UDC 004.932.2

F. Smilianets

APPLICATION OF TRANSFER LEARNING FOR ENHANCED PULMONARY DISEASE DETECTION VIA CT IMAGE EMBEDDINGS

Abstract: This paper presents a method for computed tomography imaging analysis for disease diagnosis, extending and fine-tuning a previously trained network to generate embedding vectors. A KNeighborsClassifier trained on produced embeddings achieved an accuracy of 0.987.

Keywords: multiclass classification, convolutional neural networks, computed tomography scan analysis, embedding, transfer learning, COVID-19.

INTRODUCTION

The emergence of the COVID-19 pandemic, caused by the novel coronavirus first identified in Wuhan, China in late 2019, has profoundly impacted global health and healthcare systems. As the World Health Organization marked the end of the pandemic in May 2023, the disease caused 7 million fatalities out of over 750 million cases [1].

Despite this declaration, the potential for future respiratory illness outbreaks to escalate into epidemics or pandemics remains a stark reality, posing serious challenges to healthcare infrastructures worldwide.

Computed Tomography (CT) imaging is and has been instrumental in diagnosis and research of pulmonary diseases due to its non-invasive and highly flexible process that results in a high quality three-dimensional image of patient’s body.

Another advantage of medical imaging in general and CT imaging in particular is the ability to utilize computer vision in the analysis, allowing us to implement systems that help medical professionals offset errors and speed up their diagnostic or research process.

Of computer vision methods, most prevalent now are convolutional neural networks, allowing machine learning algorithms to achieve remarkable accuracies in various applications.

Problem overview and prior research

As this paper is a continuation of our research performed in [2], and, consequentially, the study by Rahimzadeh, Attar, and Sakhaei [3], the research problem addressed in this paper is that of developing effective and precise pulmonary disease recognition methods with an emphasis on extendibility, utilizing deep learning computer vision tools. In [2], we achieved 95.086% accuracy over multi-class classification on a dataset containing three classes by extending the binary classification network from [3].

Research goal

The goal of this paper is to improve classification accuracy of the neural network, proposed in [2] by augmenting it with other algorithms. To enable this, it is proposed to convert the network to produce a mechanism for generating vector representations of CT images (embeddings).

An approach like that has not been found to be employed with CT images at the time of writing, over several searches in Google Scholar.

Datasets

As a continuation of [2], this study uses the same datasets – COVID-CTset, introduced in [3] (Dataset A) and COVID-CTset, introduced in [4].

Dataset A, originating from Neghin Medical Center in Iran, consists of CT images taken with a SOMATOM scanner, processed using syngo CT VC30-easyIQ software. These images are in TIFF format, grayscale, and have a 16-bit color depth. The dataset includes scans from 95 COVID-19 patients and 282 healthy individuals. However, it lacks images of patients with non-hospital-acquired pneumonia, necessitating the inclusion of Dataset B to produce a multi-class dataset.

Dataset B, sourced from Babak Medical Center in Iran, also used a SOMATOM scanner. The images in this dataset are in DICOM format, with the same 16-bit color depth. This set includes 169 COVID-19 patients, 76 healthy individuals, and 60 with community-acquired pneumonia. To achieve consistency with Dataset A, we converted Dataset B images into 16-bit TIFF format using a modified version of dcm2hdr [5].

The choice of 16-bit grayscale TIFF format effectively merges the benefits of DICOM and conventional image formats. The 16-bit depth ensures the preservation of the full data richness found in the original CT scans, an improvement over standard 8-bit grayscale images. This fidelity is crucial for maintaining the integrity of the diagnostic information while ensuring the ease of image processing typically associated with common graphic formats.

Furthermore, to ensure the relevance of the images from B, we applied an image filtration algorithm as described in [3]. This process involved isolating images where the lungs are prominently visible. The algorithm assesses each image’s central region; if this area is predominantly dark, it indicates the presence of lung structures, thereby confirming the suitability of individual CT slices for our analysis.

Ultimately, the processed and filtered dataset consists of 34501 individual slices, of which 16651 images belong to patients with healthy lungs, 14107 belong to patients with covid pneumonia and 3743 belong to patients with commonly acquired pneumonia.

It is important to mention that datasets employed in this study classify the images per whole 3-dimensional CT image, and not per CT slice image, since it would be an immense task to manually go through each of 34501 CT image slices in the datasets. This introduces slices which do not contain features of COVID or CAP, but are labelled as such, which may influence network performance and/or how achieved testing results correspond to actual network performance in the real world.

Network design

This study focuses on adapting the modified multi-class classification network designed and trained in [2], through the use of transfer learning, to produce 128-long CT slice embeddings.

Therefore, the classification part of the source network consisting of Dense layers gathering information from the Feature Pyramid Network (FPN) [6], was stripped down and replaced by a new structure designed to produce an embedding, as shown on Fig. 1.

