Convolutional Neural Networks for Medical Image Diagnosis and Prognosis

A. Parameswari[a], K. Vinoth Kumar [b]

[a] Department of ECE, Adithya Institute of Technology, Coimbatore, India.

[b] Department of ECE, SSM Institute of Engineering and Technology, Dindigul, India.

* Corresponding Author: parameswarieceme@gmail.com

Abstract: One of the most incredible machine learning methods is deep learning. Utilised for picture categorization, clinical archiving, item identification, and other purposes. The quantity of medical image archives is expanding at an alarming rate as hospitals employ digital photos for documentation more frequently. Digital imaging is essential for assessing the severity of a patient's illness. Medical imaging has a wide variety of uses in research and diagnostics. Due to recent developments in image processing technology, self-operating identification of medical photos is still a research area for computer vision researchers. We require an appropriate classifier in order to categorise medical photos using various classifiers. After organ prediction and classification, the research was modified to include medical picture recognition. For medical picture detection, pretrained convolutional networks and Kmean clustering techniques similar to those used for organ identification are employed. Separating the training from the test data allowed for the data's authentication. The application of this strategy has been proven to be most effective for categorising various medical images of human organs.

Keywords: Convolutional neural networks, Deep neural networks, K- Mean Clustering.

1. Introduction

The exponential growth in medical picture output has been altered by the widespread use of digital devices and camera technology. A computerized image is currently used in a modern hospital to predict the severity of a patient's disease. Classification of photographs has grown in usefulness as digital images develop quickly. Another sort of scan that appears to be capable of producing accurate data as well as finely detailed images of the brain and other bodily organs is magnetic resonance imaging (MRI). Because medical photographs include so much data, it is virtually hard for a doctor or medical professional to manually classify images. The interpretation of functions based on images is therefore fully dependent on the recognition rate.

Although these characteristics are necessary for categorization, they must be manually created in order to maintain awareness of prior information. The labelling of the data is necessary for training and tweaking of hand-crafted features, which is a challenging operation. A data abstraction approach for picture recognition and classification tests is more vivid to the range of imaging techniques and illnesses because it is less sensitive to different domain-specific intentionality. [1]. The suggested system contributed to lower prostate and some other framework costs. This method for classifying prostates is substantially more accurate. The application's use of an optimization method with the aid of texture analysis to discriminate between the bladder wall's exterior and internal limitations was, however, constrained because the technique's assessment was based on sparse data [2]. To distinguish between the boundaries of strength and weakness, the scholars adopted goals. In a segmentation technique, the bidirectional convolution repeated strands are used internal and many data from the prostate to retrieve [3]. The basic goal of medical image recognition research is to correctly categorise medical images of various body sections; as a result, training-testing technologies can address the problems. For categorization, picture interpretation is crucial. Feature extraction is a crucial step in the medical sector. The results of our learning process are based on the principles of test data and comprise a set of predetermined characteristics and images. Reduced visual elements including colour, texture, spatial arrangement, and form are frequently sent as extracted characteristics in order to describe the data. [4-5].

2. Literature Review and Related Works

a) Automatic Medical X-Ray Image Classification Using a Bag of Visual Words

Mohammad Reza Zare, Ahmed Mueen, and Woo Chaw Seng suggested an iterative classification framework in this research to address the model's accuracy [6]. This framework generates four classification models from various classes. The rapid advancements in medical technology have led to an increase in the production of medical images. The cost of manually reviewing these photographs is extremely expensive, and errors can occur. It has been improved as a result of efficient classification that makes use of information that is directly taken from the image. One common method is to divide the photographs into subgroups and separately extract the visual qualities from each area before actually combining the photos into a single application. The bag of visual words (BOW) approach has also been applied to artificial tasks and medical picture categorization. Training and testing are the two stages of the categorization process. In order to develop a classification model, the classifier is trained on extracted methods while the selected features are retrieved from training photos. A technique for feature extraction is the bow. To fulfil this assignment, each feature vector in an image must have a Euclidean metric neighbor so as to bring out the result.

