Journal of Autonomous Intelligence (2024) Volume 7 Issue 5
doi: 10.32629/jai.v7i5.1638

ORIGINAL RESEARCH ARTICLE

Mobile volume rendering and disease detection using deep learning algorithms

A. V. Krishnarao Padyala[*], Ajay Kaushik

Department of CSE, SRM University, Sonepat 131029, Haryana, India

*** Corresponding author:** A. V. Krishnarao Padyala, guntur.krishna@gmail.com

ABSTRACT

This paper introduces a system designed to convert 2D slices from Magnetic resonance imaging(MRI) and Computed Tomography(CT) scans into 3D images, facilitating mobile device-based diagnosis by medical professionals. Utilizing machine learning techniques tailored to specific image categories, the system processes Digital Imaging and Communications in Medicine(DICOM) images for disease detection. AWS cloud infrastructure, including S3 bucket, Relational Database Service(RDS), and DynamoDB, manages DICOM storage. The system delivers a final processed image displaying predicted diseases directly to the mobile screen. This innovative approach enhances medical imaging accessibility and diagnostic accuracy, offering a streamlined solution for healthcare professionals.

Keywords: COVID-19 UNet++; deep diagnosis; DICOM conversion; heart CNN; kidney CNN; stroke OzNet

ARTICLE INFO

Received: 10 March 2024
Accepted: 9 April 2024
Available online: 23 May 2024

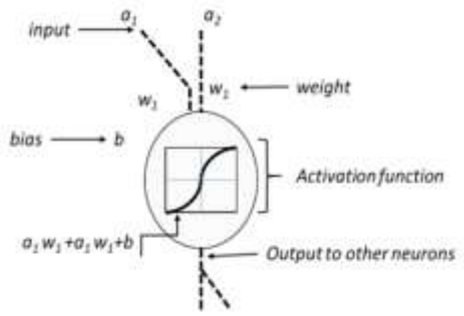
1. Introduction

In this paper, we have presented a system that can render 2D DICOM images collected as 3D images and also the disease predicted. In this system volume rendering and deep learning algorithms are applied. Volume rendering[1] is a set of techniques used to display a 2D projection of a 3D discretely sampled data set, typically a 3D scalar field. The ML algorithms are categorized as supervised, unsupervised, and semi-supervised. The most popular of them used in disease detection are decision trees (DT), support vector machines (SVM), and K Nearest Neighbor (KNN).

Most of the disease prediction techniques here used deep learning techniques. A summary of the Artificial Neural Network (ANN)[2] is as follows. The term “deep” in “deep learning”[3] originates from the multiple hidden layers characteristic of Artificial Neural Networks (ANNs). Inspired by the functioning of the brain, ANN algorithms are structured with layers comprising inputs, hidden nodes, and outputs, interconnected via links akin to synapses in biological neurons. Each neuron in the network corresponds to components of a biological neuron: dendrites for inputs, axons for outputs, and an activation function akin to the nucleus, determining the neuron’s response. Synapses in ANN represent the weights of connections between neurons.

Despite the effectiveness of ANNs, they suffer from limitations such as susceptibility to translation shifts, leading to decreased classification performance. In addressing these issues, convolutional neural networks (CNNs) emerged as an advanced solution. CNNs ensure robustness to shifts and translations by leveraging specialized architectures.

The CNN architecture comprises layers for convolution, pooling, and fully connected operations, enabling effective feature extraction and classification. This architecture illustrated in **Figure 1**, emphasizes a feed-forward approach, refining the model's ability to handle complex data patterns and improve classification accuracy.



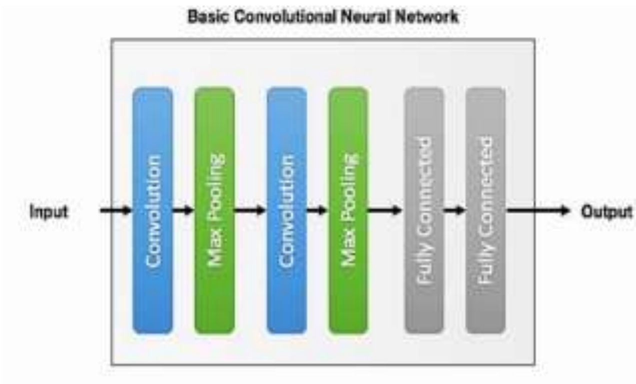


Figure 1. Input and output of the neuron, layers of convolutional neural network.

Convolution layer: In this initial layer, the input data undergo convolution with specific kernels (weights). The result is a feature map, representing the outcome of this convolution process. The degree of convolution applied to the input is regulated by the stride parameter. The convolution operation serves as the primary mechanism for feature extraction, allowing the network to learn patterns from the input signal. Subsequent layers utilize these extracted features for classification tasks.

Pooling layer: The pooling layer performs spatial down-sampling of the input while retaining essential information. Various pooling operations such as average, max, or sum are employed. In max pooling, for instance, the input is reduced by selecting only the maximum values within predefined regions determined by the stride. For instance, in a stride containing values 14 and 22, the maximum value (22) is retained, while the other (14) is discarded.

Fully connected layer: This layer represents the final stage of the architecture, where each neuron from the preceding layer connects fully to every neuron in the current layer. The number of neurons in this layer typically corresponds to the number of classes to be classified, graphically represented in **Figure 2**, each neuron in this layer is associated with specific weights. The final output target is estimated by aggregating the weighted sums obtained from all previous layers.

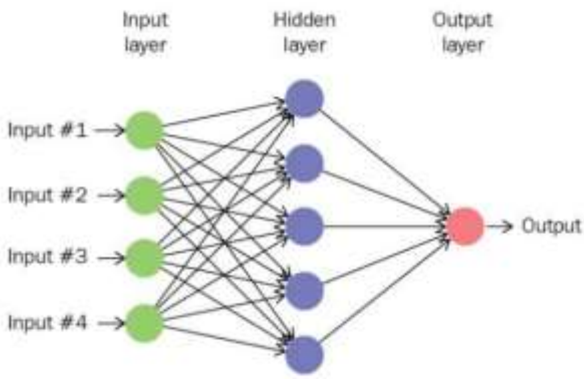


Figure2. Fully connected layer of convolutional neural network.

