

MEDICAL IMAGE ANALYSIS: REVOLUTIONIZING DIAGNOSIS THROUGH DEEP LEARNING

Abstract

Medical imaging is a crucial process in medicine, involving the acquisition of internal organ images for diagnostic and research purposes. Medical image analysis aims to enhance clinical research and treatment effectiveness. Deep learning has transformed medical image analysis, with exceptional success in tasks such as image registration, segmentation, feature extraction, and classification. The availability of processing resources and the resurrection of deep convolutional neural networks are driving this breakthrough. Deep learning excels at detecting hidden patterns in images, assisting clinicians in making correct diagnoses. It has been shown to be extremely effective in detection of cancer, organ segmentation, classification of diseases, and computer-aided diagnosis. Numerous deep-learning methods have been developed and published for analyzing medical images, catering to various diagnostic needs. Overall, deep learning has significantly improved medical image analysis, paving the way for more accurate and efficient medical diagnoses and treatments. This chapter explores the various modalities, advancements, and applications of medical imaging, showcasing how this technology has become an indispensable tool for diagnosis, treatment planning, and monitoring of various medical conditions. From the early days of X-rays to the latest cutting-edge imaging techniques, we will delve into the evolution and impact of medical imaging on modern medicine.

Keywords - Medical Imaging, Deep Learning, Diagnostic, Clinical Research, Convolutional Neural Network, Computer-assisted Diagnosis.

Authors

Anum Kamal

Assistant Professor
Department of Computer Science and Engineering
Integral University
Lucknow, India.
anumkamal22@gmail.com

Dr. Faiyaz Ahamad

Associate Professor
Department of Computer Science and Engineering,
Integral University
Lucknow, India.
faiyaz.ahamad@yahoo.com

Dr. Aisha Kamal

Professor
Department of Bioengineering
Integral University
Lucknow, India.
aishakamal04@gmail.com

I. INTRODUCTION

A very crucial component of modern healthcare, medical image analysis, enables clinicians to diagnose diseases, monitor treatment progress, and make informed decisions about patient care. Deep learning, a subgroup of artificial intelligence, has revolutionized the field of medical image analysis by providing powerful tools for automated and accurate interpretation of medical images.

In recent years, various significant advancements have been made in the field of artificial intelligence (AI), particularly in machine learning (ML). Within AI, one notable area that has found practical applications is machine learning, with a particular focus on neural networks (NN). The progression of artificial neural networks has exhibited a pattern resembling sinusoidal waves [1]. After an initial surge of interest in the late 1950s and early 1960s, there was a prolonged period of limited activity until 1986 when James McClelland and David Rumelhart published their influential book [2], reigniting enthusiasm for neural network research. However, as the 20th century drew to a close, interest in neural networks waned once again. This decline can be attributed, in part, to the insufficient computing hardware required to effectively utilize neural network models due to their data-intensive nature [3]. Only within the past decade has there been a resurgence of interest in neural networks, resulting in the successful development of applications that address real-world problems. Among the various neural network architectures, deep neural networks (DNNs) have garnered significant attention and have been applied in diverse fields such as medical image classification, electromyography recognition, disease identification, segmentation, and more.

In CAD (Computer-aided diagnosis), which is one of the most pivotal research domains within medical imaging, machine learning algorithms are frequently employed to analyze historical patient imaging data and establish a model for evaluating a patient's condition [4]. This developed model aids healthcare professionals in swiftly making diagnostic decisions. The primary imaging modalities commonly used in medical applications encompass CT or computed tomography, X-ray, PET or positron emission tomography, magnetic resonance imaging (MRI), and ultrasound. The overarching objective of medical image processing is to enhance the comprehensibility of the depicted information [5]. The primary categories of medical image analysis include enhancement, registration, segmentation, classification, localization, and detection [6]. In the past, medical images were processed using basic techniques like thresholding, region growing, and edge tracing [7]. However, the proliferation of medical imaging data in terms of volume and complexity has driven the adoption of machine-learning methods in medical image analysis. Nevertheless, these conventional machine-learning approaches rely on manually crafted features, necessitating considerable manual effort in algorithm design [8]. These limitations have spurred the adoption of artificial neural networks (ANNs). Data availability and computational processing capabilities have aided in the growth of ANNs. [9].

Deep learning techniques, such as convolutional neural networks, have increased the potential for medical image processing automation. A convolutional neural network (CNN) is a form of neural network that deals with pixel values. By implementing linear mathematical concepts, CNN makes image classification more scalable to detect patterns inside an image. In contrast to traditional CNN architectures, which typically stacked convolutional layers

sequentially, contemporary designs like Inception, ResNet, and DenseNet introduce fresh and inventive strategies for constructing convolutional layers. These novel approaches aim to enhance the efficiency of the learning process [10].

Deep learning applications and related artificial intelligence (AI) models have the potential to significantly improve human lives in a relatively short amount of time [11]. Over time, medical image processing has evolved to encompass computer vision, pattern recognition, image mining, and machine learning in various aspects [12]. Deep learning has emerged as a powerful method to achieve high accuracy, opening up new avenues for medical image analysis [13]. Applications of deep learning in healthcare address a wide range of issues, including cancer screening [14]. Today, an immense volume of data from diverse sources such as radiological imaging, genomic sequences, and pathology imaging is available to healthcare professionals [15]. To harness this wealth of information, various imaging modalities are employed, including PET (positron emission tomography), X-ray, CT (computed tomography), fMRI (functional MRI), DTI (diffusion tensor imaging), and MRI (magnetic resonance imaging) [16,17].

Deep learning involves the discovery of patterns in data structures using neural networks composed of numerous convolutional nodes of artificial neurons [18,19]. An artificial neuron functions by taking multiple inputs, performing a calculation, and producing an output similar to a biological neuron [20-24]. This calculation typically consists of a linear combination followed by a nonlinear activation function [25]. Commonly used nonlinear activation functions in networks include the sigmoid function, ReLU (rectified linear unit) and its variants, and tanh (hyperbolic tangent) [26-30]. The origins of deep learning can be traced back to the work of Warren McCulloch and Walter Pitts in 1943, and it has seen significant advancements, including the ImageNet dataset in 2008, the backpropagation model in 1961, AlexNet in 2010, convolutional neural networks (CNNs) in 1978, and long short-term memory (LSTM) networks in 1996 [31].

In 2014, Google introduced GoogleNet, which emerged as the winner of the ILSVRC 2014 challenge. It made a significant stride in reducing the computational complexity of Convolutional Neural Networks (CNNs) by introducing inception modules [32]. CNN architecture comprises numerous layers that utilize differentiable functions to transform input data into output data, often representing class scores. Deep learning essentially rejuvenates the concept of artificial neural networks by stacking artificial neurons. In CNNs, network features are created by converting kernels into layers, utilizing outputs from preceding layers. The kernels in the initial hidden layer perform convolutions on input images, with early layers capturing basic shapes, curves, and edges. In contrast, deeper layers extract more abstract and intricate features [33,34]. Refer to Figure 1 for an illustration of the applications of neural networks in the context of Medical Image Analysis.

This chapter offers an overview of the fundamental principles, methods, and practical applications of deep learning within the field of medical image analysis. It commences with an introduction to the core concepts of deep learning, including neural networks and convolutional neural networks (CNNs), which are the foundational components of numerous cutting-edge algorithms used in medical image analysis. Substantial attention is given to the utilization of deep learning techniques in this domain, encompassing topics such as the segmentation of images for organ and lesion delineation, disease classification and diagnosis,

image registration for the fusion of multi-modal data, and the employment of generative models to enhance data augmentation and synthesis. To illustrate the successful application of deep learning, the chapter incorporates real-world examples and case studies.

Additionally, it addresses the challenges and constraints associated with deploying deep learning in medical image analysis, such as limited data availability, interpretability issues, and ethical considerations. Regulatory aspects and the critical need to ensure the safety and dependability of medical image analysis systems founded on deep learning are also discussed. In conclusion, this chapter serves as a comprehensive resource for researchers, practitioners, and healthcare professionals interested in harnessing deep learning methodologies to advance medical image analysis. It underscores the transformative potential of deep learning in improving the precision and efficiency of medical diagnoses and treatments, ultimately contributing to better patient outcomes and the advancement of the medical imaging profession.

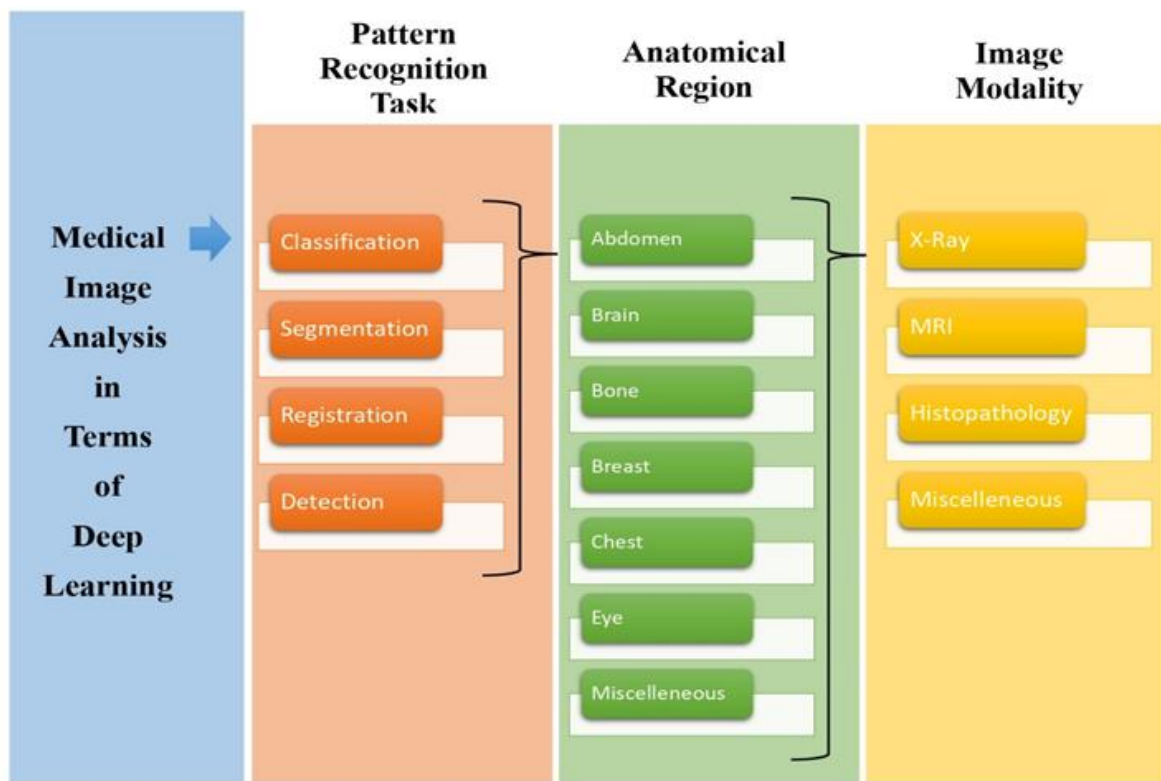


Figure 1: Application of Deep Neural Network in terms of Medical Image Analysis

II. LEARNING TYPES

This section represents 14 types of learning divided into 4 subcategories (**Figure 2**).

1. Learning Problems

- **Supervised Learning:** Supervised learning is a method of learning the relationship between input examples and a target variable [35]. In supervised learning problems, the training data consists of input vectors and their corresponding target vectors. There are two primary categories of supervised learning problems: classification and regression [36].
 - **Classification:** In classification problems, the goal is to predict a class label. For example, classifying an email as spam or not spam is a classification task.
 - **Regression:** In regression problems, the objective is to predict a numerical value. For instance, predicting the price of a house based on its features is a regression problem [37].

