Volume 1 No 1, 2025
Journal of Global Research in Multidisciplinary Studies (JGRMS)
Research Paper
Available online at https://saanvipublications.com/journals/index.php/jgrms/index

Convolutional Neural Network-Based Medical Image Classification

Mr. Vishal Joshi,
Sr. Software Engineer, John Deere,
One John Deere Place, Moline, Illinois, United States
JoshiVishal@JohnDeere.com

***Abstract*—Using imaging techniques, human body irregularities are captured. To diagnose, prognosis, and schedule treatment for inconsistency, it is necessary to comprehend the collected images. Generally, qualified medical personnel classify medical images. The inadequacy of human experts, as well as their exhaustion and imprecise estimation methods, restrict the effectiveness of image comprehension performed by qualified medical professionals. The tool for processing images effectively is a CNN. In several image interpretation competitions, they have outperformed human experts. Traditional classification techniques haven't been able to keep up with the complexity of medical images for practical application. In the past few decades, the theory of DL has developed quickly, providing technical solutions to problems with the classification of medical images. An efficient method for this is transfer learning, which involves transferring a domain among two medical image datasets or (unsupervised) adjusting pre-built CNN frameworks from a dataset of natural images to a dataset of medical images. Deep CNNs are difficult to train from scratch because convergence requires a significant amount of labeled training data as well as a skilled team. This can be enhanced, for example, by fine-tuning a CNN with lots of distinctive natural images. This investigation reveals an accuracy of 99.42 percent.**

Keywords—CNN, Transfer Learning, Medical image classification, Segmentation, Data Augmentation, Detection

I. INTRODUCTION

Deep learning, one of the top 10 breakthrough techniques in 2013, is a significant trend in the creation of automated applications. Many deep learning-based computer programs outperform humans at tasks like spotting tumors on MRI scans and cancer markers in the blood. It is a superior ANN with additional hidden layers, allowing for a higher level of abstraction and improved image processing. Due to its recent unrivaled performance in a variety of applications, including speech recognition, object recognition, medical imaging, and face recognition, this technology has gained widespread acceptance. CNNs have had a considerable impact on the field of visual interpretation. CNN-based methods are at the top of the leaderboards for the Multimodal Brain Tumour Segmentation, (BRATS), (MICCAI), Imagenet categorization, (ICPR), and an Ischemic Stroke challenge. CNN is gaining popularity as a resource for understanding medical images. Studies have extensively used CNNs for a variety of medical picture comprehension apps, including the identification of tumors & categorization of benign & malignant tumors, as well as the detection of skin blemishes. Examining internal organs to look for anomalies in the anatomy or function requires medical visualization. Medical images recording devices like CT, X-ray, PET, MRI, and ultrasound scanners show pictures or videos of the internal organs' structure and operation. To correctly identify anomalies and diagnose functional impairments, photographs, and videos must be interpreted. If an anomaly is found, its precise size, location, and shape must be identified. Traditionally, trained physicians performed these duties using their judgment and experience. The goal is for smart healthcare systems to use AI medical image understanding to do these tasks. Medical image interpretation focuses primarily on classification, detection, segmentation, and localization. Medical image classification requires picking and labelling pictures from a collection. A image is analysed by first extracting its characteristics, and then labels are assigned according to those features. CNNs represent a quantum leap in a region of picture comprehension, which includes picture categorization, partition, localization, & recognition, among other things. The effectiveness of CNNs in picture interpretation is a primary cause for their widespread adoption. Comparable to the neurons (nerve cells) of an animal, CNNs consist of convex with learnable weights and biases. CNNs' main building elements are conv surfaces, activation functions, pooling, & fully connected surfaces. Animal brains' visual cortex is made up entirely of neuronal cells that filter out image characteristics. Visual comprehension is aided by the information extracted by individual neural cells. The goal of the conv surfaces, which are modeled after neural cells, is to extract properties including edges, colors, texture, and gradient direction. CNNs heavily rely on activation functions, Conv surfaces, surfaces, and pooling with complete connections. Neurons within the visual cortex of an animal's brain extract image features. Each brain cell extracts specific information to improve visual perception. The conv surface, which is modeled over a neural cell, is responsible for extracting features from the neural cells such as edges, colors, texture, and gradient direction.