2. Existing system

A Picture Archiving and Communication System (PACS)[4,5] is a sophisticated medical imaging technology widely employed in healthcare settings to securely store, manage, and transmit electronic images along with associated reports.

PACS serves multiple purposes including the storage, retrieval, presentation, and sharing of images generated by various medical imaging devices like X-ray machines, CT scans, MRI scans, and ultrasound machines. The storage infrastructure for PACS can be either online (utilizing cloud storage) or offline (on-premises).

Primarily utilized by radiologists due to the abundance of X-ray images in radiology, PACS has also found application in other medical departments such as nuclear medicine imaging, cardiology, pathology, oncology, and dermatology. Its versatility facilitates efficient access to images for diagnostic and treatment purposes, contributing significantly to healthcare workflows.

While Picture Archiving and Communication Systems (PACS) have revolutionized medical imaging management, they are not without their drawbacks. Here are ten potential limitations:

Cost: Implementing a PACS system can be financially burdensome for healthcare organizations. The initial investment includes purchasing hardware, software, and training staff, with ongoing costs for maintenance, upgrades, and licensing fees. This expense can be prohibitive for smaller healthcare facilities.

Complexity: PACS systems are complex and require significant expertise to install, configure, and maintain. Healthcare organizations may struggle to find qualified personnel capable of managing and troubleshooting these systems effectively.

Integration challenges: Integrating PACS with existing hospital information systems (HIS), electronic health records (EHR), and radiology information systems (RIS) can be challenging. Incompatibilities between systems may lead to data silos, inefficiencies, and errors in patient care.

Interoperability issues: Despite efforts to standardize medical imaging formats like DICOM (Digital Imaging and Communications in Medicine), interoperability issues still exist between PACS systems from different vendors. This can hinder data sharing and collaboration among healthcare providers.

Data security risks: Storing sensitive patient data in digital format exposes it to cybersecurity risks such as unauthorized access, data breaches, and ransomware attacks. Healthcare organizations must invest in robust security measures to protect patient privacy and comply with regulations like HIPAA.

Workflow disruptions: Introducing a PACS system can disrupt existing workflows and routines within healthcare organizations. Staff may require time to adapt to new processes, leading to temporary decreases in productivity and patient care quality.

Downtime impact: PACS downtime, whether due to technical issues, maintenance, or upgrades, can have significant consequences for patient care. Healthcare providers may be unable to access critical imaging data, leading to delays in diagnosis and treatment.

Scalability challenges: As the volume of medical imaging data continues to grow, PACS systems must scale to accommodate increased storage and processing requirements. Scalability challenges may arise if the system architecture is not designed to handle large volumes of data efficiently.

Limited accessibility: Despite advances in mobile technology, accessing PACS data remotely from smartphones or tablets may be limited by factors such as network bandwidth, device compatibility, and security concerns. This can hinder real-time collaboration and decision-making among healthcare providers.

User training needs: Healthcare professionals require extensive training to effectively use PACS systems for image interpretation, analysis, and reporting. Inadequate training can lead to errors in diagnosis and treatment planning, compromising patient safety and quality of care.

3. Materials and methods

The system developed renders a 3D image on the mobile screen with the details about the disease predicted. This system mainly carries out two tasks, transforming the 2D image slices into 3D and image processing using machine learning/ deep learning techniques for early detection of disease.

This system majorly accepts the DICOM[6–8] images captured using CT and MRI scanners. The 2D image slices were processed and converted to 3D images. This system makes use of a branch of artificial intelligence called machine learning for image processing. While giving the input images, the user needs to mention the category of the images. The categories here include cardiac, chest, kidney, and brain-related. Based on the category, a suitable machine learning/deep learning technique is applied, and the identified features related to the disease are marked.

The **Figure 3** shows the architecture of the system for medical image processing. This includes inputs from CT/MRI machines, an image volume rendering unit, an image processing unit, temporary storage, and an AWS cloud-based storage.

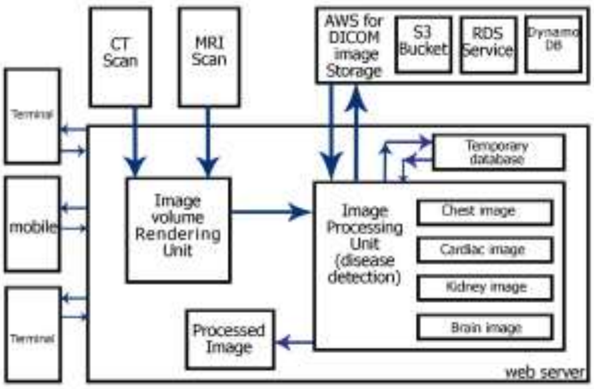


Figure 3. Architecture of mobile volume rendering and disease detection system.

Input: Medical images that were captured from CT/MRI machines stored in the form of DICOM, these were made as input to the system.

3.1. Image volume rendering unit

The image volume rendering[9] unit is responsible for converting 2D image slices to 3D images. This unit makes use of the method defined by Mamdouh et al.[10]. This method reads first all of the 2D DICOM images captured from different views. The views may include front, side, and top.

The initial process involves file conversion followed by multiple internal stages. This includes importing DICOM files, preprocessing to eliminate noise, and delineating the area for modification by drawing a series of points and lines. Executing adjustments involves identifying internal points necessary for display and manipulation. Subsequent steps include trimming, smoothing, and sculpting until achieving the desired outcome. Eventually, the focus shifts to counting the number of polygons and ensuring their smooth and accurate presentation to highlight intricate details. The DICOM image captured, its different views and the noise present are shown in **Figure 4**.

