https://doi.org/10.48047/AFJBS.6.13.2024. 2170-2191

African Journal of Biological Sciences

Different Perspective of Deep Learning: Medical Image Diagnosing

[1]Dr. C A Sathiya Moorthy, [2]Makesh Babu K, [3]Dr Mahmad Mustafa, [4]Deepeesh Vijayan, [5]Anitha Ponraj, [6]Dr. D. Hemanand

[1]Professor, Department of ECE, Sri Manakulavinayagar Engineering College, Puducherry-605107.

[2]Assistant Professor, Dept of Computer Application, Bishop Heber College (Autonomous), Trichy, India.

[3]Associate Professor, Methodist College of Engineering And Technology. An Autonomous Institution. Affiliated Osmania University. Abids, Hyderabad -500001.

[4] Department of Radio Diagnosis, Jawaharlal Institute of Postgraduate Medical Education & Research,

Puducherry – 605006, India

[4] School of Sciences and Humanities (Life Sciences), Sathyabama Institute of Science and Technology,

Chennai – 600 119, Tamil Nadu, India

[5]Associate Professor, Department Of Computer Science And Engineering, Kalasalingam Academy Of Research And Education, Kalasalingam University, Krishnakoil-626126, Tamilnadu, India.

[6]Professor, Department of Computer Science and Engineering, S.A. Engineering College (Autonomous), Poonamallee-Avadi Road, Thiruverkadu, Chennai-600077, Tamil Nadu, India.

Email-ID: [1]satya.cas@gmail.com, [2]makeshbabu.ca@bhc.edu.in, [3]mustafakmcet@gmail.com, [4]deepeesh2015@gmail.com, [5]anisainotech@gmail.com, [6]hemanand1982@gmail.com

Article History

Volume 6, Issue 13, 2024

Received: 18June 2024

Accepted: 02July 2024

doi:*10.48047/AFJBS.6.13.2024*. 2170-2191

Abstract

The main objective of this paper is to discuss various deep learning models. Deep learning algorithms are used in multiple data science fields but play a vital role in image processing, diagnosing, and predicting. In recent days, the size of data, especially in terms of images, has increased rapidly with dimensions. Earlier research works used Machine learning algorithms, but they got down while analyzing a large volume of data. Most models have been created based on deep learning methods recently, incorporating data-driven models into data learning models by eliminating human interactions. Deep learning models learn the data in-depth to identify, detect, and classify medical image data. According to domain-specific learning, extracting hierarchical and deep learning models can do high-dimensional data. Thus, deep learning models prove themselves rapidly as state-of-the-art models for various emerging applications. One of the applications is medical image diagnosing applications. This paper mainly discussed medical image diagnosis using various deep-learning algorithms. It also highlights the success stories in data/image analytics, providing reassurance about the effectiveness of deep learning models. From the experimental results, it is verified that CNN (Convolution Neural Network) obtained an average accuracy value of 98.99%, which is higher than that of other algorithms. Also, CNN consumes a processing time of 863ms, less than the other algorithms. This paper concludes by discussing research issues and future guidelines for later improvement, highlighting the potential of this field for further research and development.

Keywords: Deep Learning Algorithms, Medical Image Diagnosis, MRI, CT, Breast Mammogram, Deep Neural Networks.

1 Introduction

Medical imaging plays a major role in scientific and medical research, which provides computerized medical image division and computer-assisted design. Today, numerous tools are available for medical research, making sectional visions of human anatomy. Types of essential non-invasive imaging of human bodies are Computed Tomography (CD) and Magnetic Resonance Imaging (MRI). Here, MRI is exploited for human anatomical analysis. MRI test is the safest and painless method, and the magnetic field and radio waves are used to get detailed images of the brain and brain stem. MRI could detect brain tumors, bleeding, inflammation, growth, structural abnormalities, infections, inflammatory conditions, or blood vessel complications. It can determine if a shunt is functioning and detects damage to the brain caused by injury or stroke, creating high spatial resolution images with adequate soft-tissue contrast, which can help diagnose diseases [1]. The before and after contrast enhancement of an MRI image is shown in Fig 1.

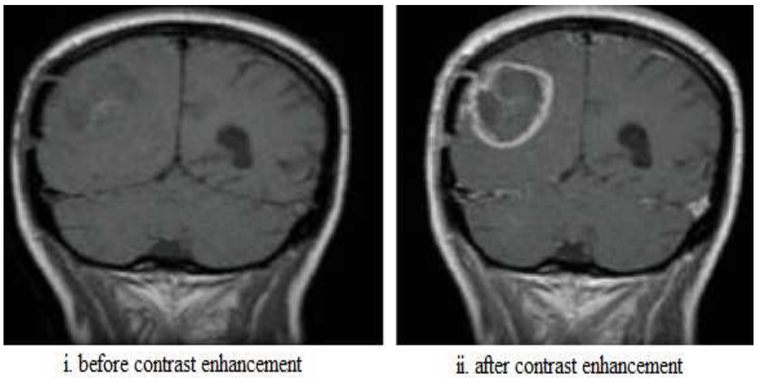


Figure-1. MRI image of before and after contrast enhancement.

Regarding anatomy and physiology, MRI techniques are characterized by a brain tumor. Determining the tumor's location, the number and size of the necrosis, the vascular presentation, and the associated edema are important factors in diagnosing the tumor. Different imaging techniques are used. These techniques are used depending on the tumor type and the diagnostic requirement [2]. Medical imaging is most important in emerging medical applications to provide better guidance to diagnose and cure diseases. Medical image analysis is one of the most important research works that needs to be designed and implemented efficiently. Some mathematical models introduced by human knowledge on medical images need to be reconstructed and handcrafted. After some time, they become data-driven models. Machine learning was a promising method for effective medical image or data analysis. The previous works show that the machine learning methods are not fully automatic in learning, extracting, and predicting the data. Since 2009, deep artificial neural networks have started outperforming benchmarks for bigdata analytics, and thus, deep learning methods are taken for deep analysis.

The brain tumor is another major cause for increasing the death rate among all other diseases in humans. As per the National Brain Tumor Society's study, 688000 people in the US survive with primary or Central Nervous System (CNS) tumors [3]. Another study from the International Association of Cancer Registries (IARC) stated that 28,000 people per month are surviving brain tumors, and approximately 24,000 people die per year in India. The brain tumor is ranked as the 10th most common kind of tumor worldwide [4]. It is necessary to detect and give proper treatment early to reduce the fatal rate unless it impacts people. MRI plays a vital role in detecting brain tumors in this brain tumor detection. However, image segmentation is the most challenging task due to low contrast, unclear boundaries, noise, variability, incomplete volume effect, and inhomogeneity. Many artificial intelligence and deep learning methods are proposed to solve the above-discussed limitations and improve MRI image processing, analysis performed, and automation [5].

Artificial intelligence (AI) seeks to mimic human intelligence and create an intelligent machine like a human, which can process information with human sense, behavior, and thought. Artificial intelligence has been used in many research fields, such as image processing, natural language processing, robotics, and tumor prediction. A core part of artificial intelligence is machine learning (ML). Machine learning is used to create mathematical models that can provide accurate and useful results by training the input data for unseen and new data. It comprises probability theory, convex analysis, statistics, and algorithm complexity concepts and is applied to cancer prediction, bioinformatics, and computer vision. Machine learning applications span the whole area of artificial intelligence. Machine learning constantly uses induction and synthesis, allowing computers to continuously gain new knowledge by simulating human behaviors and rearranging existing knowledge to improve the machine.