Deep learning is gaining popularity in a variety of fields, including medical image analysis, where it is predicted to account for $300 million in revenue by 2021. As a result, by 2021, it will have received extra funding for medicinal visualizing than the whole analytical industry had spent in 2016. It is the mainly efficient & SVM method available. This method utilizes deep neural network models, a type of neural network that provides a more precise approximation of a person's mind through the use of complex mechanisms than simple neural network models. "Deep learning" refers to the application of a DNN model. Motivated by the evolution of the human mind, the neuron is the fundamental computational unit of a neural network. It accepts multiple input signals, linearly combines them using weights, and then applies nonlinear processes to a contrasted channel to generate output channels. The primary objective is to motivate medical image comprehension specialists to make extensive use of CNNs in their investigation and diagnosis. CNNs have been described in brief detail. CNN has been discussed, as well as its multiple award-winning frameworks. The key medical image comprehension challenges have been introduced, including picture classification, segmentation, localization, and detection.

To create the described hierarchical property in a DNN, multiple layers of neurons are stacked one on top of the other. Now, there are more than one thousand surfaces! After being trained with a sufficiently large knowledge database, a DNN can produce intelligent forecasts, including interpolations and/or extrapolations, for unknown causes. The impact of deep learning on computer vision and medical visualization is therefore substantial. In reality, similar effects can be seen in a variety of contexts, such as text, voice, etc. The deep learning algorithms used in research include (deep autoencoder-DA), (DC-ELM-deep deep conventional extreme learning machine), (CNN-Conv neural networks), (Deep Belief Networks-DBN), (DBM-Boltzmann machine), (Deep Neural Networks-DNN), and their various MDLATM and BLSTM. CNN paradigm is gaining a great deal of traction in digital image processing and vision. There are numerous CNN architecture types, such as faster R-CNN3, Alexnet1, googleNEt4, Lenet2, VGGNet6, ResNEt5, ZFnet, and others.

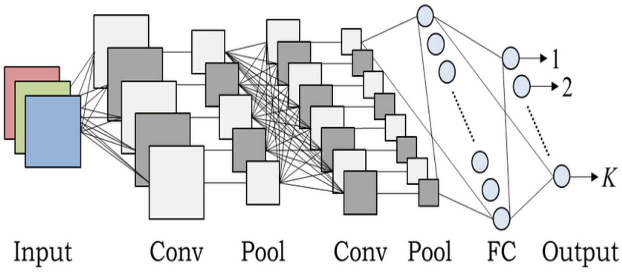

Fig. 1. Architecture of CNN

For clinical care and educational responsibilities, the categorization of medical images is crucial. Meanwhile, the conventional approach is no longer viable. Use of these also requires a substantial investment of time and energy when selecting and extracting classification components. A preliminary search is required for several image diagnostics tasks to pinpoint problems, scale measurements, and monitor changes. One of the most popular techniques for achieving cutting-edge accuracy is deep learning. It unlocked doors that had previously been locked in the field of analysis of medical images. Deep learning is used in a variety of healthcare settings, including personalized treatment recommendations as well as cancer screening and disease surveillance. Physicians have access to a wealth of data thanks to genetic sequencing, pathological imaging, and radiological imaging (MRI scans, CT scans, and X-rays). But we do not have the right equipment to turn all this information into knowledge that's useful. Next, we'll talk about advance DL applications for processing medical images. The industry of medical visualization will likely be significantly impacted by deep learning in the future, although this list is not all-inclusive [1].

II. Literature Review

Velastegui and Pedersen (2021) Classify images of colon tissue using a CNN powered by AlexNet to detect tumors. This task utilizes multiple color spaces, including XYZ, RGB, HSV, YCbCr, and CIELAB. Our analysis of the data reveals the color spaces that provide the most accurate classification of medical images[2].

Soric, Pongrac and Inza, (2020) stated that using a supervised dataset, a CNN ML model was developed in this paper. The model had to correctly classify images from the used dataset that showed both pneumonia and non-pneumonia. The accuracy, precision, and recall rates demonstrated by the model are encouraging: 90.38 percent, 87.8 percent, and 98.21 percent, respectively[3].