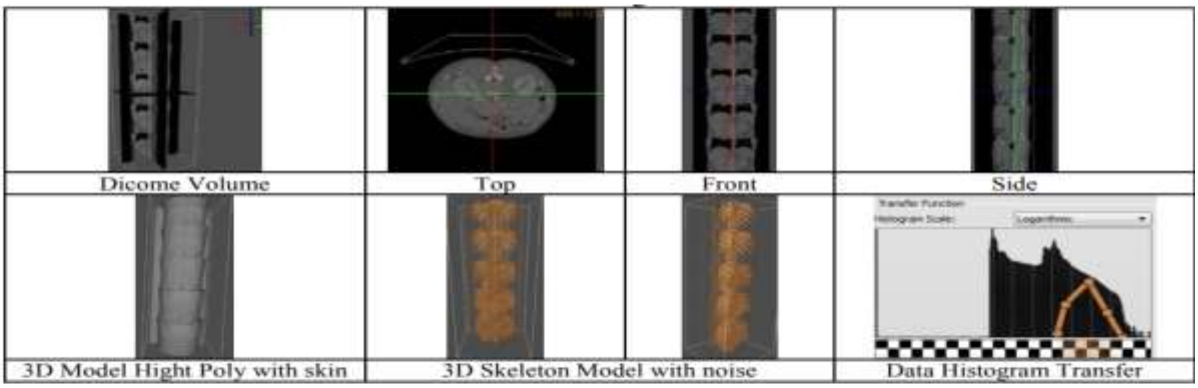
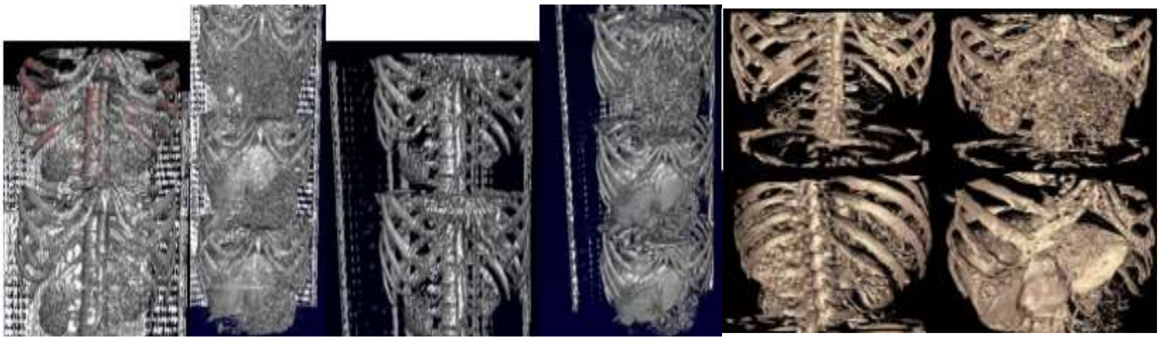


Figure 4. Captured 2D DICOM images, images with noise[2].

Ensure the final form matches the required specifications before proceeding to export. If the final form aligns with expectations, it can be exported as an OBJ file through two distinct processes: one before and another after editing segments, re-sculpting, and trimming, as illustrated in **Figure 5**. The resultant 3D model is compatible with usage in a 3D simulation engine. In instances where there is a deviation from the expected outcome, return to the designated stage to correct the desired area and revert to the previous phase.



(a) **(b)**

Figure 5. (a) front and side views (before edit segments, re-sculpting, and trimming); **(b)** front and side views (after edit segments, re-sculpting, and trimming).

3.2. Image processing unit

This is responsible for processing the medical images. The images captured may belong to different internal organs of the human. This module prompts for the type of organ image uploading and the respective machine learning/deep learning technique employed for disease detection.

To detect disease machine learning/deep learning techniques were used. These methods were summarized in machine learning-based disease diagnosis[11]. The following shows the machine learning or deep learning algorithms for coding methods used for a variety of disease detection.

3.2.1. Heart disease detection

A method developed by using a deep neural network[12] was used in this for heart disease detection[13,14]. Advancements in medical imaging have revolutionized the diagnosis and management of various diseases, particularly in the realm of cardiovascular health. With the evolution of imaging technologies, multiple options have emerged to address diagnostic challenges, including non-invasive investigations for cardiovascular disease (CVD). These modalities encompass echocardiography, computed tomography (CT), and cardiovascular magnetic resonance (CMR), each offering distinct advantages and drawbacks.

Of these techniques, CMR stands out as exceptionally efficient. It is renowned for producing high-quality images with excellent contrast for soft tissues, all without exposing patients to ionizing radiation. As a result, CMR has solidified its position as the non-invasive gold standard for assessing cardiac chamber volume and mass across a broad spectrum of cardiovascular diseases.

The left ventricle (LV), the largest chamber in the heart, plays a crucial role in maintaining cardiac ejection function. Diagnosis of cardiac diseases relies heavily on the critical index of ejection fraction (EF). This experimentation primarily focuses on the EF of the LV, as it provides accurate estimations of LV volumes, comprising end-diastolic volumes (EDV) and end-systolic volumes (ESV). Accurate estimation of LV volume is essential for detecting cardiovascular diseases (CVD). LV segmentation technology estimates volumes from cardiac magnetic resonance (CMR) images. CMR images are considered the gold standard modality for diagnosing cardiac diseases due to their high resolution and ability to differentiate soft tissues, surpassing invasive exercise standards. Leveraging these advantages, CMR images are extensively utilized in various cardiac medical image processing tasks, particularly in automated LV segmentation.

The method followed for heart disease detection is shown in **Figure 6** and is explained as follows: The initial step involves pre-processing of the data, a customary practice in research endeavors. This step not only enhances prediction accuracy but also bolsters the robustness of the framework for handling large-scale datasets. The system’s robustness is influenced by various factors such as scanning parameters, demographics (age, gender), and the level of health conditions. In this phase, a precise and robust technique for detecting Regions of Interest (ROI) was used. Additionally, applied normalization techniques to manage pixel space variance and intensity levels in the extensive CMR datasets.

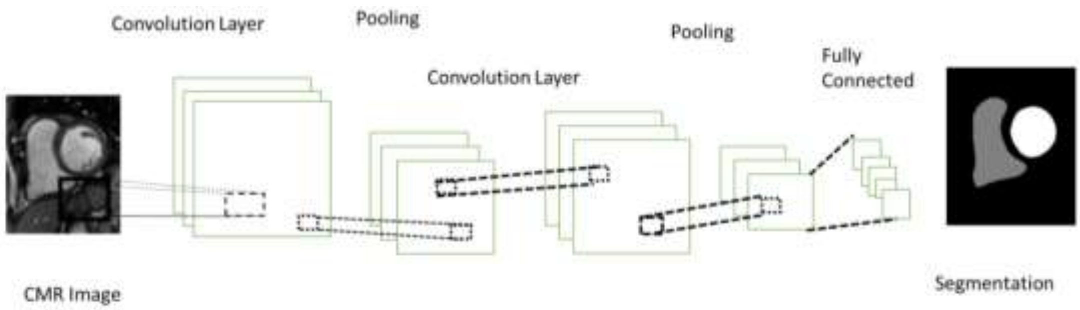


Figure 6. Method followed for heart disease detection.

