Clinical Radiology xxx (2018) 1–5

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

Machine learning "red dot": open-source, cloud, deep convolutional neural networks in chest radiograph binary normality classification

E.J. Yates [a,*], L.C. Yates [a], H. Harvey [b]

[a] Foundation Doctor, West Midlands, England, UK
[b] Kheiron Medical Technologies, Rocketspace, 40 Islington High St, London N1 8EQ, UK

ARTICLE INFORMATION

Article history:
Received 27 February 2018
Accepted 9 May 2018

AIM: To develop a machine learning-based model for the binary classification of chest radiography abnormalities, to serve as a retrospective tool in guiding clinician reporting prioritisation.

MATERIALS AND METHODS: The open-source machine learning library, Tensorflow, was used to retrain a final layer of the deep convolutional neural network, Inception, to perform binary normality classification on two, anonymised, public image datasets. Re-training was performed on 47,644 images using commodity hardware, with validation testing on 5,505 previously unseen radiographs. Confusion matrix analysis was performed to derive diagnostic utility metrics.

RESULTS: A final model accuracy of 94.6% (95% confidence interval [CI]: 94.3–94.7%) based on an unseen testing subset (n=5,505) was obtained, yielding a sensitivity of 94.6% (95% CI: 94.4–94.7%) and a specificity of 93.4% (95% CI: 87.2–96.9%) with a positive predictive value (PPV) of 99.8% (95% CI: 99.7–99.9%) and area under the curve (AUC) of 0.98 (95% CI: 0.97–0.99).

CONCLUSION: This study demonstrates the application of a machine learning-based approach to classify chest radiographs as normal or abnormal. Its application to real-world datasets may be warranted in optimising clinician workload.

Introduction

Increasing healthcare demand is this reflected in radiology departments and subsequent clinician workload.[1] Despite radiological advances, the chest radiography (CXR) remains the most commonly requested imaging technique in the UK.[2] Consequently, timely radiologist reporting of every film is not always possible, leading to a "backlog" of unreported studies. In order to deliver maximal patient benefit from this reporting backlog, a system of image abnormality prioritisation would be beneficial, allowing reporting to first focus on examination of pathology over normality.

The "red dot" system, a method of radiographer communication of potential image abnormality has been in practice for almost 40 years.[3] The modern system uses digital superimposition of the words "red dot" on such images, a tribute to the traditional method of affixing a circular, red sticker to the abnormal plain film. Systematic review of radiographer red dot usage found 78% (74–82%) sensitivity and 91% (88–93%) specificity across pooled chest and abdominal films, highlighting the role of triage.[4]

* Guarantor and correspondent: E. Yates, Foundation Doctor, West Midlands, England, UK. Tel.: +07407126966.
E-mail address: elliotyatesj@gmail.com (E.J. Yates).

https://doi.org/10.1016/j.crad.2018.05.015
0009-9260/

Although radiographers can be trained in this prospective image prioritisation methodology, it does not tackle the retrospective burden of imaging backlog.

Machine learning (ML), specifically the field of image recognition using neural networks, could provide one such avenue of pictorial classification. Deep convolutional neural networks (CNN), an architecture loosely modelled on the biological organisation of the human brain (Fig 1), represents one such machine learning approach, historically demonstrating breakthroughs in computer vision and speech recognition.[5]

Such methods have previously been applied to radiology, with advances in multiple imaging techniques. CheXNet,[6] a CNN trained to yield both probability and heat map localisation of pneumonia in chest radiographs, demonstrated an F1 score of 0.435 (95% confidence interval [CI]: 0.387–0.481), which was a statistically significant improvement on a pooled radiologist average (0.387, 95% CI: 0.330–0.442). Furthermore, Cicero *et al.* retrained the GoogLeNet CNN in the multi-label classification of chest radiograph pathology, demonstrating a peak 91% sensitivity and specificity for a single disease category.[7]

As discussed, existing ML approaches have typically focused on formal reporting of pathology or multi-label classification, rather than prioritisation based on abnormality. The aim of the present study was to explore the ability of a deep-learning, computer vision approach in binary normality classification of plain film chest radiographs to serve as a rapid screening tool to assist clinicians in prioritising scans for formal reporting.

Materials and methods

The National Institutes of Health (NIH) Clinical Center 'ChestX-ray14' chest radiography corpus[8] includes 112,120 images derived from more than 30,000 patients. The anonymised dataset of frontal-view images with corresponding labels (text-mined into 14 disease categories from the corresponding original radiology report) represents a valuable resource for neural network training. The thoracic pathology categories include: atelectasis, cardiomegaly, effusion, infiltration, mass, nodule, pneumonia, pneumothorax, consolidation, oedema, emphysema, fibrosis, pleural thickening, and hernia (Fig 2).