Disadvantages

- All of the findings pointed to an average level of classification accuracy throughout the entire dataset.The minimal precision cannot be obtained for all individual categories.

- The basic flaw with databases and other similar databases is that their highest performance on primary classes conceals their lack of accuracy in secondary classes.

- Due to the inter-class equality between the two classes and the fact that they were fetched from the same sub-body region, am is classification has occurred.

b) CNN's Article on Identifying Medical Images

CNN's article on identifying medical images the most important picture classification and segmentation challenge in the world of image analysis, ImageNet Challenger, was successfully completed using several CNN-based deep neural networks that were constructed. The task of medical classification makes extensive use of the CNN-based deep neural system. Using CNN to identify medical images might prevent difficult and expensive feature engineering because it is a great feature extractor. A customised NN with a shallow ConvLayer was provided by Qing et al. [2] to categorise lung disease picture patches. The technique may be applied generally to different medical imaging collections, the authors further discovered. According to a previous work, a CNN-based system may be trained utilising sizable chest X-ray (CXR) film datasets. These datasets include ChestX-ray8, a new CXR database with 108,948 frontal-view CXR, and the Stanford Normal Radiology Diagnostic Dataset, which comprises more than 400,000 CXR [3]. Furthermore, training a good model is challenging when there is a lack of data. As a result, jobs involving the classification of medical images utilise CNN's transfer learning extensively. 108,312 optical coherence tomography (OCT) pictures are included in the medical image dataset used by to instruct InceptionV3 using weighted transfer learning and ImageNet training. With a sensitivity of 97.8% and a specificity of 97.4%, they achieved an average accuracy of 96.6% [7]. Six human experts were also used by the authors to compare the findings. While the CNN-based approach received high scores for both sensitivity and specificity, the majority of experts received high sensitivity but low specificity. On the average weight error metric, the CNN-based system performs better than two human experts. Additionally, the researchers put their system to the test on a small group of images pertaining to pneumonia, numbering around 5,000, and discovered that it performed with an average accuracy, sensitivity, and specificity of 92.8%, 93.2%, and 90.1%. This technology could finally speed up patient diagnosis and referral, leading to early therapy introduction and a higher cure rate. Vianna has researched how to create an X-ray image categorization system utilising transfer learning, which is an essential component of a computer-aided diagnostic system. Compared to a transfer learning model with only the final classification layer retrained, and training from scratch —the authors discovered that the over fitting issue is effectively resolved by a fine-tuned transfer learning system with data augmentation, which also yields improved outcomes [4].

c) Classification of Image Net using Deep Convolutional Neural Networks

Hinton in many of the currently utilized approaches for object recognition, machine learning techniques are used. We may improve our position by amassing more datasets, learning more efficient methods, and utilising the best anti-over fitting strategies. Up until recently, databases of annotated photos contained tens of thousands of pictures. With databases this size, simple recognition tasks may be solved reasonably successfully, especially when label-preserving

transitions were added. Over 15 million high-resolution photographs with labels that are separated into roughly 22,000 categories may be found on Image Net [8]. The images were obtained from the internet and branded by users of Mechanical Turk, an Amazon crowdsourcing tool. Since 2010, the Pascal Visual Object competition has included the Image Net Large-Scale Visual Recognition Challenge (ILSVRC) [9, 10]. Despite the fact that we believe it to be helpful, we did not employ unsupervised pre-training to modify our experimental research, especially supposing we have enough processing power to greatly increase the network size without receiving an equivalent increase in the amount of tagged data. Although it will take many orders of magnitude to match the infer temporal path of the human visual system, our findings thus far have improved as we have widened our scope and spent more time preparing [11]. Last but not least, we'd like to use very large, deep convolutional networks on video streams to give information that is either missing or much less visible in static images because of the way that time is organized in the stream [12].

The amount of memory that is currently available on GPUs and how much training time we are prepared to put up with are the main factors limiting the network's capacity.

