



On the Analyses of Medical Images Using Traditional Machine Learning Techniques and Convolutional Neural Networks

Saeed Iqbal^{1,2} · Adnan N. Qureshi¹ · Jianqiang Li^{2,3} · Tariq Mahmood⁴

Received: 1 December 2022 / Accepted: 19 February 2023 / Published online: 4 April 2023
© The Author(s) 2023

Abstract

Convolutional neural network (CNN) has shown dissuasive accomplishment on different areas especially Object Detection, Segmentation, Reconstruction (2D and 3D), Information Retrieval, Medical Image Registration, Multi-lingual translation, Local language Processing, Anomaly Detection on video and Speech Recognition. CNN is a special type of Neural Network, which has compelling and effective learning ability to learn features at several steps during augmentation of the data. Recently, different interesting and inspiring ideas of Deep Learning (DL) such as different activation functions, hyperparameter optimization, regularization, momentum and loss functions has improved the performance, operation and execution of CNN. Different internal architecture innovation of CNN and different representational style of CNN has significantly improved the performance. This survey focuses on internal taxonomy of deep learning, different models of convolutional neural network, especially depth and width of models and in addition CNN components, applications and current challenges of deep learning.

1 Introduction

For many disorders, medical imaging is a crucial screening method. In 1895, Roentgen revealed that X-rays could be utilized to examine into the internal organs without causing any harm. Shortly after, X-ray radiology evolved into the earliest method for diagnosing diseases. Since then, a variety

of imaging techniques have been created, with Computed Tomography (CT) scanning, Positron Emission Tomography (PET), ultrasonography and Magnetic Resonance Imaging (MRI) being some of the most widely utilized. Additionally, increasingly intricate scanning techniques have been created. At numerous points in the healthcare system, encompassing diagnosis, characterization, grading, clinical outcome evaluation, tracking of tumor recurrence, as well as directing intervention treatments, surgeries and radiosurgery, image information is vital to judgment and finalizing decision.

Minimal two-dimensional (2D) imagegraphs are used for a specific individual instance, but several are used for 3D imaging and vast numbers are used for 4D interactive imaging. The quantity of image data that has to be evaluated is significantly increased by the use of multi-modality imaging. It is challenging for medical practitioners and radiologists and doctors to sustain operational efficiencies while employing all the diagnostic data at their disposal to increase reliability and individual treatment as a result of the rising burden. The promise and necessity of creating computerized tools to aid medical practitioners and radiologists in image interpretation and detection have been recognized as a significant field of study and advancement in medical imaging in light of current developments in computer vision and computational approaches.

Beginning in the 1960 s, attempts were made to use machines to autonomously interpret healthcare images

✉ Saeed Iqbal
saeediqbalkhattak@gmail.com

✉ Jianqiang Li
lijianqiang@bjut.edu.cn

Adnan N. Qureshi
dr.qureshi@ucp.edu.pk

Tariq Mahmood
tmsherazi@ue.edu.pk

¹ Department of Computer Science, Faculty of Information Technology & Computer Science, University of Central Punjab, Lahore, Punjab 54000, Pakistan

² Faculty of Information Technology, Beijing University of Technology, Beijing 100124, Beijing, China

³ Beijing Engineering Research Center for IoT Software and Systems, Beijing University of Technology, Beijing 100124, Beijing, China

⁴ Artificial Intelligence and Data Analytics (AIDA) Lab, College of Computer & Information Sciences (CCIS), Prince Sultan University, Riyadh 11586, Kingdom of Saudi Arabia

[1–3]. Numerous research showed that using a machine to analyze clinical data was feasible, but the research received little focus, most likely due to the lack of accessibility to slightly elevated digitized visual information and computing power. In the 1980 s, [4] at the University of Chicago's Kurt Rossmann Laboratory started systematically developing machine learning and image analysis methods for medical data with the intention of creating Computer-Aided Diagnosis (CAD) as a better solution to support radiologists in visual explanation [5]. In order to recognize microcalcifications on mammography, [6] created a CAD platform and carried out the initial spectator ability research that showed how well CAD improved breast radiologists' capability to recognize microcalcifications [7]. In 1998, the Food and Drug Administration (FDA) authorized the use of the initial professional CAD system as a backup for diagnostic radiography. Over the recent decades, one of the main areas of study and innovation in diagnostic imaging has been CAD and computer-assisted image recognition.

Multiple medical disorders in both youngsters and people can now be diagnosed and treated more accurately thanks to medical imaging. Medical imaging processes come in a variety of forms, or modalities, and each one employs a unique set of tools and methods. Ionizing radiation is used by radiography, particularly Computed Tomography (CT), mammography and fluoroscopy to provide images of the human body. The risk of acquiring disease throughout the course of one's lifetime may increase if an individual is exposed to

ionising radiation, a type of radiation with enough intensity to possibly harm DNA.

The remaining paper is organized as follows. Section 1 describes the history of artificial intelligence and usage in CAD. Section 2 explain the medical imaging modalities. In Sect. 3, we explain the development of Computer Aided Diagnosis (CAD) and usage and Sect. 4 elaborate the selection criteria of Literature review. In Sect. 5, we explain the deep learning in depth and further we identify the major CNN model in Sect. 6. In Sect. 7, we explain the major development in medical image analyses using deep learning and Sect. 8 elaborate the anatomical structure of medical images. Further, in Sect. 9, we explain the deep learning model development for medical image analysis and Sect. 10 depict the basic tool used in deep learning. Section 11 summarized the extensive survey.

2 Medical Imaging Modalities

Each imaging technique in the healthcare profession has particular data and features. As illustrated in Table 1 and Fig. 1, the various electromagnetic (EM) scanning techniques utilized for monitoring and diagnosing various disorders of the individual anatomy span the whole spectrum. Each scanning technique uses a distinct frequency and wavelength and exhibits various qualities [8]. EM waves are dispersed, mirrored, or received by an item whenever they come into contact with it. The magnetic field produced

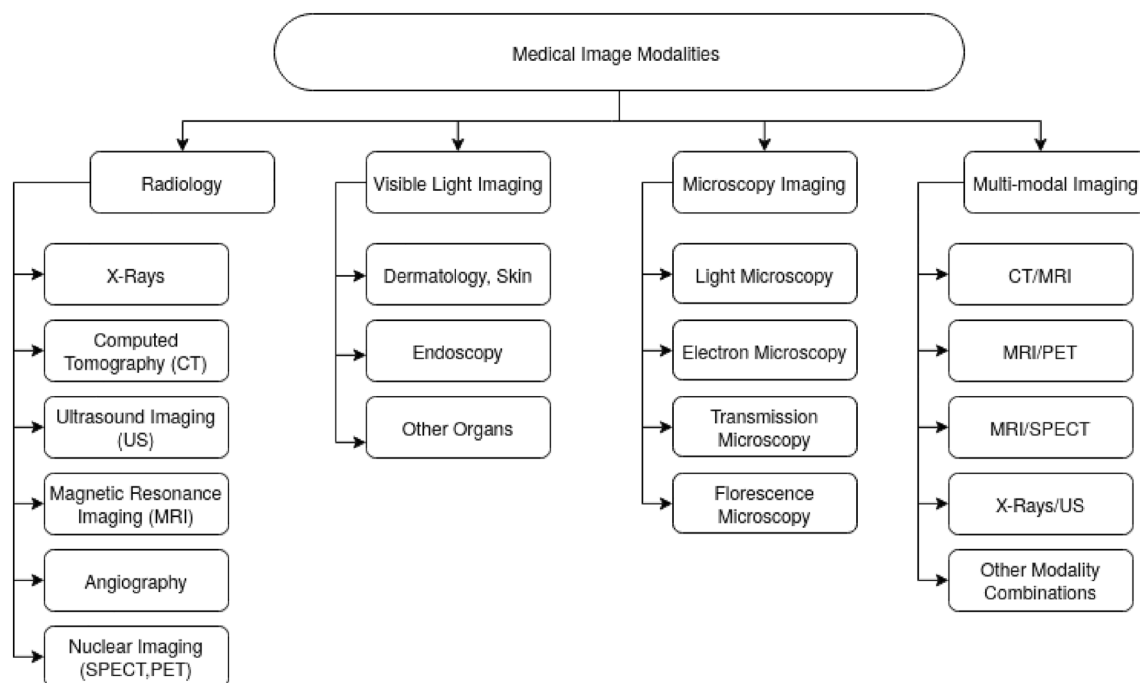


Fig. 1 Medical imaging modalities classification

Table 1 The most popular and prominent medical imaging modalities are compared

Modalities	Invasive/non-invasive	Details
X-Rays	Non-invasive	The primary method of medical scanning was radiography, which was made possible by Sir Wilhelm Roentgen's discovery of radiographs on November 8, 1895. The first radiography representation of the human anatomy was also created by Roentgen. Radiology, which is another name for radiology, established the discipline and given rise directly to radiologists, healthcare practitioners with a focus on analyzing medical images. Anatomical diagnoses of the biological anatomy are made using X-rays. Data from X-rays make it simple to identify healthy, irregular, and shattered bones [10]
Computed tomography	Non-Invasive	Midway through the 1970 s, computed tomography (CT) became widely used in hospitals and was the first medical screening tool that machines recognized. By rotating the X-beam tube all over person and directing X-beams across the person at different locations, the CT image is produced. In order to obtain data on transmissions presentation, at most one straight identifier group positions itself in before of the X-beam emitter. A machine combines the numerous data points amassed in this way to create a tomographic representation of the individual. The word 'tomography' refers to a precise image of a slice (tomo). The transmission breakthrough known as CT can produce images of various cell regions within the person's body. Although CT scans may not reveal a lot of information regarding fatty tissue, they do reveal the cranium, bone formation, significant anomalies, infarction, haemorrhage, and tumors in the brain [11, 12]
Magnetic resonance image	Non-invasive	MRI scans provide excellent information regarding blood circulation and delicate and diseased neural processes. Although MRI can detect minor lesions and strokes, it cannot detect brainwave activity or tumor cells communication [11, 12]. The magnetic flux that MRI Imaging systems employ has a magnitude that ranges from 10,000 to 60,000 roughly that of the earth's magnetic field ¹⁷ . The majority of magnetic resonance imaging employ the highly expressed hydrogen nucleus, or protons, in biological materials for their nuclear magnetic resonance features (about 10 ¹⁸ protons per cubic millimeter of tissue). There is a magnetic charge in particles. Protons typically accept electrons from radio signals with a resonant frequencies of 63 MHz while positioned in a 1.5 Tesla (T) magnetic flux. 18. The individual is positioned in a magnetic flux during magnetic resonance imaging, and a transmitter produces radio waves
Ultrasound	Non-invasive	The wavelength bands of sound energy used by the US extend from around 20 kHz to many GHz. In comparison to X-rays and CT scans, it is less expensive and better. Both operational and anatomical internal components are covered by the US [13–15]. Whenever a paperweight hits the ground, the impact sends forth pressure waves-also known as sound-that may be perceived from a range. Using specific intensity ("ultra") sonic, mechanical power can be employed to create images of the participant's body. Short sound waves are produced when the ultrasonic transmitter through is body interaction with the cell being imaged. Echoes are created when sound frequencies move through the person's cells and are mirrored by its physical structure. The transmitter then receives the sound radiation that was bounced, and it captures the sound stream that was returned. Pulse echo imaging is the name for the way that ultrasound technology operates in this manner [16]
Positron emission tomography	Invasive	PET scans show neural movement as well as glucose metabolism used as brain power. Using various hues, they also provide details about brain tumor cells [11, 12]
Single photon emission computed tomography	Invasive	By examining blood circulation, SPECT images reveal the section of the brain that is most engaged. The use of SPECT can identify conditions such as dementia, epilepsy, choked blood arteries, etc [11, 12]. Gamma rays are used in the nuclear medicine tomographic imaging technology known as single-photon emission computed tomography. While using a gamma sensor, it is extremely comparable to traditional nuclear medicine planar imaging and can produce full 3D data

by Magnetic Resonance Imaging causes the body's natural protons to coordinate. When a radio-frequency signal is delivered to the individual, protons are triggered and begin to swirl at an angle to the magnetic force. The greater region of 3×10^{16} to 3×10^{19} Hz of Computed Tomography and

X-ray imaging is extremely radiation and detrimental to public health. Gamma rays are used in nuclear imaging techniques like Positron Emission Tomography and Single Photon Emission Computed Tomography to diagnose biological functions in a person's tissues. The frequency of gamma

rays is larger than 1019 Hz, and their wavelength is fewer than 10 picometers. Magnetic Resonance Image and Ultrasound are governed by the non-ionization concept, whereas X-rays, Single Photon Emission Computed Tomography, Computed Tomography, and Positron Emission Tomography are dependent on the ionization theory [9].

3 Computer Aided Diagnosis

Despite an increase in CAD development, relatively several CAD solutions are commonly employed in clinical settings. One of the main causes might be that CAD solutions created using traditional machine learning techniques may not have attained the excellent productivity needed to satisfy radiologist and medical practitioners demands for increasing both diagnosis reliability and operational effectiveness. Given the development of deep learning in numerous computer vision and Artificial Intelligence (AI) based systems over the previous few decades, including text and signal processing, face identification, driverless cars, board games and go, there are unrealistic hopes that deep learning will lead to an innovation in CAD effectiveness and mass adoption of deep-learning-based CAD or AI for a variety of activities in the individual treatment plan. The passion has inspired a large number of deep learning projects and articles in CAD. In this section, we'll talk about a number of problems and obstacles that have arisen while trying to design CAD for diagnostic imaging that is founded on deep learning, as well as the things that should be taken into account when doing so.