According to Nkwentsha, Hounkanrin and Nicolls, (2020) in this study, we use the InceptionV3 model and anatomical and biological data to automatically classify X-ray images from Image CLEF 2009 dataset. We can convert grayscale images into 3-channel images for InceptionV3 by preparing and preprocessing X-ray images with 2 various padding techniques, 2 image improvement techniques, and layering. The optimal results were achieved by using constant padding with no changes, with an accuracy of 68.67% and a classification loss of 1.442 percent. The best results were found with constant padding with enhancement, which had a classification accuracy of 71.34 percent with a loss of only 1.608 percent[4].

In this paper, Banik, Saha and Kim, (2019) White blood cell (WBC) images can be classified using a CNN-fusion model. A fully connected network employs three max-pooling layers, five convolutional layers, and one hidden layer. The feature maps from two convolutional layers in a fully connected neural network are combined using max-pooling and then fed to a layer. The computational time and accuracy of our model are compared to those of a CNN-RNN hybrid model. Moreover, we show that our model is superior to that of CNN-RNN[5].

In the presented paper, Sadasivan and Seelamantula, (2019) used a technique based on deep learning, it is possible to automate the detection of anomalies in WCE images. Before being fed into a CNN, WCE images are segmented into patches. Patches are classified as malignant or benign using a trained deep neural network. The WCE image output highlights the abnormal patches. Using nine abnormalities from publicly available test data, we determined that the AUROC (area under the receiver-operating-characteristics curve) was approximately 98.65 percent[6].

The purpose of Wang et al., (2022) is to develop a CNN model for feature extraction and a classification system specific to medical images. Based on the mechanism of attention, we developed a synergic network that captures intra-class distinctions and inter-class similarities. The entire pipeline consists of three components: a convolutional block, attention-based, an input layer, and a synergic network. These elements were combined to train an extensive medical image classification model. Carefully designed attention models can allow networks to adaptively focus on medical image regions of interest without requiring expensive annotations like bounding boxes or part information labelling. Using the MRN et dataset and the ADNI dataset(part) and comparing our model to the current SOTA algorithm show that our algorithm achieves SOTA performance[7].

In this article Yang and Guan, (2021) stated that an enhanced CNN-based classification model is proposed. The pooling layer that is normally included in CNN is substituted with PCA. By calculating and choosing eigenvectors with high eigenvalues, the pooling layer's main goal is to rapidly lower the original data's number of dimensions. Then, using a three-layer convolutional layer, the image's primary features are extracted. To achieve the target accuracy, you may run classification tests on several datasets such as MNIST and Kaggle leaves. It demonstrates the superior recognition effect of the enhanced CNN model and its suitability for the classification processing of a wide range of data types. Additionally, it is possible to classify noisy images more accurately[8].

Jadah et al., (2022) presents a deep neural network (DNN) classification model for breast cancer, namely the CNN model AlexNet. BreakHis data set of histopathological images will be used to train the model to detect breast cancer. To improve a model's capacity to identify and classify the input image and determine whether it represents a benign or malignant tumor, changes are made to the parameters and data. Classification accuracy has been observed to increase dramatically with increasing training frequency and balancing training data, reaching 96% currently. Our goal is to achieve a higher level of precision than would be possible by repeatedly fine-tuning parameters and weights using more accurate methods[9].

III. Methodology

A relatively new machine learning method called deep neural networks has shown promise in several categorization tasks. Remarkably, a CNN outperforms other methods for many image classification tasks. On the other hand, medical image databases are scarce because labeling them requires a high level of expertise. CNN-based deep neural systems are frequently employed in the medical classification task [10]. In the 1980s, CNNs were first developed and implemented. At the time, CNN was only able to recognize handwritten numbers. Currently, it reads zip codes, PINs, and other postal-related data. Any deep learning model must be trained using a set of data and a set of computing resources, which is the most crucial concept to comprehend. CNNs were limited to the postal industry at the time due to this significant disadvantage, which prevented them from entering the machine learning field [11].

Convolutional neural networks, a subset of deep neural networks (DNNs), have excelled in image categorization applications since 2012. CNN's performance in the classification of medical images is comparable to that of human experts [12].