- Ontwo GTX5803 GBGPUs, the network trains in five to six days.

The experiment shows that the results may be improved just by waiting for faster GPUs and more datasets to become available.

3. Proposed Method

A. Image acquisition

We used a publically accessible dataset for the organ classification approach that includes pictures of the chest, breasts, and other areas of the human body. There are 12 classes in all, and it's crucial to highlight that 11 of them originate from a source of cancer picture archives that is open-access, while the remaining 12 are from Messidor. The dataset contains 3600 photos divided into 12 classes, with 300 images in each class. Given that we used a training-testing architecture, For our case, we chose a train-test ratio of 70% and 30% at random. Thus, 2520 photographs are utilised for training, whereas 1080 images are used for testing. There isn't a single image used for both testing and training. All pictures are converted from the DICOM (Digital Imaging and Communications in Medicine) format to JPG in the first step. The increasing diversity may cause the feature matrix for the neural network to categorise more accurately to become more complicated. This issue was resolved by using intensity normalisation before supplying the input to the convolutional neural network. We used the publicly accessible dataset that includes photos of the breast and brain for the medical image detection algorithm. These two classes were found in medical picture repositories. The only dataset with images for the class of brain has 353 images, and the class of breast has approximately 3000 images. In this case, different images are used for training and assessment.

B. Data Pre-processing

Deep learning is currently widely used in a variety of programmes that deal with things other than the human brain and are based on artificial neural networks. Deep structure models

such as feed-forward neural networks, which include several hidden layers, are very good examples. To determine the physical area that each pixel image represents, all of the images are scaled up to the same distance. After normalising, each image was reduced to 224*224 pixels, and a 224*224 entry image was given to the first convolution layers. Using the ReLu activation function, a kernel of four by four with equal padding, a stride of 1, and eight filters are used to create the first convolutional layer. All of the pixels were reduced in size to 224*224 pixels in order to be used as input for the InceptionV3 method.

C. Fine Tuning of Pre-trained Network

The first level of a DCNN picture type and detection method is characteristic extraction, and the second level is a type or detection module. In a stop-to-quit learning framework, the characteristic is extracted by providing the instructional images and then being identified. A classifier of the light Max layer is then employed for the educational photo data. The educational version featured several layers, five of which were convolutional and three were completely linked. In comparison to custom functions, the deep learning system instantly learns low-stage, mid-stage, and summary qualities from photographs. A training set of pictures is used as an input for the Inception-V3 preskilled approach to extract the deep functions. In our example, we employ 48 layers, including convolution layers, completely linked layers, and a set of pictures with different modalities. We moved the Inception-V3 method's capabilities by using our collection of scientific picture data for both validation and education. Soft feature's main objective is to execute re-mastering using the dataset's 12 lessons. The 0.33 in a series of Deep mastering convolutional Architectures is the version of Google's Deep studying Convolution Architectures, Inception V3. Inception V3 is taught using a dataset of 1,000 training images from the real image net dataset, which has one million training images. A deep community technique called switch learning enables us to retrain a network in accordance with new levels by using fine-tuning parameters. During the fine-tuning stage, functions are taken from the provided collection of images. To minimise the size of the input, a max-pooling layer is applied along with the four*four kernel, the same padding, the stride of one, and eight filters. The output of this pooling layer is 112 by 112 dimensions. With the purpose of re-joining the remaining three layers with the rest of the network, images are used as entry for a pre-trained Inception-V3 method and frozen. We added a global average Pooling Layer, a Dropout layer, a Batch Normalization Layer, and in our case, a Dense (tender Max) output layer to the pre-trained Inception-V3 algorithm. The length of the final totally related layer, which is 12, is equal to the variety of instructions in our dataset. In order to improve Inception-mastering V3's process, we increased the learning price components of the completely connected layer by employing a connection between the transferred layer and the ultimate community. The DCNN model is ready for usage after it has been sensitively trained to classify medical images and for the detection of most malignancies. Figure 1 illustrates the suggested approach's block diagram
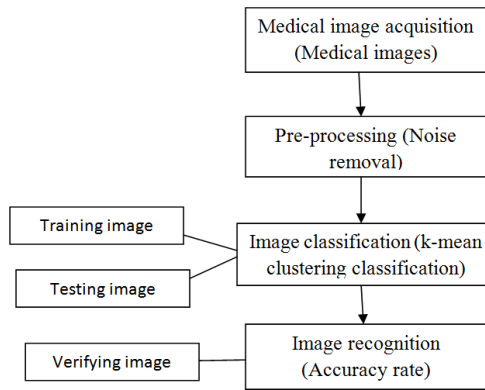
Figure 1. Block diagram of proposed method