Following pre-processing designed a deep-learning network. During this phase, the architectures of established convolutional networks improved upon the VGG model by incorporating effective technologies such as batch normalization, “Adam” training, and dropout. Understanding the relationship between the final prediction accuracy and the initial convolutional layer’s depth, optimized the CNN design accordingly. Ultimately, this method outperformed segmentation-based techniques in terms of robustness against variances when predicting volumes.

Similar to the typical flow of deep learning, this framework consists of three primary components: data pre-processing, model training using CNN, and volume prediction leading to EF computation. In addition to these components, this introduced several novel settings, notably in data pre-processing techniques, focusing on normalization and ROI detection, as well as in the design of deep CNN networks.

3.2.2. Kidney disease detection

A method developed by using a convolution neural network was adopted for kidney disease detection[15–17]. The kidneys play a crucial role in filtering waste products and toxins from the bloodstream. Abnormal cell growth in the form of tumors, or cancers, can affect individuals differently, presenting various symptoms. Early detection of kidney tumors (KT) is vital to mitigate the risk of disease progression and preserve patients’ lives.

Despite approximately one-third of KT cases being diagnosed after metastasis, many remain asymptomatic and are incidentally discovered during unrelated medical examinations. Radiography may reveal kidney masses or cysts, potentially causing abdominal pain, though symptoms are often unrelated to kidney function. However, subtle signs such as low hemoglobin, weakness, vomiting, stomach pain, blood in urine, or elevated blood sugar levels may indicate KT or associated complications like anemia, affecting roughly 30% of patients.

Unfortunately, many internal kidney tumors and solid masses are malignant, underscoring the importance of timely detection for appropriate treatment selection. Computed tomography (CT) scans of the abdomen and pelvis are pivotal diagnostic tools for assessing kidney tumors, providing detailed imaging to confirm their presence and guide treatment decisions. **Figure 7** depicts a typical KT case, illustrating a renal mass lesion in the left kidney, approximately 4 cm in size, alongside a 3D volume rendering highlighting the tumor in blue against the pink kidney background.

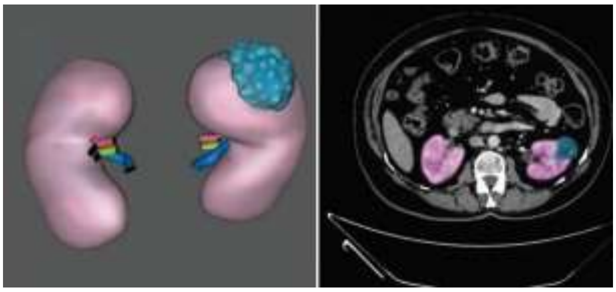


Figure 7. Renal CT from the dataset.

Given the life-threatening nature of tumors, accurate diagnosis through advanced imaging techniques is essential for prompt intervention and improved patient outcomes.

The CNN-6 detection model proposed for tumor detection comprises six deep CNN layers with fully connected ANN. The model architecture begins with a batch normalization layer, which standardizes inputs for each minibatch, enhancing classification accuracy and mitigating overfitting. Subsequently, the convolution 2D input layer applies a convolution kernel to extract features from the input image. This kernel can perform operations such as blurring, sharpening, or edge detection.

The Conv2D layer utilizes parameters to define its operation, such as filter count and kernel size. Max-pooling layers follow, reducing spatial dimensions in the output features. Dropout layers are then incorporated to prevent overfitting by randomly deactivating input units during training. Following this, a flattening layer transforms the input into a one-dimensional array. Lastly, the dense layer, serving as the hidden or output layer, employs specified neuron count and activation functions. In the output layer, the number of neurons corresponds to the number of training classes.

Figure 8 illustrates the network structure of this modified CNN architecture, showcasing the sequential flow of layers for tumor detection.

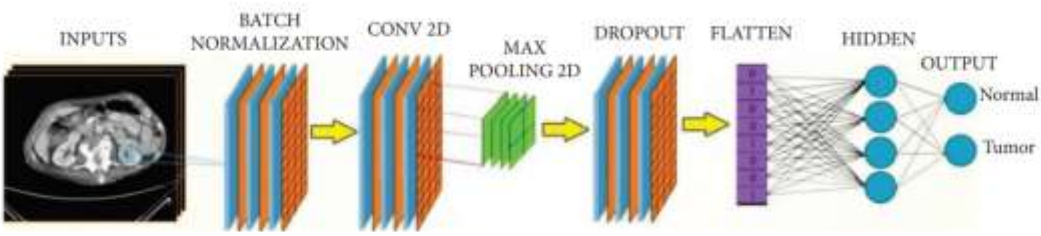


Figure 8. CNN architecture for detection of kidney tumor.

3.2.3. Brain disease detection

A method developed based on convolution neural networks (CNN) architecture called OzNet and various machine learning algorithms used for brain stroke early detection[18–20]. **Figure 9** shows a variety of brain stroke instances.

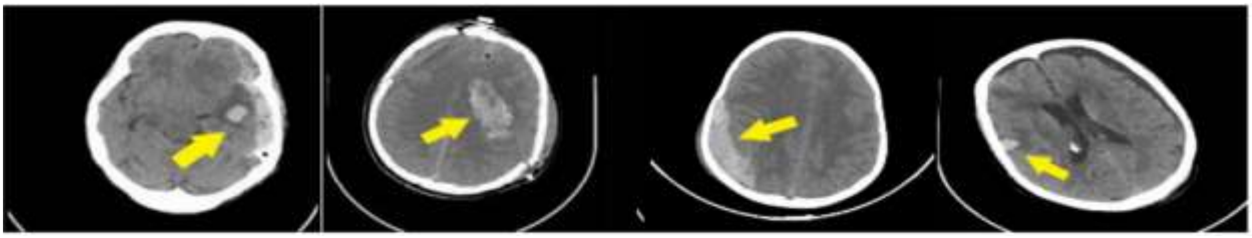


Figure 9. Stroke instances—dataset.

