthalmology is particularly suited for the application of DL due to its substantial use of medical imaging. For example, in 2018, a collaborative study between DeepMind and Moorfields Eye Hospital demonstrated that a DL algorithm could make referral recommendations for over 50 sight-threatening retinal diseases using 14,884 optical coherence tomography (OCT) scans [4]. These referral recommendations exceeded that of experts in the field, emphasising the promising future of AI in healthcare.

Despite high expectations, there are some barriers to AI that are particularly challenging for those with minimal resources. These include the limited availability of experts in the field and the huge expenses and computing power required to develop DL models [5].

Code-Free Deep Learning (CFDL)

"AI that can build AI" was a headline generated by *The New York Times* for the inception of CFDL in 2017 [6]. CFDL describes user-friendly platforms which allows the creation of DL algorithms without coding experience. This is a technique that potentially offers solutions to some of the barriers faced by bespoke AI technology.

“CFDL has the potential to democratise deep learning and could accelerate the development of clinical AI systems to achieve maximal patient benefit”

Image Classifier CFDL steps

STEP 01 Data Upload

Upload data using drag and drop interface. This can be in the format of .CSV files or directly from local computers.



STEP 02 Data Visualisation and model training

Visualise the data and complete labelling process to produce training, testing and validation data sets.



STEP 03 Model Evaluation

Performance of the model is evaluated using specific metrics.



Figure 1: Image classifier steps.

Within ophthalmology and other health specialties, this has provided healthcare professionals and researchers with the opportunity to create their own DL algorithms focusing on necessary clinical needs. CFDL is commercially available on several platforms which include Amazon Rekognition Custom Labels (Amazon); Apple Create ML (Apple); Baidu EasyDL (Baidu); Clarifai Train (Clarifai);

Google Cloud AutoML Vision; Huawei ModelArts ExeML (Huawei); MedicMind Deep Learning Training Platform (MedicMind); and Microsoft Azure Custom Vision (Microsoft) [7].

These platforms have intuitive user interfaces with simple drag and drop functions. In addition, cloud-based servers reduce the need for huge amounts of computing power. It must be noted, however, that the process of dataset curation has not been automated, and users must acquire datasets either from publicly available resources or from within their own institutions, provided strict ethical and data governance procedures are followed.

The development of image classification models using CFDL is one of the most common applications. The process can be explained using a simple three-stage approach, replicating the user interface archetype of CFDL platforms (Figure 1) [8]. Once a high-quality dataset has been acquired, the first stage is to upload this directly from the local computer or using .CSV files via cloud storage buckets. The second stage is to visualise the data and amend any labels before splitting the data into training, validation, and test sets. The data split can be done automatically or manually. At this point, the platform automates the model training process which involves data pre-processing, architecture selection and hyperparameter optimisation. The final stage is to evaluate the model. The performance of a model is measured using similar techniques to medical statistics. These measurements include confusion matrices, recall (sensitivity), precision (positive predictive value), area under the precision-recall curve (AUPRC) and F1-score (balance of precision and recall). The similarities in the analysis of DL models and statistics used in clinical science make CFDL concepts quick to grasp for clinicians, as shown in several studies [7,9,11,12].

In addition, some platforms, namely MedicMind and Google, facilitate external validation of the model. External validation is the process of testing models using independent data which was not used in the initial training or testing steps [7]. Google also facilitates batch prediction which allows efficiency testing of the algorithm with large external datasets.

Ophthalmology: Promising examples of clinically applied CFDL

Ophthalmology is one of the leading specialties in the exploration of AI in healthcare. This is evident within the clinical literature where there are several examples of CFDL applications.

