International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 9, No. 3, 2022, pp. 16-20.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Medical Image Analysis for Disease Detection and Categorisation Using Deep Learning Algorithm

Revathi Krishnan [a], Gaurav Dandekar [b]

[a] *Graph Theory and Network Analysis Lab, Madurai Institute of Analytics, India*.
[b] *Graph Theory and Network Analysis Lab, Madurai Institute of Analytics, India*.

Abstract

Medical image classification is crucial for clinical treatment and educational purposes. The conventional approach has attained its performance limit. Utilizing them requires considerable time and effort to extract and select categorization characteristics. The deep neural network (DNN) is an innovative Deep Learning (DL) technique that has demonstrated its efficacy in many categorization problems. The Convolutional Neural Network (CNN) excels, achieving superior outcomes in diverse picture categorization tasks. Compiling medical picture collections is challenging due to the requisite professional competence for accurate labeling. This paper investigates the application of a CNN algorithm on a chest X-ray database for pneumonia classification. Three methodologies are assessed by experimentation. This text describes a linear Support Vector Machine (SVM) classification utilizing local rotation and orientation-invariant features, employing transfer learning on two convolutional neural network examples: Visual Geometry Group (VGG16) and InceptionV3, as well as a capsule network trained from the ground up. Enhancing data is a preprocessing technique utilized across all three approaches. The experimental results indicate that data enrichment is usually successful in improving the efficiency of all three methods. Transfer learning is a more effective classification technique for short datasets than an SVM utilizing Oriented Fast and Rotated Binary (ORB) resilient dependent elementary variables and capsule networks. In machine learning, it is crucial to retrain particular features on a new target database to enhance efficiency. The second critical factor is an appropriate network connection corresponding to the dataset's size.

Keywords: Medical Image Analysis, Disease Detection, Categorisation, Deep Learning.

DOI: 10.71086/IAJSE/V9I3/IAJSE0923
Received: 13 June 2022/Revised: 15 July 2022/Accepted: 22 August 2022/Published: 30 September 2022

Introduction

Medical imaging is a crucial diagnostic instrument for numerous disorders (Anwar et al.,2018). In 1895, Roentgen found that X-rays could non-invasively visualize the human body, leading to the emergence of X-ray radiology as the first imaging modality shortly thereafter. Since that time, numerous imaging techniques have been created, including Computer Tomography (CT), ultrasound imaging, Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET), alongside more advanced imaging processes (Filippou & Tsoumpas, 2018). Image information is essential for decision-making throughout multiple stages of care for patients, involving detection, characterization, setting, response to treatment evaluation, disease recurrence tracking, and the guidance of interventional treatments, surgeries, and radiation therapy treatments. The quantity of images for a particular patient case escalates significantly from several 2D pictures to hundreds with 3D photography and hundreds with 4D dynamic photography. The utilization of multi-modality imaging enhances the volume of picture data requiring interpretation. The escalating workload hinders radiologists and doctors from sustaining efficient workflows while leveraging every piece of imaging data to improve accuracy and care for patients. Recent advancements in Deep Learning (DL) and computer science have underscored the need to create automated tools to aid radiologists in image interpretation and evaluation, marking it as a crucial medical imaging domain of study and creation.

Utilizing computers to analyze medical images automatically originated in the 1960s. Numerous studies have shown the viability of using computers for medical image processing; this research has not garnered significant attention, perhaps due to restricted access to high-quality digitized image information and computational capabilities (Razzak & Zaib, 2017). A study initiated the systematic advancement of Machine Learning (ML) (Barragán-Montero et al., 2021) and image processing methodologies for medical imaging in the 1980s, aiming to create Computer-Aided Design (CAD) (Fay, 2021) as a supplementary opinion to aid radiographers in translating images.

Background

DL (Mishra et al., 2020) has become the preeminent ML technique in numerous applications. DL is a representation learning technique wherein a sophisticated multi-layer Neural Network (NN) design autonomously learns data models by converting input data into various levels of abstraction. Deep Convolutional NN (DCNN) is the predominant DL architecture for picture pattern recognition applications. Given an adequately extensive training set, a DCNN can autonomously identify appropriate characteristics from the training examples for a specific task by iteratively modifying its weights through backpropagation. DCNN hence identifies representations of characteristics via training and does not necessitate manually crafted features given input. When adequately trained on a substantial and relevant dataset, the features derived by a DCNN are anticipated to surpass handmade characteristics due to their enhanced selectivity and consistency. Significantly, due to the computerized nature of the learning method, DL can efficiently assess dozens or millions of instances that human experts cannot perceive or retain during their careers. DL can be more resilient to the extensive feature variances among distinct classes, provided the training set is vast and diverse for analysis.