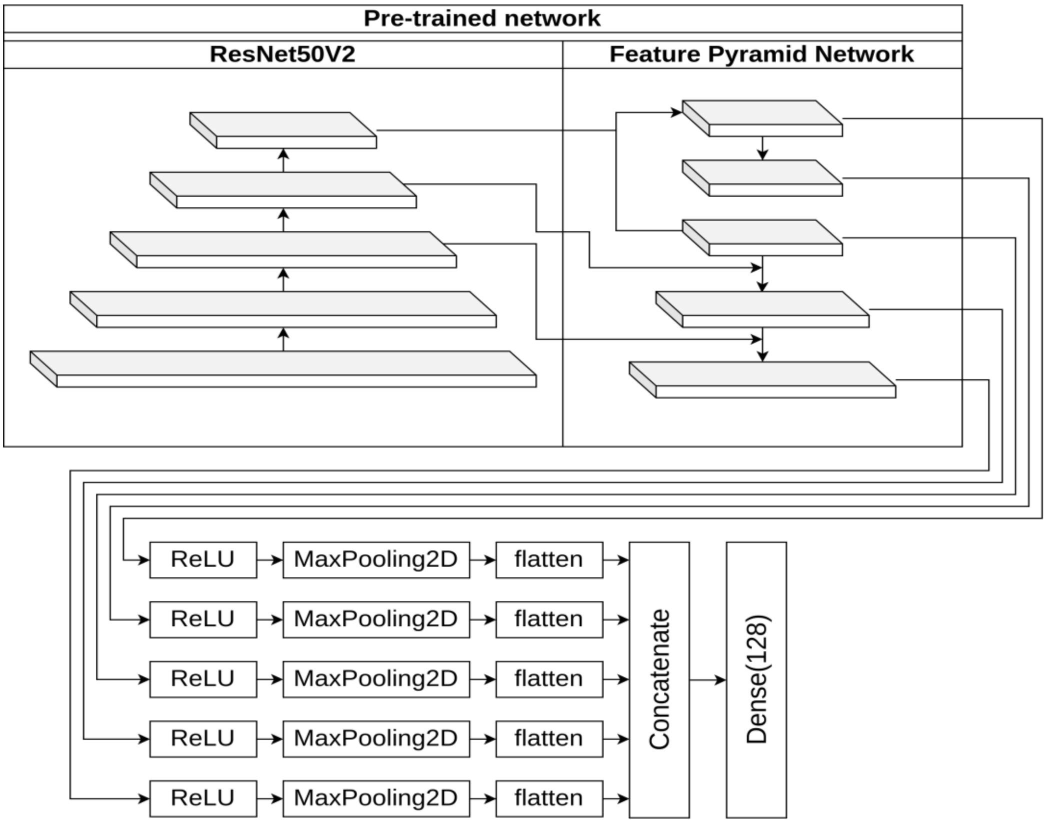

Fig. 1. Architecture of proposed neural network

As seen in the picture above, the dense part of the prior network was removed. The outputs of FPN layers were then put through ReLU activation and MaxPool layers to stratify the features and reduce the dimensionality of the data going through. After that, all outputs were concatenated and connected to a single Dense layer that serves as the output for the network and produces the vector representation of the CT slice.

Training procedure

To implement transfer learning and utilize the features learned by the convolutional part of the network in [2], the weights for them from the network, trained in this study, were applied to the network, and the layers were frozen, fixing the acquired knowledge. Therefore, only the output layer was trained to produce a vector. The dataset was split into three parts:

- main training dataset, consisting of 25440 images;

ISSN 1560-8956

- validation dataset, consisting of 5175 images, unused in this study;
- testing dataset, consisting of 5431 images.

Data batches were built using a mechanism, inspired by [7]. Each data batch consisted of 18 sets of anchor, positive, negative triplets, 6 sets per class, each item in each of those triplets would be randomly selected from the training dataset. To ensure good generalization, images were randomly augmented using Keras ImageDataGenerator functions get_random_transform and apply_transform.

At each training step, as in [7], the embeddings for all the anchor, positive and negative samples would be computed, and pairwise cosine similarities (1) between anchor and positive and anchor and negative would be calculated from the resulting vectors, and scaled down by a temperature factor of 0.2:

$$cosinesimilarity = \frac{A \cdot B}{AB}, \quad (1)$$

where A and B are the compared samples. To incentivize the network to output distinctive vectors for images from different classes, Triplet Loss [8] is used as described in (2), with an obvious correction that in the case of this study, vectors compared in the loss function are generated by a single network, and not by a triplet of siamese or other networks:

$$loss = max\left(P - N + m, 0\right), \quad (2)$$

where P is positive cosine similarities, N is negative cosine similarities and m is the margin. The selection of triplet loss for this case was invited by the fact that triplet loss forces the network to separate the comparison pairs by a specified margin, which incentivises the output layer we are training with the weights we transferred from a classification network to find exploit distinct features from the existing convolutional output.