The development of CAD based applications uses techniques for conventional machine learning. Image processing techniques were utilized in the traditional machine learning methodology to CAD in diagnostic imaging to identify abnormalities of cancer and discriminate between several categories of features, such as healthy or aberrant, cancerous or mild, on the images. Depending on subject matter expertise, CAD designers create image analysis and separation of features algorithms to describe the image features that may differentiate between the numerous states. The competence of the theoretical conceptions or practical image processing approaches that are meant to convert the image features to quantitative data, as well as the subject experience of the CAD designers, are frequently factors that affect how well the feature classifiers perform. A method employs the extracted features as incoming response variable and a forecast strategy is created by varying the weights of the different characteristics based on numerical characteristics of a series of training instances to determine the likelihood that a image corresponds to a particular categories. Although if they have observed a plethora of cases from the user community, the individual programmer

might not be capable of converting the complicated illness trends into a limited amount of feature descriptors using the traditional machine learning technique. The hand-engineered characteristics could also have trouble standing up to the public's wide range of ordinary and atypical behaviors. The generated CAD system frequently performs poorly in terms of classifier or generalization, leading to a large percentage of wrongful convictions at great sensitivity or conversely.

In several domains, deep learning has become the cutting-edge machine learning technique. Deep learning is a kind of representation learning technique that employs a sophisticated multi-layer neural network topology autonomously trains data interpretations by abstracting the raw data into several layers. Deep convolutional neural networks (DCNN) represent the most widely utilised deep learning systems for sequence identification applications in images. By continuously modifying its parameters through training algorithm, DCNN may be taught to autonomously retrieve pertinent features from the training instances for a specific job. CNN model does not demand explicitly generated features as input because feature representations are discovered during training. The DCNN features are anticipated to outperform hand-engineered features in terms of selectivity and in-variance if adequately trained using a sizable training set that is representative of the population of interest. Deep learning will rapidly examine dozens or hundreds of examples that even individual specialists would not be capable of seeing and memories in their lifetimes since the training procedure is mechanized. As provided as the training dataset is big and diversified sufficient for it to evaluate, deep learning can consequently be highly resilient to the vast scope of changes in characteristics across distinct groups to be discriminated [17]. Machine learning algorithms are well-known to discover and learn the relationship between data and explore to retrieve hidden information from the data. Machine learning have influential techniques and methods that can learn from previous actions. The machine learning algorithms observe and interact with environment to improve the efficacy of objective functions.

Image recognition, tracking and identification is an essential research area in machine learning that is used in a wide range of technologies such as gesture identification, driver-less cars, medical assessment such as medical image processing, tumor cell categorization, imagegrammetry and so on. Object spotting is the task of locating many objects in images and recognizing their locations and categorization. Object identification is regarded to be one of the challenging tasks in the area of machine sensing since the appearance of objects varies substantially based on a range of factors [18, 19].

All machine learning technologies and associated artificial intelligence (AI) models, medical evidence and image analysis may offer the highest opportunity for

creating a significant, long-term impact on individual experiences in a relatively short period of time [17]. Image searching, production, computer vision and image-based visualization are all components of the software evaluation and interpretation of medical images [20]. In numerous dimensions, medical image analysis has evolved to encompass image preprocessing techniques, pattern matching, data mining especially image based dataset and deep learning [21]. Deep learning is a popular way for determining the correctness of the future situation. This offered up additional possibilities for interpretation of medical images. Computational intelligence approaches in healthcare handle a diverse range of concerns, from early diagnosis to infection tracking to individualized therapy recommendations. A vast amount of material is now available to clinicians via multiple media platforms such as mri scans, genetic sequencing and pathological imaging [22]. To convert each of this input into valuable advice, the particularly related utilized for patient data include X-ray, Ultrasound, Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) [23, 24]. Deep learning is the process of discovering trends in complex objects by utilizing neural network models comprised of multiple convolution layers (comprising several nodes) of artificial neurons [25]. An artificial neuron is a sort of cell that act like a real brain, accepts numerous input images, does a computation and then delivers the optimal outcome [26–28]. This straightforward approach uses a nonlinear activation function before a linear input pattern matching form [26]. Numerous often deployed nonlinear activation functions of a system include the sigmoid transition, ReLU and their variations and tanh (hyperbolic tangent) [29–31].

In recent years, advances in health science and image processing have been made thanks to deep learning algorithms [26, 32–34]. Deep learning techniques using CNNs and computational techniques, in especially, execute better when analyzing large amounts of data, and this has received a lot of study interest. Recent studies imply that the application of deep-learning-based computer-aided identification in medical practice can dramatically decrease the amount of time needed to analyze films and increase diagnosing effectiveness [35, 36].

For healthcare system, numerous research have explored CADe/CADx technology [37–39]. Studies that have already been published examined CADe/CADx methods for medical images using deep learning algorithms [40, 41]. All of these investigations, meanwhile, are missing explanations of the imaging procedures diagnostic procedures and supporting documentation used in the different engineering solutions. Additionally, released assessments must offer a much more thorough assessment of contemporary literature. As a result, we thoroughly review the most recent state-of-the-art

implementations of cutting-edge CADe/CADx ideas that address medical images using CNN deep learning algorithms and computational techniques. Due to the dearth of papers using additional models from deep learning in CADe/CADx, we did not include them. We also creatively divide the four phases in the conventional CADe/CADx of medical images into two phases by gaining knowledge from the diagnostic imaging methodology in order to straightforwardly relate their pivotal technology superior properties, taking into account that the primary technology components of these research findings lie in various medical steps. By discussing the implementation, benefits, and drawbacks of CNN in the identification of medical images, as well as potential approaches for investigators to address these challenges, we may indicate the path of future study in this area and potentially other healthcare domains.

4 Literature Review Selection Criteria

For this analysis benchmark, there were four steps: (1) keyword search queries in the IEEE Xplore, ACM Digital Library, Scopus, Google Scholar, Science Direct, PubMed, and Web of Science libraries and databases; (2) gathering relevant things and eliminating multiple copies; (3) choosing the benchmark of use; only the lung nodule detection technology based on CT image deep learning was maintained; and (4) evaluating detection systems utilizing the established metrics. The database search terms and logical expressions that we used are listed below: ‘deep learning’ or ‘deep convolutional neural network’ or ‘convolutional neural network’ or ‘CNN’ or ‘DCNN’ and ‘healthcare’ or ‘health-care’ or ‘health care’ or ‘medical’ or ‘clinical’ or ‘image’ or ‘images’ or ‘brain’ or ‘brain injury’ or ‘head’ or ‘head injury’ or ‘skin’ or ‘breast’ or ‘chest’ or ‘pulmonary’ or ‘lung’ or ‘lungs’, and ‘nodule’ or ‘nodules’ or ‘tumor’ or ‘tumors’ or ‘cancer’ and ‘detection’ or ‘detect’ or ‘detected’ or ‘detecting’ or ‘computer-aided detection (CADe)’ or ‘computer-aided diagnosis (CADx)’ or ‘CAD’ or ‘CADx’ or ‘CAD’ or ‘CADe’ and ‘histology’ or ‘histopathology’ or ‘histopathological’ or ‘X-ray’ or ‘Xray’ or ‘CXR’ or ‘MRI’ or ‘Magnetic Resonance Imaging’ or ‘Computed Tomography Scan’ or ‘CT Scan’ or ‘CT-Scan’ or ‘Computed Tomography’ or ‘CT’.

5 Deep Learning Methodologies

Deep learning-based medical image segmentation is a popular topic in image classification, registration, segmentation and tumor detection research and has great use in the medical field. Deep learning technology can improve computer-aided diagnosis accuracy and efficacy

while also easing resource constraints in healthcare, decreasing doctor stress, and reducing reliance on expert knowledge. An overview of some of the most well-known deep learning frameworks is provided below.

The purpose of this part is to formally introduce and define the deep learning ideas, methods, and frameworks that we discovered in the articles on medical image interpretation that were analyzed for this study.

5.1 Neural Network

The fact that neural networks are generic function capable of approximating, or that they are able estimate any mathematical expression to any degree of precision, is a crucial characteristic of neural networks. In other sense, if any process—whether biological or not—can be conceptualized as a function of a collection of variables, then that behavior may be simulated to any unlimited amount of precision, constrained only by the magnitude or complication of the system. Although the aforementioned concept of general approximation is not theoretically precise, it does illustrate one factor that has contributed to the long-standing curiosity in neural networks. This assurance, however, does not offer a mechanism to identify the neural network model's ideal characteristics that will yield the closest estimation for a particular dataset. Additionally, there is no assurance that the system will deliver precise forecasts for fresh data. All artificial neural models' underlying components are synthetic neurons. A mathematical expression that specifically translates sources into outcomes is all that constitutes an artificial neuron. Any amount of input numbers are accepted by a single artificial neuron, which then processes them using a particular mathematical expression to produce an outcome [42].

A network system which is used in machine learning is known as Neural Network. It took inspiration from human brain and works similar to human brain. The network architecture of Neural Network is made up of Artificial Neurons. It is a network that has weights on it, you can adjust those weights so that it can learn from it. A neural network has a number of layers which groups the number of neurons together. Each of them has its own function. Network's complexity depends on the number of layers. That is why the Neural Network is also known as multi-layer perceptron. There are three types of neural network layers. (1) Input Layer, (2) Hidden Layer and (3) Output Layer. Each of them has its own specific purpose. These layers are made up of nodes and each of them has its own domain of knowledge. Neural Networks are highly efficient because they can learn very quickly.

5.2 Multi-layer Perceptrons

An neural network model's another very basic structure is layer upon layer of densely integrated and interconnected neurons. In this structure, a set quantity of "input neurons" stand in for the input feature values that are determined from the records and transmitted to the sub net, and each linkage among a couple of neurons stands in for one learnable weight parameter. Artificial neural learning refers to the process of maximizing these components, which are the primary variables that may be changed in a neural network [43–45].

The network's ultimate results are represented by a quantity of output nodes at the opposite end of the network. When properly set up, a network of this kind may be used to create hierarchical, sophisticated judgments regarding the input since each neuron in a particular layer obtains information from every neuron in the layer below. The earliest networks helpful for bioinformatics implementation were layers of neurons arranged in this straightforward manner; these layers are sometimes referred to as "multilayer perceptrons."

5.3 Feed Forward Neural Networks

The earliest and most fundamental model of a neuron is the perceptron. A group of inputs are taken, added, and then an activation function is applied before sending the results to the output layer. A fundamental class of neural networks is called a feed forward neural network (FFNN). The intermediate levels are buried, with the input layer at the top and the output layer at the bottom. There is no feedback in the entire network as the signal propagates unidirectional from the input layer to the output layer [46–48].

Multilayer perceptrons is another name for Feed Forward Neural Network depicted in Eq. 1.

$$y_l = \sum_{i=0}^n (W_i * X_i) + b \quad (1)$$

5.4 Recurrent Neural Networks (RNNs)

When processing data with time series characteristics, RNN excels. It can also help with data analysis and mining for information about time series features and semantic change. The weight of the single unit that makes up the RNN layer is shared. Each training sample example will only go through one unit in the state of the various time series, after which its weight will be changed continuously [49–51].

5.4.1 Long-Short Term Memory (LSTM)

The vanishing/exploding gradient issue is solved with gates and an explicitly designated memory cell in Long-Short Term Memory (LSTM) networks. The main driving force behind these is electronics, not biology. Input, output, and forget gates, together with a memory cell, are all present in every neuron. These gates' function is to safeguard information by controlling the flow of it or blocking it [52–54]. The Eq. 2 show the LSTM's feedforward calculation procedure in more detail.

$$\begin{aligned}
 \mathbf{I}_t &= \sigma(\mathbf{X}_t \mathbf{W}_{xi} + \mathbf{H}_{t-1} \mathbf{W}_{hi} + \mathbf{b}_i), \\
 \mathbf{F}_t &= \sigma(\mathbf{X}_t \mathbf{W}_{xf} + \mathbf{H}_{t-1} \mathbf{W}_{hf} + \mathbf{b}_f), \\
 \mathbf{O}_t &= \sigma(\mathbf{X}_t \mathbf{W}_{xo} + \mathbf{H}_{t-1} \mathbf{W}_{ho} + \mathbf{b}_o), \\
 \tilde{\mathbf{C}}_t &= \tanh(\mathbf{X}_t \mathbf{W}_{xc} + \mathbf{H}_{t-1} \mathbf{W}_{hc} + \mathbf{b}_c),
 \end{aligned}
 \tag{2}$$

Among them, W_{hi} , W_{hf} , W_{ho} , W_{hi} represent the matrix parameters related to the input and 3 Gates, and then b_i , b_f , b_o , b_c represent the bias parameters related to the input and the three Gates, represents the Sigmoid function, and finally, 0 represents the same position of the two vectors Elements are multiplied together.

5.4.2 Gated Recurrent Units (GRU)

Although the gradient problem in RNN can be greatly reduced by LSTM, a single LSTM unit has four times as many parameters as a single RNN unit due to the addition of three extra Gates. Greater parameters necessitate more computing power. Although it has fewer parameters than LSTM because it lacks an output gate, GRU is similar to LSTM with forget gate [55].

A very powerful LSTM neural network variant is the GRU. GRU makes the structure shallower and computationally less expensive while maintaining the LSTM effect. It is also the most used type of neural network at the moment because it has a simpler structure and better results than an LSTM network. Due to the fact that GRU is a variation of LSTM, it can alleviate the lengthy reliance issue in RNN [56]. The GRU feedforward calculation is displayed in Eq. 3.

$$\begin{aligned}
 \mathbf{R}_t &= \sigma(\mathbf{X}_t \mathbf{W}_{xr} + \mathbf{H}_{t-1} \mathbf{W}_{hr} + \mathbf{b}_r), \\
 \mathbf{Z}_t &= \sigma(\mathbf{X}_t \mathbf{W}_{xz} + \mathbf{H}_{t-1} \mathbf{W}_{hz} + \mathbf{b}_z),
 \end{aligned}
 \tag{3}$$

\mathbf{W}_{xr} , $\mathbf{W}_{xz} \in \mathbb{R}^{d \times h}$ represent the input training parameters and \mathbf{W}_{hr} , $\mathbf{W}_{hz} \in \mathbb{R}^{h \times h}$ are bias parameters, and the two Multiply elements of the same position of the vector, respectively.