Convolutional NN (CNN) (Anwar et al., 2018) originates from the neocognitron introduced in the early 1980s. A CNN was initially trained using backpropagation to recognize written digit patterns. CNNs were employed in multiple fields, including object identification and characters and recognition of faces, in the early 1990s. The initial use of CNN to analyze medical images occurred in 1993 when a CNN was trained to identify lung nodules in chest radiography. The previous year, CNN was utilized for microcalcification identification in mammograms, followed by bulk identification in the subsequent year (Kang et al., 2021).

Proposed DL-based Medical Image Analysis

CNNs are the most extensively studied DL techniques in medical imaging analysis. This occurs because CNNs maintain spatial associations when processing input images. Spatial interactions are critically significant in radiography, such as the junction of bone and muscle or the interface between healthy lung tissue and malignant tissue. As seen in Fig. 1, a CNN processes an input picture with raw pixel values through Convolutional Layering, Rectified Linear Unit (ReLU) Layering, and Pooling Levels. This culminates in the last Fully Connected Layer (FCL) that allocates the grade or odds, thereby categorizing the data provided into the lesson with the most significant probability.

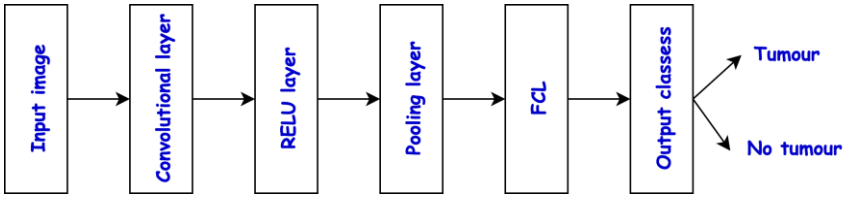


Fig. 1: Layered Architecture

3.1 Convolution Level

A convergence is an operation involving two functions. In image processing, one function comprises the parameters (e.g., pixel counts) at a specific place in the image, while the other function is a filtering (or kernel); both can be expressed as arrays of numbers. The computation of the dotted product of the two variables yields a result. The filter is moved to the following location in the image as determined by the stride duration.

The calculation is reiterated until the complete image is encompassed, resulting in a mapping of attributes (or activation). This visualization indicates areas where the detector is active and detects features like a straight path, a dot, or a curved border. When an image of a face is input into a CNN, its filters first identify low-level features like lines and corners. These accumulate into increasingly complicated characteristics in future levels, such as a nose, eye, or ears, as the maps of features serve as inputs for the following layer in the CNN structure.

3.2 Rectified Linear Units

The ReLU level is a function of activation that nullifies unfavorable values entered. This streamlines and expedites computations and training while mitigating the disappearing slope issue. It is formally defined as:

$$f(x) = max(0, x) \quad (1)$$

Let x represent the input to the neurons. Additional functions for activation comprise the sigmoid, tanh, leaky, randomized, and parameterized RELUs.

3.3 Pooling Level

The pooling level is positioned between the convexity and RELU levels to diminish the number of variables to be computed and decrease the image dimensions (width and elevation excluding depth). Max-pooling is the more prevalent method; alternative pooling layers are Average pooling and L2-normalization pools. Max-pooling extracts the maximum input value in a filter while disregarding the remaining values, summarizing the most significant stimulation within an area. The reasoning follows that the relative positioning of a highly active feature concerning another is more important than its precise location.

3.3 Dense Layer

The last component of a CNN is the FCL, which indicates that each neuron in the previous layer is interconnected with every neuron in the FCL. Like convolution, ReLU, and layer pooling, the number of wholly connected layers varies based on the desired degree of feature abstractions. This stage receives input from the previous layer (CNN, RELU, or combination) and calculates the likelihood score for categorization into several accessible categories. This section analyzes a combination of the most prominently activated elements that suggest the image corresponds to a specific class.

For instance, on histological slides made of glass, tumor cells exhibit a higher DNA-to-cytoplasmic ratio than typical cells. If DNA characteristics were prominently identified from the previous layer, the CNN would have been more inclined to forecast the existence of cancerous cells. CNN development techniques like backpropagation and random gradient descent enable CNN to acquire significant correlations from training pictures.

Results and Findings

4.1 Equipment

An average high-performance computer suffices for Orientated fast and Rotated Binary (ORB) and Support Vector Machine (SVM) categorization, such as one with 8GB of memory, an i5 processor (2.4 GHz), and a 256GB Solid-State Drive (SSD). A Graphical Processing Unit (GPU) should be utilized to expedite the training procedure when developing Deep Neural Networks (DNN). The report in question uses a Google GPU. A virtual machine example equipped with a four-core CPU, 8 GB of memory, and a Tesla K80 is used.