The inner part of machine learning is the concept of deep understanding. Conventional machine learning methods provide the results based on the trained input features from manually designed source data or other simple machine learning models. Alternatively, Deep learning models automatically learn useful features from source data and avoid these manual and complicated steps. The profound learning concept is efficiently used in medical data for this automatic feature learning. Deep learning has enhanced the concept of artificial neural networks. The idea of artificial neural networks inspires the human brain, which has interconnections of neurons. Compared to artificial neural networks (ANN), the deep learning model has many hidden layers, allowing it to create concise high-level features for different classes [7].

2. Literature Survey

Many research studies are being done to predict brain tumors using MRI. Inappropriately, low-variance enhanced MRI generally does not have adequate diagnostic image quality. With this low variance MRI, the results of the deep network for predicting a 100% contrast dose image from a study obtained with 10% of the standard contrast dose. This example MRI is obtained from a patient with meningioma. Such methods help give many patients the most accurate diagnostic images [9].

A furious and robust intelligent hybrid machine learning model is proposed for automatically detecting brain tumors. It is based on the feedback pulse-linked neural network for segmentation, the Discrete Wavelet Transform (DWT) for feature extraction, the Principal Component Analysis (PCA) for dimensional reduction of bandwidth coefficients, and the Feedforward backpropagation neural network for classifying inputs as healthy or pathological cases. Experiments showed that the classification accuracy was 99% [10]. The authors in [11] developed the Probabilistic Neural Network (PNN) classifier to train and test accuracy in locating tumor locations in brain MRI images. The proposed technique has proven effective when test results record almost 100% and 95% accuracy in training and test data sets, respectively, when identifying normal and abnormal tissues from brain MR images [11]. In the standardized method proposed for pre-processing, identical intensities depict the same tissue type in an image. Two CNN structures are trained, one for high-grade glioma (**HGG**) and one for low-grade glioma (LGG) (HGG deeper than LGG), in which the **Leaky Rectified Linear Activation** (LReLU) activation function is applied to whole layers except the last one. In addition to using small kernels, convolutional layers are proposed with a few weights that must be trained. Furthermore, regulation and data augmentation are used to reduce redundancy. The Dice metrics BRATS 2013 and BRATS 2015 for the two data sets are 0.88, 0.83, 0.77, 0.78, 0.65, and 0.75 for the complete, central, and upgraded areas, respectively [12].

Another work proposed a fully automatic deep convolution network-based U-net method for brain cancer segmentation. Here, the show results are suitable for the main area and comparable to the complete area of the tumor [13]. A new Deep Neural Network is proposed for unbalanced tumor labels. A two-stage training procedure has been used to train CNNs effectively. The authors have explored two types of manners: two-track and layered architecture—two-way configuration models with simultaneous local and global contextual features. The layered structure considered the labels' dependence on the pixels. Two CNNs are stacked, one of which is a CNN's output taken as the following CNN input, which helped achieve greater efficiency [14]. The proposed clustering algorithm, Fast generalized fuzzy C-means (FGFCM), in which the splitting time is considerably reduced, depends only on the gray values rather than the image size. It is also used to reduce the uncertainty of the right segment by local similarity measures. This method is suitable when no noise exists, which usually happens [15].

The authors in [19] presented a detailed survey of various deep learning methods for breast cancer prediction. The approach is to actualize ML algorithms with the necessary parameters for training and testing for better execution. Right now, a thorough overview of AI algorithms (Support Vector Machine (SVM), Decision Trees (DT-C4.5), Naïve Bayes, k-NN, and Artificial Neural Network (ANN)) is finished considering the ongoing exploration papers on the expectation of breast cancer growth tumors on the Wisconsin Database from the UCI Repository. The authors in [20] proposed Inception_V3 and Inception_ResNet_V2 architectures for the binary and multi-class breast cancer histopathological image classification issues using transfer learning techniques. The balanced subclasses with Ductal Carcinoma are the baseline for balancing the imbalanced histopathological images in subclasses. The images turn up and down, right and left, and rotate for balancing. The experimental results showed that the Inception_V3 and Inception_ResNet_V2-based histopathological image classification of breast cancer is superior to the existing methods. The authors in [21] discussed various earlier research works on machine learning-based breast cancer detection and classification. According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, some research searched PubMed, Embase, and Web of Science core databases from November 30, 2020. The extracted information included literature, database, data preparation, modeling process information, model construction and performance evaluation information, and candidate predictor information.

2.1 Objectives of the Paper

The main objective of this paper is to discuss various deep learning algorithms to understand their architecture. To understand how deep learning algorithms are used in diagnosing different medical images. It helps to understand the core functionalities of deep learning algorithms and their variation in efficiency in terms of various medical image processing. It is easy to identify a suitable deep-learning algorithm for diagnosing a particular medical image from the comparison.

2.2 Limitations and Motivation

i. Intensity distribution in an image for various tissue types does not allow adaptation to scanners or individuals' variations or time.
ii. Immediate neighboring pixels are used to regain the intensity inhomogeneity and recognize the pixels' labels. However, the computation of these neighborhood terms at each iteration takes a lot of time.
iii. In MRI images, gray values exceed the image size. So, we need to balance between sensitivity and image preservation information.
iv. MRI images have low contrast, unclear boundaries, noise, variability, and incomplete volume effect. It has made the situation tougher when classifying data.

Several proposed methods are used in the experiment to overcome the above-said issues. However, deep learning models are rapidly proving to be sophisticated for reducing these risk factors from clinical images.

2. Deep Neural Networks Review

Deep Neural Networks (DNNs) are now great tools for managing complex data. The significant variations of DNN and conventional machine learning are the integrated nonlinearity of DNNs and endwise training. Deep neural networks are very active methods for extracting the most relevant features for a particular task. Nowadays, DNNs are helpful in various medical imaging processes [16].

3.1 ResNet

Residual Network (ResNet) [17, 18] is the most popular feature of DNN in computer vision. ResNet architecture is depicted in Figure-2. It is also defined by mathematical equation (1) as

$$\boldsymbol{u}_{k+1} = \boldsymbol{u}_{\boldsymbol{k}} + F_{k(uk)} \qquad (1)$$

From the above equation, $F_{k(uk)}$ denotes the input (resp. output) feature of the $\boldsymbol{k}$-th layer of the residual network. It indicates a nonlinear residual block with maximum trainable parameters. In stable training facilitation, the skip connection of ResNet is crucial where the network is too deep. DenseNet, U-Net, and learned diffusion model TRD have skip connections similar to ResNet.