Additionally, the Indiana University hospital network chest radiograph database,[9] previously the largest public dataset before the NIH release, contains 7,470 frontal and lateral view images with corresponding full radiologist narrative report. Both datasets were combined to facilitate neural network training across both normal and abnormal images (Fig 3).

Image preparation and metadata parsing

Images were obtained from the two aforementioned public repositories. Indiana University images were manually separated into frontal (n=3,819) and lateral (n=3,651)

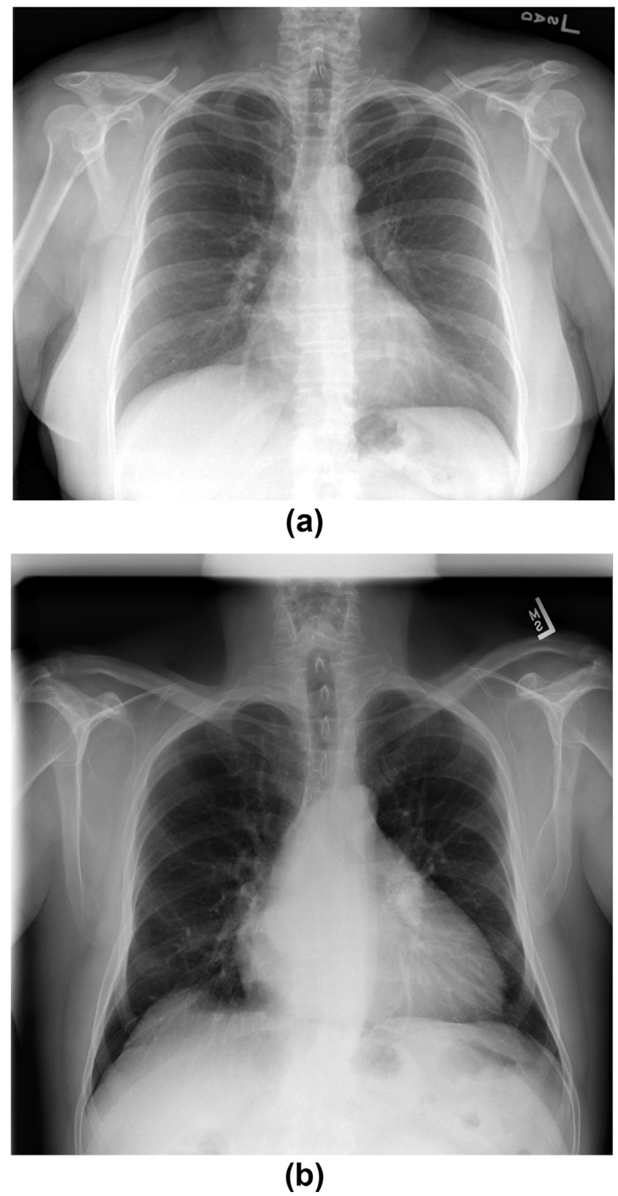

Figure 2 (a) Example of a "normal" image from the Indiana University Chest X-ray Collection. Image reference: *CXR100_IM-0002-1001*. CC BY-NC-ND 4.0 licence. (b) Example of an "abnormal" (cardiomegaly, emphysema) image from the ChestX-ray14 corpus. Image reference: *00000001_001*. Licence usage unrestricted.

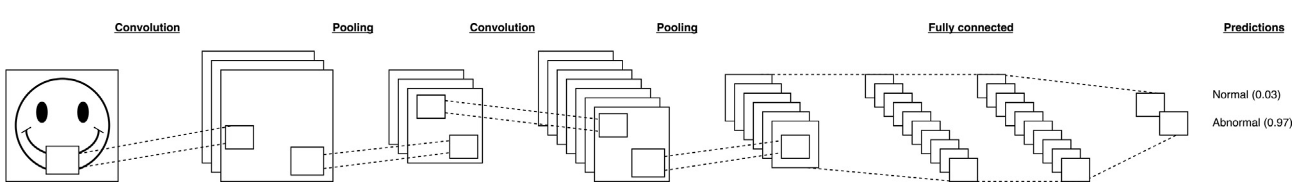

Figure 1 A pictorial representation of the "layers" of a deep CNN.

Please cite this article in press as: Yates EJ, et al., Machine learning "red dot": open-source, cloud, deep convolutional neural networks in chest radiograph binary normality classification, Clinical Radiology (2018), https://doi.org/10.1016/j.crad.2018.05.015