5.4.3 Bidirectional Recurrent Neural Networks

Bidirectional Long Short-Term Memory Networks (BiLSTM), Bidirectional Gated Recurrent Units (BiGRU), and Bidirectional Recurrent Neural Networks (BiRNN) all resemble their unidirectional counterparts in appearance. On the other hand, conventional RNNs cannot process data for the future and can only process input in one direction. The forward, backward, and output Eq. 4 indicate how these bidirectional networks can extract complete temporal information at time t by combining past and future data, improving the model's performance on sequence issues [57, 58].

$$\begin{aligned}
 \bar{\mathbf{H}}_t &= \phi(\mathbf{X}_t \mathbf{W}_{xh}^{(f)} + \bar{\mathbf{H}}_{t-1} \mathbf{W}_{hh}^{(f)} + \mathbf{b}_h^{(f)}), \\
 \bar{\mathbf{H}}_t &= \phi(\mathbf{X}_t \mathbf{W}_{xh}^{(b)} + \bar{\mathbf{H}}_{t+1} \mathbf{W}_{hh}^{(b)} + \mathbf{b}_h^{(b)}), \\
 \mathbf{O}_t &= \mathbf{H}_t \mathbf{W}_{hq} + \mathbf{b}_q.
 \end{aligned}
 \tag{4}$$

5.5 Unsupervised Models

Unsupervised learning research's major objective is to pre-train a deep learning model (also known as a "discriminator" or "encoder") that will be utilized for many other challenges. The encoder characteristics must be broad sufficient to be applied to classification techniques, such as training on ImageNet and producing outcomes that are as near to supervised models as feasible [59].

As of right now, supervised models consistently outperform unsupervised trained models. This is to ensure that the model can more effectively incorporate the dataset's features thanks to the supervision. However, if the model is subsequently extended to other activities, monitoring may likewise become less effective. Unsupervised training is hoped to be able to offer more universal characteristics for learning to complete any work in this respect [60, 61].

5.5.1 Autoencoders (AEs)

In that they are an adaptation of FFNNs rather than a fundamentally unique design, autoencoders (AEs) are similar to FFNNs in this regard. The only thing to keep in mind is that the number of input characteristics (number of neurons) in the input layer and the output layer should match (the number of neurons). Check to see if the input and output are both equal. The basic idea behind autoencoders is to automatically compress data rather than encrypt it, hence the name. The entire network has hidden

layers that are thinner than the input and output layers and is structured like an hourglass [62, 63]. The Variational Autoencoders (VAEs) is depicted in equation ??.

$$\partial(q) = E_{zq(z|x)}(\lg p(x|z)) - KL(q(z|x)||p(z)) \quad (5)$$

A unique network called Sparse Autoencoders (SAEs) adds sparsity to the network. Sparse Autoencoder wants to use the original data to create a low-dimensional representation. The preparation of the image typically involves sparse automatic coding to decrease the size of the data and retrieve possibly useful information.

5.5.2 Boltzmann Machine (BM)

Unbalanced connections make up a Boltzmann machine. In terms of graph theory, it is a full graph. Any device is linked. Neurons and other components will make the decision to switch on or off. Initially, BM was solely employed to refer to models that contained just binary variables. The limited Boltzmann machine primarily adds “restriction” as compared to the Boltzmann machine. To make the complete graph a bipartite graph is the alleged restriction. In addition to other things, restricted Boltzmann machines may be employed to automatically learn (hidden state outcomes are features), develop deep belief networks, and decrease complexity (fewer hidden layers) [64, 65].

5.5.3 Deep Belief Network (DBN)

Multi-layer Restricted Bolt Machine (RBM) based neural networks are called Deep Belief Networks (DBN). It can be categorized as either a discriminative or generative model. To put on weight, the training strategy adopts the unsupervised hungry layer-by-layer approach. Deep Belief Network learning has been finished layer by layer. After being employed to estimate the hidden layer in each level, the array of objects is then utilised as the data vector for the following (higher) layer. Numerous Restricted BMs must be “connected in series” to form a DBN, with the outcome of one BM serving as the intake of the following and the hidden layer of the first BM serving as the feature map of the second. During the learning phase, it is essential to thoroughly train the Restricted BM of the former layer before learning the Restricted BM of the current layer to the previous layer. Hua et al. [66, 67].

5.5.4 Generative Adversarial Network (GAN)

Generative Adversarial Networks (GANs) are essentially a training mode and not a final network structure. The fundamental tenet of Generative Adversarial Network is that the discriminator and generator work in tandem. The

discriminator must attempt to differentiate among real samples (such as real pictures) and fake samples produced by the generator, or “Fake images.” The ideal competitive environment is one where both sides continuously improve, where the capability to differentiate between them becomes increasingly stronger, and where the capacity to produce deception becomes greater as well. The outcome of the opposition is unimportant. What matters most in the conclusion is the generator’s capacity to produce instances that are sufficiently equivalent to the real instances, resemble the sample data in the training set, and have a dispersion that is substantially close to that of the training examples distribution. Xin et al. [68, 69]. Define the objective function that GAN must optimize, which is represented by the expression in Eq. 6.

$$\min_D \{-y \log D(\mathbf{x}) - (1 - y) \log(1 - D(\mathbf{x}))\}, \quad (6)$$

5.6 Convolutional Neural Networks (CNNs)

An uncommon variety of neural network is called a convolutional neural network (CNN) or a deep convolutional neural network (DCNN). They can be used for other types of input, such as audio, though image processing is where they are most frequently used. Convolutional neural networks (CNNs/ConvNets) use neurons with biases and weights that can be learned. The classification score is generated after each neuron computes the dot product using the data it has been given. By incorporating specific characteristics into the network structure using CNN, we can improve the effectiveness of the feedforward function and pass parameters [70]. People who share reduce many variables. Equation 7 explain the input and output capabilities of the CNN as well as the network’s overall non linearity and cost function are depicted in Eqs. 8 and 9 respectively.

$$x_{ij}^{\ell} = \sum_{a=0}^{m-1} \sum_{b=0}^{m-1} \omega_{ab} y_{(i+a)(j+b)}^{\ell-1} \quad (7)$$

$$y_{ij}^{\ell} = \sigma(x_{ij}^{\ell}). \quad (8)$$

$$\frac{\partial E}{\partial \omega_{ab}} = \sum_{i=0}^{N-m} \sum_{j=0}^{N-m} \frac{\partial E}{\partial x_{ij}^{\ell}} \frac{\partial x_{ij}^{\ell}}{\partial \omega_{ab}} = \sum_{i=0}^{N-m} \sum_{j=0}^{N-m} \frac{\partial E}{\partial x_{ij}^{\ell}} y_{(i+a)(j+b)}^{\ell-1} \quad (9)$$

5.7 Convolutional Layer

Convolutional Layer is a first layer of Convolutional Neural Network. This layer consists of sets of Filters or Kernel. Their job is to use a Convolutional operation to the input and passing the result to the succeeding layer. The filter takes a

subset of the input data. The territorial relationship between pixels by learning image options using tiny squares of input data ensures by this layer. The Convolutional layer's vector is the final output. This layer performs linear multiplications with the goal of extracting the input image's high-level characteristics as a convolution operational activity. Due to the linear nature of the convolutional process at this layer, the final size is also produced by layering the activation maps of all filters on the depth dimension. Similar to an old neural network, linear operation primarily entails the multiplication of weights with the input [71, 72].

5.8 Deconvolution

Deconvolution also known as transposed convolution is a mathematical operation by which the effect of convolution reverses. It is exactly the multivariate Convolutional function's inverse. For example, giving input to the Convolutional layer and getting output then give the same output to deconvolutional layer can get you to the same input you given first. Deconvolution is just to reverse the input back to large size [73, 74].

5.9 Dilated Convolution (Atrous Convolution)

A type of convolution by which the kernel inflates by inserting holes between the kernel elements. Another parameter to the Convolutional layers has been introduced by dilated convolution called the dilation rate. Additional parameter of dilation rate as indicates that how much the kernel is expended. It's targeted to increase the size of reception field by avoiding the increase in parameter sizes. There are normally spaces inserted between the elements of kernel [75].

5.10 Striding

Striding defines the step size of the kernel while sliding through the image. Stride of one defines that the kernel slides through the image pixel by pixel. Stride of two defines that the kernel slides through image by moving 2 pixels per step i.e., skipping of 1 pixel. Stride (≥ 2) can be used for down-sampling an image [76].

5.11 Padding

How the border of an image is handled is defined by padding. Spatial output dimensions kept equal to the input image by padded convolution, by padding 0 around the input boundaries (if necessary). On the other side, without adding 0 around the input boundaries, unpadded convolution only perform convolution on the pixels of the input image [77]. The size of output is smaller than the input size. For an input

image, size i , kernel size k , padding p and stride s , from convolution the output image has size o :

$$o = \text{lower_bound}((i + 2p - k)/s) + 1 \quad (10)$$

5.12 Pooling Layer

A new layer added after the Convolutional layer is a pooling layer. Specially, after the Convolutional layer applies a nonlinearity to the feature maps output. The inclusion of pooling layer right after the Convolutional layer is a usual pattern used in the ordering of layers in a convolutional neural network and can be repeated once or more then once in a given model. The pooling layer separately operates upon each feature map so that it can create a new set of pooled feature maps of same number. Pooling involves selection of pooling operation just like a filter going to be applied on feature maps. Size of pooling filter or operation is smaller in size than the size of the feature map. Mainly, it is always 2×2 pixels applied with a 2 pixels stride. It means that pooling layer always going to reduce the size by a factor of 2 of each feature map [78]. Two of the common functions used in pooling operation are given below:

1. *Average Pooling* A convolutional neural network's layers repeatedly apply learnt filters to input images to produce feature maps that list the features present in the image. For each region on the feature map, determine the average value. In order to construct a down-sampled (pooled) feature map, the average value for regions of a feature map is calculated using the average pooling method. It is frequently applied following a Convolutional layer. It provides a tiny bit of translation in-variance, which means that changing the image's size slightly has little impact on the numbers of the majority of pooled outcomes. Max Pooling collects higher obvious characteristics like edges, although it collects characteristics more uniformly.
2. *Maximum Pooling* A feature map is down-sampled (pooled) by using the Max Pooling pooling procedure, which determines the highest value for regions of the feature map. It is frequently used following a Convolutional layer. It adds a tiny bit of translation in-variance, which means that changing the image's size slightly has little impact on the numbers of the majority of pooled outcomes [79].

5.13 Fully Connected Layer

Fully connected layer is the very important component of the neural networks. It is very successful in the field of recognizing and classifying the images. Fully connected

layer is the classic neural network architecture. Fully connected layers are those layers where all the inputs from one layer is connected to every activation unit of the next layer. In this layer, all the input units have a separable weight to each output unit. There are two of fully connected layer one is for input and other is for output. The fully connected input layer flattens the input and change it into the vector for the input of next stage. Then the next stage analyzes it and apply weight to project the right label. Then the fully connected output provides the expected labels to each label [80].

5.14 Activation Functions

Mathematical equations that determine the output of a neural network are known as activation functions. The function is attached with each and every neuron in the network. It determines whether it should be fired (activated) or not based on whether the input of each neuron is relevant for the model's prediction. Activation functions are also help normalize the output of each neuron to a specific range between -1 and 1 or between 1 and 0. Activation functions must be computationally efficient because for each data sample they are calculated across thousands or millions of neurons. Modern day neural networks use a backpropagation technique to train the model which places increased computational strain on activation function and its derivative function [81].

5.15 Batch Normalization

Batch normalization is also known as batch norm. It's a layer which allows each and every layer of network to do learning more independently. Its use is to normalize the previous layers output. The input layer is scaled by the activation in normalization. By using batch normalization, learning becomes more efficient. Moreover, to avoid over-fitting of a model it can be used as a regularization. To standardize the input or the outputs, the layer is added to the sequential model. It can be used at some several points in between layers of model. It is placed right after defining sequential model and after the convolution layers and pooling layers [82].

5.16 Dropout

The dropout term relates to "dropping out units" in a neural network. The thing dropping a unit out means removing the unit temporarily from the network and all the connections whether incoming or outgoing all are removed. To prevent neural networks from over-fitting the Dropout method is used. In simple way dropout ignores the units during the time period of training of the system [83].

5.17 Softmax

SoftMax is a function, it is additionally referred as soft argmax or normalized exponential function. This function which is used as the activation function in output layer of the neural network models that predict the multinomial probability distribution. It is a function that changes the K real value into the vector of K real value whose sum is 1. The input value can be anything like positive, negative less than or greater than one, but this function changes the value to 0 and 1, for the probability. If the input is greater than this function will change the value in large probability or if the input is lesser than the function will change the value in small probability, but the value will always lie between 0 or 1 [83].

5.18 Optimizer

Optimizer is the techniques or algorithms which are used to change the attributes of the neural networks. They are used to solve optimization issues by minimizing the function. Optimizer is the one which is used to reduce the losses and to provide us with accurate results [84]. Some strategies are used to initialize the weight and with the period of time it is updated by using the following equation:

$$W_{new} = W_{old} - lr * (\nabla W.L)W_{old} \quad (11)$$

This equation is used to update the weight and to get the accurate results.