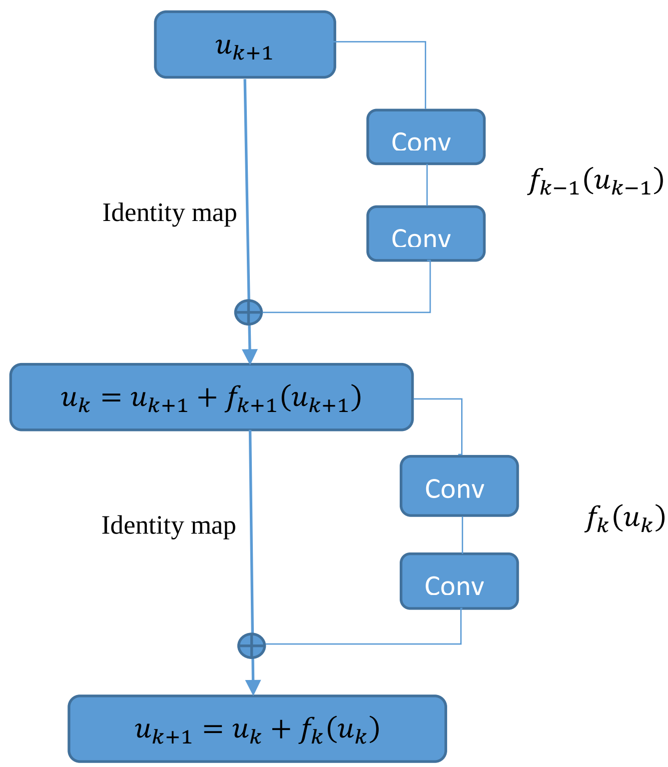


Figure-2. ResNet

4.2 Autoencoder

The autoencoder is one of the ANN models used to learn different data representations. It learns as an encoder from data collected with a decoder. It is done so as not to lose any vital information at the time of the encoding and decoding process. The autoencoder is defined in the following figure-3. For a given image *X*, the parameterized mapping f_θ is typically an encoder, used as a feature maps extraction from X. $\boldsymbol{Y = f_\theta(X)}$ denotes encoded multi-channel feature maps. Another parameterized mapping then decodes the encoded Y feature map $\boldsymbol{g_{\theta'}}$ for obtaining reconstructed data, $\boldsymbol{Z}$**.** Parameters $\boldsymbol{\theta}$ and $\boldsymbol{\theta'}$ are highly optimized on a given data set. It is done to choose the proper loss function, measured to minimize average discrepancies between X and Z. AE. It is mostly similar to linear representations such as Fourier and Wavelet transform. Encoding is regarded as decomposition, and decoding as reconstruction. Feature maps are represented as coefficients.

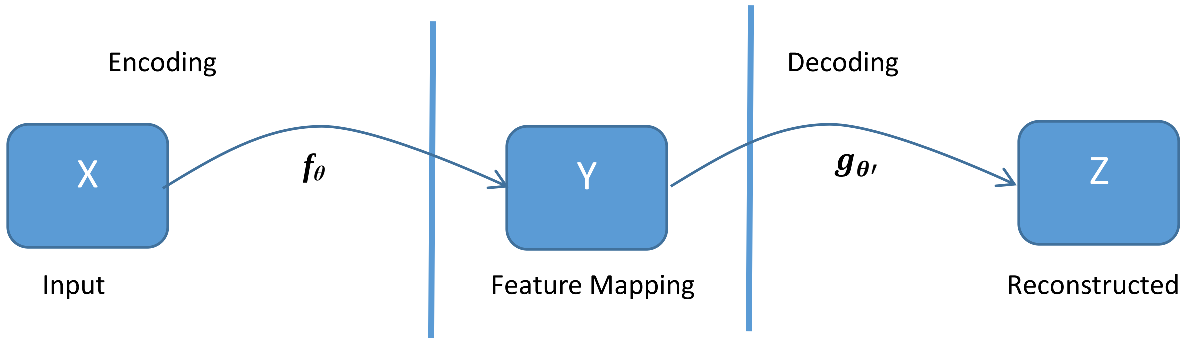


Figure- 3. Autoencoder

However, AE representation is nonlinear and learned from a data set. A stacked denoising autoencoder (SDAE) is used to learn a more effective and robust model. From a noisy data X, the encoder and decoder are used to recover Z ≈ X, using DNN. For image segmentation, a special DNN called SegNet is used based on the encoder and decoder of the framework. The encoder and decoder framework are adopted for existing scenarios for image denoising and super-resolution. A residual encoder/decoder CNN lowers the noise level and contains exact features in low-dose CT images. These are later reconstructed using the filtered back-projection (FBP) algorithm.

4.3 U-Net

U-Net is nothing but a U-shaped deep neural network. It is mainly proposed for biomedical image segmentation. This is one of the most successful deep image segmentation models. The architecture of U-Net resembles AE architecture, where the left half view of U-Net acts as an encoder and the right half view acts as the decoder. The difference between U-Net and AE is the use of skip connections of U-Net. Like U-Net, V-Net is a DNN for 3D volumetric medical image segmentation. The multi-resolution deep convolution framework is developed from the U-Net and convolution framework. Recently, U-Net has been extended to image analysis tasks. U-Net is more successful than conventional models in architecture and pixel-based image segmentation formed from convolutional neural network layers. It's even effective with limited dataset images. The presentation of this architecture was first realized through the analysis of biomedical images. The architecture of U-Net is depicted in Figure-4. U-Net takes its name from the architecture, which, when visualized, appears similar to the letter U, as shown in the figure above. Input images are obtained as a segmented output map—the most special aspect of the architecture in the second half. The network does not have a fully connected layer. Only the convolution layers are used. A ReLU activation function activates each standard convolution process.

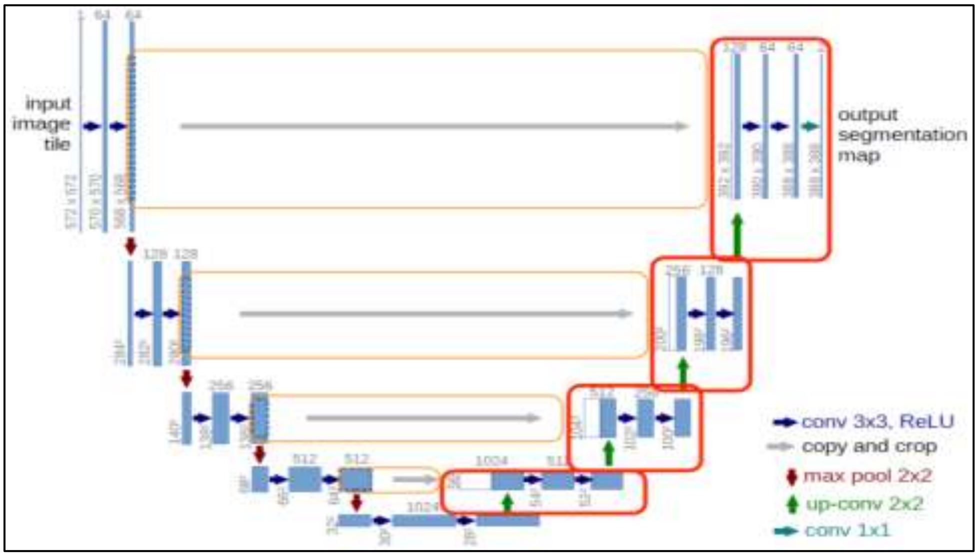


Figure-4. U-Net Architecture

5. Interpretations of Deep Neural Networks

Traditional machine learning methods, such as decision trees, random forests, and support vector machines, have been developed in recent years, which benefit theoretical studies in machine learning. PAC, Rademacher complexity, and PAC are existing machine learning theories that are unsuitable for analyzing DNN. DNN comprises simple pooling, convolutions, and element-wise activation functions. Sometimes, the entire network becomes difficult to analyze. Because of this issue, deep theoretical learning is becoming increasingly popular in the machine learning platform. DNN is interpreted from two perspectives: representation learning and differential equations. Function approximation is considered a powerful tool in characterizing the efficacy of the given representation. This provides a deep analysis of the capacity of DNN and how efficiently the approximate function is used in various function spaces. A differential equation is more challenging than a function approximation, which can explicitly guide the architectures of DNN and its corresponding training algorithms. Apart from these perspectives, several other views are based on theoretical interpretations of deep learning.