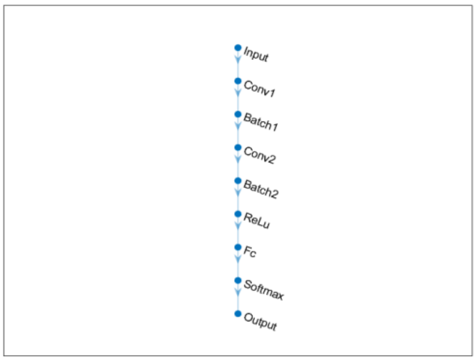

Fig. 2. Visual Representation of a Neural Network

The names of CNN architectures are derived from the conv surfaces that compose them. In their input images, conv surfaces are tasked with finding specific local characteristics. Each conv surface node is connected to a few spatially coupled neurons in the input image channels, allowing it to recognize nearby structures. Nodes of the conv surface are assigned link weights to facilitate the search for comparable local characteristics across all input channels. The term "kernel" may refer to either a set of common weights or a convolution kernel [13]. The n local characteristics across all input images whose intensities are reflected in the output n feature maps are identified using a conv surface with n kernels. A pooling surface follows each sequence of convolutional surfaces, mimicking the basic and complex cells seen in the primary visual cortex; this reduces computing complexity and allows for the creation of a hierarchical collection of picture properties[14]. By erasing the exact locations of the largest response within overlapping or non-overlapping local regions, the max-pooling surface simplifies characteristic charts. Therefore, maximum pooling may further increase translation invariance [15]. CNNs consist of a large number of pairs of pooling surfaces and To generate the desired outputs, Conv is followed by a large number of fully connected surfaces and then softmax or regression surfaces [16]. Modern CNN systems achieve computational efficiency by removing the pooling layer and replacing it with a conv surface whose stride is greater than 1. Throughout the training, the input of each secreted surface changed dynamically due to a parameter change in a previous surface at each training epoch [17]. The network's search for the ideal hyperparameter becomes more difficult if the changes are significant, and finding the ideal value might be computationally expensive. Two researchers describe batch normalization as a technique that can be used to address this issue (Engle, 1992). Batch normalization enables the use of a faster learning rate, which accelerates the optimal value's attainment. It makes training of deeper network architectures quicker and more efficient. The essence of data normalization is calculating the mean and variance of sub-batch data points and normalizing them so that they have a mean of 0 and a variance of 1. The CNN architecture supports a wide variety of layer types. The output layer, the last layer of the CNN architecture, determines the final classification (Doi, 2007). Loss functions are used in the output layer of the CNN model to calculate the expected error that the training samples will produce.

Utilizing multiple CNN layers (Alswayed A.S., 2020)

- **Input Layer:** The first convolutional neural network (CNN) layer is depicted in this image; this is the "data layer," where input images are scaled and transformed into output features.
- **Convolution Layer:** These layers, which are now referred to as "Convolution layers," filter images, identify image functions, and measure real-world metrics.
- **Pooling Layer:** In addition, this layer reduces the file size of large images while maintaining the most essential details. It keeps each window functioning as desired and matches each window feature appropriately[18].
- **ReLU (Rectified Linear Unit Layer):** During pooling, the Rectified Linear Unit Layer(ReLU) layer replaces all negative layer numbers with 0. As a result, CNN is stabilized, and observed values stay relatively close to 0.
- **Fully Connected Layer (FCL):** High-level filter images are identified in the final layer. [19].

IV. SIMULATION

The model will be implemented with separate layers due to its advantageous presentation in image classification frameworks. Both Epochs will be used to construct a model with around 18,000 photos; each Epoch has 140 iterations in addition to 30 support iterations, and the learning cost is 0.0001. With this configuration, the framework can be created in 2 minutes and 46 seconds. There is a maximum usage limit of two Epochs. To create complex, data-driven, hierarchical image layers, CNNs may need large amounts of training data. If there is enough training data, CNNs may help learn data-driven, layered, hierarchical imagery features. Medical image categorisation using CNNs often involves one of three methods: either training the CNN from scratch, utilising pre-prepared CNN features that are commercially accessible, or doing unsupervised pre-training with supervised fine-tuning. Although it is possible to transfer domains between medical image datasets, a more successful method is transfer learning, which involves fine-tuning pre-prepared CNNframeworks by a natural image dataset tomedical image issues.