5.19 Momentum

Neural network momentum is a simple technique or method which improves accuracy and training speed both. Momentum helps in flattening the variations if there is continuous change in the direction of the gradient. The momentum value is used to avoid the situation of getting stuck in local minima. It is the value which is between 0 and 1. If the value of momentum is greater than the learning is kept smaller. The larger value of the momentum is also considered as that the convergence will occur rapidly. The small value of the momentum cannot avoid local minima and can also result in the delay of systems training [85].

5.20 Learning Rate

Learning rate are the hyper-parameters in the configuration of the neural networks. It used in the training of neural networks that has small positive value. It controls the adaptation of the model to the problem. Learning rate controls how much change in the model is required in response to estimate the error when the values of weights are updated. It is very

challenging to choose learning rate because its values are too small which may result in a lengthy training process [86].

6 CNN Model Zoo

6.1 LeNet

LeNet architecture is very compact and simplified. LeNet was introduced by Yann LeCun in 1989 [87]. It is simplified to such an extent that it can be trained on CPU if

you do not have any kind of GPU support but if you have any kind of GPU support for your computer, a lot faster results could be achieved. LeNet has many versions which are LeNet-5, LeNet-4, LeNet-1, Boosted LeNet-4. LeNet is used for handwritten words recognition. Mostly it was used with the applications related to MNIST dataset. It consists upon basic CNN parts like pooling layers, Convolutional layers and fully connected layers [88]. (Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 and 24)

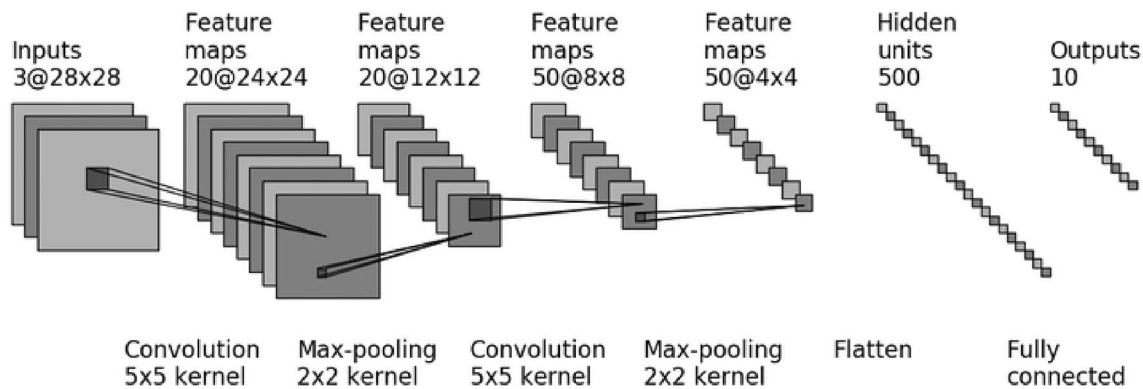


Fig. 2 LeNet is a classic convolutional neural network employing the use of convolutions, pooling and fully connected layers. It was used for the handwritten digit recognition task with the MNIST

dataset (dataset of 60,000 small square 28×28 pixel gray-scale images of handwritten single digits between 0 and 9) [87]

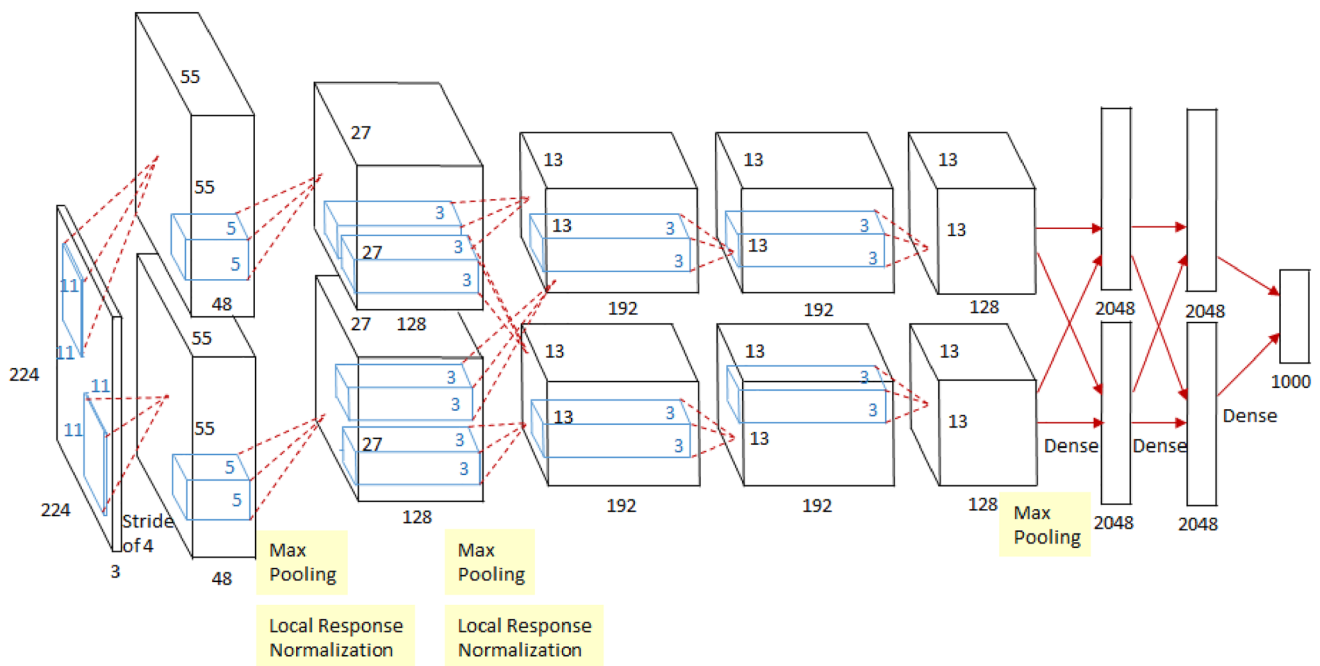


Fig. 3 AlexNet is a CNN architecture and its basic building blocks are max pooling and dense layer. It has imageNet classification with the deep learning techniques. This architecture will be use in pre trained model for the object detection in computer vision [89]

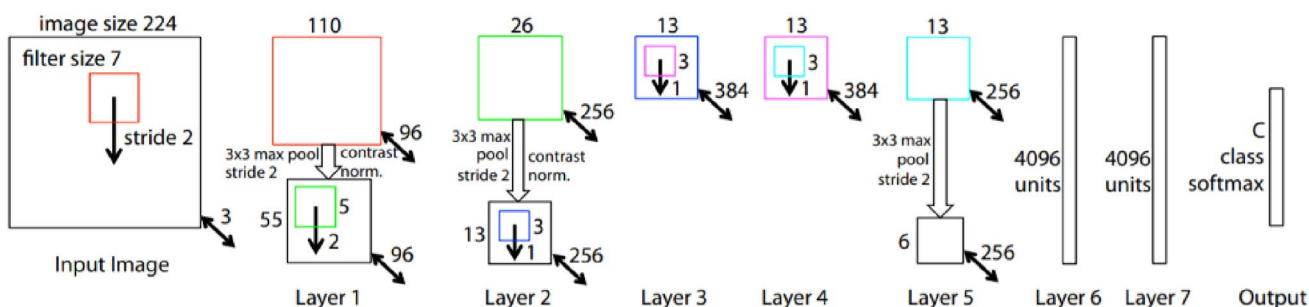


Fig. 4 ZfNet is a classic convolutional neural network. The design was motivated by visualizing intermediate feature layers and the operation of the classifier [90]

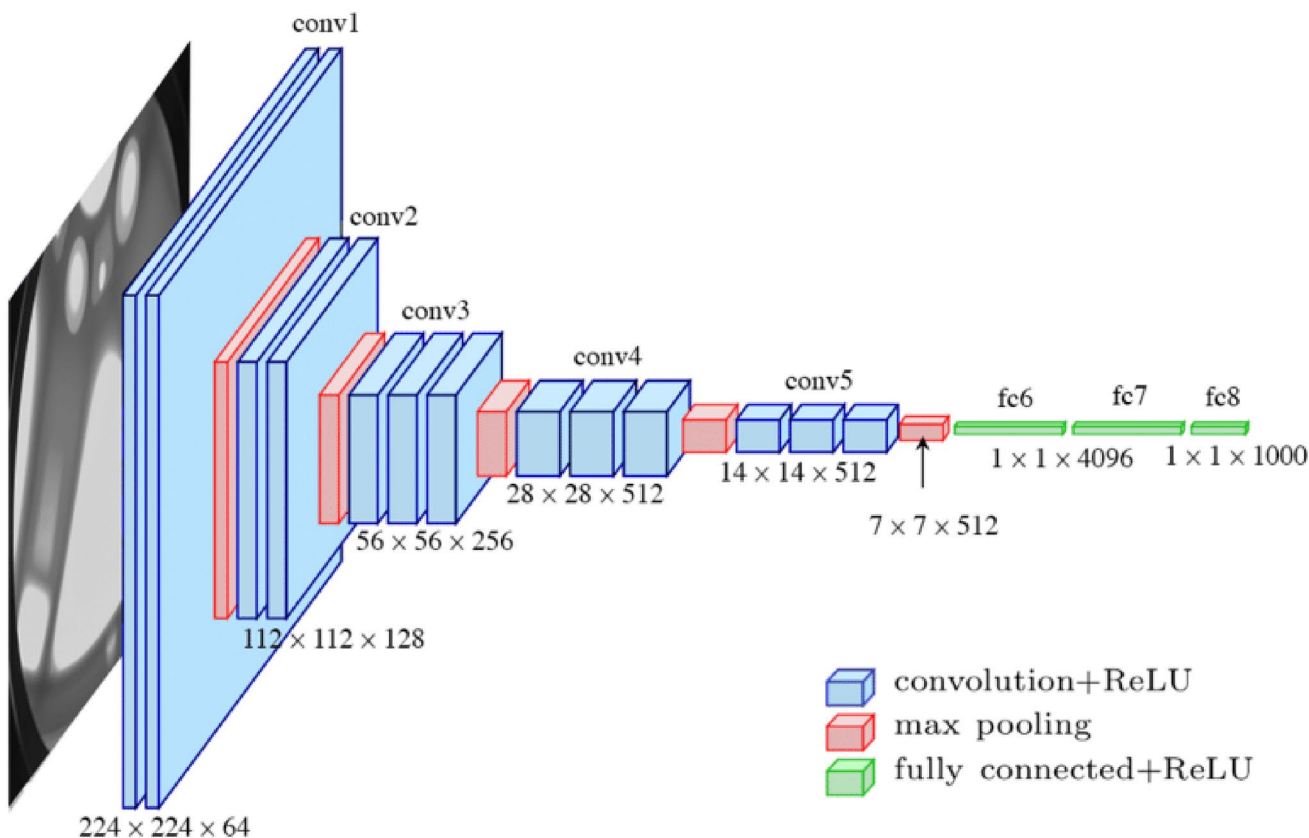


Fig. 5 The convolution neural network receives a fixed-size RGB image with a dimension of 224 by 224. The primary preprocessing it performs is removing out of each pixel the average RGB numbers calculated from the training sample. The image is then sent via a series

of convolutional (Conv.) layers that contain filters with a tiny 3 x 3 input patch, the smallest number necessary to retain the concepts of up/down, left/right and center portion [91]

LeNet-5 CNN architecture is made up of 7 layers. The layer composition consists of 3 Convolutional layers, 2 subsampling layers and 2 fully connected layers.

- The first layer is the input layer. In this layer you have a simple digit range from 0-9. It is a 32 x 32 grayscale image and each grayscale image have a digit range from 0-9. You need to classify what digit image has in it. The

first you used a Convolutional technique in which you are using a filter size of 5 x 5 and stride is 1. The next layer is of size 28 x 28 x 6 where 6 is the no. of filters you've used.

- Next the subsampling technique in which you used an avg pooling where the filter size is 2 x 2 and stride is of 2. The next layer is of size 14 x 14 x 6 where 6 is the no. of filters you've used.

Fig. 6 GoogLeNet is a deep CNN and it has a 22-layer architecture and researchers at Google's motivated by visualizing intermediate feature layers and the operation of the classifier [92]

- Further they employed Convolutional with 5×5 filter and stride is of 1. The next layer is of size $10 \times 10 \times 16$ where 16 is the no. of filters you've used.
- They applied average pooling where the filter size is 2×2 and stride is of 2. The next layer is of size $5 \times 5 \times 16$ where 16 is the no. of filters you've used. Here $5 \times 5 \times 16$ means 400 neurons. Number of layers are 16 and each layer have 25 neurons in it.
- For dense connection, they employ fully connection each 400 neurons to first hidden layer having 180 neurons which have 180×400 connections.
- These 180 neurons are fully connected to the next hidden layer having 84 neurons size of 84×180 connections.
- Finally the output layer will be of 10×10 which is fully connected to the previous layer size of 10×84 connections.

The main disadvantage of LeNet-5 is at time padding technique is not used to it takes some extra effort to retain the size of the image.

6.2 AlexNet

AlexNet CNN architecture is more complex as compared to the LeNet architecture. AlexNet was introduced by Alex Krizhevsky. In 2012, AlexNet competed in ImageNet Large Scale Visual Recognition Challenge. AlexNet won the competition and was at par then the opponent. AlexNet was able to get 15.3% top-5 error which was 10.8% lower than the opponent. AlexNet requires more depth in its architecture as compare to LeNet. AlexNet consist of pooling layers, Convolutional layers and fully connected hidden layer [89].