5.1 Representation Learning Perspective

Generally, medical images and natural images usually have low-dimensional structures. Various transformations efficiently extract these low-dimensional structures. Fourier, curve, and wavelet transform methods are successful examples of transformations. They can produce an efficient representation of images by transformation. They are predesigned linear transformations and are independent of the given image data. DNNs are also used to extract sparse or low-dimensional features from the given images. DNN can learn from a set of images, and they are nonlinear.

The quality of the given image is measured by its exact ability to approximate functions living in a certain function space. For example, let $\boldsymbol{\Phi} := \{\boldsymbol{\varphi i}(\boldsymbol{x}) : \boldsymbol{x} \in \boldsymbol{R}^{\boldsymbol{n}}, \boldsymbol{i} \in \boldsymbol{N}_+\}$ be a set of atoms and an element in function space $\boldsymbol{F}$ is the function f(x), equipped with norm ‖ . ‖. The most basic and important approximation properties state as follows: for any given ε > 0, there exists $\tilde{f}_{\alpha,N} := \sum_{i=1}^{N} \alpha i \varphi i(x) \;\; with\ N \in N_+ \;\; and\ \alpha = \{\alpha 1,\cdots,\alpha N\} \in R_N$ such that

$$\left\| f - \tilde{f}_{\alpha,N} \right\| < \varepsilon \qquad (2)$$

A perfect representation requires fewer atoms (i.e., smaller N) to achieve ε-approximation. The model of various types of Φ has been utilized in literature, such as splines, polynomials, Fourier bases, and wavelets. The neural network is one of the most powerful tools for approximating a function. Factors affecting approximation power are a neural network's depth and width. Consider a shallow network with only one hidden layer.

$$\tilde{f}_N(x;\Theta) = \sum_{i=1}^{N} a_i\, \sigma(w_i^T\, x + b_i) \qquad (3)$$

where $x \in R^n$ is the input data, the trainable parameters are Θ ={ a_i, w_i, b_i}, i = 1,···, N. σ(z) is an element applied for nonlinear activation function. Examples of σ(z) are ReLU(z) = max(0,z), $tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}, sigmoid(z) = \frac{1}{1+e^{-z}}$ and more generally a sigmoidal function [114] that has the property.

$$\sigma(z) = \begin{cases} 1, & if\ z \to +\infty \\ 0, & if\ z \to -\infty \end{cases} \qquad (4)$$

DNN is a typical neural network with multiple hidden layers. It can be considered as a successive composition of multiple shallow networks. DNN used for regression analysis must have depth L and width N = (N_1, N_2, …, N_L), and it is denoted as

$$\tilde{f}_{L,N}(x;\Theta) : R^n \to R, \qquad (5)$$

this equation can be recursively defined as $\Theta^l = (\Theta^{l-1}, \Theta^l),\ \tilde{f}_{\Theta^l} = (\Theta^l \circ \sigma \circ \tilde{f}_{\Theta^{l-1}}), \Theta^l : R^{N_l} \to R^{N_{l+1}}$ with Θ^l (x) $= w^l x + b^l$, and $\tilde{f}_{L,N} := \tilde{f}_{\Theta^L}$. Recent results on the approximation property, i.e., universal approximation, are a wide class of functions that can be approximated by several neural networks with exactly one hidden layer through a total number of neurons (N). It will increase exponentially by decreasing ε. While approximating a target function, increasing the depth L produces benefits in a neural network. While maintaining the same level of approximation accuracy, the polynomial reduction in the number of neurons is made. Several examples give concrete evidence that DNN is more efficient in approximation than shallow networks. DNN with O(L^3) layers and constant width could not be approximated by networks with O(L) depth of less than 2^L depth.

The efficiency of the depth of ReLU activated Researchers investigate DNNs. This research is done by proving the availability of existing classes of wide neural networks. A narrow network cannot realize this wide neural network since the depth is merely a polynomial bound. While comparing to the previous results, deep networks cannot be recognized by any external network. The size of the shallow network is merely an exponential bound. Previous works notify that dept is slightly more effective than width. Hanin proves that the minimum width of ReLU activated DNNs ensures continuous function approximation, even though depth is more important than width. Typical DNN cannot be too narrow. It is because we cannot approximate continuous function even with infinite depth. Yarotsky analyzed the depth of ReLU activated DNNs with the dependence of optimal approximation. In approximating a multivariate polynomial, the number of variables of the polynomial and the total number of neurons in DNNs should grow linearly. Shen provided a fundamental analysis on ReLU activated DNN. It is done through a nonlinear approximation of composite dictionaries. By comparing N-term approximation order with one hidden layer neural network, they proved the advantage of depth over width. Theoretical analysis on the ResNet is provided other than the generic DNN.

Researchers investigated the possible connection between linear finite element functions and ReLU deep neural networks. An efficient ReLu activated DNNs are proposed for representing linear finite element functions. It is theoretically established that a minimum of 2 hidden layers is needed for representing any linear finite element functions in $\Omega \subseteq R^d$ when $d \geq 2$. By using this relationship, they developed a straightforward error estimation as $O(N^{\frac{1}{d}})$. It is done for a particular type of ReLU activated DNNs, and the nonzero parameters O(N) involves the h-adaptive linear finite element approximation theory. Apart from the approximation model, a unified model MgNet is developed. The model's main advantage is that this model recovers and extends some CNN simultaneously for image classification and multi-grid methods. Multi-grid techniques are used to solve discretized PDEs by combining deep learning and multi-grid methodologies.

5.2 Differential Equation and Control Perspective

DNN $\tilde{f}\,(x;\Theta) : R^n \rightarrow R^m$ and its composite functions are explained in previous sections. $\tilde{f}\,(x;\Theta)$ as an iterative mapping between R^n and R^m. Generic DNN is generally viewed as a certain dynamic system. A generic DNN is quite complex to analyze because it does not contain many special structures. It has been proved that many DNNs have special structures in their in-build architecture. Designing special structures of DNNs is to make the structure easily trainable and for better generalization. Developing neural architecture search (NAS) is a subfield of Automated machine learning (AutoML). It was created to search for an effective DNN architecture for different datasets and tasks. ResNet is popular among DNNs. Ultra-deep networks are efficiently trained by using ResNet shortcuts. It also provides higher accuracy in multilevel tasking. This success inspired several similar but robust neural architectures. Almost all the designs are based on empirical studies. Previously, NAS was deployed to search for new neural network architectures. NAS possesses its own risk in searching; it is of a high standard for new researchers who do not have access to computational resources. NAS could not always help find new and interpretable neural networks. To seek the principles for the architecture design, we need alternate solutions for ResNet interpretation. The methods of ResNet solve the Ordinary Differential Equation (ODE) called Forward Euler Scheme, while the controls are optimal. ResNet acts as a dynamic system. Designing more efficient algorithms requires a detailed analysis of ResNet's dynamics and control perspective. By observing the multiple-state-of-the-art deep network architecture, the authors suggested the general bridge is acts between numerical

ODE and deep neural architecture. ResNet approximate stochastic differential equations get weaker with incomplete stochastic training strategies. Through numerical (stochastic) differential equation methods, new perspectives of systematic design deep neural architectures are developed.

5.2.1 Numerical Difference Equations and Architecture Design

We first show how ResNet is related to the forward Euler scheme in numerical ODEs. Considering a building block of ResNet (2.1), as shown in Fig. 1, it can be rewritten as

$$\boldsymbol{u}_{k+1} = \boldsymbol{u}_{\boldsymbol{k}} + \Delta t_k F(u_k, t_k) \tag{6}$$

where, Δt_k is the step size and $F(u_k, t_k) = \frac{1}{\Delta t_k} f_k(u_k)$, The above formula is the forward Euler scheme solving the following ordinary differential equation (ODE):

$$\frac{du}{dt} = \mathrm{F(u,t)} \tag{7}$$

Therefore, the ResNet can be viewed as the forward Euler discretization of the ODE with a step size $\Delta t_k = 1$ for every k. It was first observed by Weinman [86]. More recently, Zhang et al. [144] showed that there are benefits of considering ResNet with $0 < \Delta t_k < 1$.