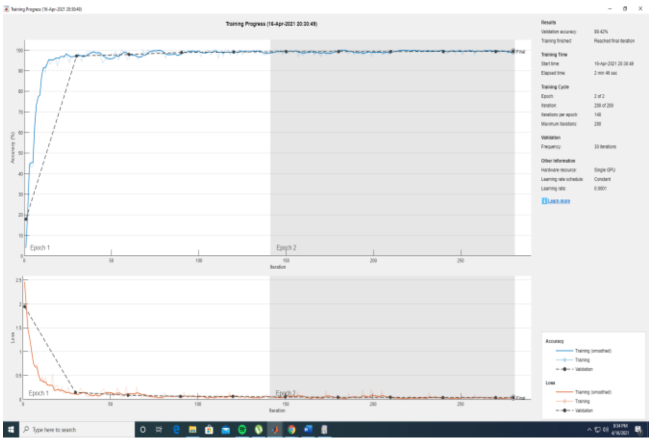

Fig. 3. The Final Result Graph

V. RESULT OBTAINED

Each category defined during the preparation phase enabled a neural network to extract individual properties. These characteristics may include, among others, the size, shape, number of curves, and corners of a physical fraction. These characteristics allowed the neural network to achieve an approximate 99.42% degree of accuracy by significantly reducing its failure function with each iteration. The task at hand, the available data and computing power, as well as the model's intended use, all influence the architecture choice. The initial sample size for legal annotation is insufficient. Radiologists and doctors invest a lot of resources on medical image labelling, however current picture archives do not have enough coverage. The second step is to standardize the datasets and evaluation metrics.

- During training, we achieve 99.42% accuracy, the highest percentage we've ever achieved.
- The x-axis of Figure 3 indicates the number of iterations, whereas the y-axis indicates the percentage of precision.

CNNs have the potential to generate hierarchical picture characteristics with layers if provided with sufficient training data. CNNs may aid in the acquisition of data-driven, hierarchical, layered visual qualities that are highly descriptive if trained with sufficient data. There are three primary ways to effectively deploy CNNs for medical image classification: use of commercially available pre-prepared CNN characteristics, preparation from scratch, and infeasible pre-preparation with supervised fine-tuning. Transfer learning is a powerful method for adapting pre-built CNN models from datasets of natural images to medical image issues. It's also possible to transfer domains across medical image datasets.

VI. CONCLUSION

Despite their popularity, CNN-based algorithms can only process two-dimensional images, whereas the vast majority of medical image data is stored in three-dimensional volumes in clinical practice. Dropout and batch normalisation are two state-of-the-art neural network optimisation methods that they use in their study. A new linear nexus design is also developed by the authors using a nonlinear activation and inception module. These systems' primary objectives are to save computational time, reduce parameter space, and address 3D modalities. To determine the optimal network architecture for specific medical image analysis applications, additional research is required. In this study, a simple CNN was applied to a medical dataset. Importantly, our calculations show that it is possible, albeit unlikely given the enormous disparity between the source and target databases, to transfer information by natural images to medicinal images. In addition, it was found that previously unrecognized effects on DCNN, including CNN design and transfer learning, varied depending on the application and the amount of fine-tuning needed. The results show that all subfields of rehabilitative image analysis, including lesion segmentation, classification, and detection, are increasingly using CNN-based methods. Despite their limitations, data augmentation and transfer learning may be used efficiently to make up for inadequate training data.

REFERENCES

[1] G. J. Tearney *et al.*, "In vivo endoscopic optical biopsy with

optical coherence tomography," *Science (80-.).*, 1997, doi: 10.1126/science.276.5321.2037.

[2] R. Velastegui and M. Pedersen, "The Impact of Using Different Color Spaces in Histological Image Classification using Convolutional Neural Networks," in *Proceedings - European Workshop on Visual Information Processing, EUVIP*, 2021. doi: 10.1109/EUVIP50544.2021.9484035.