Supervised learning tasks may encompass one or multiple input variables, with the input data taking on diverse formats, such as numerical and categorical data [37]. Take, for instance, the MNIST dataset, comprising images of handwritten digits as inputs (encoded as pixel data); it serves as a well-known illustration of a classification problem. In this scenario, the objective is to categorize each image into one of the ten available digits [38].

Several machine learning algorithms are specifically designed for supervised learning tasks. Some examples include:

- **Decision Trees:** Decision trees are a type of supervised learning algorithm employed in both classification and regression tasks. They create a structure like a tree for making decisions based on input features [38, 39].
 - **Support Vector Machines (SVMs):** SVMs are another supervised learning method used for classification and regression. They aim to find a hyperplane that best separates different classes or predicts numerical values [38, 39]. These algorithms are referred to as "supervised" because they learn from input data by making predictions and are fine-tuned and improved through an iterative process that guides their performance [40]. Some methods are well-suited for specific tasks, such as logistic regression for classification or linear regression for regression. Others, like artificial neural networks, can be adapted for both types of problems with minor modifications [41, 43].
- **Unsupervised Learning:** Unsupervised learning addresses challenges related to modeling and understanding data relationships without the presence of explicit data relationships or target variables, setting it apart from supervised learning [42]. Unlike supervised learning, unsupervised learning solely relies on input data, without any

predefined outputs or target values. Consequently, unsupervised learning lacks a guiding instructor to correct and guide the model's learning process.

Unsupervised learning encompasses several methods, but two primary challenges commonly encountered by practitioners are clustering and density estimation:

- **Clustering:** Clustering is an unsupervised learning task that involves grouping data points based on their inherent similarities or patterns. Discovering clusters or natural groupings within the data is the main objective [44].
- **Density Estimation:** Density estimation is another unsupervised learning task that aims to summarize and model the distribution of data points in the dataset. It involves estimating how data points are distributed in the feature space [45]. One well-known clustering technique is K-Means, where "k" represents the number of cluster centers to be identified within the data. Kernel Density Estimation is a density estimation technique that uses small groups of closely related data samples to estimate the distribution of new data points in the problem space [46]. Clustering and density estimation techniques are valuable for uncovering underlying trends and patterns in data.

Additionally, unsupervised learning encompasses other approaches, such as data visualization and dimensionality reduction:

- **Data Visualization:** Data visualization is an unsupervised learning task that involves creating visual representations of data to aid in understanding patterns and relationships within large datasets. It helps highlight trends, associations, and patterns that might be challenging to discern through raw data alone [47].
- **Dimensionality Reduction:** Dimensionality reduction is another unsupervised learning task that focuses on creating lower-dimensional representations of data while preserving essential information. Random projection is an efficient dimensionality reduction technique, especially in cases where datasets have too many dimensions for direct principal component analysis [48].

These unsupervised learning techniques play a crucial role in exploring and making sense of complex data without the need for explicit labels or target values, enabling practitioners to uncover hidden structures and insights within the data.

- **Reinforcement Learning:** Reinforcement learning is a collection of problems in which an agent needs to learn how to navigate and make decisions within a specific environment by utilizing feedback [49]. While it shares similarities with supervised learning, the key distinction lies in the nature of the feedback, which can be delayed and noisy. In reinforcement learning, the agent receives feedback intermittently, and because the environment is inherently stochastic, it can be challenging for the agent to establish a clear causal relationship between its actions and the outcomes it observes [50].

Some well-known examples of reinforcement learning algorithms include:

- **Deep Reinforcement Learning:** This approach combines deep learning techniques with reinforcement learning principles to enable agents to learn complex decision-making tasks, often involving high-dimensional input data.
- **Q-Learning:** Q-learning is a reinforcement learning algorithm that is model-free. It is used for estimating the value of taking a specific action in a particular state. It helps agents learn optimal policies by iteratively updating Q-values based on observed rewards and actions.
- **Temporal-Difference Learning:** Temporal-difference learning is a reinforcement learning method that focuses on estimating the expected future rewards an agent can obtain. It involves updating value functions based on the difference between current and predicted future values.

Reinforcement learning is particularly useful in scenarios where an agent must learn how to interact with an environment to achieve specific goals, even when the environment is uncertain and dynamic. It has applications in various fields, including robotics, game playing, autonomous systems, and more, where agents learn to make sequential decisions through trial and error.

2. Hybrid Learning Problems

- **Semi-supervised Learning:** The concept you're describing is known as semi-supervised learning. In semi-supervised learning, the training dataset contains only a small portion of labeled instances, while the majority of the data remains unlabeled [51]. The goal is to make efficient use of this limited labeled data while leveraging the vast amount of unlabeled data available [52].

Semi-supervised learning approaches can employ various techniques to effectively utilize the unlabeled data. This can include the use of unsupervised methods like clustering and density estimation to discover patterns or groups within the data [52, 53]. Once these patterns or groups are identified, supervised learning strategies or techniques can be applied to assign labels to the previously unlabeled instances or augment the dataset with additional labels for more accurate predictions [54].

Semi-supervised learning is particularly relevant in scenarios where obtaining labeled data is expensive or time-consuming. It has applications in various domains, including:

- **Audio Data:** For tasks like automated speech recognition, where labeled data is limited but large amounts of unlabeled audio data are available.

- **Text Data:** In natural language processing, semi-supervised learning can be used to improve tasks like sentiment analysis or text classification when labeled text data is scarce.
- **Image Data:** In computer vision, when dealing with image data, semi-supervised learning methods can help in scenarios where labeling a large number of images is impractical [43, 54].

Semi-supervised learning bridges the gap between supervised and unsupervised learning, allowing for more efficient utilization of data resources when labeled examples are limited.

- **Multi-Instance Learning:** In multi-instance learning (MIL), a set of examples is labeled as either containing at least one instance of a class or not containing any instances of the class, but the individual members or instances within that collection are not individually marked or labeled [55]. This differs from traditional supervised learning, where each individual data point is associated with a specific label. In MIL, the focus is on learning patterns or relationships within collections of instances, making it particularly applicable in scenarios where precise instance-level labeling is challenging or expensive.
- **Self-Supervised Learning:** Self-supervised learning is a machine learning paradigm that uses unlabeled data to build a pretext learning problem, such as context prediction or picture rotation, for which a target objective may be determined without the requirement for external supervision [56]. Self-supervised learning algorithms, like autoencoders, exemplify this approach.
 - **Autoencoders:** Autoencoders represent a type of neural network model employed to create a condensed representation of input data [56, 57]. They consist of two primary components: an encoder and a decoder, with a bottleneck layer situated between them, which serves as the internal compact representation of the input data [58]. In the training phase, autoencoders are presented with the input data both as the input and the target output. This compels the model to learn the process of encoding the input into a compressed representation and subsequently decoding it to reconstruct the original input data [57]. Following the training process, the decoder component is typically no longer required, and the encoder is used to generate compact input representations whenever necessary. Historically, autoencoders have found utility in tasks like reducing dimensionality and learning distinctive features from data [58].
 - **Generative Adversarial Networks (GANs):** Self-supervised learning is also exemplified through generative adversarial networks, commonly known as GANs [58, 59]. GANs are generative models often employed to produce synthetic data, such as images, using only an assortment of unlabeled instances from the intended domain [59]. GANs comprise two key components: a generator and a discriminator. These components are trained in a competitive manner, where the generator's goal is to produce realistic data samples, while the discriminator's task

is to differentiate between real and generated data. This adversarial training process ultimately results in the creation of high-quality synthetic data.

Self-supervised learning, including methods like autoencoders and GANs, has gained prominence for its ability to leverage unlabeled data effectively and learn meaningful representations from it, making it a valuable technique for various machine learning tasks.

3. Statistical Inference

- **Transductive Learning:** In statistical learning theory, transduction, or transductive learning, refers to the process of predicting specific examples from a given domain [60]. It differs from induction, which is learning general laws from observable cases [61]. In transduction, the focus is on making predictions for specific instances within the domain.

Transduction comes into play when the objective is to estimate the value of a function at a particular point of interest, and it becomes particularly relevant when one seeks to obtain the best possible outcomes with limited knowledge or data [62].

A classic example of a transductive algorithm is the k-nearest neighbors (k-NN) algorithm. In k-NN, the transductive approach involves using the algorithm directly whenever a prediction is needed, without creating a formal model of the training data. Instead, it relies on the proximity or similarity of the nearest neighbors to make predictions for specific instances [50, 63].

In essence, transductive learning focuses on making predictions for specific cases within the existing dataset, rather than learning general rules that can be applied universally to unseen data. It can be particularly useful in scenarios where the goal is to make specific predictions based on the available data without building a comprehensive model of the entire dataset.

- **Inductive Learning:** Inductive learning involves the use of generalization and proof to assess results. In the context of machine learning, inductive learning refers to the process of using specific instances, often moving from specific to general, to make general predictions or conclusions [64]. Many machine learning algorithms employ inductive reasoning, where they learn from specific past instances to derive general rules or a model [64, 65]. It's essentially an induction approach applied to machine learning models.

In the context of machine learning:

- The model is a generalization of the concrete examples present in the training dataset.
- The training data is utilized to build a model or hypothesis about the problem.
- This model is then assumed to be capable of making accurate predictions or generalizations on new, unseen data in the future [65].

Inductive learning is a fundamental concept in machine learning, as it involves the process of deriving general principles or rules from specific examples, enabling the model to make predictions or decisions about new, unseen data based on what it has learned from the training data.

- **Deductive Learning:** Deduction, also known as deductive inference, entails employing broad, general concepts to assess particular and concrete outcomes. To gain a clearer grasp of induction, it's beneficial to distinguish it from deduction. Deduction represents the antithesis of induction [66]. Whereas induction advances from particular instances to form general conclusions, deduction moves from overarching principles or premises to specific results [67]. Induction constitutes a bottom-up style of reasoning, utilizing existing evidence to substantiate a conclusion, whereas deduction takes the top-down approach, aiming to satisfy all premises before reaching a specific outcome [68].

In the context of machine learning:

- Induction corresponds to the process of building a model or hypothesis from specific examples in the training dataset, moving from particular instances to general principles.
- Deduction, on the other hand, involves using the established model to make predictions or draw specific conclusions based on the general principles encoded in the model [69].

In this sense, the machine learning model serves as a deductive method, where it applies the general rules it has learned to make predictions or decisions for specific instances. Deduction is a crucial aspect of machine learning, allowing models to use the knowledge gained during the induction phase to new, unseen data to make specific predictions or inferences.

4. Learning Techniques

- **Active Learning:** Active learning is a methodology where a machine learning model interacts with a human user or operator by posing questions during the learning process to address uncertainty [70]. It falls within the realm of supervised learning and seeks to achieve results that are on par with or better than passive supervised learning, even when the model has access to limited or more efficient data [71]. At the heart of active learning is the fundamental principle that when a machine learning algorithm is allowed to choose the data from which it learns, it can attain higher accuracy with fewer training labels [72]. In active learning, the model takes an active role in asking questions, typically by presenting unlabeled data instances, which are subsequently labeled by an oracle, often a human annotator [73].

Active learning is especially valuable in situations where there is a scarcity of data, and acquiring or labeling new data is costly or resource-intensive [74]. The active learning process enables strategic sampling of data from the domain, reducing the number of required samples while enhancing the model's effectiveness [75]. It is a

powerful approach for optimizing the use of resources and improving the efficiency of the machine learning training process.

- **Transfer Learning:** Transfer learning is a learning approach in which a model acquires knowledge or expertise from one problem and then applies that learned knowledge as a reference point for another, related task [76]. This approach is particularly useful in situations where there is a process closely related to the main problem, and the associated task requires a substantial amount of data [77].