5.2.2 Stochastic Training and Optimal Control

Stochastic training includes dropout and noise injection and is mainly used to train DNNs. ResNet, stochastic depth, and shake-shake training methods are alternatively viewed as stochastic control.

$$\min_{\Theta} \; E_{X(0)\sim data}[\,E[L(X(T))] + \int_0^T R(\Theta) \tag{8}$$

$$\text{s.t. } dX = F(X,\theta), dt + G(X+\theta)dB_t \tag{9}$$

The above equation is the weakest link of the ResNet with the shake-shake mechanism or stochastic depth, respectively. It shows the connection between stochastic training and control and offers the relationship between DNNs with randomness and stochastic differential equations. Optimal control of backward Kolmogorov's equation is closely related to the variants of stochastic training of ResNet. The introduction of viscosity to the equations is done by popular dropout regularization.

6. Deep Models in Medical Image Reconstruction

For CT imaging, FBP and algebraic reconstruction methods are considered classical medical image reconstruction methods. Though these methods are widely used in practice and are highly efficient, they are vulnerable to noise and inaccuracy in measured data. Many regularization-based models and algorithms have been developed to obtain high-quality images. By combining traditional image reconstruction methods with deep learning, medical science is well-developed in image reconstruction. While integrating the conventional approach into deep neural networks, two criteria are considered: post-processing and raw-to-image. The validation process by these two methods is still incomplete. Pre-processing methods do map between the reconstructed low-quality image and its high-quality image. It is done since DNNs can approximate generic functions or mappings. This method is highly efficient in providing results during the initial reconstruction of the images. Initial reconstructed images contain noise and complex artifacts, which makes it harder even for deep neural networks. Missing information from the initial reconstructed images is hard to recover by any post-processing methods. This is why initial reconstruction methods are different from post-processing methods.

In converting raw to image data, the mapping is done between the raw data and the reconstructed images. Raw data includes projection data of CT and K-space data for MRI. Data distribution in the domain between the raw data is quite challenging because the distribution in raw data is entirely different from that in the image domain. Applying direct mapping using DNN without special characters is difficult since it requires more training data and is computationally exhaustive. It highly depends on impressive initializations of model parameters (AUTOMAP). It is well known that handcraft modeling and mapping require having several dynamic structures that a designed partial differential equation can represent. It is more conventional to combine handcraft methods with deep neural networks. Studies show the connection between DNN and dynamic systems and the advantages of recognizing such connections. Handcraft modeling in image restoration includes tools for freely selecting for mapping $\boldsymbol{F}$ through a general technique called unrolling dynamics. Unroll optimization algorithms are introduced to solve handcrafted models in feed-forward networks. Domain knowledge of the handcrafted techniques is incorporated into the best suitable parameters in an end-to-end fashion. The deep model technique through unrolling dynamics is more understandable than regular deep models, and the number of parameters is less in the unrolling dynamics method. It is well suitable for small sample learning. It paves the way for combining general domain knowledge with deep learning, and the decision is based on the importance of the component needed for the model. It can be handcrafted without losing the exponential power of the existing model. The major difference between medical image reconstruction and restoration is the quality. The quality analysis is applied to the reconstructed images. The quality metrics are better task-based than generic metrics such as PSNR and SSIM. The main question lies between the realizations of the generic metrics because of conventional methods. Researchers proposed a task-based quality metric technique by hooking an image reconstruction method from unrolling dynamics. Similar approaches are made possible in computer vision in creating the generic quality metrics for image denoising. New works include raw-to-task, which provides promising results because it is more effective in fundamental analysis and reconstruction sort of all works.

6.1 Post - Processing

The initial reconstructed images are improved, and the qualities of the images are enhanced with post-processing techniques. This section CT is considered an example due to the incomplete measured data and the limitations in the viewing angle. FTB reconstructed images, and the noise caused by low tube currents are often referred to as a source of degradation in CT images. The degraded and clean image mapping is approximated using a residual encoder-decoder CNN. It is quite efficient in removing noise from the reconstructed images. Generative adversarial methods (GAN) protect the structures of the CT images. It combines Wasserstein distance and the perceptual difference between the input degraded image and the corresponding clean image. The radiation dose and the acquisition time are reduced by decreasing the number of projections of X-rays. It is known as the sparse view or limited-angle CT. It leads to streaking artifacts because of their incomplete measurements, and it is relatively simple in FBP-reconstructed images. Capturing global patterns is done using DNN with multi-scale architecture. The artifacts in FBP reconstructed U-net reduce CT images. The high-quality CT image is obtained by removing the waste-learned artifacts from U-Net degraded images. U-Net also takes a role in residual learning.

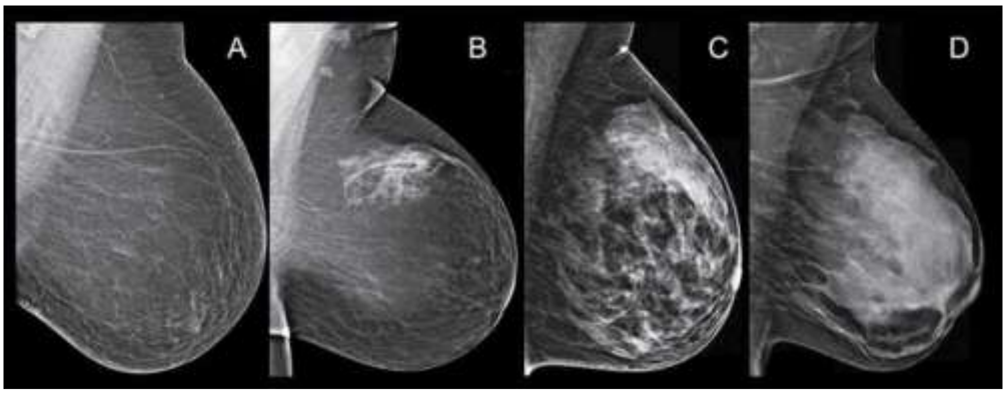
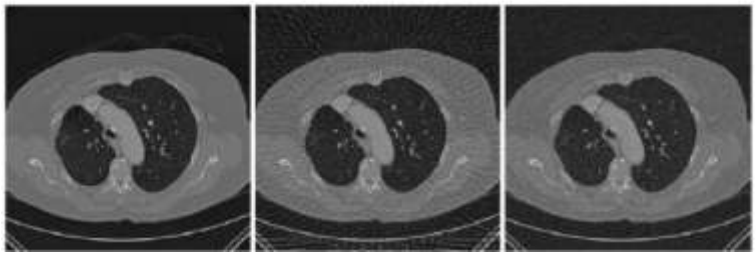


Figure-5. Breast Mammogram Image Classification Using Density

Density-based breast imaging provides breast images and reports to doctors. It is called BI-RADS, which groups various types of breast images with multiple densities. For example., in Figure-5, (A) shows Almost entirely fatty breast tissue, found in about 10% of women, (B) shows the scattered areas of dense glandular tissue and fibrous connective tissue (scattered fibroglandular breast tissue) found in about 40% of women, (C) shows the Heterogeneously dense breast tissue with many areas of glandular tissue and fibrous connective tissue, found in about 40% of women, and (D) shows the Extremely dense breast tissue, found in about 10% of women. The above image is classified as "heterogeneously dense" (C) or "extremely dense" (D) breasts.