Stroke stands as the leading neurological cause of death and disability worldwide. This cerebrovascular event occurs due to either ischemia or hemorrhage in the brain arteries, resulting in diverse motor and cognitive impairments that compromise functionality. Globally, around 16 million individuals suffer from strokes. Early detection plays a pivotal role in managing strokes effectively.

Brain imaging techniques, notably CT and Magnetic Resonance Imaging (MRI), are crucial in diagnosing strokes. While MRI is typically preferred for stroke diagnosis due to its detailed imaging capabilities, in emergencies, CT scans are favored due to time constraints. Hence, rapid assessment becomes paramount for patient care.

Researchers are focusing on leveraging artificial intelligence, particularly deep learning algorithms, to facilitate the swift recognition of stroke from MRI or CT scans. This technology aims to assist doctors in efficiently diagnosing strokes, thereby improving patient outcomes. The OzNet convolutional neural network (CNN) architecture was developed to effectively classify brain stroke CT images. To enhance its performance, we integrated machine learning methods such as decision trees (DT), k-nearest neighbors (kNN), linear discriminant analysis (LDA), naive bayes (NB), and support vector machines (SVM). Initially, OzNet was employed for binary classification of the CT dataset into stroke and normal cases, yielding acceptable results.

Subsequently, utilized OzNet for deep feature extraction from the images, resulting in 4096 features extracted from its fully connected layer. To reduce feature dimensionality, applied the minimum Redundancy Maximum Relevance (mRMR) method, selecting 250 significant features.

These important features were then classified using various machine learning algorithms. Hemorrhagic stroke CT images, depicting stroke lesions highlighted by arrows, were utilized as the “stroke” class in this study. The primary objective of this research is to determine the optimal structure for detecting strokes from brain CT images. It hypothesizes that OzNet outperforms previous methodologies in achieving this goal.

In the realm of healthcare, swift and accurate results hold paramount importance, driving the widespread adoption of deep learning algorithms by researchers. In this study, introduced a novel deep learning approach named OzNet tailored specifically for 2D biomedical images. Despite its apparent resemblance to a conventional CNN, OzNet incorporates specialized parameters, filter sizes, numbers, padding, stride configurations, and layers to ensure robust performance on biomedical imagery.

OzNet is structured as a novel CNN architecture comprising 34 layers, organized into seven blocks. Each block includes a convolutional layer, a maximum pooling layer, a Rectified Linear Unit (ReLU) activation function, and a batch normalization layer. Following these blocks, two fully connected layers, a dropout layer, a SoftMax layer, and a classification layer are sequentially linked. The ReLU activation function is chosen for its faster computation compared to alternative activation functions.

When OzNet serves as a classifier, it employs a cross-entropy approach. Additionally, due to its multiple convolutional layers, OzNet is also utilized for feature extraction from images, demonstrating effectiveness in this regard.

Conducted comparative analyses between OzNet and established architectures such as GoogleNet, Inceptionv3, and MobileNetv2 for stroke detection from brain CT images. Applied 10-fold cross-validation for these architectures and implemented data augmentation techniques on the brain stroke CT images dataset. Furthermore, Stochastic Gradient Descent with Momentum (SGDM) is chosen as the optimization method, with a momentum parameter set at 0.95 and a learning rate of 0.0001.

3.2.4. Chest scan-based disease detection

A method developed by using the UNet++ model was used for COVID-19 disease detection[21,22]. **Figure 10(1a–1d)** shows the lesions predominantly resembled ground glass, with thickened blood vessels coursing through, exhibiting gas-bronchial signs within. **Figure 11(2a–2c)** the lesions primarily presented as ground glass changes, with an additional observation of paving stone-like alterations in **Figure 11(2d)**. **Figure 12(3a–3c)** the lesions transitioned to a solid state with a wide spectrum, exhibiting air-bronchial signs within, 4- the lesion, predominantly grid-like in nature with ground glass features, is situated in the lower lobe of both lungs. **Figure 13(5a,5b)** shows the lesions primarily manifest as consolidations. **Figure 13(6a,6b)** shows that the lesions predominantly consist of large ground glass shadows, exhibiting changes akin to white lung, with evident air-bronchial signs.

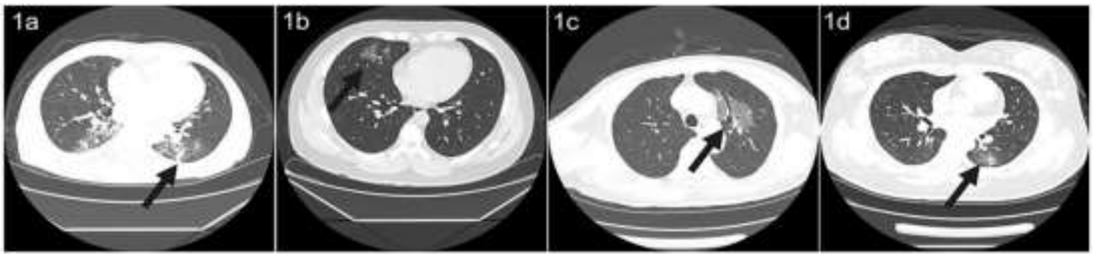


Figure 10. (1a–1d) the lesions were mainly ground-glass-like, with thickened blood vessels walking and including gas-bronchial signs in **1c**.