Key characteristics of transfer learning include:

- **Sequential Learning:** In transfer learning, the model learns sequentially, first on one task and then on another. It leverages the knowledge gained from the initial task to improve performance on the subsequent task.
- **Reuse of Model Weights:** During transfer learning, model weights learned from one task can be used as a starting point or initialization for another task. This reuse of previously learned knowledge helps accelerate training and improve performance on the new task.

For example, consider the task of image classification. A prediction model, such as an artificial neural network, can be initially trained on a large dataset containing a wide range of images. Once the model has learned general features like lines and patterns from this broader task, it can then be fine-tuned on a simpler, more specific dataset, such as images of cats and dogs. The knowledge and features that the model has already acquired during the larger mission, such as recognizing lines and patterns, can be beneficial for the subsequent task of distinguishing between cats and dogs [76].

Transfer learning is a valuable technique in machine learning, allowing models to leverage pre-existing knowledge and expertise to improve their performance on new, related tasks, often with less data and training effort.

- **Multi-Task Learning:** Multi-task learning is a machine learning technique aimed at enhancing the generalization of models by incorporating information from multiple tasks, which can be viewed as soft constraints applied to the model's parameters [78]. This approach proves particularly valuable when there is an abundance of labeled input data available for one task that can be shared with another task that has limited labeled data [79, 80].

Key characteristics of multi-task learning include:

- **Shared Input Patterns:** In multi-task learning, the same input patterns are utilized for multiple different outputs or supervised learning tasks [81]. This means that the model processes the same input data to predict various outputs simultaneously.
- **Task-Specific Predictions:** Different parts of the model are dedicated to predicting each of the tasks. While the core of the model processes the input data,

task-specific modules or branches are responsible for making predictions for each task [80].

By jointly training on multiple tasks, multi-task learning allows the model to leverage shared information and improve generalization across all tasks. It enables the model to learn from the abundant data available for one task and apply that knowledge to improve performance on other related tasks with fewer labeled examples. This approach is particularly valuable in scenarios where tasks have shared underlying patterns or dependencies.

- **Online Learning:** Online learning is a machine learning approach that differs from traditional offline learning, where a batch of data is used to refine a model equation [82]. In online learning, the model is updated incrementally as each new data point arrives, rather than waiting until the end of a batch (which may never occur) [83]. This approach is particularly valuable when dealing with streaming data, where the data can change rapidly over time [84–86].

Key points about online learning include:

- **Streaming Data:** Online learning is essential when dealing with data streams, where new data points continuously flow in, and waiting for a complete batch update is impractical.
- **Continuous Knowledge Expansion:** It is valuable for applications that involve a continuously expanding knowledge base, even if changes are incremental [87].
- **Performance Comparison:** Online learning aims to minimize the discrepancy between how well the model is performing and how well it would have performed if it had access to all the available knowledge as a batch [88].

One common technique used in online learning is stochastic or online gradient descent, particularly when training artificial neural networks [89]. In this context, stochastic gradient descent helps minimize generalization error, especially when examples or mini-batches are drawn from a data stream, as seen in the online training scenario [90].

Online learning is crucial in scenarios where timely updates and adaptation to changing data are essential, such as real-time analytics, online recommendation systems, and monitoring evolving trends.

- **Ensemble Learning:** Ensemble learning is a machine learning technique in which at least two models are trained on similar data and their predictions are combined to make a final prediction or decision [91]. Unlike individual models, the primary goal of ensemble learning is to achieve better overall performance by leveraging the collective knowledge of the models in the ensemble [92].

Key characteristics of ensemble learning include:

- **Multiple Models:** Ensemble learning involves training and combining predictions from multiple models, often using the same dataset.
- **Improved Performance:** The aim is to enhance the overall performance compared to what any individual model can achieve on its own.
- **Model Diversity:** Ensembles benefit from diversity in the models used, as it helps capture different aspects of the data and reduces the risk of overfitting [93–96].

Ensemble learning is a valuable technique for improving predictive capabilities and reducing the uncertainty associated with stochastic learning algorithms, such as artificial neural networks. Common ensemble learning algorithms include Bootstrap, weighted averaging, and stacking (stacked ensemble) [97]. By combining the strengths of multiple models, ensemble learning can lead to more robust and accurate predictions in various machine learning applications.

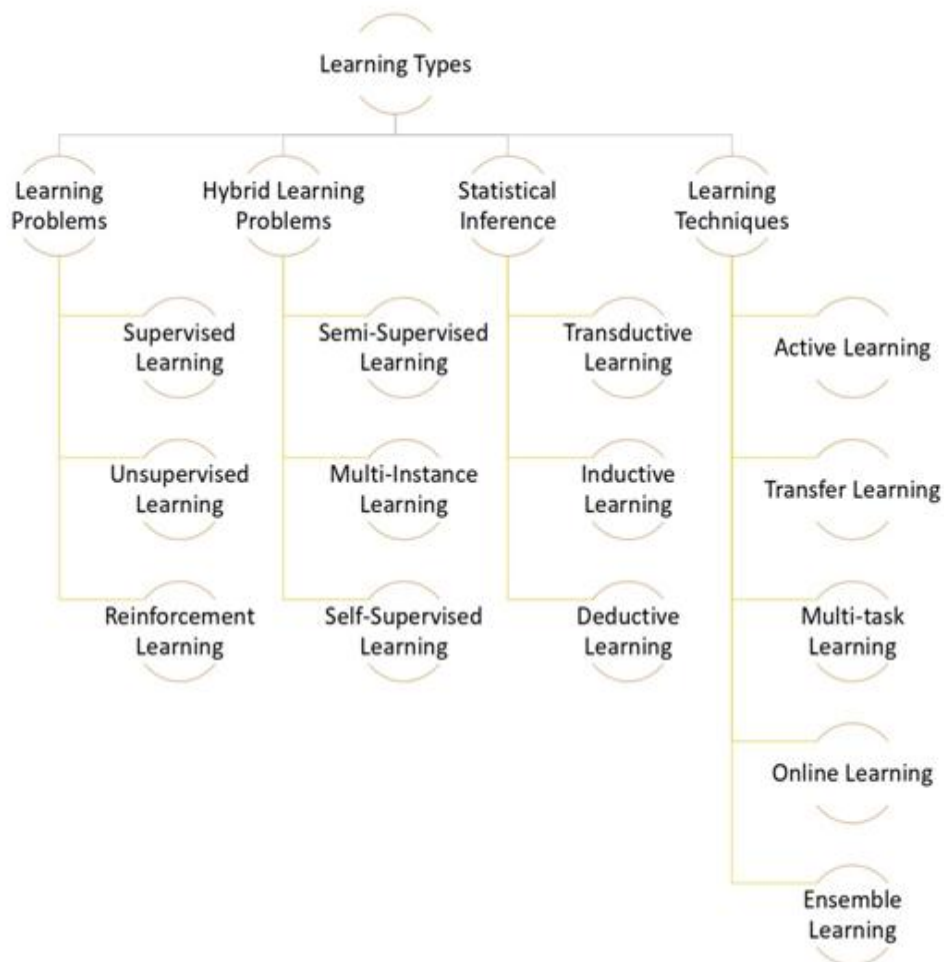


Figure 2: Learning Types

III. SUMMARY OF DATA SETS OF MEDICAL IMAGES USED IN MEDICAL ANALYSIS

In the domain of medical imaging, neural networks, regardless of their architectural variations, rely heavily on extensive datasets for effective training. A significant challenge in medical imaging is the need for substantial training data, and this compilation aims to address this need. **Table 1**, represents a summary of datasets that encompass various organs and imaging modalities, catering to diverse medical applications. These datasets serve as a valuable resource for deep neural network (DNN) training, with detailed information about the specific modalities they represent.

One notable application discussed in this chapter is the machine learning and image processing-based approach to categorizing and identifying breast cancer. This approach integrates a comprehensive pipeline, including image preprocessing, feature extraction using AlexNet, feature selection using the relief algorithm, and the utilization of various machine learning methods such as least square support vector machines, K-nearest neighbors, random forests, and Naive Bayes for illness classification and detection. The inclusion of the geometric mean filter further enhances image quality. The datasets used in this study, particularly data from the MIAS dataset, contribute significantly to the accurate diagnosis of breast cancer, showcasing the benefits of this proposed method.

The datasets included in this compilation have been carefully selected and prepared by clinical experts, offering anonymized medical images of patients. Several of these datasets have been central to various challenges organized to assess the effectiveness of automated classification, detection, or segmentation algorithms, ultimately reducing the time required for medical diagnoses. Transfer Learning (TL) architectures have emerged as a powerful tool in automated medical image processing, with applications spanning segmentation, object recognition, illness classification, severity grading, and more. TL models leverage their prior training on large generic datasets to adapt to specific medical imaging tasks, offering high-quality decision support while requiring less task-specific training data.

Furthermore, the chapter highlights the significance of information fusion in medical imaging, where data from multiple sources and modalities contribute to a holistic understanding. Image registration, the process of aligning multiple images, plays a pivotal role in (semi-)automatic medical image analysis. While traditional methods based on image intensity and manually crafted characteristics have been prevalent, recent advancements have seen the integration of supervised and unsupervised deep learning algorithms for image registration, enhancing the accuracy of information transfer across medical images. These developments reflect the ongoing evolution and innovation in the field of medical imaging, driven by the convergence of machine learning and advanced image processing techniques [93,94].

Table 1: Summary of the medical datasets.

S. No.	Ref.	Datasets	Application	Modality
1.	[98]	CHAOS	Abdominal	CT and MRI
2.	[99]	TADPOLE	Alzheimer’s Disease	MR and PET
3.	[100]	MRBrainS18	Brain	3T MRI Scans
4.	[101]	MUSHAC	Brain	MRI
5.	[102]	MITOSTAPIA	Breast Cancer	H and E stained
6.	[103]	DREAM	Breast Cancer	Digital mammograms
7.	[104]	Rсна-bone-age	Child age prediction	X-ray
8.	[105]	MIMIC-CXR	Chest X-ray	X-ray
9.	[106]	ECG Arrhythmia	Cardiac Diagnosis	ECG
10.	[107]	IDRID	Diabetic Retinopathy	Fundus Images
11.	[108]	18F-FDG PET	Head and Neck	PET Scans
12.	[109]	MRNet-v1.0	Knee	MRI
13.	[110]	LUNA	Lung Cancer	CT Images
14.	[111]	LIDC-IDRI	Lung Cancer	CT Images
15.	[112]	CLUST	Liver	Ultrasound
16.	[113]	IVDM3S	Low back pain	MRI
17.	[114]	MTOP	Brain Injury	MRI
18.	[115]	VFA, DXA	Vertebral Fracture Analysis	X-ray
19.	[116]	Warwick-QU	Tumors	H and E stained
20.	[117]	KITS	Kidney	CT Scan
21.	[118]	ISCI	Skin Cancer	Dermo.S.
22.	[119]	AIDA	Gastroenterological Diseases	Endoscopy
23.	[120]	RETOU	Retinal	OCT Scans

IV. STRUCTURES OF DEEP LEARNING

Over the past two decades, we have witnessed a remarkable evolution in deep learning models, which has greatly expanded the scope of problems that can be effectively addressed using neural networks. Deep learning is not a single technique but rather a versatile class of algorithms and architectures that can be applied to a wide variety of problem domains. Although the concept of connectionist structures, which underpin neural networks, has been around for over 70 years, it has been propelled to the forefront of artificial intelligence through modern architectural innovations and the utilization of graphical processing units (GPUs). Figure 3 provides an overview of the general architecture of neural networks.