6.2 Raw-to-image

End-to-end training can be set up as deep feed-forward networks with the help of unrolled optimization algorithms. Unrolling dynamics prove to have better convergence than the conventional optimization algorithms. In this case, machine learning techniques are used to improve optimization algorithms. The dual aspect of optimization algorithms inspires new and more effective architectures for image reconstruction. The FBP reconstruction images are depicted in Figure-6 for CT images and Figure-7 for breast mammogram images.



(a). Ground Truth, (b). Low Dose CT, (C). Low Tube CT

Figure-6. FBP reconstructed images

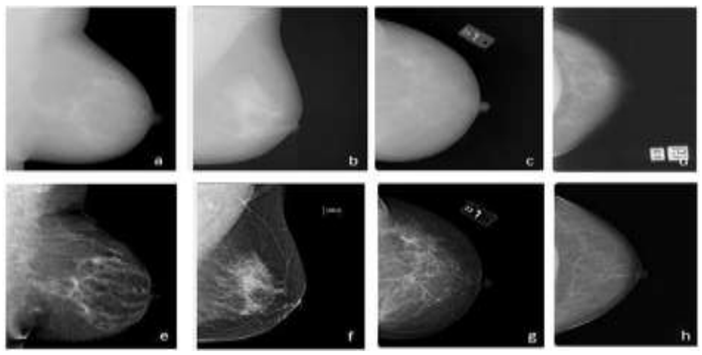


Figure-7. Raw To Image of Brest Mammogram Images

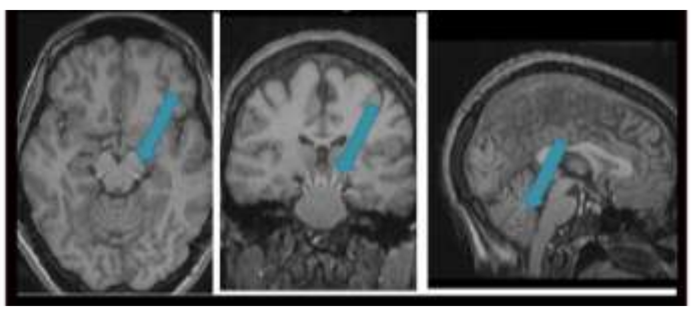


Figure-8. Raw To Image of MRI

The image conversion from raw to medical digital images is applied to various medical image modalities and creates the input images. The input images obtained for several types of medical images are shown in Figure-6 (CT images), Figure-7 (X-ray images), and Figure-8 (MRI images).

ADMM-net

ADMM-net potentially benefits the design of a deep neural network for inverse problems with the help of an unrolling dynamics optimization algorithm. The iteration scheme of the ADMM algorithm has tuning parameters such as $\boldsymbol{\mu}$**,** $\boldsymbol{\lambda}$**,** and a special handcrafted operator called ***W. W*** is difficult to determine adaptively for a given data set. ADMM-Net is nothing but the unrolling method of the ADMM algorithm. It gives the tuning parameters, and the predefined linear operators W are made available from the training data. The sparsity-promoting function Φ

= **|| . ||₁** parameterized by a piecewise linear function with learnable parameters. The thresholding operator F λ(·) is learned from the ADMM algorithm from the training data. In the basic version of the ADMM algorithm $\boldsymbol{d}^{k+1}$ is updated by

$$d^{k+1} = T_{\Theta_1}(W_{\Theta_2}(u^{k+1}) + b^k \quad (10)$$

The term T_{Θ_1} (·) is parameterized piecewise linear function, and Θ_1 is the parameter and the notation W_{Θ_2} is a parameterized convolution layer based on the parameters Θ_2.

The primary ADMM-Net is later upgraded to generic-ADMM-Net. In this technique, a different variable splitting strategy is induced in ADMM algorithmic derivation. The results obtained from Generic-ADMM-Net are more significant than the BM3D-based algorithm.

Primal-Dual Networks (PD-Net)

Iteration schemes are unrolled PDHG to design a new model for CT image reconstruction. It is also called as Primal-Dual Networks (PD-Net). The main ideology of this model is to approximate each resolvent/proximal operator into the sub-problem of PDHG with the help of a neural network. Choosing optical forms such as Φ and ***F*** is complex, and this model prevents these difficulties. Each layer in the model takes the form:

$$w^{k+1} = N_w([w^k, Wu^k]; \Theta_w^k) \quad (11)$$

$$u^{k+1} = N_u([u^k, \mathrm{W}^T \mathrm{w}^{k+1}, \mathrm{A}^T \mathrm{f}]; \Theta_w^k), \quad (12)$$

Where *f* is the measured data, the imaging operator is denoted by *A*, $Nw(\cdot;\Theta_w^k)$ and $Nu(\cdot;\Theta_w^k)$ are the neural networks, and the parameters are Θ_w^k and Θ_w^k . $[v_1,\cdots,v_m]$ denotes the concentration of the components. *W* is the linear operator that can be fixed or learned from the data. In comparing with FBP, PD-Net is more significant in boosting performance.

Joint Spatial-Radon (JSR)-Net

The artifacts contain noise and incomplete data. Joint Spatial-Radon (JSR)-Net is a domain reconstruction model for sparse view CT imaging.

$$\min_{u,f} \; F(u,f,Y) + R(u,f), \quad (13)$$

where the data fidelity term $F(u,f,Y)$ is defined by

$$F(u,f,Y) = \frac{1}{2} \|R_{\Gamma^c}(f - Y)\|^2 + \frac{\alpha}{2} \|R\Gamma \; (Au - f)\|^2 + \frac{\gamma}{2} \|R\Gamma^{\,c}(Au - Y)\|^2 \quad (14)$$

and the regularization term defined by

$$R(u,f) = \|\lambda 1 \cdot W_{1u}\|_{1,2} + \|\lambda 2 \cdot W_{2f}\|_{1,2} \quad (15)$$

Notation R_{Γ} denotes a restriction operator concerning the data missed region by Γ. The value of R_{Γ} is 1, if the index of the element is Γ otherwise, it is 0. The region of available measured data is indicated by Γ^{c} and the complement of the region is denoted by Γ. A. Y. denotes the random transform's discrete form and the measured projected data. In the JSR model, u and f are underlying CT images, and it restores the high-quality projection data. W_i,i =1, 2, are tight frame transforms, and λ_i, i = 1, 2, are the regularization parameters.

The Random domain and image domain consistency are enforced simultaneously by the handcrafted JSR model. Thus, the reconstructed image quality is enhanced. The positron emission tomography model is adopted in similar data fidelity. The JSR model is improved by learning how the tight frame transforms $\boldsymbol{W}_i$. Metal artifacts in multi-chromatic energy CT are reduced by improving the JSR model. The potential of the JSR framework is seen in existing research. Solving the JSR model is done using the unrolling method. JSR-Net is designed for sparse viewing and limited-angle CT. Proximal operators are approximated using JSR-Net by adopting neural networks similar to the PD-Net. The efficient utilization of multi-domain features improves the quality of the reconstructed image. The result obtained using JSR-Net is given in Figure-9.