AlexNet CNN architecture is made up of 8 layers. The first 5 are Convolutional layers, some of them followed by 3 max pooling layer and last 3 are fully connected layers. It also uses ReLU Nonlinearity, multiple GPU's (on two parallel GPU's) and overlapping pooling.

- The first layer is the input layer. It is a $224 \times 224 \times 3$ RGB image. The first you used a Convolutional technique in which you are using a filter size of 11×11 and stride is 4 and pool size of 2. The next layer is of size $55 \times 55 \times 96$ where 96 is the no. of filters you've used.
- Next the pooling technique in which you used a max pooling where the filter size is 3×3 and stride is of 2. The next layer is of size $27 \times 27 \times 96$ where 96 is the no. of filters you've used.



Fig. 7 The really quite initial point we can see in the accompanying graphic is that there is a direct link that bypasses various model levels. The core of leftover blocks is a link known as a “skip connection.” This skip connection causes the outcome to differ. Input 'X is calculated by the layers weight in the absence of the skip connection, and then a bias factor is added [94]

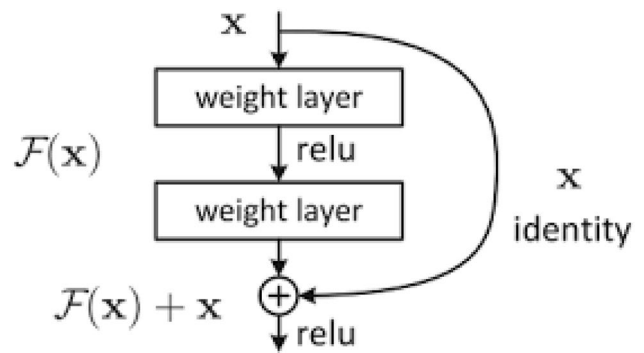
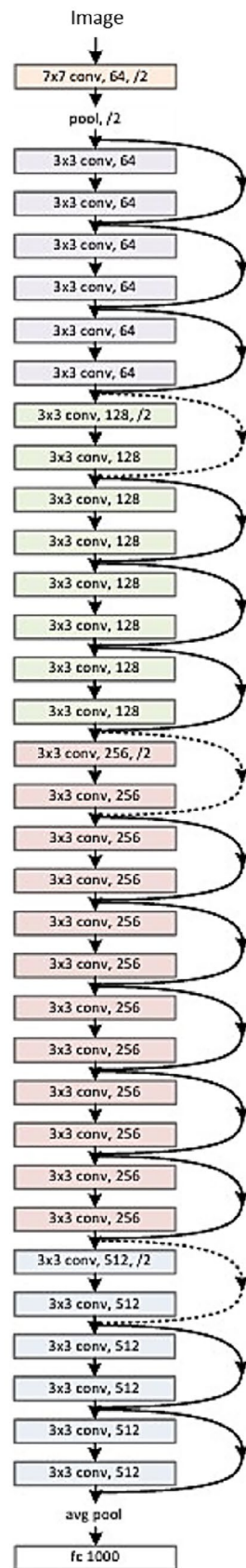


Fig. 8 Residual Block [94]

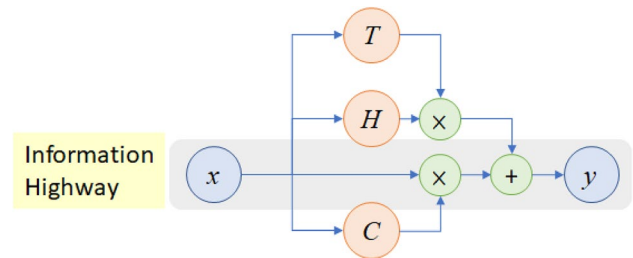


Fig. 9 It enables direct communication between neurons in various levels. Data transmission is regulated by a gating technique. Data can go through various levels of neural networks thanks to gating processes [95]

- They use same Convolutional with $5 \times 5 \times 96$ filter. The next layer is of size $27 \times 27 \times 256$ which means there are total of 256 filters and each filter have $5 \times 5 \times 96$ filters in it.
- Furthermore, they employ max pooling where the filter size is 3×3 and stride is of 2. The next layer is of size $13 \times 13 \times 296$ where 256 is the no. of filters you've used.
- They apply same Convolutional with $3 \times 3 \times 256$ filter. The next layer is of size $13 \times 13 \times 384$ which means there are total of 384 filters and each filter have $3 \times 3 \times 256$ filters in it.
- They engage same Convolutional with $3 \times 3 \times 384$ filter. The next layer is of size $13 \times 13 \times 354$ which means there are total of 384 filters and each filter have $3 \times 3 \times 384$ filters in it.
- Next the same Convolutional with $13 \times 13 \times 384$ filter is used. The next layer is of size $13 \times 13 \times 256$ which means there are total of 256 filters and each filter have $3 \times 3 \times 384$ filters in it.
- They enforce max pooling where the filter size is 3×3 and stride is of 2. The next layer is of size $6 \times 6 \times 256$. Here $6 \times 6 \times 256$ means 9216 neurons. Number of layers are 256 and each layer have 36 neurons in it.

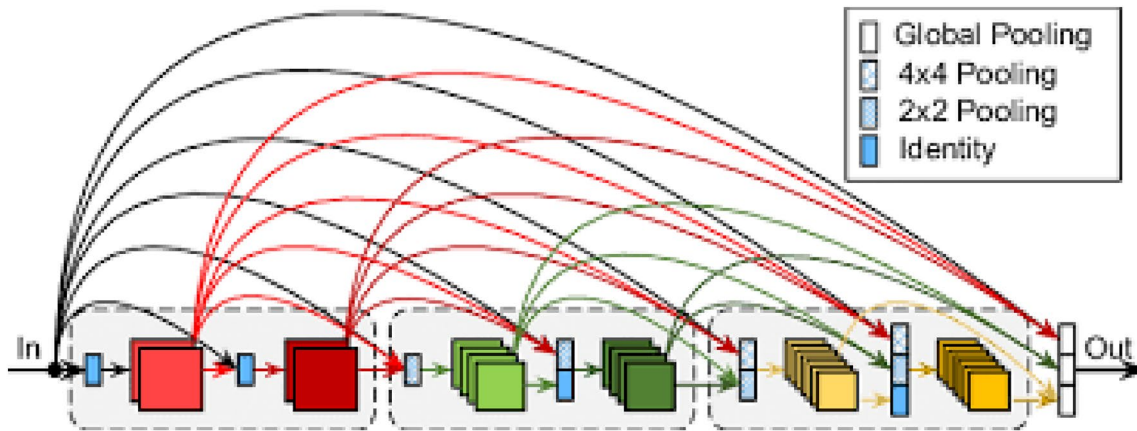


Fig. 10 ResNet and DenseNet are relatively similar, however there are several key distinctions. While DenseNet combine the outcome of the prior level to the next level, it employs the addition technique(+) that combines the previous layer with the subsequent layer [96]

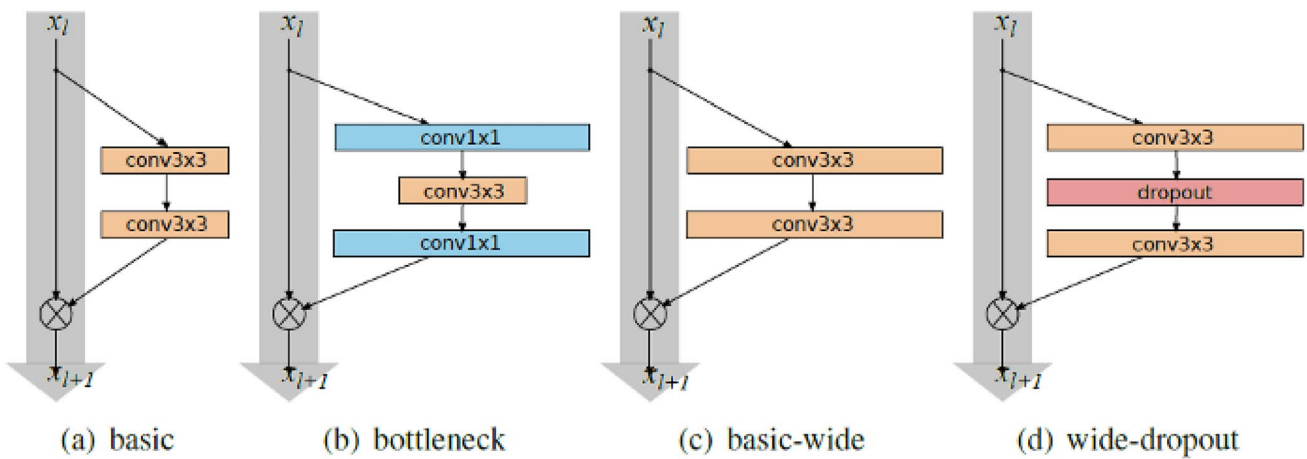


Fig. 11 It has same accuracy as ResNet but number of layers is reduced and training time is shorter. WRN-22-8 outperform ResNet-1001 by 0.92% on CIFAR-10 and 3.46% on CIFAR-100. WRN-40-4 is 8 times faster than ResNet-1001 [97]

Fig. 12 Pyramidal Net: [98]
Two popular approaches top down and bottom-up are used in pyramidal network of CNN

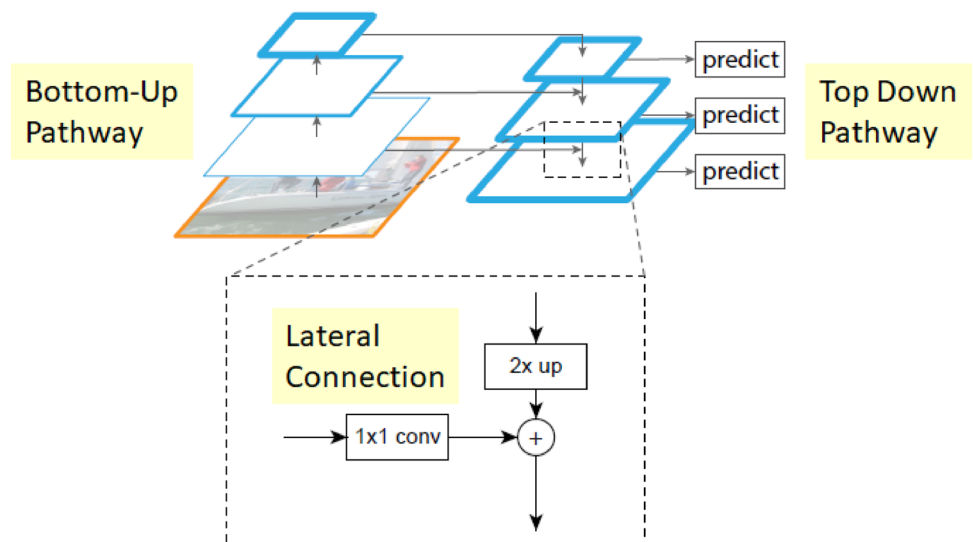


Fig. 13 Inception V1 [99]

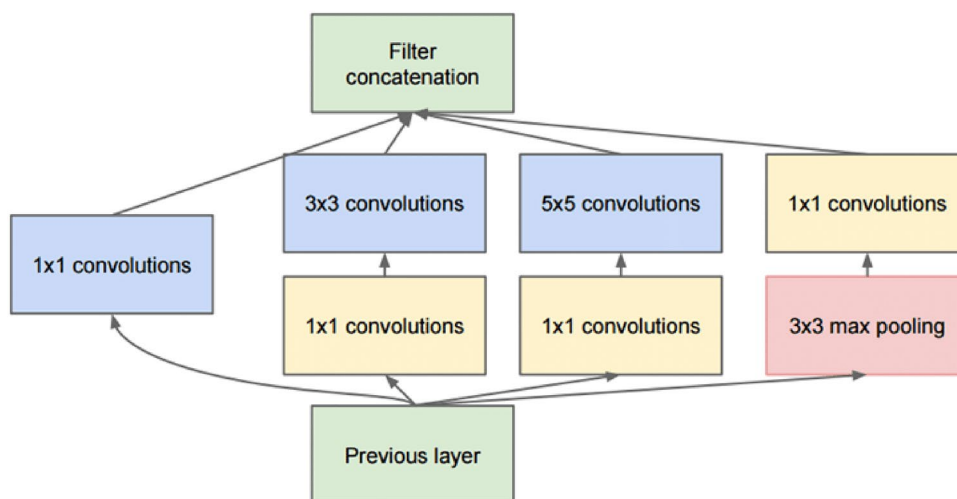
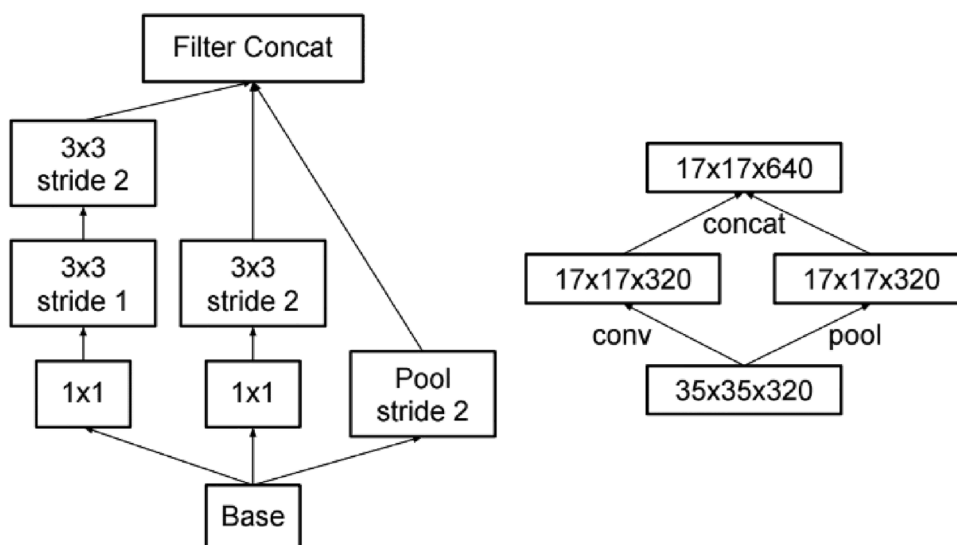


Fig. 14 Inception V2 [100]



- For dense connection, they employ fully connect each 9216 neurons to first hidden layer having 4096 neurons which have 4096×9216 connections.
- These 4096 neurons are fully connected to the next hidden layer having 4096 neurons size of 4096×4096 connections.
- Finally the output layer will be of 1000 neurons which is fully connected to the previous layer size of 100×4096 connections.