The swift advancement of deep learning techniques can be credited to the synergy of deeply layered neural networks and the adoption of GPUs to expedite their computations. In this chapter, we delve into a comparative examination of diverse deep learning architectures. A typical deep learning architecture comprises a range of layers, each designed for specific tasks. These layers encompass input layers, convolutional and fully connected layers,

sequence layers, activation layers, normalization, dropout, cropping layers, pooling and unpooling layers, combination layers, object detection layers, generative adversarial network layers, and output layers.

The true essence of a neural network lies in its hidden layer(s), which can effectively model complex data patterns. These hidden layers consist of nodes or neurons whose actual values remain unknown during training, as only the input and output values are accessible. While every neural network includes at least one hidden layer, there is no strict rule mandating the multiplication of hidden units by the number of inputs. The ideal number of hidden units can often be less than the number of inputs. In situations where ample training examples are available, multiple hidden units may be employed, but in scenarios with limited data, as few as two hidden units can suffice for effective learning. The adaptability of neural networks, coupled with their capacity to accommodate the intricacies of the problem at hand, renders them a potent tool in the realm of deep learning [95].

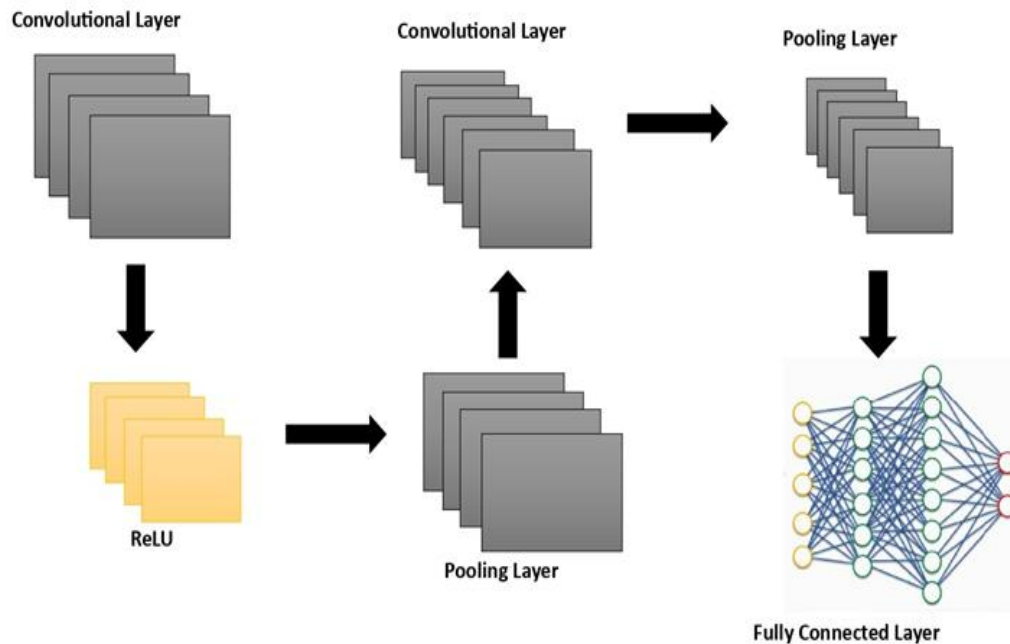


Figure 3: General Convolutional Neural Network Architecture (CNN)

- 1. Deep Neural Network:** A Deep Neural Network (DNN) represents a type of machine learning model that draws inspiration from the structure and functioning of the human brain. It is also commonly referred to as a neural network or artificial neural network. DNNs serve as a foundational element in deep learning, a branch of machine learning that concentrates on training models with multiple layers, often termed "deep" networks, to execute a wide array of tasks, including image recognition, natural language processing, and various others.

This architecture typically incorporates a minimum of two layers that facilitate the handling of nonlinear complexities, making it suitable for tasks such as classification and regression. Its widespread usage is primarily attributed to its high accuracy in solving

complex problems [121]. However, it does come with its set of drawbacks. Training this model can be challenging because errors are propagated backward through the layers, which can result in slow convergence. Additionally, the learning behavior of the model may exhibit delayed improvements, posing certain limitations [122].

- 2. Convolutional Neural Network:** A Convolutional Neural Network (CNN) is a specialized deep learning model tailored for tasks like image classification and computer vision. CNNs consist of layers that perform convolution operations to extract local patterns, apply activation functions for non-linearity, use pooling to reduce spatial dimensions and employ fully connected layers for high-level feature learning. During training, CNNs adjust their weights to make accurate predictions. Popular CNN architectures, such as AlexNet and ResNet, have revolutionized computer vision tasks. These networks excel at recognizing intricate patterns in grid-like data and are crucial in various domains beyond image analysis.

This particular model is most suitable for handling 2D data. It comprises a convolutional filter that efficiently transforms 2D data into 3D, demonstrating robust performance and rapid learning capabilities. However, it should be noted that this model requires a substantial amount of labeled data for effective classification [123]. CNNs, like this one, do encounter some challenges, including issues related to local minima, relatively slow convergence rates, and the potential for significant human interference. Following the notable success of AlexNet in 2012, CNNs have increasingly been employed to enhance the efficiency of human clinicians in the field of medical image processing [124].

- 3. Recurrent Neural Network:** A Recurrent Neural Network (RNN) is a type of artificial neural network explicitly designed for processing sequential data. Unlike conventional feedforward neural networks, RNNs possess connections that loop back on themselves, allowing them to retain the memory of prior inputs. This inherent recurrent structure makes RNNs particularly well-suited for tasks that involve sequences, such as natural language processing, speech recognition, and time series prediction. Nonetheless, RNNs come with certain limitations, including the challenge of the vanishing gradient problem, which can impede their ability to capture long-range dependencies in data [125]. In response to these limitations, more advanced RNN variants like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks have been developed, offering enhanced performance in managing sequential data.

Recurrent Neural Networks (RNNs) possess the unique capability to recognize sequences. They achieve this by distributing weights of the neurons across all time steps, allowing them to process information in a sequential manner. RNNs come in various variants, including LSTM, BLSTM, MDLSTM, and HLSTM, each with its own strengths and applications [126]. These models have demonstrated state-of-the-art accuracy in tasks such as character recognition, speech recognition, and various natural language processing-related problems, as they are well-suited for learning sequential patterns and temporal dependencies [127].

- 4. DC-ELM - Deep conventional-extreme Learning Machine:** Deep Conventional-Extreme Learning Machine (Deep C-ELM) is an advanced machine learning approach that combines elements of deep learning and extreme learning machines. It aims to leverage the strengths of both techniques to enhance the efficiency and effectiveness of neural network models. In Deep C-ELM, the initial layers of the neural network are shallow and randomly initialized, similar to the extreme learning machine concept. These layers act as feature extractors, transforming the input data into a higher-level representation. Then, deeper layers are added, which can be fine-tuned using traditional gradient-based methods like backpropagation. This combination allows Deep C-ELM to capture complex hierarchical features while benefiting from the speed and simplicity of extreme learning machines

To effectively extract high-level features from input images, it employs multiple alternating convolution layers and pooling layers [128]. These extracted features are then fed into an ELM classifier, which leads to improved results with faster learning speeds [12]. In the final hidden layer, a deep conventional-extreme learning machine is utilized to implement stochastic pooling, significantly reducing the dimensionality of the feature space. This reduction in dimensionality helps save valuable training time and computational resources [125]. It has found applications in various fields, including image and speech recognition, where it offers improved performance and faster training times compared to traditional deep learning architectures.

- 5. DBM - Deep Boltzmann Machine:** A Deep Boltzmann Machine (DBM) is a complex neural network model used for unsupervised learning and generative tasks. It comprises multiple layers of stochastic binary units connected by weighted edges, allowing it to capture intricate data patterns and hierarchies. DBMs are trained using Markov Chain Monte Carlo methods, making them proficient at tasks like image generation and feature learning.

A Deep Boltzmann Machine (DBM) is a generative model consisting of three layers. It shares similarities with a deep belief network but differs in that it allows bidirectional connections in the bottom layers. In a DBM with N hidden layers, unidirectional connections are established among all hidden layers. The top-down feedback mechanism is employed to enhance the accuracy of inference by integrating ambiguous results [129]. However, optimizing the parameters of a DBM can be challenging, particularly when dealing with large datasets. Their training complexity has led to the popularity of simpler deep learning models for many applications, while DBMs remain valuable for specific tasks that demand unsupervised learning and generative modeling.

- 6. DBN – Deep Belief Networks:** Deep Belief Networks (DBNs) are a type of neural network architecture used in machine learning and deep learning. They consist of multiple layers of hidden variables and are trained in a layer-wise fashion, starting with Restricted Boltzmann Machines (RBMs). DBNs have been employed for various tasks, including image and speech recognition, by learning hierarchical representations from data. While they played a significant role in the early development of deep learning, more recent architectures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have become more prevalent for specific applications due to their ease

of training and effectiveness. Nevertheless, DBNs remain a fundamental concept in the history of deep learning.

Deep Belief Networks (DBNs) are a graphical model primarily used for generative purposes. They combine elements of probability and statistics with neural networks and artificial intelligence [126]. DBNs consist of multiple layers with nodes, where the layers are interconnected but not the nodes themselves. The primary objective is to help the machine classify data into different categories. However, one drawback of this architecture is that the initialization process can make training computationally expensive [126].

7. **Deep Autoencoders:** Deep autoencoders are a class of artificial neural networks used for unsupervised learning and feature extraction. These networks consist of an encoder and a decoder, both composed of multiple layers, creating a deep architecture. The encoder reduces the size of the input data to a lower-dimensional representation, and the decoder reconstructs the original input from this compressed representation. Deep autoencoders are particularly useful for dimensionality reduction, data denoising, and feature learning tasks. They can capture complex patterns in data by learning hierarchical representations through their deep structure. Variants like stacked autoencoders and denoising autoencoders have been employed for various applications in machine learning, including image and text data, where they have proven effective for feature extraction and data compression.

Autoencoders are particularly useful in unsupervised learning processes, aiding in dimensionality reduction and feature extraction. In this type of model, the number of inputs is equal to the number of outputs, making it suitable for various applications [130]. One of the significant advantages of autoencoders is that they don't require labeled data. Different types of autoencoders, including denoising autoencoders, sparse autoencoders, and conventional autoencoders, are used for various purposes such as enhancing robustness [12-14]. While they require a pre-training step, training can still be carried out efficiently.

8. **DSN – Deep Stacking Networks:** Deep Stacking Networks (DSNs) represent a powerful fusion of deep neural networks and ensemble learning. These architectures stack multiple layers of neural networks to create a hierarchical structure for feature extraction and prediction. Each layer progressively captures more abstract and refined features from the data, enhancing the model's ability to handle complex tasks like image recognition and natural language processing. While DSNs offer improved performance, they demand substantial computational resources and extensive data for effective training, making them well-suited for challenging machine learning problems where accuracy and robustness are paramount.

A deep stacking network, sometimes referred to as a deep convex network, represents a distinct architecture in the realm of deep learning [16]. Unlike conventional deep learning systems where there is a single deep network, a deep stacking network is essentially a collection of individual networks, each with its hidden layers, even though it forms a deep network when combined. This architectural model addresses one of the challenges in deep learning: the difficulty of training [17]. Training becomes significantly

more complex with each layer in a deep learning design. Therefore, the deep stacking network approaches training as a series of individual training problems rather than a single, overarching issue [18].

9. **LSTM/GRU – Long short-term Memory/Gated Recurrent Unit Network** : Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are both specialized variants of recurrent neural networks (RNNs) designed to address the vanishing gradient problem and improve the modeling of long-range dependencies in sequential data.