If we believe that AI should be used in clinical practice alongside the medical expertise of clinicians in the future, this will require education for those without technical experience. CFDL provides the opportunity to learn and understand how AI can impact clinical

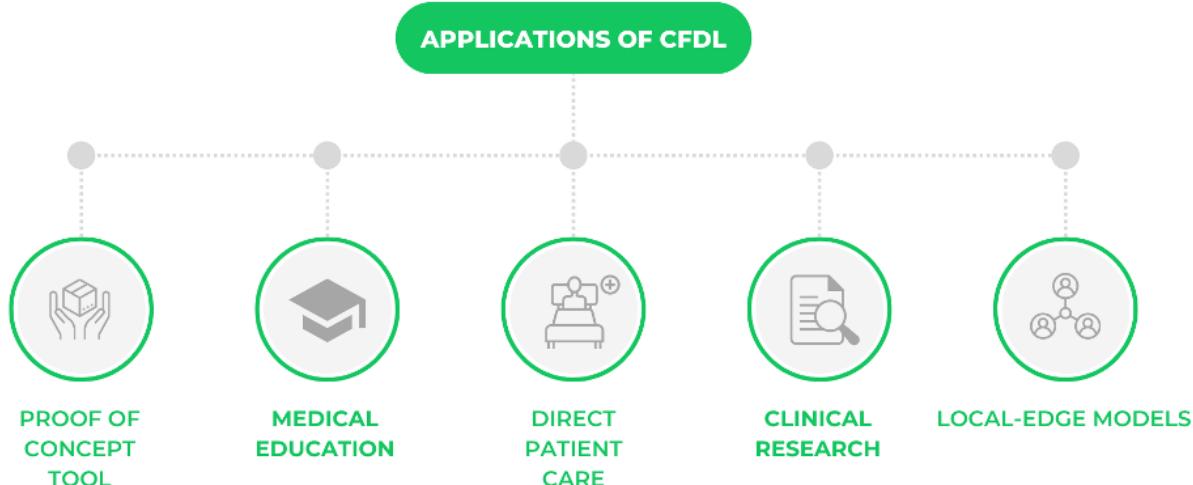
decision-making, whilst practically creating CFDL algorithms. This was demonstrated by Faes et al. in 2019, whereby two physicians with no coding or machine learning experience successfully created deep learning image classification models using Google Cloud AutoML Vision [9]. These models used publicly available datasets, such as retinal fundus and OCT images. What was exciting about this study was the fact that the performance (AUPRC) of the ophthalmology-based models was comparable to bespoke deep learning models using the same datasets. These encouraging results showed the feasibility of non-technical clinicians developing DL algorithms with minimal training. There are other image classification successes using CFDL, for example, classifying retinal pathologies from ultra-widefield pseudocolour fundus images [10]. The performances of the models had very good diagnostic accuracy (sensitivity range 84.91% to 89.77%) in differentiating retinal vein occlusion, retinitis pigmentosa and retinal detachment from normal fundi. The significance of this study is that performance was once again comparable to bespoke DL models, emphasising that CFDL offers an opportunity for clinicians to develop their own DL solutions.

Fares et al. carried out a similar study in 2020, whereby two ophthalmologists with no coding experience created a prediction model for postoperative proliferative vitreoretinopathy (PVR) following rhegmatogenous retinal detachment (RDD) repair [11]. This was done using an interactive application in MATLAB and was compared with bespoke manually coded models. Performances (F1 scores) were again comparable (and marginally higher), further emphasising the feasibility of CFDL, and the top model boasted an impressive specificity of 97.8%. This shows that prediction models can also be developed using CFDL which could help with identifying prognosis of disease and likelihood of success after treatment.

Another interesting study was carried out by Kim et al. Performance of human experts and CFDL models were compared in classifying pachychoroid disease [12]. Two models were generated, and the results were comparable to two retina specialists, with very similar precision and recall. It was also interesting to note that the CFDL models had achieved a superior performance compared to ophthalmic residents, further emphasising the potential of CFDL (89.19% precision and recall for top-performing model versus 68.18% and 92.50% for ophthalmic residents).

Code-free deep learning can also be applied to tabular data. Abbas et al. introduced a Jupyter notebook extension to Google Cloud AutoML Tables using the What-If Tool (WIT) [13]. This is an open-source tool which allows users to see how model predictions change if input features are hypothetically changed, i.e., different patient subgroups can be analysed to test the bias of the models

Figure 2: Applications of CFDL.



generated. This study found that a model predicting visual acuity outcomes in neovascular age-related macular degenerations had varied performances across different ethnicities. The use of the WIT tool in CFDL gives the opportunity for models to be tested for bias to increase clinical applicability.