The main difference between LeNet-5 and AlexNet is that it uses ReLU activation function that will output the input directly if it is positive, otherwise, it will output zero.

6.3 ZfNet

ZfNet CNN is an improvement over the AlexNet. ZfNet was introduced in 2013 in ILSVRC (ImageNet Large Scale Visual Recognition Challenge). In ZfNet the size of Filter is reduced and Convolutional strides are also reduced. The design of ZfNet came across the motivation of visualizing intermediate feature layers and classifiers operation [90].

- The first layer is the input layer. It is a $224 \times 224 \times 3$ RGB image. The first you used a Convolutional technique in which you are using a filter size of 7×7 and stride is 2 and pool size of 3. The next layer is of size $55 \times 55 \times 96$ where 96 is the no. of filters you've used.
- Used the pooling technique in which you used a max pooling where the filter size is 3×3 and stride is of 2.

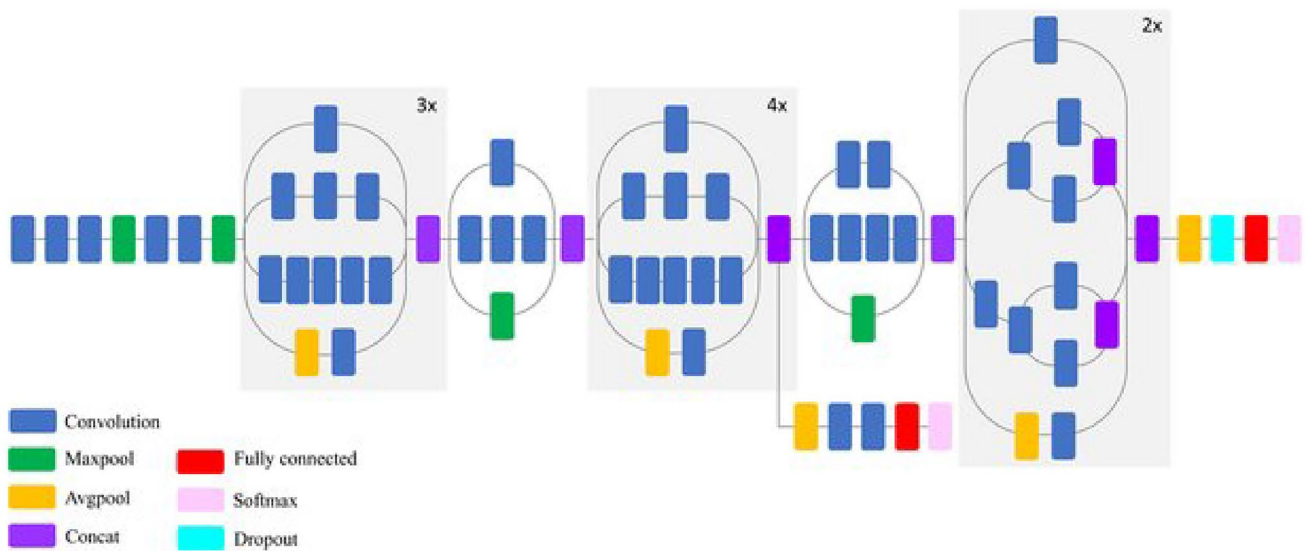
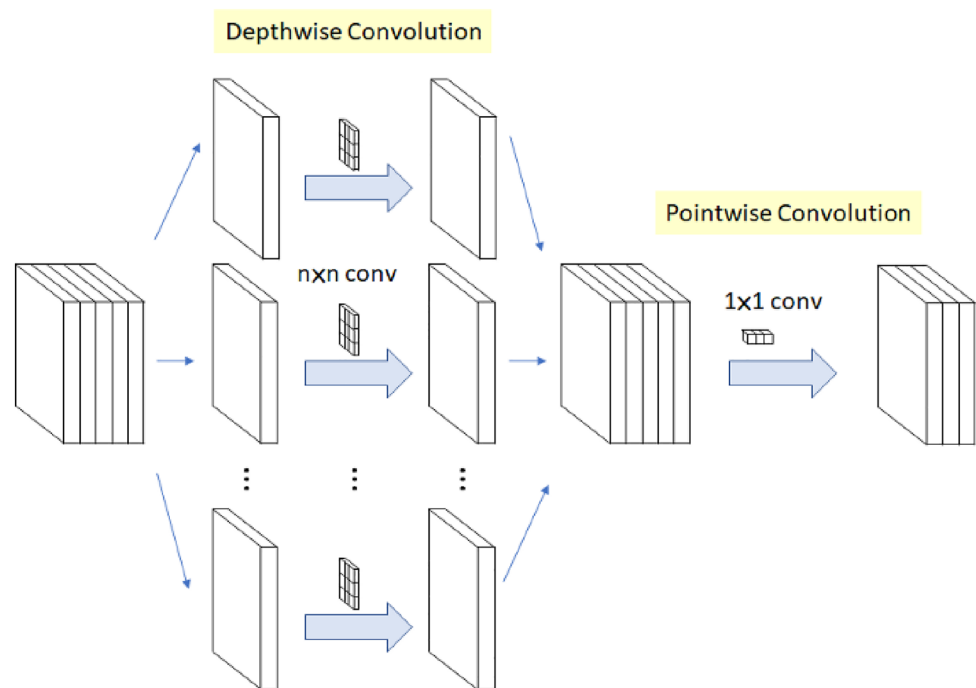


Fig. 15 Inception V3 [101]

Fig. 16 Xception is a convolutional neural network architecture that relies solely on depth wise separable convolution layers [102]



The next layer is of size $27 \times 27 \times 96$ where 96 is the no. of filters you've used.

- Next they employ Convolutional with $5 \times 5 \times 96$ filter. The next layer is of size $7 \times 7 \times 256$ which means there are total of 256 filters and each filter have $5 \times 5 \times 256$ filters in it.
- Further, they apply max pooling where the filter size is 3×3 and stride is of 2. The next layer is of size

$13 \times 13 \times 256$ where 256 is the no. of filters you've used.

- They used interactively used same Convolutional with $3 \times 3 \times 256$ filter. The next layer is of size $13 \times 13 \times 384$ which means there are total of 384 filters and each filter have $3 \times 3 \times 256$ filters in it.
- They employ same Convolutional with $3 \times 3 \times 384$ filter. The next layer is of size $13 \times 13 \times 348$ which

Fig. 17 A straightforward, extremely modular design network framework called ResNeXt is used to classify images. Recurring a building block that combines a group of transformations with the identical structure forms the foundation of our network [103]

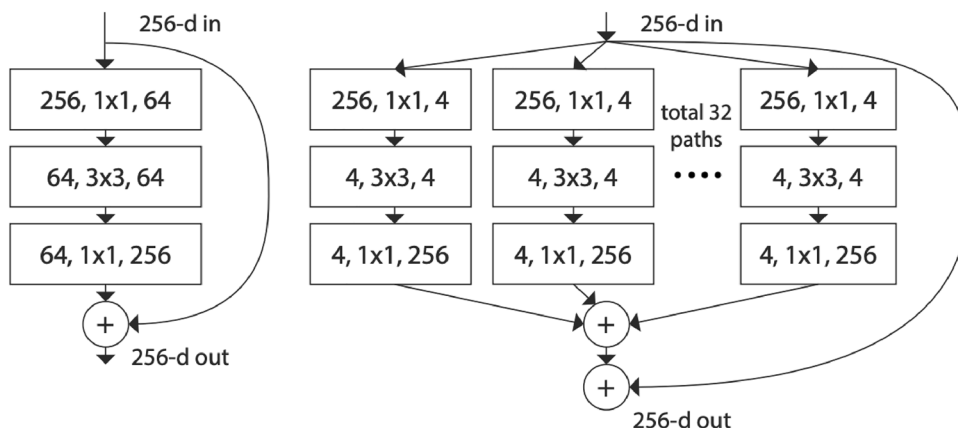
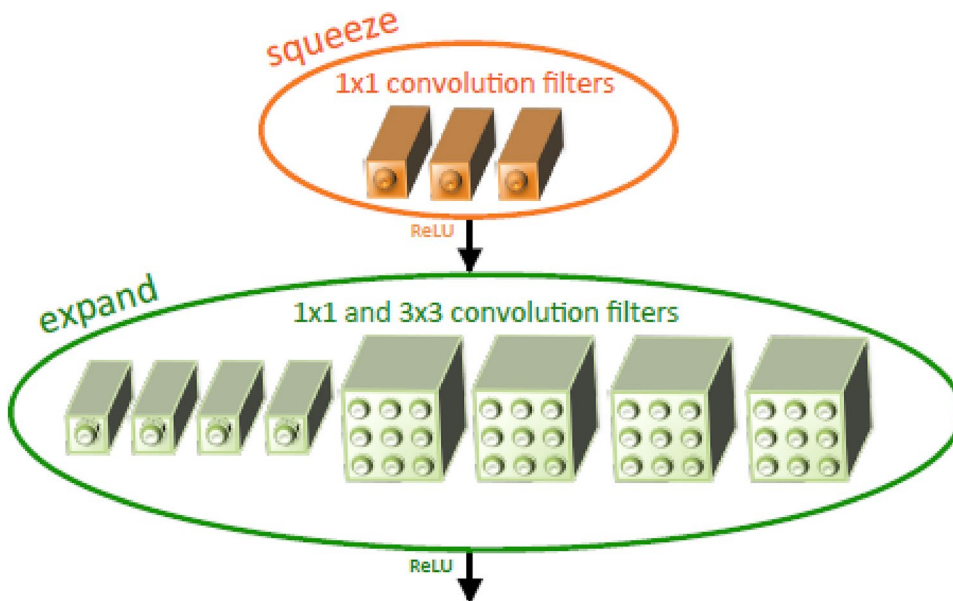


Fig. 18 Fire module is comprised of squeeze convolutional layer feeding into an expand layer and that has a mix of 1×1 and 3×3 convolution filters [104]



- means there are total of 384 filters and each filter have $3 \times 3 \times 384$ filters in it.
- Further, they employ same Convolutional with $3 \times 3 \times 384$ filter. The next layer is of size $13 \times 13 \times 256$ which means there are total of 256 filters and each filter have $3 \times 3 \times 384$ filters in it.
- They apply max pooling where the filter size is 3x3 and stride is of 2. The next layer is of size $6 \times 6 \times 256$. Here $6 \times 6 \times 256$ means 9216 neurons. Number of layers are 256 and each layer have 36 neurons in it.
- For dense connection, they employ fully connect each 9216 neurons to first hidden layer having 4096 neurons which have 4096×9216 connections.
- These 4096 neurons are fully connected to the next hidden layer having 4096 neurons size of 4096×4096 connections.

- Finally the output layer will be of 1000 neurons which is fully connected to the previous layer size of 1000×9216 connections.

6.4 VGG

The convolutional neural network architecture called the VGG framework, or VGGNet, that covers 16 layers is also known as VGG16. It was developed by A. Zisserman and K. Simonyan from the University of Oxford. The study article titled “Very Deep Convolutional Networks for Large-Scale Image Recognition” contains the framework that these authors released. In ImageNet, the VGG16 model outperforms top-5 accuracy results of about 92.7%. A resource called ImageNet has over 14 million photos that fall into about 1000 categories. It was also among the most well-liked models presented at ILSVRC-2014 [91].