The Gated Recurrent Unit network, commonly referred to as GRU, was initially introduced by Hochreiter and Schmidhuber in 1997. However, it has gained significant popularity in recent years as a prominent RNN architecture for various applications [13,14]. In contrast to conventional neuron-based neural network models, LSTM (Long Short-Term Memory) introduced the concept of a memory cell [14]. This memory cell can retain information for a short or extended period, allowing it to remember what's important, not just the most recent input [16]. In 2014, a variant of LSTM called the gated recurrent unit (GRU) was introduced. This model incorporates two gates while omitting the output gate found in the LSTM model [17]. For specific applications, the GRU demonstrates performance similar to LSTM but with the advantage of being simpler, requiring fewer parameters, and running faster. The GRU integrates two critical gates: the update gate, which determines how much of the previous cell content to preserve, and the reset gate, which governs how to combine the new input with the previous cell content. By setting the reset gate to 1 and the update gate to 0, a GRU can emulate a standard RNN. This adaptability renders GRUs versatile and suitable for a broad spectrum of tasks in sequence modeling and natural language processing, among others.

Both LSTM and GRU networks have significantly advanced the field of deep learning for sequential data, offering solutions to the challenges posed by traditional RNNs and enabling more accurate and efficient modeling of temporal data patterns. The choice between them often depends on the specific task, available computational resources, and the trade-off between model complexity and performance.

V. IMPLEMENTING DEEP LEARNING IN MEDICAL IMAGE ANALYSIS

1. **Image Classification:** Image classification in medical image analysis is a pivotal task that involves categorizing medical images into predefined classes or diagnoses. This application is particularly valuable for automating the process of identifying diseases or conditions within medical images, such as X-rays, MRI scans, or CT scans. Deep learning techniques, notably Convolutional Neural Networks (CNNs), have revolutionized image classification in the medical field. By training these networks on large datasets of labeled medical images, they can learn to distinguish between different pathologies, making them indispensable tools for automated diagnosis. Notable achievements include studies demonstrating deep learning's ability to detect skin cancer with accuracy comparable to dermatologists and the development of AI systems capable of identifying various eye diseases from retinal scans. One of the primary challenges is the availability of annotated medical images for training deep learning models. Ensuring that models generalize well across different patient populations and image acquisition settings is essential. Deep

learning models often lack transparency, making it difficult for clinicians to understand the reasoning behind a particular diagnosis.

- 2. Image Segmentation:** Image segmentation in medical image analysis involves the precise delineation of regions or objects within an image. This application is essential for identifying and isolating specific structures or abnormalities within medical images, such as tumors, organs, or blood vessels. Deep learning techniques have significantly advanced image segmentation tasks in the medical field. By leveraging neural networks, particularly convolutional and semantic segmentation networks, accurate and automated segmentation can be achieved. Successful applications include the segmentation of brain tumors in MRI scans, where deep learning models enable surgeons to plan interventions more effectively, and the segmentation of cardiac structures in MRI, crucial for diagnosing and treating heart conditions. Achieving pixel-level accuracy, especially in complex structures like blood vessels, remains challenging. Variations in imaging techniques and patient anatomy make it difficult to create models that work universally.
- 3. Object Detection:** Object detection in medical image analysis focuses on locating and classifying specific objects or anomalies within medical images. This application is vital for tasks such as identifying tumors, lesions, or organs of interest. Deep learning-based object detection techniques have proven highly effective in enhancing the accuracy of disease identification and localization within medical images. Methods such as Faster R-CNN and YOLO (You Only Look Once) enable automated and precise detection. Examples of successful applications include the detection of lung nodules in chest X-rays and CT scans, aiding in early diagnosis of lung cancer, and the identification of polyps in colonoscopy images, improving the accuracy of colon cancer screening. Object detection's impact on medical diagnosis and treatment is profound, as it allows for faster and more accurate identification of anomalies and reduces the workload on healthcare professionals, enabling them to focus on more complex cases.
- 4. Disease Prediction and Risk Assessment:** Deep learning models play a crucial role in predicting diseases and assessing patient risk by analyzing patterns in medical images and patient data. These models can predict the likelihood of a patient developing a specific disease or assess their risk based on factors such as genetic data, medical history, and imaging results. For instance, deep learning models can predict cardiovascular disease from coronary artery CT scans or assess a patient's risk of developing cancer based on genetic markers and imaging data. The potential for personalized medicine is substantial, as deep learning enables tailored treatments and interventions based on individual patient profiles, optimizing healthcare outcomes. These models hold promise in revolutionizing preventive medicine and early intervention, ultimately improving patient care.
- 5. Image Registration:** Image registration in medical image analysis involves aligning and overlaying images from different modalities or time points to enhance the accuracy of diagnosis and treatment planning. Deep learning techniques have improved the precision of image registration, particularly in cases involving non-rigid deformation or complex anatomical variations. This advancement is crucial in various medical applications, such as intraoperative image-guided surgery, where real-time registration ensures accurate navigation during procedures. Challenges in image registration with deep learning include handling noisy data, ensuring real-time performance, and validating registration accuracy.

in clinical settings. Nonetheless, these techniques have the potential to revolutionize surgical procedures and treatment planning, ultimately benefiting patient outcomes.

To recap, the applications of deep learning in medical image analysis span a diverse array of tasks, including image classification, segmentation, object detection, disease prediction, and image registration. Despite their impressive achievements, tackling challenges such as data availability, interpretability, and ensuring broad applicability continues to be an ongoing pursuit within the realm of medical imaging. These efforts hold the potential to profoundly influence healthcare and enhance patient outcomes.

VI. APPLYING DNN FOR DIAGNOSIS AND DETECTION USING MEDICAL IMAGES

Recent developments and research efforts in various anatomical regions have harnessed deep neural network (DNN) architectures to support medical practitioners. These advancements span multiple areas, including brain, breast, kidney, liver, chest, eye, cardiac, abdominal, spine, dermatology, gastrointestinal, neuroimaging, musculoskeletal, pulmonary, and dental imaging. In these domains, DNNs have been applied for tasks such as disease detection, lesion identification, segmentation, and classification, using diverse medical imaging modalities like MRI, CT scans, X-rays, and endoscopic images. These technologies facilitate quicker and more accurate diagnosis and assist in treatment planning and monitoring patient outcomes.

1. Applications of DNN in Lung Disease

Examples of DNN Techniques in Lung Diseases:

- Abdelhamid et al. [131] utilized the U-Net model to diagnose lung diseases from CT images, achieving a Dice-Coefficient index of 95.02%.
- Ferreira et al. [132] proposed a regularized V-Net model for segmentation tasks, yielding an average Dice Coefficient of 93.6% per lobe and 76.2% inter-lobar Dice Coefficient.

DNN Approaches during the COVID-19 Pandemic:

- Ucaret et al. [133] introduced a deep Bayes-Squeeze Net model for COVID-19 diagnosis, addressing dataset imbalances and reporting a remarkable 98.26% accuracy.
- Mittal et al. [134] developed ICC and ECC models, combining convolutions with capsules, achieving accuracies of 95.33% and 95.90%, respectively.

Chest X-ray Diagnosis:

- Chest X-rays are the preferred diagnostic method for lung diseases due to their speed compared to lab testing.
- Accuracy rates for binary and three-class classification based on chest X-ray scans were 99.48% and 97%, respectively.

Researchers have effectively employed a variety of DNN architectures to achieve high accuracy in the classification, detection, registration, and segmentation of lung diseases.

2. Applications of Dnn in Eye Disease: DNN techniques have been applied in ophthalmology for various purposes alongside other organ diagnostics. These applications include the analysis of retinal fundus images for diabetic retinopathy classification, age and gender prediction, and detection of retinopathy, macular edema, and glaucoma-like disk.

- Nagasato et al. [135], implemented the Visual Geometry Group VGG-16 model. It achieved an impressive AUC (Area Under the Curve) of 98.6% in its analysis.
- Wu, Xia et al. [136], developed the NFN+ model for retinal vessel mapping. The model achieved high AUC scores for three different datasets: DRIVE (98.30%), STARE (98.75%), and CHASE (98.94%).
- Liu et al. [137], implemented a self-adaptive deep learning method based on Inception-V3. It achieved an impressive AUC of 99.50% in their analysis.

DNN techniques have been used for classification, segmentation, detection, and prediction in ophthalmology using various image modalities. These techniques address retinopathy caused by different factors and rely on diverse imaging methods.

3. Applications of Dnn in Bone Age Disease: Historically, bone age assessment relied on visual evaluation of a single reference X-ray image, leading to interpretation variability. Recent research focuses on using deep learning to predict bone age, typically using left-hand radiology images for children.

- S. S. Halabi et al. [138], investigated pediatric bone age assessment using various techniques, including Inception V3, ResNet-50, Ice Module Architecture, ML with handcrafted features, and Gabor texture energies. Based on mean absolute distance (MAD), results ranged from 4.2 to 4.5 months.
- S. H. Tajmir et al. [139], developed a CNN model based on LeNet-5 for bone age assessment. It achieved 68.2% accuracy overall and 98.6% accuracy within a year of the model.
- Transfer Learning Impact by Kandel et al. [140], investigated the impact of transfer learning on musculoskeletal image classification. It Found that transfer learning outperformed training CNNs from scratch in terms of accuracy.

4. Applications of Dnn in Osteosarcoma Diagnostics: The complexity of tissue structure limits digital automation in histopathology. Identifying tissue types, like osteoblasts, osteocytes, and osteoclasts, in bone biopsy samples is crucial. Malignant characteristics include nuclear membrane abnormalities, pleomorphism, large multinucleated cells, hyperchromatic nuclei, and aberrant mitosis. Hematoxylin and eosin (H&E) staining is commonly used in cancer research but may result in under-representation due to slide preparation and staining issues.

- O. Daescu et al. [141], studied 64,000 osteosarcoma image patches resized to 128x128 and proposed a CNN model. It achieved a classification accuracy of approximately 92.4%.
- XAI techniques applied to medical image analysis using deep learning [142]. XAI framework has been presented for categorizing deep learning-based medical image analysis techniques.
- M. D Acunto et al. [143], implemented a Faster R-CNN model. It attained an accuracy of 97%.

Testing various architectures with relatively small-sized datasets due to structural complexity and difficulty in distinguishing normal and disease cells.

V. RECENT ADVANCEMENTS IN MEDICAL IMAGE ANALYSIS FOR DEEP LEARNING

In recent years, the field of medical image analysis has witnessed remarkable progress, driven by innovative approaches and advanced technologies. This chapter delves into some of the most noteworthy recent advancements, shedding light on their impact and potential applications in healthcare.

- 1. Self-Supervised Learning: Pioneering the Path to Unsupervised Medical Insights:** Self-supervised learning has emerged as a groundbreaking paradigm shift in medical image analysis. This approach leverages the inherent structure and content within medical images to train deep learning models without the need for extensive labeled datasets. Recent advancements in self-supervised learning techniques have led to significant improvements in diagnostic accuracy and disease detection. By capitalizing on the rich information contained within medical images, these methods enable the extraction of latent features and representations, ultimately enhancing the performance of models in tasks such as image segmentation, disease classification, and anomaly detection. Moreover, self-supervised learning has proven invaluable in scenarios where labeled data is scarce or difficult to obtain, empowering healthcare professionals with more accessible and accurate diagnostic tools.
- 2. Generative Adversarial Networks (GANs): Shaping the Future of Synthetic Medical Imagery:** Generative Adversarial Networks (GANs) have revolutionized the generation of synthetic medical images and found diverse applications in the healthcare domain. GANs excel in creating realistic, high-fidelity medical images, which are invaluable for training and augmenting datasets. This technology plays a crucial role in addressing the data scarcity issue often encountered in medical imaging. By generating synthetic data that closely mimics real-world patient cases, GANs facilitate the development of more robust and generalizable deep-learning models. Moreover, GANs have extended their utility beyond image generation to tasks such as data augmentation, domain adaptation, and anomaly detection. As a result, GANs have become an indispensable tool for researchers and healthcare professionals striving to improve the accuracy and reliability of medical image analysis systems.