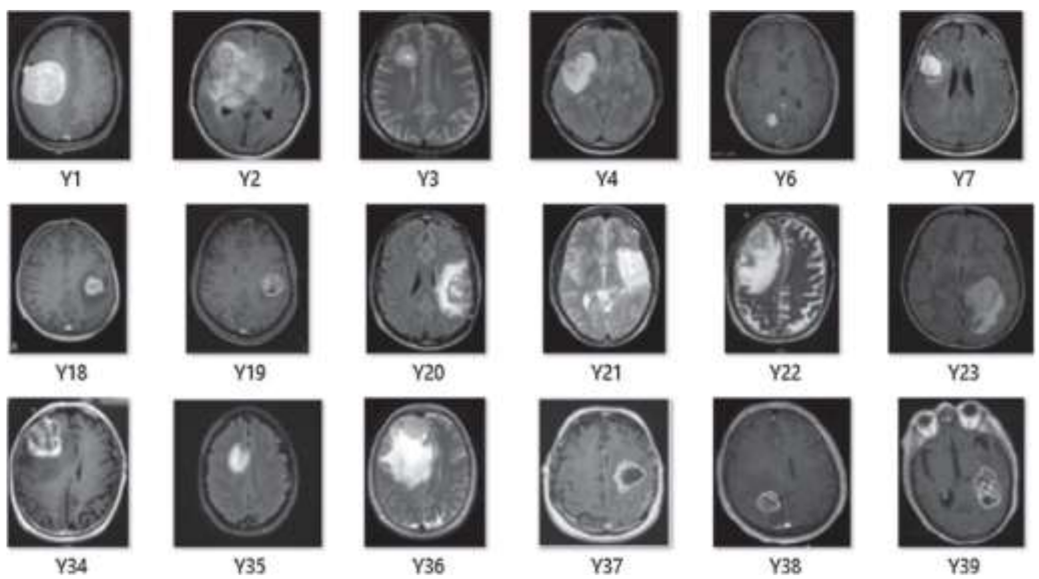


Figure-9. Results Obtained Using JSR-Net

6.3 Raw-To-Task

Two separate stages are carried out in a traditional medical image analysis workflow, shown in Figure-10. The high-quality image is reconstructed from the raw data, and the next step is to make a diagnosis based on the high-quality reconstructed image.

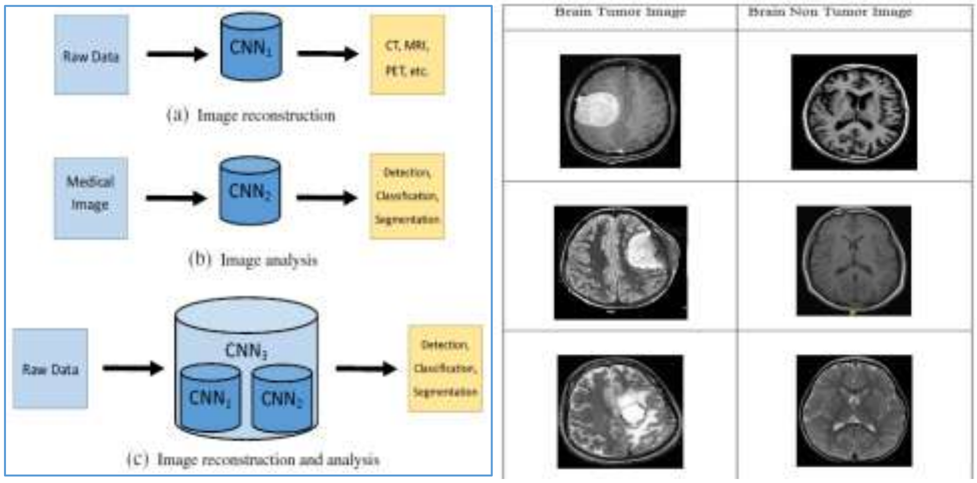
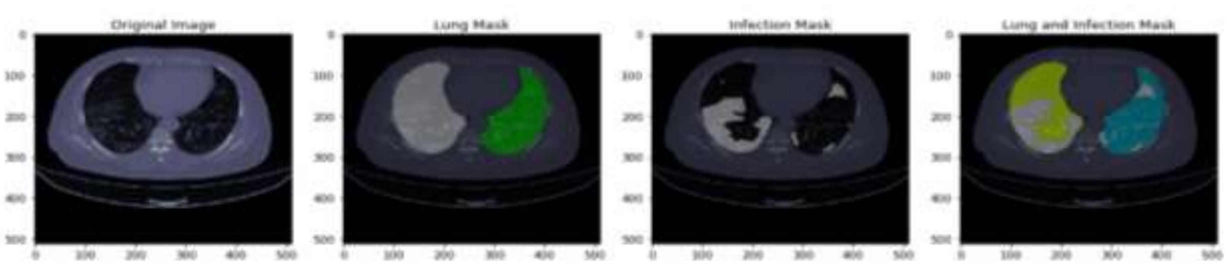


Figure-10. CNN Workflow and Experimental Result

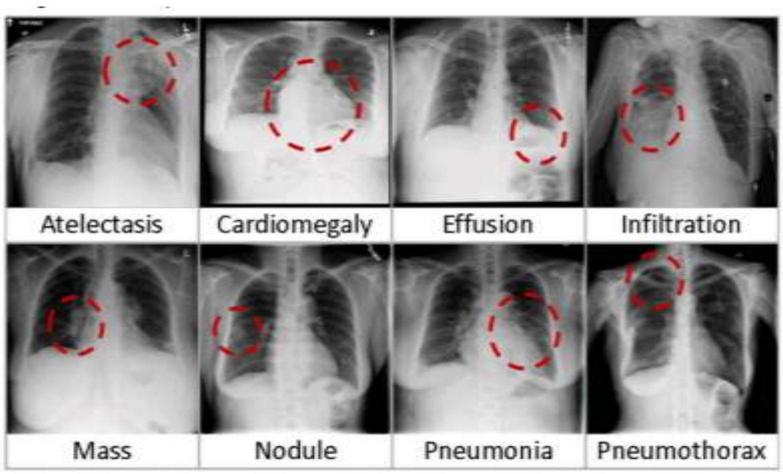
As discussed above, designing feed-forward deep networks for image reconstruction takes two steps. There are more choices for in-feed-forward neural networks for image reconstruction. Connecting the two networks and conducting end-to-end training (from fine-tuning or scratch) is the easiest image reconstruction method. Task-based image quality metrics always define the second network for image reconstruction learned from obtained data. From the image above, the task taken is lung nodule recognition, where the learned image quality is automatically updated within the lung areas. Such quality metric is specific to lung nodule recognition since the exact image quality present outside of the lung areas is irrelevant for tasking.

Experiment Results and Discussion

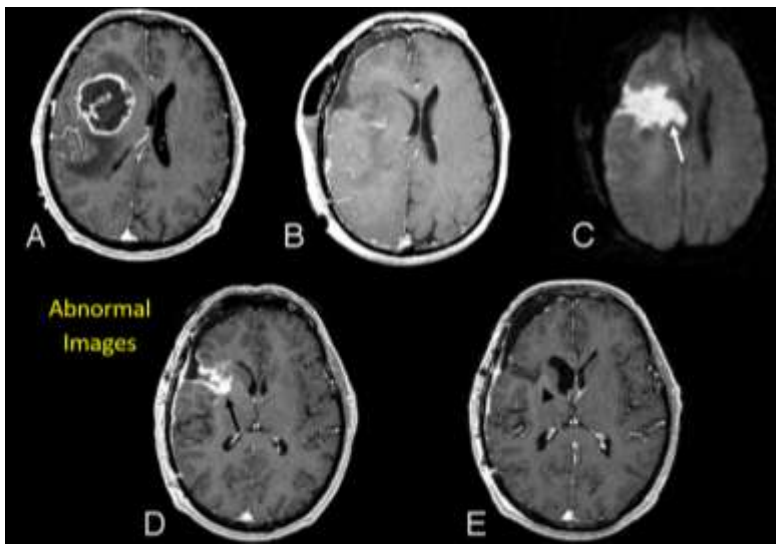
All the methods discussed in the paper are experimented with in MATLAB software and the results are verified. Experimenting with the above-said methods, various kinds of images (multi-modalities) such as 200 CT images [22], 5800 X-ray images [23], and 500 mammogram images [24] are downloaded from various websites, which permits all to download the images freely. A total of 6500 images were collected [21-23] and experimented with. The performance is compared in terms of accuracy and time complexity obtained by all the methods, which is given in Table-1. The accuracy of the methods is carried out for the classification of the images as, normal or abnormal. Based on this common classification, all the methods are verified on three kinds of datasets, and performance is verified. The obtained abnormal images for the three different datasets are given in Figure-11.