One of the most exciting results showing the potential of CFDL came from a study by Korot et al. [14]. The CFDL model managed to replicate a study by Poplin et al. which identified sex from retinal fundus images - a pattern which had never been clinically documented [15]. This feat demonstrates that CFDL can be used for clinical image classification, as well as in research to identify new patterns.

Limitations

CFDL offers great promise within the world of AI, however, parts of the process still require refining and regulation.

As discussed, platforms currently do not automate the process of labelling and formatting datasets which is a crucial stage in the development of robust deep learning algorithms. The initial manipulation of data must be done manually which can be time-consuming and expensive. In addition, datasets must comply with ethical and data governance protocols especially when using hospital data [7,8]. This could lead to an element of discriminatory bias within the developed models (due to the specific selection of data available), which could increase health disparities, resulting in worse health outcomes, particularly for the minority populations. NHS England is planning a pilot to eradicate biases in AI to make software more ethical. The system designed by the Ada Lovelace Institute will allow an assessment of the possible risks of biases of an algorithm before being able to access NHS data [16].

A further limitation is the 'black-box phenomenon', which is the inability to understand how the algorithm is determining results [7]. This means that the model created could be focussing on features which would not be clinically significant. This problem is not unique to CFDL, however, due to its automated nature, model explainability could be further reduced in comparison to bespoke DL models. A potential solution could be the integration of saliency maps which are currently available on MedicMind. Saliency maps help to map out on the image the sections that are targeted during the training process, which could help to provide some clinical interpretability [8].

A critical step that is often overlooked in CFDL platforms is the ability to carry out external validation. This is key to test, evaluate performance and further train algorithms to ensure that the system works for data which has not previously been seen. Currently, Google Cloud AutoML Vision allows users to carry out batch prediction using the command line interface as mentioned before [8]. Many platforms do not offer any form of external validation. If they were to include a function to automate this process, this would greatly support the feasibility of CFDL.

Future of CFDL and its potential applications

CFDL has the potential to democratise deep learning and could accelerate the development of clinical AI systems to achieve maximal patient benefit. Wider availability for non-technical clinicians and researchers without the need for a deeper understanding of mathematics and programming behind AI technology is an attractive proposition. Nevertheless, comparable results are currently limited, indicating the need for bespoke DL algorithms for a wider range of clinical AI systems. In these early stages, CFDL could be used as a tool for proof-of-concept projects prior to investing in the development of expensive and more advanced deep learning models [8].

The barriers of expensive running costs and the need for highly specialised programming expertise for AI systems is even greater

for underprivileged healthcare systems across the world. CFDL provides a platform for those with limited resources to create DL systems. Screening programmes with the use of image classifiers could be implemented in under-resourced communities [17]. In addition, CFDL models can be downloaded and run on local edge models, which are low power models that do not require a continuous internet connection, easing the burden of poor internet infrastructure [18,19].

Many of the studies used clinicians to produce CFDL models. This shows the potential of these platforms to be used in medical education [20]. The practical experience would teach medical students and clinicians the skills needed to grasp an understanding of AI in healthcare, whilst demonstrating the ethical principles and limitations which need to be considered. Integration within medical education would give confidence to adapt to changing times in healthcare with huge focuses on the development of AI technology.

Conclusion

CFDL offers promising solutions to barriers of healthcare AI technology. With regulation and clinical validation, CFDL models will be key to leading the revolution of the democratisation of AI with applications in direct patient care, clinical research, and medical education.

TAKE HOME MESSAGE

- Deep learning has been shown to achieve promising results in image classification, object detection and natural language processing.
- Code-free deep learning offers an opportunity for non-technical clinicians and researchers to produce deep learning models, thereby representing a first step towards democratisation of AI-enabled healthcare.
- Ophthalmology is a field which has been successful in producing CFDL models, with results comparable to bespoke algorithms.
- Applications of CFDL could include direct patient care, although robust clinical validation would be required; prior to this CFDL is most likely to be used in clinical research and medical education.
- CFDL could be used as a proof-of-concept tool before investment for expensive bespoke AI.