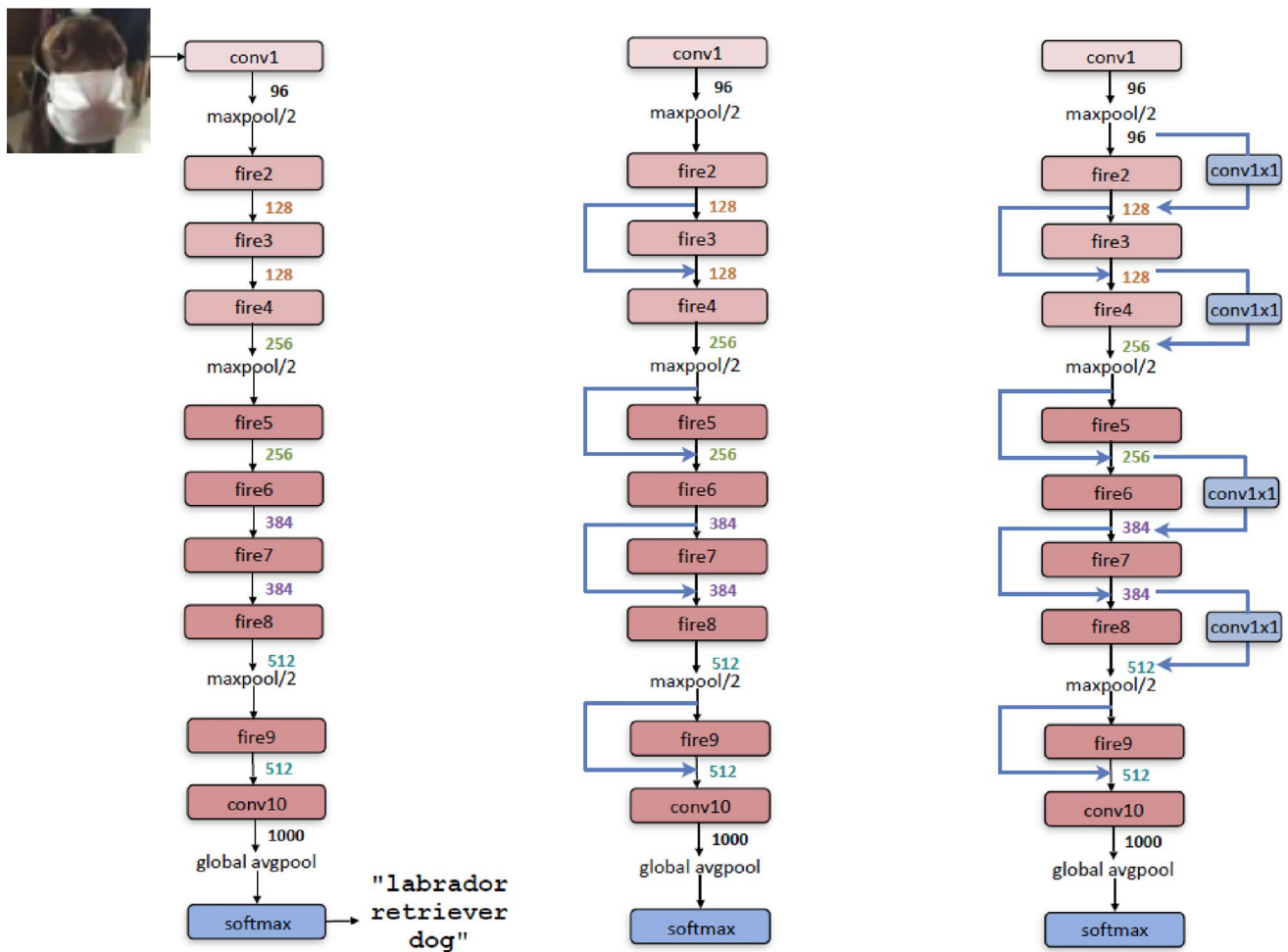


Fig. 19 Squeeze Net is smaller network architecture that was designed as a more compact replacement for AlexNet and it perform 3 times faster and it is used in mage classification [104]

VGGNet architecture is made up of 16 layers. The layer composition consists of 13 Convolutional layers, 5 pooling layer and 3 fully connected layers.

- The first layer is the input layer. It is a $224 \times 224 \times 3$ RGB image. The first you used a Convolutional technique in which you are using a filter size of 3×3 and stride is 1 and same padding size. You Convolutional this layer 2 times with the same filter size. The next layer is of size $224 \times 224 \times 64$ where 64 is the no. of filters you've used.
- Further, they employ the pooling technique in which you used a max pooling where the filter size is 2×2 and stride is of 2. The next layer is of size $112 \times 112 \times 64$ where 64 is the no. of filters you've used.
- Regularly, they use the same Convolutional filter size of 3×3 and stride is 1 and same padding size. Again Convolutional this layer 2 times. The next layer is of size

$112 \times 112 \times 128$ which means there are total of 128 filters.

- Repeatedly they use max pooling where the filter size is 2×2 and stride is of 2. The next layer is of size $56 \times 56 \times 128$ where 128 is the no. of filters you've used.
- They continuously employ the same Convolutional filter size of 3×3 and stride is 1 and same padding size. But Convolutional this layer 3 times. The next layer is of size $56 \times 56 \times 256$ which means there are total of 256 filters.
- They use max pooling where the filter size is 2×2 and stride is of 2. The next layer is of size $28 \times 28 \times 256$ where 256 is the no. of filters you've used.
- Repeatedly use same Convolutional filter size of 3×3 and stride is 1 and same padding size. Again Convolutional this layer 3 times. The next layer is of size $28 \times 28 \times 512$ which means there are total of 512 filters.
- Regularly apply max pooling where the filter size is 2×2 and stride is of 2. The next layer is of size

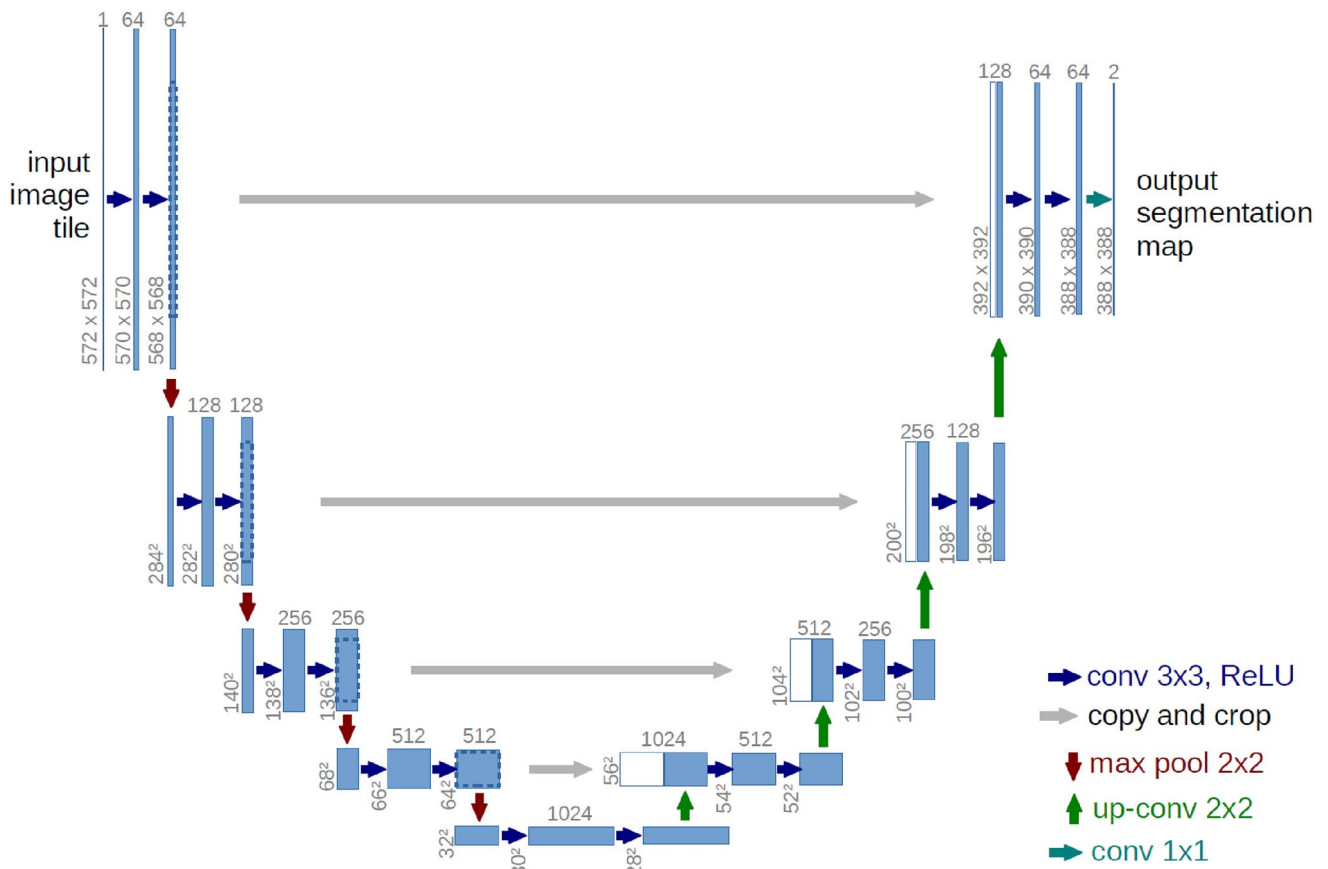


Fig. 20 UNet is a CNN network architecture that expanded with few changes in CNN architecture. This model was invented to deal with biomedical field I this image can take input and this model target was

not only classifying the image it can also identify the infection as well as infection location or area [105]

$14 \times 14 \times 512$ where 512 is the no. of filters you've used.

- They employ the same Convolutional filter size of 3×3 and stride is 1 and same padding size. Again Convolutional this layer 3 times. The next layer is of size $14 \times 14 \times 512$ which means there are total of 512 filters.
- Use max pooling where the filter size is 2×2 and stride is of 2. The next layer is of size $7 \times 7 \times 512$ where 512 is the no. of filters you've used. Here $7 \times 7 \times 512$ means 25088 neurons. Number of layers are 512 and each layer have 49 neurons in it.
- For dense connection, fully connect each 25088 neurons to first hidden layer having 4096 neurons which have 4096×25088 connections.
- These 4096 neurons are fully connected to the next hidden layer having 4096 neurons size of 4096×4096 connections.
- Finally the output layer will be of 1000 neurons which is fully connected to the previous layer size of 1000×4096 connections.

6.5 GoogleNet

GoogleNet was introduced in 2014 by a team at google. GoogleNet was the 2014 winner of ILSVRC (ImageNet Large Scale Visual Recognition Challenge). GoogleNet achieve top-5 error rate of 6.67% which was very close of the error rate of human level. GoogleNet used the convolutional neural network inspired by the LeNet CNN. GoogleNet implemented Inception module in it. GoogleNet used RMSprop, image distortions and batch normalization [92].

The GoogLeNet architecture consists of 22 layers (27 layers including pooling layers) and part of these layers are a total of 9 inception modules.

- The first layer is the input layer. It is a $224 \times 224 \times 3$ RGB image. The first you used a convolutional technique in which you are using a filter size of 7×7 and stride is 2. The next layer is of size $112 \times 112 \times 64$ where 64 is the no. of filters you've used.
- Further, they employ the pooling technique in which you used a max pooling where the filter size is 3×3

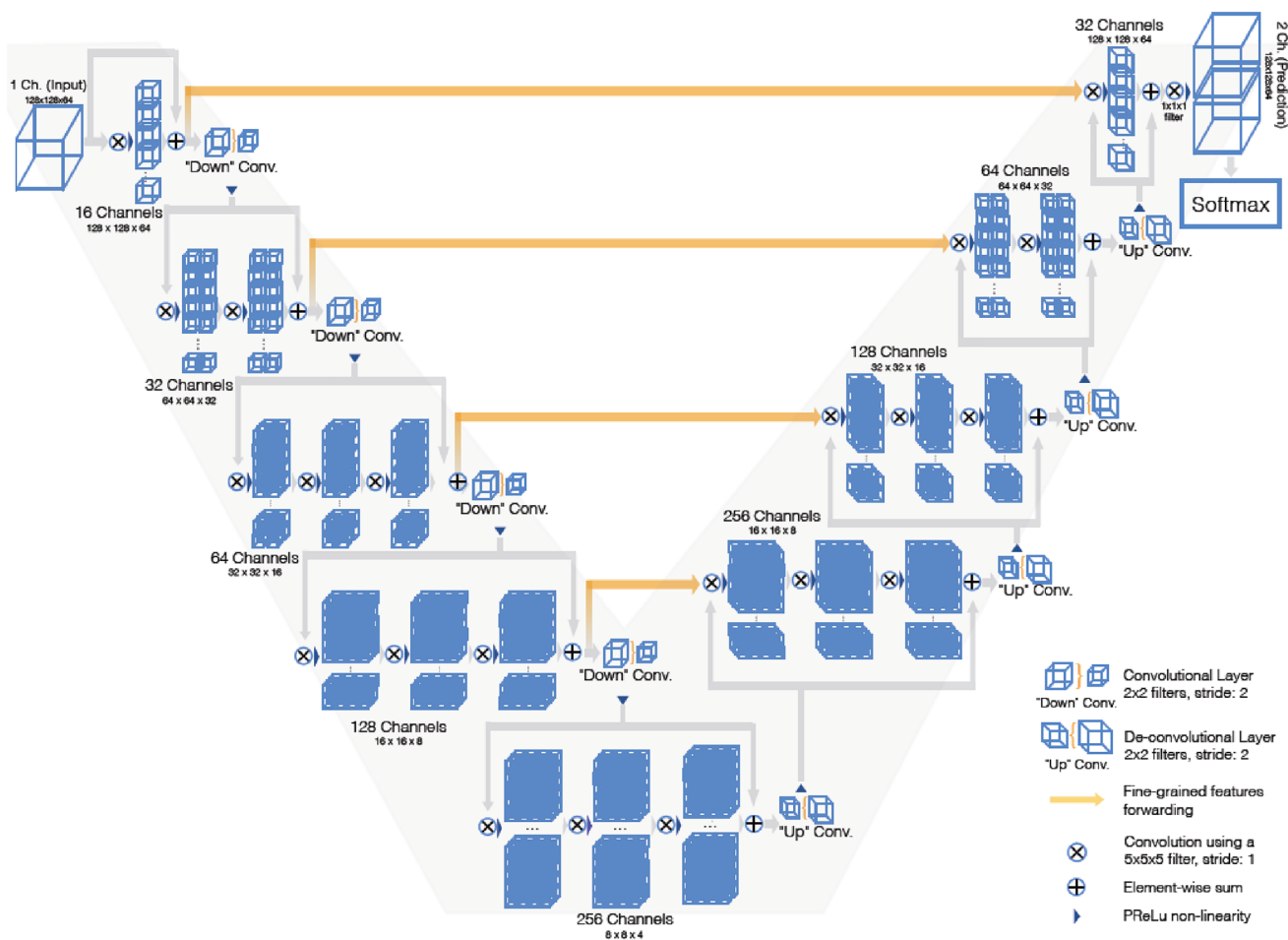


Fig. 21 It is convolutional neural network architecture that can handle biomedical field and help out to perform the segmentation so it called volumetric network because detailed medical terms are not easy so

experts required for detail which is high cost and automatic segmentation can help to reduce the cost in other words we can say that to achieve the high accuracy [106]