- 3. Explainable AI (XAI): Bridging the Gap Between Algorithms and Healthcare Professionals:** Explainable AI (XAI) has gained prominence in the medical image analysis community due to its pivotal role in making deep learning models more interpretable for healthcare professionals. Recent developments in XAI techniques have paved the way for greater transparency and trust in AI-driven diagnostic systems. These methods provide insights into the decision-making process of deep learning models, allowing clinicians to better understand and validate the predictions made by AI systems. By visualizing feature importance, highlighting salient regions in images, and offering detailed explanations for model outputs, XAI empowers healthcare practitioners to make more informed and confident decisions. Furthermore, XAI contributes to regulatory compliance and ethical considerations, ensuring that AI-driven medical image analysis remains accountable and aligned with the highest standards of patient care.

Recent advancements in self-supervised learning, Generative Adversarial Networks (GANs), and Explainable AI (XAI) are shaping the landscape of medical image analysis. These technologies hold the promise of more accurate diagnoses, enhanced data availability, and improved collaboration between AI systems and healthcare professionals. As the field continues to evolve, it is evident that these innovations will play a pivotal role in revolutionizing healthcare and improving patient outcomes.

VI. FUTURE DIRECTION AND CONCLUSION

The future of medical image analysis holds promise through the integration of emerging technologies. Combining natural language processing (NLP) with deep learning can enable a more comprehensive understanding of patient conditions by extracting insights from clinical records and medical images. Reinforcement learning (RL) offers the potential for adaptive treatment planning, and tailoring therapies to individual patient responses. Multi-modal data fusion, incorporating diverse information sources, can revolutionize disease diagnosis and treatment by providing a holistic view of patient health. These trends signify a dynamic future where deep learning plays a pivotal role in enhancing healthcare outcomes, though challenges like interpretability and data integration must be addressed for successful implementation.

In conclusion, this chapter underscores the transformative impact of deep learning on medical image analysis, enhancing diagnostic accuracy and patient care. Deep learning's significance in addressing healthcare challenges, from image segmentation to disease classification, cannot be overstated. However, it is crucial to emphasize the responsible and ethical use of AI in healthcare, ensuring patient privacy, data security, and model transparency. By maintaining a balance between innovation and ethical considerations, deep learning promises to revolutionize healthcare while preserving trust among healthcare professionals and patients.

REFERENCES

- [1] Mall, P. K., Singh, P. K., Srivastav, S. K., Narayan, V., Paprzycki, M., Jaworska, T., & Ganzha, M. (2023). A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities. *Healthcare Analytics*, 4, 100216. <https://doi.org/10.1016/j.health.2023.100216>.
- [2] J.L. McClelland, D.E. Rumelhart, P.D.P.R. Group, et al., *Parallel Distributed Processing*, Vol. 2, MIT Press, Cambridge, MA, 1986.
- [3] S. Suganyadevi, V. Seethalakshmi, K. Balasamy, A review on deep learning in medical image analysis, *Int. J. Multimed. Inf. Retr.* 11 (1) (2022) 19–38.
- [4] H. P. Chan, L. M. Hadjiiski, and R. K. Samala, “Computer-aided diagnosis in the era of deep learning,” *Medical Physics*, vol. 47, no. 5, pp. e218–e227, 2020.
- [5] F. Ritter, T. Boskamp, A. Homeyer et al., “Medical image analysis,” *IEEE Pulse*, vol. 2, no. 6, pp. 60–70, 2011.
- [6] J. Ker, L. Wang, J. Rao, and T. Lim, “Deep learning applications in medical image analysis,” *IEEE Access*, vol. 6, pp. 9375–9389, 2017.
- [7] M. Jena, S. P. Mishra, and D. Mishra, “A survey on applications of machine learning techniques for medical image segmentation,” *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 4489–4495, 2018.
- [8] Sistaninejad, B., Rasi, H., & Nayeri, P. (2023). A Review Paper about Deep Learning for Medical Image Analysis. *Computational and Mathematical Methods in Medicine*, 2023, 1–10. <https://doi.org/10.1155/2023/7091301>
- [9] S. Niyas, S. J. Pawan, M. Anand Kumar, and J. Rajan, “Medical image segmentation with 3D convolutional neural networks: a survey,” *Neurocomputing*, vol. 493, pp. 397–413, 2022.
- [10] P. Dutta, P. Upadhyay, M. De, and R. G. Khalkar, “Medical image analysis using deep convolutional neural networks: CNN architectures and transfer learning,” in *2020 International Conference on Inventive Computation Technologies (ICICT)*, pp. 175–180, Coimbatore, India, February 2020.
- [11] LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444. <https://doi.org/10.1038/nature14539>
- [12] Pang S, Yang X (2016) Deep Convolutional Extreme learning Machine and its application in Handwritten Digit Classification. *Hindawi Publ Corp Comput Intell Neurosci* 2016:3049632. <https://doi.org/10.1155/2016/3049632>
- [13] Chollet F et al (2015) Keras. <https://github.com/fchollet>
- [14] Zhang Y, Zhang S, et al (2016) Theano: A Python framework for fast computation of mathematical expressions, *arXiv e-prints*, abs/1605.02688. <http://arxiv.org/abs/1605.02688>
- [15] Vedaldi A, Lenc K (2015) Matconvnet: convolutional neural network for matlab. In: *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, pp 689–692. <https://doi.org/10.1145/2733373.2807412>
- [16] Guo Y, Ashour A (2019) Neutrosophic sets in dermoscopic medical image segmentation. *Neutrosophic Set Med Image Anal* 11(4):229–243. <https://doi.org/10.1016/B978-0-12-818148-5.00011-4>
- [17] Merjulah R, Chandra J (2019) Classification of myocardial ischemia in delayed contrast enhancement using machine learning. *Intell Data Anal Biomed Appl*, pp 209–235. <https://doi.org/10.1016/B978-0-12-815553-0.00011-2>
- [18] Oliveira FPM, Tavares JMRS (2014) Medical Image Registration: a review. *Comput Methods Biomech Biomed Eng* pp 73–93. <https://doi.org/10.1080/10255842.2012.670855>
- [19] Wang J, Zhang M (2020) Deep FLASH: an efficient network for learning-based Medical Image Registration. In: *Proceedings of 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp 4443–4451. <https://doi.org/10.1109/cvpr42600.2020.00450>
- [20] Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X (2020) Deep learning in medical image registration: a review. *Phys Med Biol* 65(20). <https://doi.org/10.1088/1361-6560/ab843e>
- [21] Haskins G, Kruger U, Yan P (2020) Deep learning in medical image registration: a survey. *Mach Vis Appl* 31:1–18. <https://doi.org/10.1007/s00138-020-01060-x>
- [22] De Vos BD, Wolterink JM, Jong PA, Leiner T, Viergever MA, Isgum I (2017) ConvNet-based localization of anatomical structures in 3D medical images. *IEEE Trans Med Imaging* 36(7):1470–1481. <https://doi.org/10.1109/TMI.2017.2673121>

- [23] Song Y, Cai W, Huang H et al (2013) Region-based progressive localization of cell nuclei in microscopic images with data-adaptive modelling. *BMC Bioinformatics* 14:173. <https://doi.org/10.1186/1471-2105-14-173>
- [24] Sharma H, Jain JS, Gupta S, Bansal P (2020) Feature extraction and classification of chest X-ray images using CNN to detect pneumonia. 2020 In: *Proceedings of the 10th international conference on Cloud Computing, data science & engineering (confluence)*, pp 227–231. <https://doi.org/10.1109/Confluence47617.2020.9057809>
- [25] Hassan M, Ali S, Alquhayz H, Safdar K (2020) Developing intelligent medical image modality classification system using deep transfer learning and LDA. *Sci Rep* 10(1):12868. <https://doi.org/10.1038/s41598-020-69813-2>
- [26] Abbas A, Abdelsamea MM, Gaber MM (2021) Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Appl Intell* 51:854–864. <https://doi.org/10.1007/s10489-020-01829-7>
- [27] Kowsari K, Sali R, Ehsan L, Adorno W et al (2020) Hierarchical medical image classification, a deep learning approach. *Information* 11(6):318. <https://doi.org/10.3390/info11060318>
- [28] Singh SP, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyas B (2020) 3D deep learning on medical images: a review. *Sensors* 20(18):5097. <https://doi.org/10.3390/s20185097>
- [29] Shen D, Wu G, Suk H (2017) Deep learning in medical image analysis. *Annu Rev Biomed Eng* 19:221–248. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- [30] Wang SH, Phillips P, Sui Y, Bin L, Yang M, Cheng H (2018) Classification of Alzheimer’s disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. *J Med Syst* 42(5):85. <https://doi.org/10.1007/s10916-018-0932-7>
- [31] Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 60(6):84–90. <https://doi.org/10.1145/3065386>
- [32] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S et al (2015) Going deeper with convolutions. In: *Proceedings of the IEEE Computer Society conference on computer vision and pattern recognition*; IEEE Computer Society, 7(12), pp 1–9.
- [33] Lowekamp BC, Chen DT, Ibanez L, Blezek D (2013) The design of SimpleITK. *Front Neuroinform* 45(7). <https://doi.org/10.3389/fninf.2013.00045>
- [34] Avants B, Tustison N, Song G (2009) Advanced Normalization Tools (ANTs). *The Insight Journal*. <http://hdl.handle.net/10380/3113>
- [35] Classification using faster R-CNN. *Malays J Comput* 4(1):225–236. <https://doi.org/10.24191/mjoc.v4i1.6095>
- [36] Jain G, Mittal D, Takur D, Mittal MK (2020) A deep learning approach to detect Covid-19 coronavirus with X-ray images. *Elsevier-Biocybern Biomed Eng* 40(4):1391–1405. <https://doi.org/10.1016/j.bbe.2020.08.008>
- [37] Togacar M, Ergen B, Comert Z (2020) COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Elsevier-Comput Biol Med* 121:103805. <https://doi.org/10.1016/j.compbimed.2020.103805>
- [38] Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ (2020) Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning. *Elsevier-Med Image Anal* 65:101794. <https://doi.org/10.1016/j.media.2020.101794>
- [39] Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Springer-Phys Eng Sci Med* 43(2):635–640. <https://doi.org/10.1007/s13246-020-0865-4>
- [40] Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R, Singh V (2020) Application of deep learning for fast detection 138(3):109944. <https://doi.org/10.1016/j.chaos.2020.109944>
- [41] Ouchicha C, Ammor O, Meknassi M (2020) CVDNet: a novel deep learning architecture for detection of coronavirus (Covid-19) from chest X-ray images. *Elsevier-Chaos Solitons Fractals* 140(5):110245. <https://doi.org/10.1016/j.chaos.2020.110245>
- [42] Sethy PK, Behera SK, Ratha PK (2020) Detection of coronavirus disease (COVID-19) based on deep features and support vector machine. *Int J Math Eng Manag Sci* 5(4):643–651. <https://doi.org/10.33889/IJMEMS.2020.5.4.052>
- [43] Jaiswal AK, Tiwari P, Kumar S, Gupta D, Khanna A, Rodrigues JJ (2019) Identifying pneumonia in chest X-rays: a deep learning approach. *Measurement* 145(2):511–518. <https://doi.org/10.1016/j.measurement.2019.05.076>