[3] M. Soric, D. Pongrac, and I. Inza, "Using convolutional neural network for chest x-ray image classification," in *2020 43rd International Convention on Information, Communication and Electronic Technology, MIPRO 2020 - Proceedings*, 2020. doi: 10.23919/MIPRO48935.2020.9245376.

[4] X. Nkwentsha, A. Hounkanrin, and F. Nicolls, "Automatic classification of medical X-ray images with convolutional neural networks," in *2020 International SAUPEC/RobMech/PRASA Conference, SAUPEC/RobMech/PRASA 2020*, 2020. doi: 10.1109/SAUPEC/RobMech/PRASA48453.2020.9041052.

[5] P. P. Banik, R. Saha, and K. D. Kim, "Fused Convolutional Neural Network for White Blood Cell Image Classification," in *1st International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2019*, 2019. doi: 10.1109/ICAIIC.2019.8669049.

[6] V. S. Sadasivan and C. S. Seelamantula, "High accuracy patch-level classification of wireless capsule endoscopy images using a convolutional neural network," in *Proceedings - International Symposium on Biomedical Imaging*, 2019. doi: 10.1109/ISBI.2019.8759324.

[7] S. Wang *et al.*, "A Synergic Neural Network For Medical Image Classification Based On Attention Mechanism," in *Proceedings - 2022 Asia Conference on Algorithms, Computing and Machine Learning, CACML 2022*, 2022. doi: 10.1109/CACML55074.2022.00022.

[8] Y. Yang and C. Guan, "A Novel Convolutional Neural Network for Image Classification," in *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer, ICFTIC 2021*, 2021. doi: 10.1109/ICFTIC54370.2021.9647344.

[9] Z. Jadah, A. Alfitouri, H. Chantar, M. Amarif, and A. A. Aeshah, "Breast Cancer Image Classification Using Deep Convolutional Neural Networks," in *Proceedings - 2022 International Conference on Engineering and MIS, ICEMIS 2022*, 2022. doi: 10.1109/ICEMIS56295.2022.9914251.

[10] W. R. Hendee, "The impact of future technology on oncologic diagnosis: Oncologic imaging and diagnosis," *Int. J. Radiat. Oncol. Biol. Phys.*, 1983, doi: 10.1016/0360-3016(83)90353-X.

[11] A. Heidenreich, F. Desgrandschamps, and F. Terrier, "Modern approach of diagnosis and management of acute flank pain: Review of all imaging modalities," *Eur. Urol.*, 2002, doi: 10.1016/S0302-2838(02)00064-7.

[12] R. E. Bunge and C. L. Herman, "Usage of diagnostic imaging procedures: A nationwide hospital study," *Radiology*, 1987, doi: 10.1148/radiology.163.2.3550886.

[13] "https://www.kaggle.com/andrewmvd/medical-mnist".

[14] L. G. B. A. Quekel, A. G. H. Kessels, R. Goei, and J. M. A. Van Engelshoven, "Miss rate of lung cancer on the chest radiograph in clinical practice," *Chest*, 1999, doi: 10.1378/chest.115.3.720.

[15] F. Li, S. Sone, H. Abe, H. MacMahon, S. G. Armato, and K. Doi, "Lung cancers missed at low-dose helical CT screening in a general population: Comparison of clinical, histopathologic, and imaging findings," *Radiology*, 2002, doi: 10.1148/radiol.2253011375.

[16] Q. Li *et al.*, "Computer-aided diagnosis in thoracic CT," *Semin. Ultrasound, CT MRI*, 2005, doi: 10.1053/j.sult.2005.07.001.

[17] K. Suzuki *et al.*, "An MTANN CAD for detection of polyps in false-negative CT colonography cases in a large multicenter clinical trial: preliminary results," in *Medical Imaging 2008: Computer-Aided Diagnosis*, 2008. doi: 10.1117/12.769824.

[18] A. Mansoor *et al.*, "Segmentation and image analysis of abnormal lungs at CT: Current approaches, challenges, and future trends," *Radiographics*, 2015, doi: 10.1148/rg.2015140232.

[19] K. Suzuki, "Computer-aided detection of lung cancer," in *Image-Based Computer-Assisted Radiation Therapy*, 2017. doi: 10.1007/978-981-10-2945-5_2.