(a). CT Image Diagnosis



(b). Classification on X-Ray Images



(c). Classification on MRI Images

Figure-11. Experimental Classification

Performance Comparison

The performance comparison in terms of classification accuracy and time is given in the following table for the above-discussed deep learning methods.

Table-1. Performance Comparison in Terms of Accuracy and Time

Methods	Accuracy (%)	Time (ms)
ResNet	89.96%	1896
Autoencoder	86.45%	956
U-Net	93.1%	1569
DNN	97.61%	1981
ADMM-Net	95.34%	1894
PD-Net	96.89%	1367
JSR-Net	97.42%	1234

CNN (Proposed)	98.99%	863

The comparison shows that the Convolution Neural Network (CNN) obtained better results than the other deep learning models. The values given in Table-1 are the average values obtained from multiple medical image diagnoses. Thus, this paper decided that the CNN algorithm is highly suitable for any medical image diagnosis.

7. Challenges and opportunities

However, deep models completely dominate in image reconstruction. Still, there is plenty of room to develop the model further since it has some clinical practice limitations.

- The everlasting hunger for labeled data: only limited labeled information is available to develop new models in medical image reconstruction techniques. Examining medical images is more exaggerated in terms of both time and expert knowledge.
- The limited number of observations: gathering a large amount of medical data for a specific task is quite difficult because of privacy concerns. Common causes are less valuable than rare cases, and large amounts of data are needed for rare-case scenarios.
- The decision made by radiologists does not rely only on medical images. More and more information is observed from the patients, and the examiner's experience is also essential in decision-making. Thus, gathering data from various sources can also be useful in developing a system hierarchy.
- The reasoning is merely essential, but not more than inference. Reasoning procedures are mainly hidden in deep network models. Based on the wrong rationale, there is a chance that the models will make the wrong prediction, and hence, the model becomes unreliable.

Conclusion

The main objective of this paper is to provide a deep knowledge of various deep learning algorithms. By analyzing various deep learning algorithms, it helps to choose a better deep learning algorithm for several types of medical image diagnosis. To do that, MRI, CT, and X-ray images are taken and analyzed using various algorithms. This paper provides the overall functionalities of the deep learning algorithms. All the algorithms have been experimented with, and the results have been verified regarding accuracy and time complexity. The CNN obtained an average accuracy value of 98.99%, higher than other algorithms. Also, CNN consumes a processing time of 863ms, less than the other algorithms. From the comparison, it is found that the CNN algorithm is suitable for medical image diagnosis on any imaging.

In future work, the CNN algorithm will be elaborately explained with detailed architecture and experiments on a larger number of medical images.

References

[1] https://kidshealth.org/en/parents/mri-brain.html

[2] Kavitha Angamuthu Rajasekaran and Chellamuthu, "Advanced Brain Tumour Segmentation from MRI Images", Basic Physical Principles and Clinical Applications, 2017.

[3] https://braintumor.org/brain-tumor-information/

[4] https://www.narayanahealth.org/brain-tumour

[5] Anjali Wadhwaa, Anuj Bhardwaja, Vivek Singh Vermab, "A review on brain tumor segmentation of MRI images", Magnetic Resonance Imaging, Elsevier, 61, 247–259, 2019.

[6] Jin Liu, Yi Pan, Min Li, Ziyue Chen, Lu Tang, Chengqian Lu, and Jianxin Wang, "Applications of Deep Learning to MRI ImagesW A Survey", Big Data Mining and Analytics, IEEE, pp 1– 18, Volume 1, March 2018,

[7] Alexander Selvikvag Lundervolda, Arvid Lundervold, "An overview of deep learning in medical imaging focusing on MRI", Z Med Phys, pp.102–127, Elsevier, 2019.

[8] Debnath Bhattacharyya, and Tai-hoon Kim, "Brain Tumor Detection Using MRI Image Analysis", Springer-Verlag Berlin Heidelberg, , CCIS 151, pp. 307–314, 2011.

[9] Gong E, Pauly JM, Wintermark M, Zaharchuk G. Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI. J Magn Reson Imaging. 48:330–40. doi: 10.1002/jmri.25970, 2018.

[10] El-Dahshan E-S A, Mohsen H M, Revett K, Salem A-B M. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. Exp Syst Appl, 41(11):5526–45, 2014

[11] Georgiadis P, Cavouras D, Kalatzis I, Daskalakis A, Kagadis G C, Sifaki K. Improving brain tumor characterization on MRI by probabilistic neural networks and nonlinear transformation of textural features. Comput Methods Prog Biomed 89(1):24–32, 2008.

[12] Nyul L G, Udupa J K, Zhang X., "New variants of a method of MRI scale standardization".IEEE Trans Med Imaging;19(2):143–50, 2000.

[13] Dong H, Yang G, Liu F, Mo Y, Guo Y., "Automatic brain tumor detection and segmentation using u-net based fully convolutional networks". Annual Conf. Medical Image Understanding and Analysis. Springer; p. 506–17, 2017.

[14] Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio, "Brain tumor segmentation with deep neural networks". Med Image Anal, 35:18–31, 2017.

[15] Cai W, Chen S, Zhang D. "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation". Pattern Recogn, 40(3):825–38, 2007.

[16] Hai-Miao Zhang, Bin Dong, "A Review on Deep Learning in Medical Image Reconstruction", Springer, 2020.

[17] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016).

[18] He, K., Zhang, X., Ren, S., Sun, J.: Identity mappings in deep residual networks. In: European Conference on Computer Vision, pp. 630–645 (2016).

[19] Prachi Jain, Syeda Zoya Kulsum, Rohan, Srivastava, Aryan Singh, Dr. Rekha PM, (2020), "Analysis of Deep Learning Algorithms for Breast Cancer Diagnosis on WBC", International Journal of Engineering Research & Technology (IJERT), Vol. 8, No. 15.

[20] Juanying Xie, Ran Liu, Joseph Luttrell and Chaoyang Zhang, (2019), "Deep Learning Based Analysis of Histopathological Images of Breast Cancer", Frontiers in Genetics, Vol. 10, No.80. DOI: 10.3389/fgene.2019.00080.

[21] Li J, Zhou Z, Dong J, Fu Y, Li Y, Luan Z, et al. (2021) Predicting breast cancer 5-year survival using machine learning: A systematic review. PLoS ONE 16(4): e0250370. https://doi.org/10.1371/ journal.pone.0250370.

[22] https://www.kaggle.com/kmader/siim-medical-images?select=overview.csv

[23] https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

[24] https://www.kaggle.com/kmader/mias-mammography