- [44] Civit-Masot J, Luna-Perejon F, Dominguez Morales M, Civit A (2020) Deep learning system for COVID-19 diagnosis aid using X-ray pulmonary images. *Appl Sci* 10(13):4640. <https://doi.org/10.3390/app10134640>
- [45] Punn NS, Sonbhadra SK, Agarwal S (2020) COVID-19 epidemic analysis using machine learning and deep learning algorithms. <https://doi.org/10.1101/2020.04.08.20057679>
- [46] Dahl GE, Yu D, Deng L, Acero A (2012) Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Trans Audio Speech Lang Process* 20(1):30–42. <https://doi.org/10.1109/TASL.2011.2134090>
- [47] JC, Niemeijer M (2016) Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 57(13):5200–5206. <https://doi.org/10.1167/iovs.16-19964>
- [48] Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S, Navab N (2016) Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Trans Med Imaging* 35:1313–1321. <https://doi.org/10.1109/TMI.2016.2528120>
- [49] Anavi Y, Kogan I, Gelbart E, Geva O, Greenspan H (2015) A comparative study for chest radiograph image retrieval using binary texture and deep learning classification. In: *Proceedings of the IEEE engineering in Medicine and Biology Society*, pp 2940–2943. <https://doi.org/10.1109/EMBC.2015.7319008>
- [50] Anavi Y, Kogan I, Gelbart E, Geva O, Greenspan H (2016) Visualizing and enhancing a deep learning framework using patients age and gender for chest X-ray IMAGE Retrieval. In: *Proceedings of the SPIE on medical imaging*, 9785, 978510
- [51] Andermatt S, Pezold S, Cattin P (2016) Multi-dimensional gated recurrent units for the segmentation of biomedical 3D-data. In: *Proceedings of the deep learning in medical image analysis (DLMIA). Lecture Notes in Computer Science*, 10 0 08, pp 142–151
- [52] Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S (2016) Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Trans Med Imaging* 35(5):1207–1216. <https://doi.org/10.1109/TMI.2016.2535865>
- [53] Antony J, McGuinness K, Connor NEO, Moran K (2016) Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. arxiv: 1609.0246951.
- [54] Apou G, Schaadt NS, Naegel B, Forestier G, Schönmeier R, Feuerhake F, Wemmert C, Grote A (2016) Detection of lobular structures in normal breast tissue. *Comput Biol Med* 74:91–102. <https://doi.org/10.1016/j.compbimed.2016.05.004>
- [55] Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Match Intell* 35(8):1798–1828. <https://doi.org/10.1109/TPAMI.2013.50>
- [56] Baumgartner CF, Kamnitsas K, Matthew J, Smith S, Kainz B, Rueckert D (2016) Real-time standard scan plane detection and localization in fetal ultrasound using fully convolutional neural networks. In: *Proceedings of the medical image computing and computer-assisted intervention. Lecture Notes in Computer Science*, 9901, pp 203–211. https://doi.org/10.1007/978-3-319-46723-8_24
- [57] Balasamy K, Ramakrishnan S (2019) An intelligent reversible watermarking system for authenticating medical images using wavelet and PSO. *Clust Comput* 22(2):4431–4442. <https://doi.org/10.1007/s10586-018-1991-8>
- [58] Bengio Y (2012) Practical recommendations for gradient-based training of deep architectures. *Neural networks: tricks of the trade. Lecture Notes in Computer Science*, 7700, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-35289-8_26
- [59] Han XH, Lei J, Chen YW (2016) HEp-2 cell classification using K-support spatial pooling in deep CNNs. In: *Proceedings of the deep learning in medical image analysis (DLMIA). Lecture Notes in Computer Science*, 10 0 08, pp 3–11. <https://doi.org/10.1007/978-3-319-46976-81>
- [60] Haugeland J (1985) *Artificial intelligence: the very idea*. The MIT Press, Cambridge. ISBN: 0262081539
- [61] Havaei M, Davy A, Warde Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin PM, Larochelle H (2016) Brain tumor segmentation with deep neural networks. *Med Image Anal* 35:18–31. <https://doi.org/10.1016/j.media.2016.05.004>
- [62] Havaei M, Guizard N, Chapados N, Bengio Y (2016) HeMIS: hetero-modal image segmentation. In: *Proceedings of the medical image computing and computer-assisted intervention. Lecture Notes in Computer Science*, 9901, pp 469–477. https://doi.org/10.1007/978-3-319-46723-8_54
- [63] Benou A, Veksler R, Friedman A, Raviv TR (2016) Denoising of contrast enhanced MRI sequences by an ensemble of expert deep neural networks. In: *Proceedings of the deep learning in medical image analysis*

- (DLMIA). *Lecture Notes in Computer Science*, 10 0 08, pp 95–110. <https://doi.org/10.1146/annurev-bioeng-071516-044442>
- [64] BenTaieb A, Hamarneh G (2016) Topology aware fully convolutional networks for histology gland segmentation. In: *Proceedings of the medical image computing and computer-assisted intervention. Lecture Notes in Computer Science*, 9901, pp 460–468. https://doi.org/10.1007/978-3-319-46723-8_53
- [65] BenTaieb A, Kawahara J, Hamarneh G (2016) Multi-loss convolutional networks for gland analysis in microscopy. In: *Proceedings of the IEEE international symposium on biomedical imaging*, pp 642–645. <https://doi.org/10.1109/ISBI.2016.7493349>
- [66] Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. *J Mach Learn Res* 13(10):281–305
- [67] Birenbaum A, Greenspan H (2016) Longitudinal multiple sclerosis lesion segmentation using multi-view convolutional neural networks. In: *Proceedings of the deep learning in medical image analysis (DLMIA). Lecture Notes in Computer Science*, 10 0 08, pp 58–67. https://doi.org/10.1007/978-3-319-46976-8_7
- [68] Cheng X, Zhang L, Zheng Y (2015) Deep similarity learning for multimodal medical images. *Comput Methods Biomech Biomed Eng* pp 248–252. https://doi.org/10.1080/21681163.2015.11352_99
- [69] Kallen H, Molin J, Heyden A, Lundstr C, Astrom K (2016) Towards grading gleason score using generically trained deep convolutional neural networks. In: *Proceedings of the IEEE international symposium on biomedical imaging*, pp 1163–1167. <https://doi.org/10.1109/ISBI.2016.7493473>
- [70] Balasamy K, Suganyadevi S (2021) A fuzzy based ROI selection for encryption and watermarking in medical image using DWT and SVD. *Multimedia Tools Appl* 80:7167–7186. <https://doi.org/10.1007/s11042-020-09981-5>
- [71] Lekadir K, Galimzianova A, Betriu A, Vila MDM, Igual L, Rubin, DL, Fernandez E, Radeva P, Napel S (2017) A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. *IEEE J Biomed Health Inform* 21(1):48–55. <https://doi.org/10.1109/JBHI.2016.2631401>
- [72] Li R, Zhang W, Suk HI, Wang L, Li J, Shen D, Ji S (2014) Deep learning based imaging data completion for improved brain disease diagnosis. *Med Image Comput Comput-Assist Interv* 17(Pt3):305–312. https://doi.org/10.1007/978-3-319-0443-0_39
- [73] Li W, Manivannan S, Akbar S, Zhang J, Trucco E, McKenna SJ (2016) Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks. In: *Proceedings of the IEEE international symposium on biomedical imaging*, pp 1405–1408. <https://doi.org/10.1109/ISBI.2016.7493530>
- [74] Moeskops P, Viergever MA, Mendrik AM, Vries LSD, Benders MJNL, Isgum I (2016) Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans Med Imaging* 35(5):1252–1262. <https://doi.org/10.1109/TMI.2016.2548501>
- [75] Pinaya WHL, Gadelha A, Doyle OM, Noto C, Zugman A, Cordeiro Q, Jackowski AP, Bressan RA, Sato JR (2016) Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. *Nat Sci Rep* 6:38897. <https://doi.org/10.1038/srep38897>
- [76] Sarraf S, Tofighi G (2016) Classification of Alzheimer’s disease using fMRI data and deep learning convolutional neural networks. *arxiv: 1603.08631*
- [77] Schaumberg AJ, Rubin MA, Fuchs TJ (2016) H&E-stained whole slide deep learning predicts SPOB mutation state in prostate cancer. *arxiv: 064279*. <https://doi.org/10.1101/064279>
- [78] He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. *arxiv: 1512.03385*
- [79] Janowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J Pathol Inform* 7:29. <https://doi.org/10.4103/2153-3539.186902>
- [80] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: convolutional architecture for fast feature embedding. In: *Proceedings of the twenty-second ACM international conference on multi-media*, pp 675–678. <https://doi.org/10.1145/2647868.2654.889>
- [81] Kainz P, Pfeiffer M, Urschler M (2015) Semantic segmentation of colon glands with deep convolutional neural networks and total variation segmentation. *arxiv: 1511.06919*
- [82] Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, Johnson HJ, Paulsen JS, Turner JA, Calhoun VD (2014) Deep learning for neuroimaging: a validation study. *Front Neurosci* 8:229. <https://doi.org/10.3389/fnins.2014.00229>
- [83] Poudel RPK, Lamata P, Montana G (2016) Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. *arxiv: 1608.03974*
- [84] Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Nielsen M (2013) Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: *Proceedings of the medical image*

- computing and computer-assisted intervention. *Lecture Notes in Computer Science*, 8150, pp 246–253. https://doi.org/10.1007/978-3-642-40763-5_31
- [85] Rajkumar A, Lingam S, Taylor AG, Blum M, Mongan J (2017) High-throughput classification of radiographs using deep convolutional neural networks. *J Digit Imaging* 30:95–101. <https://doi.org/10.1007/s10278-016-9914-9>
- [86] Ravi D, Wong C, Deligianni F, Berthelot M, Andreu Perez J, Lo , Yang GZ (2017) Deep learning for health informatics. *IEEE J Biomed Health Inform* 21(1):4–21. <https://doi.org/10.1109/JBHI.2016.2636665>
- [87] Ravishankar H, Prabhu SM, Vaidya V, Singhal N (2016) Hybrid approach for automatic segmentation of fetal abdomen from ultrasound images using deep learning. In: *Proceedings of the IEEE international symposium on biomedical imaging*, pp 779–782. <https://doi.org/10.1109/ISBI.2016.7493382>
- [88] Sahiner B, Chan HP, Petrick N, Wei D, Helvie MA, Adler DD, Goodsitt MM (1996) Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE Trans Med Imaging* 15(5):598–610. <https://doi.org/10.1109/42.538937>
- [89] Samala RK, Chan HP, Hadjiiski LM, Cha K, Helvie MA (2016) Deep-learning convolution neural network for computer-aided detection of microcalcifications in digital breast tomosynthesis. In: *Proceedings 9785, medical imaging, computer-aided diagnosis, 97850Y*. <https://doi.org/10.1117/12.2217092>
- [90] Samala RK, Chan HP, Hadjiiski L, Helvie MA, Wei J, Cha K (2016) Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography. *Med Phys* 43(12):6654. <https://doi.org/10.1118/1.4967345>
- [91] Suganyadevi S, Shamia D, Balasamy K (2021) An IoT-based diet monitoring healthcare system for women. *Smart Healthc Syst Des Secur Priv Asp*. <https://doi.org/10.1002/9781119792253.ch8>
- [92] Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R (2016) Deep learning for automated skeletal bone age assessment in X-ray images. *Med Image Anal* 36:41–51. <https://doi.org/10.1016/j.media>
- [93] Balasamy K, Shamia D (2021) Feature extraction-based medical image watermarking using fuzzy-based median filter. *IETE J Res* 1–9. <https://doi.org/10.1080/03772063.2021.1893231>
- [94] Stern D, Payer C, Lepetit V, Urschler M (2016) Automated age estimation from hand MRI volumes using deep learning. In: *Proceedings of the medical image computing and computer-assisted intervention. Lecture Notes in Computer Science*, 9901, pp 194–202. https://doi.org/10.1007/978-3-319-46723-8_23
- [95] Suk HI, Shen D (2013) Deep learning based feature representation for AD/MCI classification. In: *Proceedings of the medical image computing and computer-assisted intervention. Lecture Notes in Computer Science*, 8150, pp 583–590. https://doi.org/10.1007/978-3-642-40763-5_72
- [96] Sun W, Seng B, Zhang J, Qian W (2016) Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. *Comput Med Imaging Gr* 57:4–9. <https://doi.org/10.1016/j.compmedimag>
- [97] Teikari P, Santos M, Poon C, Hynynen K (2016) Deep learning convolutional networks for multiphoton microscopy vasculature segmentation. arxiv: 1606.02382
- [98] A.E. Kavur, N.S. Gezer, M. Barış, S. Aslan, P.-H. Conze, V. Groza, D.D. Pham, S. Chatterjee, P. Ernst, S. Özkan, et al., CHAOS challenge-combined (CT-MR) healthy abdominal organ segmentation, *Med. Image Anal.* 69 (2021) 101950.
- [99] Tadpole.grand-challenge, 2021, [Online]. Available: <https://tadpole.grandchallenge.org>. [Accessed: 27-Sep-2021].
- [100] Grand challenge on MR brain segmentation at MICCAI 2018, 2021, [Online]. Available: <https://mrbrains18.isi.uu.nl/data/>. [Accessed: 13-Dec-2021].
- [101] Multi-modality, lower spine 3D, 2020, [Online]. Available: <https://projects.iq.harvard.edu/cdmri2018/data> [Accessed: 11-May-2020].
- [102] D. Racoceanu, F. Capron, Semantic integrative digital pathology: Insights into microsemiological semantics and image analysis scalability, *Pathobiology* 83 (2–3) (2016) 148–155.
- [103] icar, 2018-Challenge.grand-challenge, 2021, [Online]. Available: <https://icar2018-challenge.grand-challenge.org/Dataset/>. [Accessed: 11-Oct-2021].
- [104] S.S. Halabi, L.M. Prevedello, J. Kalpathy-Cramer, A.B. Mamonov, A. Bilbily, M. Cicero, I. Pan, L.A. Pereira, R.T. Sousa, N. Abdala, et al., The RSNA pediatric bone age machine learning challenge, *Radiology* 290 (2) (2019) 498–503.
- [105] A.E.W. Johnson, T.J. Pollard, N.R. Greenbaum, M.P. Lungren, C. Deng, Y. Peng, Z. Lu, R.G. Mark, S.J. Berkowitz, S. Horng, MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs, 14, (2019) 1–7.