D. Neural Network on Medical Image Classification

Clustering Methods:

By calculating fresh centroids equal to the mean of all the points assigned to each previous centroid, K-means (**KM**) iteratively assigns each feature point to its nearest centroid. If there is a significant difference between the old and new centroids but it is less than a certain threshold, iteration stops. The centroids are initially either selected at random or given to the algorithm. The initialization of centroids is crucial to the algorithm's convergence. The clustering pictures of the provided training photos are displayed in Figure 2.
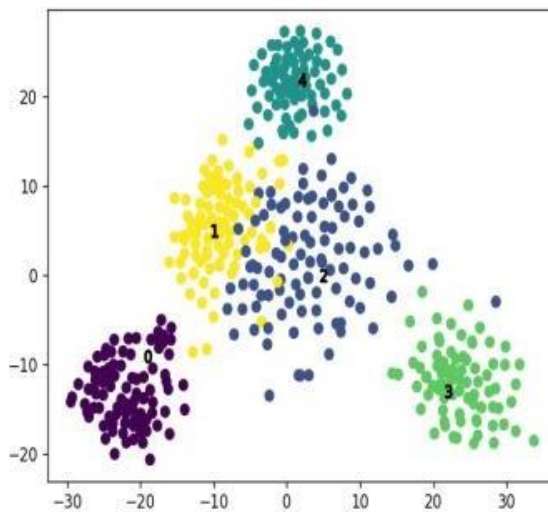

Figure2. Cluster Corresponding to Image

4. Experimental Results

The proposed deep convolution neural network has been created and trained in this research utilising well-known and widely-applied deep learning methodology for categorising medical

images and detecting medical images. In this instance, pre-processing the images was done first, then the model was tuned to increase classification accuracy while short in graining time, and last it was used to identify the given medical images. The pre-processing of the images and subsequent fine-tuning are done in the medical image detection model as well. Pre-processing cleans up the photos of the imperfections. The remaining 30% of the photos are then saved for validation and testing while the remaining 70% are used for training. Ten epochs were used to train the network. After training and validation, we were able to increase the net accuracy of classifying organs to about 97.95% compared to test accuracy and validation accuracy. The below Figure. 3, Figure. 4, Figure. 5 shows the experimental view of medical images.
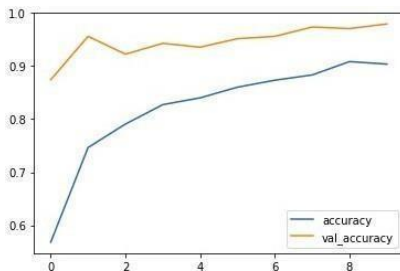

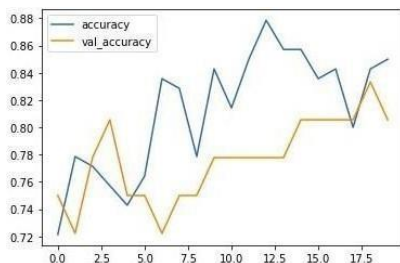**Figure 3.**Validation and training accuracy of organ classification.