4.2 Software

A Python software initially designed for identifying plants has been adapted to evaluate ORB and SVM classifications. It was altered to utilize the new dataset and executed on a laptop. Each round of the test requires around four hours. The CNN-based technique is computationally expensive, necessitating execution on a Virtual Machine (VM) in Google GPU Clouds. A Python version of a Capsules network designed for brain tumor detection was adapted with the pneumonia datasets. It must also be executed on the GPU virtual machine.

4.3 Outcomes

Table 1: Data Augmentation Analysis

Aug methods	Training pictures	Accuracy (%)
0	5200	87.3
1	5200	89.5
2	5200	88.4
1	8000	91.3
2	8000	90.5

Table 1 presents the augmentation methods employed in the evaluation in the first section, the total number of training pictures generated through augmenting in the second section, and the mean accuracy attained by VGG16 transferring learning algorithms with all the standard settings in the last row. The results indicate that Aug-l is a superior augmentation paradigm to Aug-2; Aug-l will be utilized as the default enhancement framework for subsequent trials. Table 2 demonstrates that VGG16 outperforms InceptionV3 and Model3, establishing VGG16 as the superior classification system. In the subsequent test, VGG16 and the model will undergo more fine-tuning.

Table 2: Classification Analysis

Model	VGG16	Inception V3
1	87.4	68.4
2	75.6	82.3
3	83.4	90.2
4	89.5	87.4
5	86.9	89.4

Conclusion

This research investigates the application of CNN-based categorization to a short chest X-ray database, given the significance of medical image categorization and the specific challenges posed by limited medical image datasets, and evaluates its results. The following findings were provided from the studies. CNN-based learning through transfer is the superior method among the three approaches. The capsule networks outperform the ORB and SVM classifiers.

CNN-based techniques generally outperform conventional methods because they can learn and choose features dynamically efficiently; the optimal results are achieved using transfer learning with VGG16, incorporating one retrained convolutional layer, which marginally exceeds the state-of-the-art performance. The unfrozen convolutional layer can acquire specific features from the fresh information. The particular feature is crucial for enhancing accuracy; balancing a model's expressive capacity and overfitting is imperative. An overly simplistic network typically fails to extract sufficient information from the information, resulting in inadequate accuracy. A highly intricate network is challenging to train and is prone to rapid overfitting. Precision is insufficient. A model of networks of appropriate scale, combined with practical techniques to mitigate overfitting, such as an adequate dropout rate and suitable addition of data, is essential for optimal results. Due to time constraints, further investigation is necessary. In transfer training and learning, a fine-tuned deep neural system with thawed convolutional layers often results in overfitting.

References

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. *Journal of medical systems*, *42*, 1-13. https://doi.org/10.1007/s10916-018-1088-1

Filippou, V., & Tsoumpas, C. (2018). Recent advances on the development of phantoms using 3D printing for imaging with CT, MRI, PET, SPECT, and ultrasound. *Medical physics*, *45*(9), e740-e760. https://doi.org/10.1007/s10916-018-1088-1

Razzak, M. I., Naz, S., & Zaib, A. (2017). Deep learning for medical image processing: Overview, challenges and the future. *Classification in BioApps: Automation of decision making*, 323-350. https://doi.org/10.1007/978-3-319-65981-7_12.

Barragán-Montero, A., Javaid, U., Valdés, G., Nguyen, D., Desbordes, P., Macq, B., ... & Lee, J. A. (2021). Artificial intelligence and machine learning for medical imaging: A technology review. *Physica Medica*, *83*, 242-256. https://doi.org/10.1016/j.ejmp.2021.04.016.

Fay, C. D. (2020). Computer-aided design and manufacturing (CAD/CAM) for bioprinting. *3D Bioprinting: Principles and Protocols*, 27-41. https://doi.org/10.1007/978-1-0716-0520-2_3.

Mishra, S., Tripathy, H. K., & Acharya, B. (2020). A precise analysis of deep learning for medical image processing. In *Bio-inspired neurocomputing* (pp. 25-41). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-5495-7_2.

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. *Journal of medical systems*, *42*, 1-13. https://doi.org/10.1007/s10916-018-1088-1.

Kang, D., Gweon, H. M., Eun, N. L., Youk, J. H., Kim, J. A., & Son, E. J. (2021). A convolutional deep learning model for improving mammographic breast-microcalcification diagnosis. *Scientific reports*, *11*(1), 23925. https://doi.org/10.1038/s41598-021-03516-0.