- [106] ECG rhythm classification, 2021, [Online]. Available: https://irhythm.github.io/cardiol_test_set/ . [Accessed: 23-May-2021].
- [107] idrid.grand-challenge.org. [Online]. Available: <https://idrid.grand-challenge.org/Data>.
- [108] Pet-radiomics-challenges, 2021, [Online]. Available: <https://www.kaggle.com/c/pet-radiomics-challenges/data>. [Accessed: 28-Dec-2021].
- [109] N. Bien, P. Rajpurkar, R.L. Ball, J. Irvin, A. Park, E. Jones, M. Bereket, B.N. Patel, K.W. Yeom, K. Shpanskaya, S. Halabi, E. Zucker, G. Fanton, D.F. Amanatullah, C.F. Beaulieu, G.M. Riley, R.J. Stewart, F.G. Blankenberg, D.B. Larson, R.H. Jones, C.P. Langlotz, A.Y. Ng, M.P. Lungren, Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet, *PLoS Med.* 15 (11) (2018) 1–19.
- [110] Lung nodule analysis-2016, 2020, [Online]. Available: <https://luna16.grandchallenge.org/>. [Accessed: 11-May-2020].
- [111] S.G. Armato III, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, B. Zhao, D.R. Aberle, C.I. Henschke, E.A. Hoffman, et al., The lung nodule database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans, *Med. Phys.* 38 (2) (2011) 915–931.
- [112] Challenge on liver ultrasound tracking, 2021, [Online]. Available: <https://clust.ethz.ch/>. [Accessed: 15-Aug-2021].
- [113] Lower spine 3D multi-modality MRI data sets, 2021, [Online]. Available: <https://ivdm3seg.weebly.com/data.html>. [Accessed: 20-Oct-2021].
- [114] Mild traumatic brain injury outcome prediction, 2021, [Online]. Available: <https://tbichallenge.wordpress.com/data/>. [Accessed: 11-Dec-2021].
- [115] L. Deleskog, N. Laursen, B.R. Nielsen, P. Schwarz, Vertebral fracture assessment by DXA is inferior to X-ray in clinical severe osteoporosis, *Osteoporos. Int.* 27 (7) (2016) 2317–2326.
- [116] S. Rezaei, A. Emami, N. Karimi, S. Samavi, Gland segmentation in histopathological images by deep neural network, 1–5.
- [117] N. Heller, N. Sathianathan, A. Kalapara, E. Walczak, K. Moore, H. Kaluzniak, J. Rosenberg, P. Blake, Z. Rengel, M. Oestreich, J. Dean, M. Tradewell, A. Shah, R. Tejjpaul, Z. Edgerton, M. Peterson, S. Raza, S. Regmi, N. Papanikolopoulos, C. Weight, The KiTS19 challenge data: 300 kidney tumor cases with clinical context, CT semantic segmentations, and surgical outcomes, 2019, pp. 1–13.
- [118] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, A. Schlaefer, Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data, *MethodsX* 7 (2020) 100864.
- [119] Analysis of images to detect abnormalities in endoscopy accessed, 2021, [Online]. Available: <https://isbi-aida.grand-challenge.org/>. [Accessed: 20-Oct-2021].
- [120] Retouch.grand-challenge, 2021, [Online]. Available: <https://retouch.grandchallenge.org>. [Accessed: 11-May-2021].
- [121] Tran PV (2016) A fully convolutional neural network for cardiac segmentation in short axis MRI. arxiv: 1604.00494.abs/1604.00494
- [122] Xie Y, Xing F, Kong X, Su H, Yang L (2015) Beyond classification: structured regression for robust cell detection using convolutional neural network. In: Proceedings of the medical image computing and computer-assisted intervention. Lecture Notes in Computer Science, 9351, pp 358–365. https://doi.org/10.1007/978-3-319-24574-4_43
- [123] Xie Y, Zhang Z, Sapkota M, Yang L (2016) Spatial clockwork recurrent neural network for muscle perimysium segmentation. In: Proceedings of the international conference on medical image computing and computer-assisted intervention. Lecture Notes in Computer Science, 9901. Springer, pp 185–193. https://doi.org/10.1007/978-3-319-46723-8_22
- [124] Xu T, Zhang H, Huang X, Zhang S, Metaxas DN (2016) Multimodal deep learning for cervical dysplasia diagnosis. In: Proceedings of the medical image computing and computer-assisted intervention. Lecture Notes in Computer Science, 9901, pp 115–123. https://doi.org/10.1007/978-3-319-46723-8_14
- [125] Zhang Q, Xiao Y, Dai W, Suo J, Wang C, Shi J, Zheng H (2020) Deep learning based classification of breast tumors with shearwave elastography. *Ultrasonics* 72:150–157. <https://doi.org/10.1016/j.ultras.2016.08.004>
- [126] Yang D, Zhang S, Yan Z, Tan C, Li K, Metaxas D (2015) Automated anatomical landmark detection on distal femur surface using convolutional neural network. In: Proceedings of the IEEE international symposium on biomedical imaging, pp 17–21. <https://doi.org/10.1109/isbi.2015.7163806>

- [127] Zhao J, Zhang M, Zhou Z, Chu J, Cao F (2017) Automatic detection and classification of leukocytes using convolutional neural networks. *Med Biol Eng Comput* 55(8):1287–1301. <https://doi.org/10.1007/s11517-016-1590-x>.
- [128] Zhang H, Li L, Qiao K, Wang L, Yan B, Li L, Hu G (2016) Image prediction for limited-angle tomography via deep learning with convolutional neural network. *arxiv: 1607.08707*
- [129] Yu L, Yang X, Chen H, Qin J, Heng P (2017) Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D MR images. In: *Proceedings of the thirtyfirst AAAI conference on artificial intelligence*, pp 66–72
- [130] Razzak MI, Naz ZS, Zaib A (2018) Deep learning for medical image processing: overview, challenges and the future. *Classification in BioApps: automation of decision making*. Springer, Cham, Switzerland, pp 323–350. https://doi.org/10.1007/978-3-319-65981-7_12
- [131] B.A. Skourt, A. El Hassani, A. Majda, Lung CT image segmentation using deep neural networks, *Procedia Comput. Sci.* 127 (2018) 109–113.
- [132] F.T. Ferreira, P. Sousa, A. Galdran, M.R. Sousa, A. Campilho, End-to-end supervised lung lobe segmentation, in: *2018 International Joint Conference on Neural Networks, IJCNN, 2018*, pp. 1–8.
- [133] F. Ucar, D. Korkmaz, COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images, *Med. Hypotheses* 140 (2020) 109761.
- [134] A. Mittal, D. Kumar, M. Mittal, T. Saba, I. Abunadi, A. Rehman, S. Roy, Detecting pneumonia using convolutions and dynamic capsule routing for chest X-ray images, *Sensors* 20 (4) (2020) 1068.
- [135] D. Nagasato, H. Tabuchi, H. Masumoto, H. Enno, N. Ishitobi, M. Kameoka, M. Niki, Y. Mitamura, Automated detection of a nonperfusion area caused by retinal vein occlusion in optical coherence tomography angiography images using deep learning, *PLoS One* 14 (11) (2019) e0223965.
- [136] Y. Wu, Y. Xia, Y. Song, Y. Zhang, W. Cai, NFN+: A novel network followed network for retinal vessel segmentation, *Neural Netw.* 126 (2020) 153–162.
- [137] B.H.M. van der Velden, H.J. Kuijf, K.G.A. Gilhuijs, M.A. Viergever, Explainable artificial intelligence (XAI) in deep learning-based medical image analysis, 2021, *arXiv Prepr. arXiv:2107.10912*.
- [138] S.S. Halabi, L.M. Prevedello, J. Kalpathy-Cramer, A.B. Mamonov, A. Bilbily, M. Cicero, I. Pan, L.A. Pereira, R.T. Sousa, N. Abdala, et al., The RSNA pediatric bone age machine learning challenge, *Radiology* 290 (2) (2019) 498–503.
- [139] S.H. Tajmir, H. Lee, R. Shailam, H.I. Gale, J.C. Nguyen, S.J. Westra, R. Lim, S. Yune, M.S. Gee, S. Do, Artificial intelligence-assisted interpretation of bone age radiographs improves accuracy and decreases variability, *Skeletal Radiol.* 48 (2) (2019) 275–283.
- [140] I. Kandel, M. Castelli, A. Popovič, Musculoskeletal images classification for detection of fractures using transfer learning, *J. Imaging* 6 (11) (2020) 127.
- [141] S. Mahore, K. Bhole, S. Rathod, Machine learning approach to classify and predict different osteosarcoma types, in: *2021 8th International Conference on Signal Processing and Integrated Networks, SPIN, 2021*, pp. 641–645.
- [142] B.H.M. van der Velden, H.J. Kuijf, K.G.A. Gilhuijs, M.A. Viergever, Explainable artificial intelligence (XAI) in deep learning-based medical image analysis, *Med. Image Anal.* 79 (2022) 102470.
- [143] M. D’Acunto, M. Martinelli, D. Moroni, Deep learning approach to human osteosarcoma cell detection and classification, in: *International Conference on Multimedia and Network Information System, 2018*, pp. 353–361.