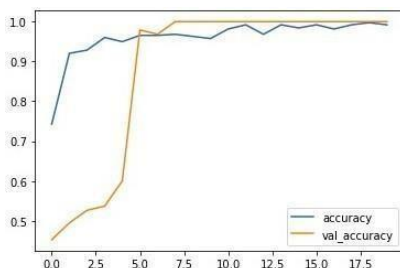**Figure 4.** Accuracy of medical image detection in brain.

Figure 5. Accuracy of medical image detection in breast.

5. Conclusion and Future Work

This technique for categorising picture data uses features and a pre-trained CNN and may be applied to large-scale platforms and used in real-time by mobile apps in the field. It shows how using deep learning for supervised classification and unsupervised clustering helps in the creation of a model to automatically categorise these pictures at the source with little help from professionals. Most crucially, this provides assurance that the deep CNN's high-level features, which were learned on generalise well to pictures on which it was not trained, and a vast diversity of images. We also want to use more classifiers to improve the system's precision and effectiveness, examine picture pre-treatment and segmentation approaches, investigate the effects of modifying the parameters used by clustering algorithms, and expand the database in terms of classes and diversity.

References

[1] Q. Zhu, B. Du, B. Turkbey, P. Choyke, & P. Yan, (2018). Exploiting interslice correlation for **MRI** prostate image segmentation, from recursive neural networks aspect. *Complexity,*2018. https://doi.org/10.1155/2018/4185279

[2] Q. Li, W. Cai, X. Wang, Y. Zhou, D.D. Feng, M. Chen (2014). Medical image classification with convolutional neural network. *13th International Conference on Control Automation Robotics & Vision (ICARCV),* IEEE, Singapore. https://doi.org/10.1109/ICARCV.2014.7064414

[3] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, Chestx-ray8: hospital-scale chest X-ray database andbenchmarks on weakly-supervised classification and localization of common thorax diseases. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* IEEE, USA. https://doi.org/10.1109/CVPR.2017.369

[4] V.P. Vianna, (2018). Study and development of a computer-aided diagnosis system for classification of chest X-ray images using convolutional neural networks pre-trained for imagenet and data augmentation. *arXiv preprint.*

[5] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, I. Eric, & C. Chang, (2014). Deep learning of feature representation with multiple instance learning for medical image analysis. *IEEE international conference on acoustics, speech and signal processing (ICASSP),* IEEE, 1626-1630.

[6] M. Reza Zare, A. Mueen, & W. Chaw Seng, Automatic classification of medical X-ray images using a bag of visual words. *IET Computer Vision,* 7(2) (2013). 105-114. https://doi.org/10.1049/iet-cvi.2012.0291

[7] D.S. Kermany et al., Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell,* 172(5) (2018) 1122-1131. https://doi.org/10.1016/j.cell.2018.02.010

[8]   R. Ashraf et al., Deep Convolution Neural Network for Big Data Medical Image Classification," in IEEE Access, 8 (2020) 105659-105670,.https://doi.org/10.1109/ACCESS.2020.2998808

[9]   J. Wan, D. Wang, S.C.H. Hoi, P. Wu, J. Zhu, Y. Zhang, & J. Li, Deep learning for content-based image retrieval: A comprehensive study. In *Proceedings of the 22nd ACM international conference on Multimedia.* (2014) 157-166. https://doi.org/10.1145/2647868.2654948

[10]   A. Krizhevsky, & I. Sutskever, Geoffrey E. hinton, ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing System 25* (2012).

[11]   S.H. Dar T. Mahmood, M. Yasir, Z. Abbas, Deep convolution neural network for big data medical image classification, *IEEE Access, 8* (2020) 105659-105670. https://doi.org/10.1109/ACCESS.2020.2998808

[12]   A. Parameswari, K.V. Kumar, & S. Gopinath, Thermal analysis of Alzheimer's disease prediction using random forest classification model, *Materials Today: Proceedings.* 66(3) (2022) 815-821. https://doi.org/10.1016/j.matpr.2022.04.357

Funding

Conflict of interest

The Authors have no conflicts of interest to declare that they are relevant to the content of this article.

About The License