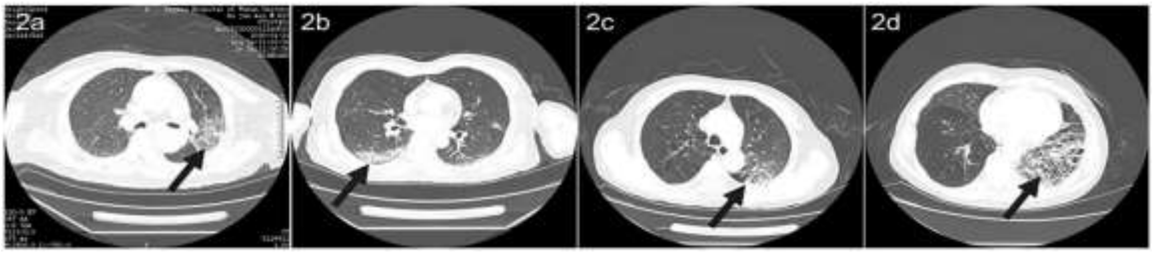


Figure 11. (2a–2d) the lesions were mainly ground glass changes, and paving stone-like changes were observed on **2(d)**.

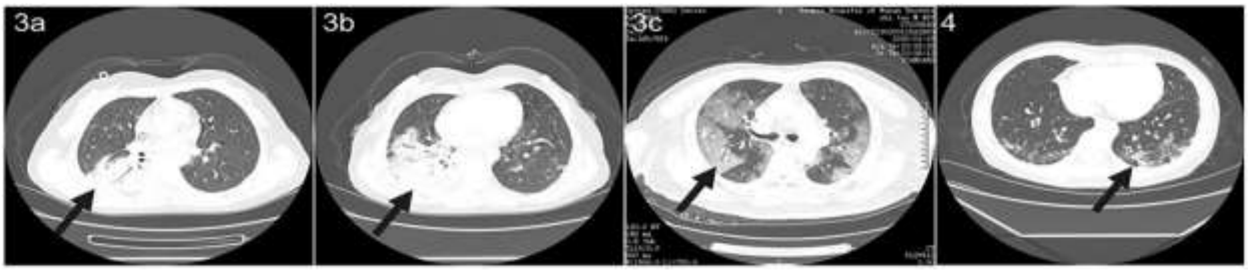


Figure 12. (3a–3c) the lesions become solid with a large range, and air-bronchial signs are seen inside; **(4)** the lesion is located in the lower lobe of both lungs and is mainly grid-like change with ground glass lesion.

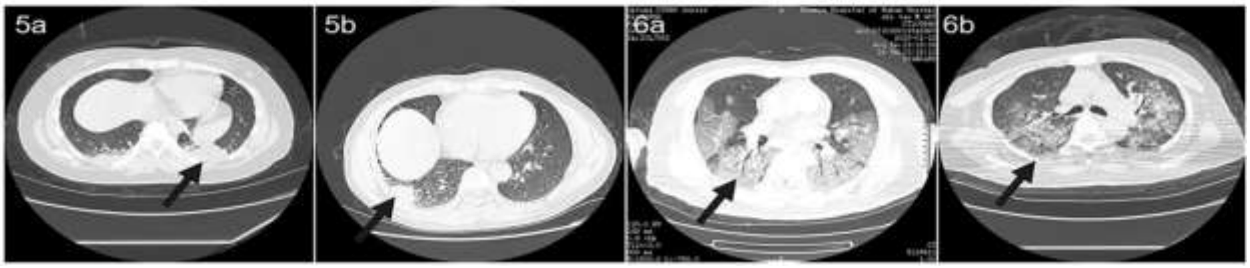


Figure 13. (5a,5b) the lesions are mainly consolidation; **(6a,6b)** the lesions are mainly large ground glass shadows, showing white lung-like changes, with air-bronchial signs.

The model processes raw CT images, generating prediction boxes highlighting suspicious lesions. It then extracts valid areas and filters out unnecessary fields to reduce false positives. To predict on a per-case basis, consecutive image predictions are logically linked. CT images are divided into quadrants, and results are output only when three consecutive images show lesions in the same quadrant. UNet++ architecture for processing the CT images is shown in **Figure 14**.

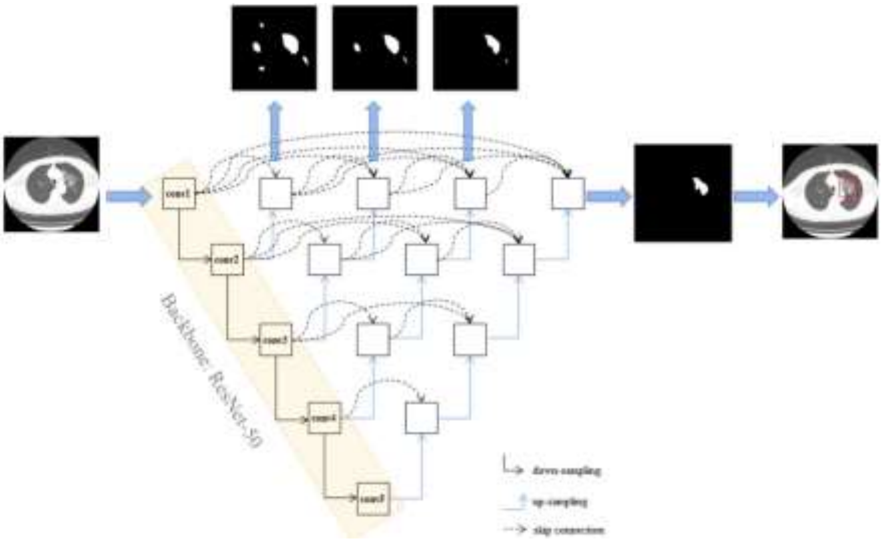


Figure 14. UNet++ architecture for processing CT images[7].

The method here used builds upon the UNet++ architecture, which is known for its efficacy in medical image segmentation. ResNet-50 serves as the backbone of UNet++, pretrained on the ImageNet dataset, with its parameters transferred to UNet++. The UNet++ architecture, comprises encoder and decoder segments connected through nested dense convolutional blocks. It addresses the semantic gap between encoder and decoder feature maps before fusion. The encoder downsamples to extract features, while the decoder upsamples to map features back to the original image, enabling pixel-level classification and segmentation.

Initially, UNet++ was trained to identify valid areas in CT images using 289 randomly selected images for training and 600 for testing. These images were labeled by researchers with rectangles encompassing all valid areas. UNet++ achieved 100% accuracy in extracting valid areas from the testing set.