Results

After 4300 batches of training, the resulting network was then used to vectorize the entirety of the dataset into embeddings. The embeddings from the main and supplementary training sets were used to train a KNeighborsClassifier [9] with 6 neighbours, which achieved a multi-class classification accuracy score of 0,987, supported by confusion matrix (Fig. 2)

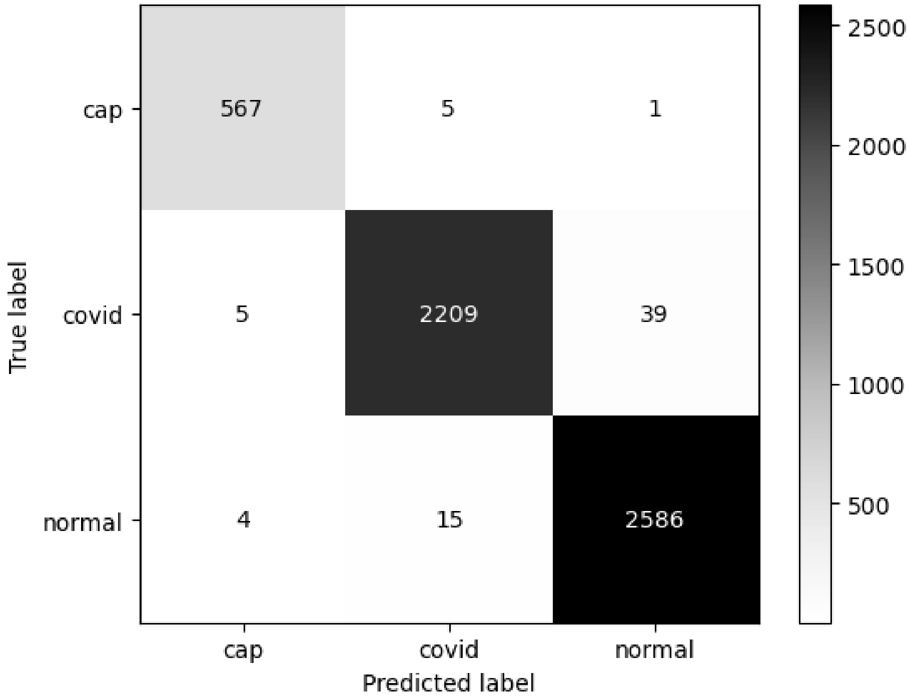

Fig. 2. Confusion matrix of a KNeighborsClassifier trained on embeddings produced by proposed neural network

This confusion matrix emphasizes the potential of classic machine learning algorithms aided by dimensionality reduction even in tasks as complicated as CT image analysis for disease classification.

Conclusions.

Previously proposed neural network for multi-class classification was restructured into one that would generate embeddings of CT images. Using embeddings generated with new neural network, a simpler algorithm was trained to achieve an accuracy of 0.987. Presented approach is a novel one, as other attempts at CT image embeddings have not been found. The neural network designed and trained in this study potentially paves the way to more accurate and useful medical decision support systems through better extensibility and modularity, as well as allowing vector database storage of CT scans to aid in similar clinical case search.

REFERENCES

1. WHO Coronavirus (COVID-19) Dashboard [Web resource]. – Resource access mode: https://covid19.who.int.

2. Multi-class classification of pulmonary diseases using computer tomography images / F. Smilianets and O. Finogenov // Adaptive systems of automatic control – 2023 – T. 43 – ISSN: 2522-9575 – URL:doi.org/10.20535/1560-8956.43.2023.292255.

3. A fully automated deep learning-based network for detecting covid-19 from a new and large lung ct scan dataset / M. Rahimzadeh, A. Attar, and S. M. Sakhaei // Biomedical Signal Processing and Control – 2021 – T. 68 – P. 102588 – URL:doi.org/ 10.1016/j.bspc.2021.102588.

4. COVID-CT-MD, a COVID-19 computed tomography scan dataset applicable in machine learning and deep learning / P. Afshar, S. Heidarian, N. Enshaei, et al. // Scientific Data – 2021 – T. 8 – URL:doi.org/10.1038/s41597-02100900-3.

5. DCM2HDR: DICOM to HDR image conversion. [Web resource]. – Data access mode: https://doi.org/10.5281/zenodo.1246724, accessed at 14.01.2024.

6. Feature pyramid networks for object detection / T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie – 2017 – URL:doi.org/10.48550/arXiv.1612.03144.

7. Metric learning for image similarity search / M. Kelcey – 2020 [Web resource]. – Resource access mode: keras.io/examples/vision/metric_learning/.

8. Deep metric learning using triplet network / E. Hoffer and N. Ailon – 2014 –URL: doi.org/10.48550/ARXIV.1412.6622.

9. k-Nearest neighbour classifiers - A Tutorial / Padraig Cunningham, Sarah Delany // ACM Computing Surveys – 2021 – T. 54 – URL: doi.org/10.1145/3459665.