References:

1. What is Artificial Intelligence (AI)? <https://www.ibm.com/uk-en/cloud/learn/what-is-artificial-intelligence> Last accessed July 2022.
2. Artificial Intelligence (AI) in Healthcare Market Size 2022-2030 <https://www.precedenceresearch.com/artificial-intelligence-in-healthcare-market> Last accessed July 2022.
3. Keane PA, Topol EJ. AI-facilitated health care requires education of clinicians. *Lancet* 2021. Epub ahead of print.
4. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;24(9): 1342-50.
5. Code-free deep learning: The next phase of AI-enabled healthcare? <https://europe.ophthalmologytimes.com/view/code-free-deep-learning-the-next-phase-of-ai-enabled-healthcare-> Last accessed July 2022.
6. Building A.I. That Can Build A.I. *The New York Times*. <https://www.nytimes.com/2017/11/05/technology/machine-learning-artificial-intelligence-ai.html> Last accessed July 2022.
7. Korot E, Guan Z, Ferraz D, et al. Code-free deep learning for multi-modality medical image classification. *Nat Mach Intell* 2021;3(4):288-98.
8. O'Byrne C, Abbas A, Korot E, Keane PA. Automated deep learning in ophthalmology: AI that can build AI. *Curr Opin Ophthalmol* 2021;32(5):406-12.

9. Faes L, Wagner S, Fu D, et al. Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. *Lancet Digit Heal* 2019;1(5):e232-42.
10. Antaki F, Coussa RG, Kahwati G, et al. Accuracy of automated machine learning in classifying retinal pathologies from ultra-widefield pseudocolour fundus images. *Br J Ophthalmol* 2021. Epub ahead of print.
11. Antaki F, Kahwati G, Sebag J, et al. Predictive modeling of proliferative vitreoretinopathy using automated machine learning by ophthalmologists without coding experience. *Sci Reports* 2020;10(1):1-10.
12. Kim IK, Lee K, Park JH, et al. Classification of pachychoroid disease on ultrawide-field indocyanine green angiography using auto-machine learning platform. *Br J Ophthalmol* 2021;105(6):856-61.
13. Abbas A, O'Byrne C, Fu DJ, et al. Evaluating an automated machine learning model that predicts visual acuity outcomes in patients with neovascular age-related macular degeneration. *Graef's Arch Clin Exp Ophthalmol* 2022;260(8):2461-73.
14. Korot E, Pontikos N, Liu X, et al. Predicting sex from retinal fundus photographs using automated deep learning. *Sci Rep* 2021;11(1):10286.
15. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2018;2(3):158-64.
16. UK to pilot world-leading approach to improve ethical adoption of AI in healthcare. GOV.UK: <https://www.gov.uk/government/news/uk-to-pilot-world-leading-approach-to-improve-ethical-adoption-of-ai-in-healthcare> Last accessed July 2022.
17. Bellemo V, Lim ZW, Lim G, et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. *Lancet Digit Health* 2019;1(1):e35-44.
18. Greco L, Percannella G, Ritrovato P, et al. Trends in IoT based solutions for health care: moving AI to the edge. *Pattern Recognit Lett* 2020;135:346-53.
19. Keane PA, Topol EJ. Medicine and meteorology: Cloud, connectivity, and care. *Lancet* 2020;395(10233):1334.
20. Korot E, Guan Z, Ferraz D, et al. Code-free deep learning for multi-modality medical image classification. *Nat Mach Intell* 2021. Epub ahead of print.

[All links last accessed July 2022].

AUTHORS



Rohan Misra,

Foundation Year 1 Doctor, West Hertfordshire Teaching Hospitals NHS Trust, Hertfordshire, UK.



Ciara O'Byrne,

Ophthalmology ST1, North London Deanery Honorary Research Fellow, Moorfields Eye Hospital NHS Foundation Trust, London, UK.



Prof Pearse Keane,

Consultant Ophthalmologist, Moorfields Eye Hospital NHS Foundation Trust, Professor of Artificial Medical Intelligence, Institute of Ophthalmology, UCL, London, UK.

Declaration of competing interests:

Professor. Keane has acted as a consultant for Google, DeepMind, Roche, Novartis, Apellis, and BitFount and is an equity owner in Big Picture Medical. He has received speaker fees from Heidelberg Engineering, Topcon, Allergan, and Bayer.