For detecting suspicious lesions on CT scans, 691 images of COVID-19 pneumonia lesions labeled by radiologists and 300 images from non-COVID-19 pneumonia patients were used. The model was trained in Keras using an image-to-image approach, taking raw CT scan images (512 × 512 resolution) as input and expert-labeled maps as output. A confidence cutoff of 0.50 and a minimum prediction box size of 25 pixels were applied. Training curves for extracting valid areas and detecting suspicious lesions are provided in **Figure 15**.

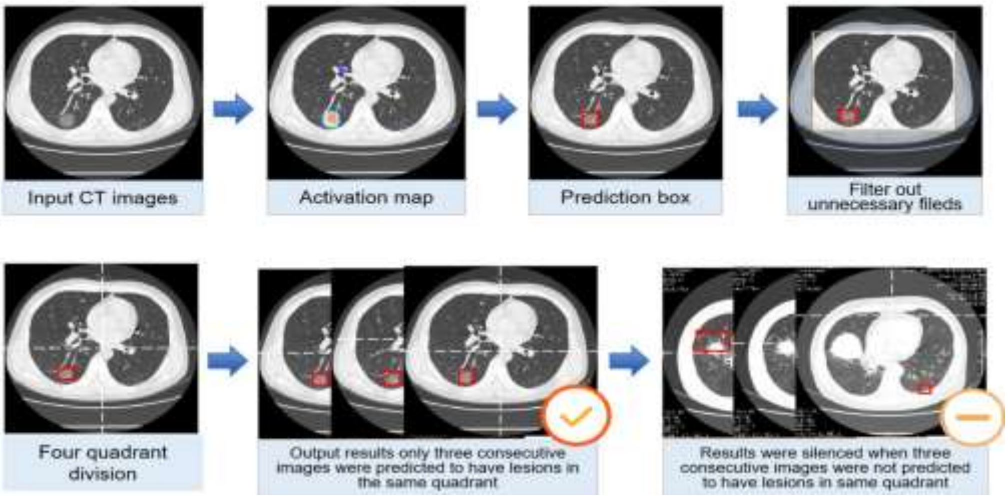


Figure 15. Processing CT images for detecting COVID-19 disease[7].

The prediction process involves inputting raw images into the model, which outputs prediction boxes framing suspicious lesions. Valid areas are extracted, and unnecessary fields are filtered out to mitigate false positives. To predict on a per-case basis, consecutive image predictions are logically linked. Results are output only when three consecutive images predict lesions in the same quadrant.

3.3. AWS for DICOM image storage

This is a set of cloud services offered by Amazon Web Services[23–25]. The services used in this include S3 bucket, RDS, and DynamoDB. S3 is a Simple Storage Service useful for storing images. RDS is a Relational Data Service that can support a variety of relational databases. MySQL database was used in this system as part of the RDS service. DynamoDB is a NoSQL database service offered by AWS. While processing the images these cloud storage services are used by the image processing unit.

2D DICOM images, that were fed as input to the Image Volume Rendering Unit and 3D image generated were stored in S3 buckets of AWS service. The data related to diseases detected by the Image Processing Unit were stored in the AWS RDS and DynamoDB services.

4. Results and discussion

In this study, we investigated the efficacy of employing convolutional neural network (CNN) architectures for disease detection from CT or MRI images. Our findings indicate that the CNN architecture developed specifically for heart and kidney disease detection yielded promising results, demonstrating high accuracy in classifying medical images. By leveraging deep learning techniques, the model effectively analyzed intricate patterns within the images, allowing for the precise identification of signs indicative of heart or kidney disease.

Furthermore, our investigation into the OzNet CNN architecture, tailored for classifying brain stroke CT images, revealed significant success in accurately diagnosing strokes. The incorporation of specialized features within the OzNet architecture contributed to its exceptional performance in detecting stroke-related abnormalities. This highlights the importance of designing CNN architectures with features optimized for specific medical domains to achieve superior diagnostic accuracy.

Additionally, our study explored the application of UNet++ for COVID-19 detection from CT scans, demonstrating its effectiveness in identifying suspicious lesions associated with COVID-19 pneumonia. The adaptability of deep learning architectures, such as UNet++, proved invaluable in addressing the unique challenges posed by infectious disease detection in medical imaging.

Moreover, we investigated the potential of rendering 2D DICOM images as 3D DICOM images in enhancing medical imaging analysis. Our results showed that this innovative approach enabled the capture of additional spatial information, resulting in more comprehensive visualization of anatomical structures and pathological features. The enhanced spatial context provided by 3D images has the potential to improve the accuracy of diagnostic algorithms and facilitate more informed treatment planning.

However, our study underscores the importance of rigorous validation of rendered 3D images to ensure accuracy and reliability. Quality assurance measures are imperative to maintain diagnostic integrity and prevent the introduction of artifacts or distortions that could compromise the diagnostic quality of the images.

In conclusion, our findings highlight the potential of advanced deep-learning architectures and innovative techniques in enhancing diagnostic accuracy and patient care in the field of medical imaging. Continued research and development in this area are essential to further refine these techniques and maximize their clinical utility, ultimately improving outcomes for patients.

5. Conclusion and future work

In conclusion, advanced convolutional neural networks (CNNs) have shown promising results in disease detection from medical imaging modalities such as CT and MRI. For heart and kidney disease detection, CNN architectures have been developed specifically to analyze these organs' images effectively. Additionally, the OzNet CNN architecture has been tailored for classifying brain stroke CT images with high accuracy. Moreover, the UNet++ network has been successfully utilized for detecting COVID-19 from medical images, particularly CT scans.

Future work in this area could focus on several aspects: Enhancement of CNN Architectures: Continuously refining CNN architectures to improve accuracy and efficiency in disease detection tasks. This could involve experimenting with different network architectures, such as integrating attention mechanisms or utilizing transfer learning from pre-trained models.

Multi-modal integration: Investigating methods to incorporate information from multiple imaging modalities, such as combining CT and MRI data, to enhance disease detection and diagnosis accuracy.

Clinical validation and deployment: Conducting comprehensive clinical validation studies to assess the performance of these CNN models in real-world settings and ensuring their seamless integration into clinical workflows.

Automated segmentation and quantification: Developing algorithms for automated segmentation and quantification of disease-related regions from medical images, which could assist radiologists in precise diagnosis and treatment planning.

Interpretability and explainability: Exploring techniques to enhance the interpretability and explainability of CNN models' predictions to provide clinicians with insights into how the models arrive at their decisions.

Longitudinal analysis: Extending analysis beyond single images to longitudinal studies, tracking changes in disease progression over time, which could provide valuable insights into treatment efficacy and patient prognosis.

By addressing these areas, future research endeavors can further advance the capabilities of CNN-based approaches in disease detection from medical imaging and ultimately contribute to improving patient outcomes and healthcare delivery.

Author contributions

Conceptualization, AVKP and AK; methodology, AVKP and AK; software, AVKP; validation, AVKP and AK; formal analysis, AVKP; investigation, AVKP; resources, AK; data curation, AVKP; writing—original draft preparation, AVKP; writing—review and editing, AVKP and AK; visualization, AVKP; supervision, AK; project administration, AVKP and AK; funding acquisition, AK. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The Department of Computer Science and Engineering at SRM University, Sonepat has supported this research work.

Conflict of interest

The authors declare no conflict of interest.

References

1. Ljung P, Krüger J, Groller E, et al. State of the Art in Transfer Functions for Direct Volume Rendering. Computer Graphics Forum. 2016; 35(3): 669-691. doi: 10.1111/cgf.12934
2. Madhiarasan M, Louzazni M. Analysis of Artificial Neural Network: Architecture, Types, and Forecasting Applications. Journal of Electrical and Computer Engineering. 2022; 2022: 1-23. doi: 10.1155/2022/5416722
3. Kim M, Yun J, Cho Y, et al. Deep Learning in Medical Imaging. Neurospine. 2019; 16(4): 657-668. doi: 10.14245/ns.1938396.198
4. Costanza D, Coluccia P, Castiello E, et al. Description of a low-cost picture archiving and communication system based on network-attached storage. Veterinary Radiology & Ultrasound. 2022; 63(3): 249-253. doi: 10.1111/vru.13061
5. Tadayon H, Nafari B, Khadem G, et al. Evaluation of Picture Archiving and Communication System (PACS): Radiologists' perspective. Informatics in Medicine Unlocked. 2023; 39: 101266. doi: 10.1016/j.imu.2023.101266
6. Mantri M, Taran S, Sunder G. DICOM Integration Libraries for Medical Image Interoperability: A Technical Review. IEEE Reviews in Biomedical Engineering. 2022; 15: 247-259. doi: 10.1109/rbme.2020.3042642
7. Aiello M, Esposito G, Pagliari G, et al. How does DICOM support big data management? Investigating its use in medical imaging community. Insights into Imaging. 2021; 12(1). doi: 10.1186/s13244-021-01081-8
8. Natsheh Q, Sălăgean A, Zhou D, et al. Automatic Selective Encryption of DICOM Images. Applied Sciences. 2023; 13(8): 4779. doi: 10.3390/app13084779
9. Mady AS, Abou El-Seoud S. An Overview of Volume Rendering Techniques for Medical Imaging. International Journal of Online and Biomedical Engineering (iJOE). 2020; 16(06): 95. doi: 10.3991/ijoe.v16i06.13627
10. Mamdouh R, El-Bakry HM, Riad A, et al. Converting 2D-Medical Image Files "DICOM" into 3D- Models, Based on Image Processing, and Analysing Their Results with Python Programming. WSEAS TRANSACTIONS ON COMPUTERS. 2020; 19: 10-20. doi: 10.37394/23205.2020.19.2
11. Ahsan MdM, Luna SA, Siddique Z. Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. Health Care. 2022.
12. Zakariah M, AlShalfan K. Cardiovascular Disease Detection Using MRI Data with Deep Learning Approach. International Journal of Computer and Electrical Engineering. 2020; 12(2): 72-82. doi: 10.17706/ijcee.2020.12.2.72-82
13. Arooj S, Rehman S ur, Imran A, et al. A Deep Convolutional Neural Network for the Early Detection of Heart Disease. Biomedicines. 2022; 10(11): 2796. doi: 10.3390/biomedicines10112796
14. Alarfaj A, Hosni Mahmoud HA. Deep Learning Prediction Model for Heart Disease for Elderly Patients. Intelligent Automation and Soft Computing. 2023.
15. Alzu'bi D, Abdullah M, Hmeidi I, et al. Kidney Tumor Detection and Classification Based on Deep Learning Approaches: A New Dataset in CT Scans. Journal of Healthcare Engineering. 2022; 2022: 1-22. doi: 10.1155/2022/3861161
16. Hannan SA, Pal P. Detection and classification of kidney disease using convolutional neural networks. J Neurol Neurorehab Res. 2023.
17. Rajkumar K, Sri Ramoju RT, Balelly T, et al. Kidney Cancer Detection using Deep Learning Models. In: Proceedings of the 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI). doi: 10.1109/icoei56765.2023.10125589
18. Ozaltin O, Coskun O, Yeniay O, et al. A Deep Learning Approach for Detecting Stroke from Brain CT Images Using OzNet. Bioengineering. 2022; 9(12): 783. doi: 10.3390/bioengineering9120783

19. Khairandish MO, Sharma M, Jain V, et al. A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images. IRBM. 2022; 43(4): 290-299. doi: 10.1016/j.irbm.2021.06.003
20. Ganesan M, Sivakumar N, Thirumaran M. Internet of medical things with cloud-based e-health services for brain tumour detection model using deep convolution neural network. Electronic Government, an International Journal. 2020; 16(1/2): 69. doi: 10.1504/eg.2020.105240
21. Aggarwal P, Mishra NK, Fatimah B, et al. COVID-19 image classification using deep learning: Advances, challenges and opportunities. Computers in Biology and Medicine. 2022; 144: 105350. doi: 10.1016/j.compbiomed.2022.105350
22. Chen J, Wu L, Zhang J, et al. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. Scientific Reports. 2020; 10(1). doi: 10.1038/s41598-020-76282-0
23. Available online: https://docs.aws.amazon.com/AmazonS3 (accessed on 9 January 2024).
24. Available online: https://docs.aws.amazon.com/AmazonRDS (accessed on 9 January 2024).
25. Available online: https://docs.aws.amazon.com/amazondynamodb (accessed on 9 January 2024).