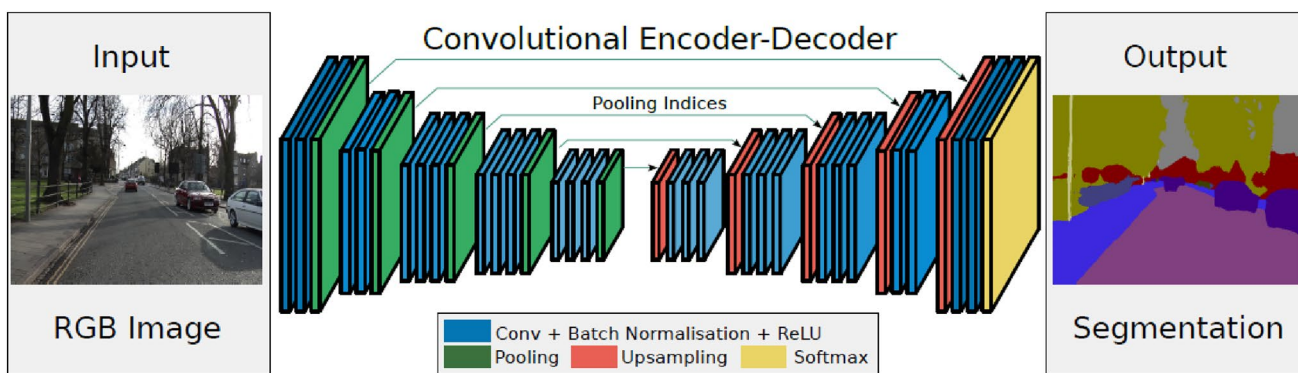


Fig. 22 It is deep encoder and decoder architecture for multi class and in this we can segmentation pixel wise [107]

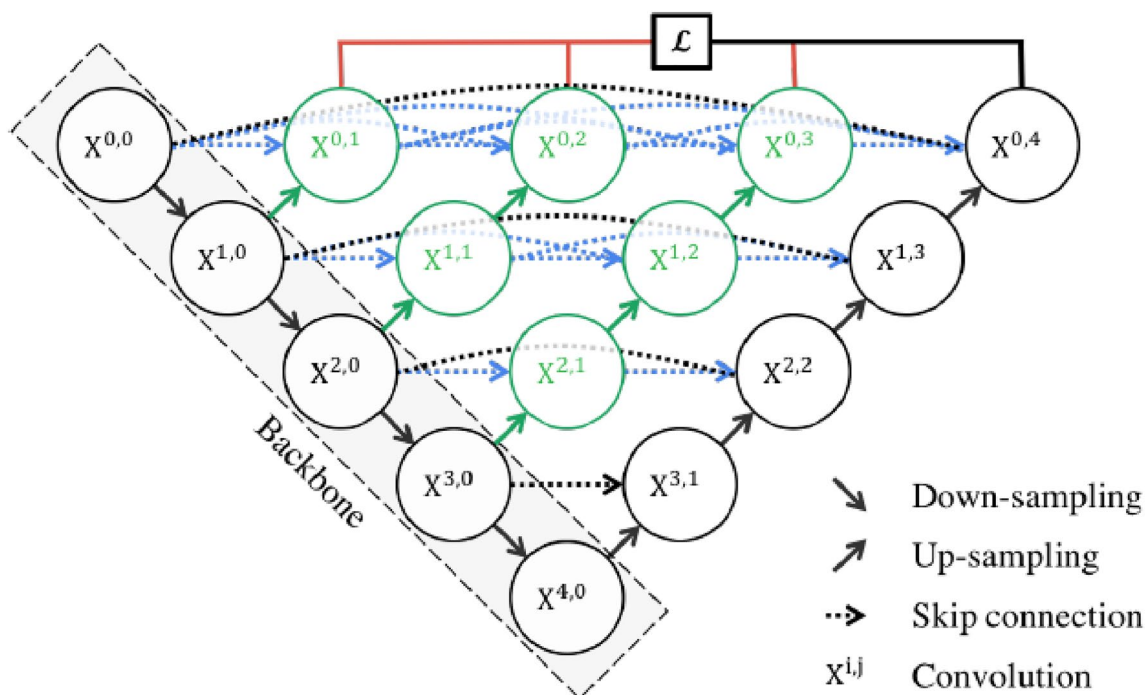


Fig. 23 U-net is a CNN architecture that expanded with few changes in the CNN. This architecture was invented to segment the bio medical images like not only classify the image it can also identify the infection as well as infection area [108]

and stride is of 2. The next layer is of size $56 \times 56 \times 54$ where 64 is the no. of filters you've used.

- They apply Convolutional with 3×3 filter and stride 1. The next layer is of size $56 \times 56 \times 192$ where 192 is the no. of filters you've used.
- Frequently apply max pooling where the filter size is 3×3 and stride is of 2. The next layer is of size $28 \times 28 \times 192$ where 192 is the no. of filters you've used.
- The inception technique is used in which you perform 3 filters (1×1), (3×3), (5×5) and then the max pooling. The next layer will be of size $28 \times 28 \times 256$.
- Regularly they employ the inception technique, the next layer will be of size $28 \times 28 \times 480$.
- They apply max pooling where the filter size is 3×3 and stride is of 2. The next layer is of size $14 \times 14 \times 480$ where 480 is the no. of filters you've used.
- The inception technique is used 5 times and you have the next layer of size $14 \times 14 \times 832$.
- Further, they apply max pooling where the filter size is 3×3 and stride is of 2. The next layer is of size $7 \times 7 \times 832$ where 832 is the no. of filters you've used.
- Do the inception $2 \times$ times and you found the next layer size of $7 \times 7 \times 1024$.
- Regularly, they apply average pooling with the filter size of 7×7 and stride 1. You have the next layer of size

$1 \times 1 \times 1024$ and drop the 40% and you have finally $1 \times 1 \times 100$ output.

GoogleNet has the benefit of training more quickly than VGG. Pre-trained GoogleNets are less in size than VGGs. GoogleNet does have a volume of only 96MB, but a VGG model might have > 500 MBs.

6.6 ResNet

ResNet is an application of Keras. ResNet is a short form of Residual Network. It was introduced in 2015 and it changed the research community forever by ground breaking results. There are many version of ResNet and each version is different from other and serve a different purpose. The different version of ResNet are ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-110, ResNet-152, ResNet-164, ResNet-1202 [93, 94].

ResNet based on two intuitions:

1. Error rate shouldn't decline as we add more layers and go deeper.
2. To reconcile the expected with the real, continue to training the residuals.

These are the functions of a Residual Network.

$$y = F(x, W_i) + x \quad (12)$$

$$y = F(x, W_i) + W_{sx} \quad (13)$$

These two are the equations used where x & y are input and output vectors.

6.7 Highway Networks

Highway networks are used to increase the depth of the neural network and it also does networks optimization. Highway networks uses gating functions approach to regulate information flow. Highway networks are inspired by Long-Short Term Memory recurrent neural networks (LSTM). Highway networks are mostly used in speech recognition tasks and sequence labeling [95].

In highway network, two non-linear transforms T and C are introduced:

$$y = H(x, W_H) \cdot T(x, W_T) + x \cdot C(x, W_C) \quad (14)$$

where T is the Transform Gate and C is the Carry Gate. In particular, $C = I - T$,

$$y = H(x, W_H) \cdot T(x, W_T) + x \cdot (1 - T(x, W_T)) \quad (15)$$

We can have below conditions for particular T values:

$$\begin{cases} x, & \text{if } T(x, W_T) = 0, \\ H(x, W_H), & \text{if } T(x, W_T) = 1 \end{cases} \quad (16)$$

When $T = 0$, we pass the input as output directly which creates an information highway. That's why it is called Highway Network. When $T = 1$, we use the non-linear activated transformed input as output.

6.8 DenseNet

DenseNet is used in object recognition. DenseNet has proven better than the ResNet because it outperforms ResNet in object Recognition. The architecture of DenseNet and ResNet are almost similar but there is a slight change that plays an important part in outperforming ResNet. DenseNet uses concatenation between layers while ResNet uses additive method. DenseNet requires GPU support because it uses concatenation [96].

Its main advantage is that it doesn't allow data to vanish from the input layer to the output layer.

6.9 Wide ResNet

Wide ResNet is a modified version of ResNet. Wide ResNet is called wide Residual Network because there is increase in feature map size per each layer. WRN architecture is quite

identical to the ResNet architecture but there is increase in the feature map size per layer it means that there is increase in the number of channels created in per convolutional layer [97].

In WideResNet, order changed to BN-RELU-CONV. It add more convolutional layers per block and Increases filter size in convolutional layers. It has widening factor k and network includes 40 layers with $k = 2$ times wider than the original would in WRN-40-2 and 4 times in WRN-40-4.

6.10 Pyramidal Net

It has two approaches (1) Bottom Up Pathway is the feedforward calculation of the spine Convolutional Net. It is characterized that one pyramid level is for each stage. (2)Top down pathway and Lateral Connection—The higher goal highlights are up inspected spatially coarser, yet semantically more grounded, include maps from higher pyramid levels. All the more explicitly, the spatial goal is up examined by a factor of 2 utilizing the closest neighbor for straight-forwardness [98].

Specifically, the element maps from base up pathway goes through 1×1 convolutions to lessen the channel measurements and the element maps from the base up pathway and the top-down pathway are converged by component astute expansion. Finally, a 3×3 convolution is annexed on each combined guide to create the last component map, which is to diminish the associating impact of up inspecting. This last arrangement of highlight maps is called P2, P3, P4, P5, relating to C2, C3, C4, C5 that are separately of similar spatial sizes.

6.11 Inception

Inception is a convolutional neural network (CNN) that is used for object detection and image analysis. Contrarily, the Inception structure was intricate (heavily engineered). It employed several strategies to increase speed and accuracy of execution. Numerous variations of the network were produced as a result of its ongoing development. Deep neural networks are costly to compute. The researchers reduce the cost by including an additional 1×1 convolution w preceding 3×3 and 5×5 convolution layers. This reduces the amount of input vectors. Contrary to what would seem logical, 1×1 convolutions are significantly more affordable than 5×5 convolution layers, and the fewer input streams also aid. But keep in mind that the 1×1 convolution is added after the max pooling layer, not before [99].

- *Inception V1* It performs convolution on the input with 3 different size i.e. (1×1 , 3×3 , 5×5).
- It also performed max pooling.

- The output will be concatenated and send to the next inception module.
- *Inception V2*
- It performs convolutional on the input with 2 different size i.e. $(1 \times 1, 3 \times 3)$ as the major change is that 5×5 replaced to two 3×3 convolution.
- This decrease computational time and thus increase computational speed.
- 3×3 convolutional is 2.78 lesser than 5×5 .
- It also converts $n \times n$ factorization to $1 \times n$ and $n \times 1$.
- It is 33% cheaper than the $n \times n$ factorization.
- *Inception V3* It is similar to V2 with the following changes:
 - RMSprop optimizer.
 - Batch normalization in the fully connected layer.
 - 7×7 factorized conv.
 - Label Smoothing Regularization.

6.12 Xception

Xception is a convolutional neural network that was introduced by Francois Chollet. Xception net is the further extension of Inception neural network. Its architecture is similar to the architecture of the inception model but it is further extended by using depth wise separable convolutions. In Xception network the model parameters are used efficiently and that's why it outperforms inception model. The channel-wise $n \times n$ spatial convolution is known as depth-wise convolution. If there are five channels in the aforementioned diagram, we will get five $n \times n$ spatial convolution. The 1×1 convolution used to adjust the scale via pointwise convolution [102].

Although Xception uses the same amount of parameter estimation as Inception V3, it does so more effectively. This framework, known as Xception, considerably surpasses Inception V3 on a bigger image classification sample that consists of 350 million images and 17,000 classes, outperforming Inception V3 on the ImageNet dataset (about which Inception V3 was built). The model is substantially lightweight and has fewer components, which is its major benefit.

6.13 ResNeXt

ResNeXt is designed for efficient image classification. The architecture of ResNeXt consist of stacks of topology blocks. Width and Filter size is also shared which are also known as hyper-parameters. It uses cardinality which refers to the size of the set of the transformations. It is an important part in the addition to dimensions [103].

Comparable to an Inception Module, it collects a collection of transformations using a “split-transform-merge” technique. Our straightforward approach yields a homogeneous, multi-branch framework with a minimal amount of hyper-parameters. In contrast to the measurements of depth and breadth, this technique also reveals a new aspect, which we refer to as “cardinality” (the magnitude of the collection of transformations).

6.14 Squeeze Net

Squeeze Net is a deep convolutional neural network that was introduced in 2016. Squeeze Net was introduced over Caffe framework but later on it has been formed on different networks. It is a CNN model that uses different design strategies to minimize the parameters. Smaller network so less communication across serves within training phase. Due to the lesser so its require minimum bandwidth to export a new model from the cloud to an autonomous car [104]. More flexible to deploy on FPGAs and other types of hardware with limited memory. It uses the following strategy for uniqueness:

- Replace filters (replace the 3×3 filters with 1×1)
- Decreases the number of input (decreases the number of input channels to 3×3 filters if layer has 3×3 filters so total parameters in layer is $(\text{number of input channels}) \times (\text{number of filters}) \times (3 \times 3)$)
- Down Sampling (Down sample late in the networks do convolutional layers have large activation function maps.

When we use Fire modules so we set 1×1 to be less than $(e1 \times 1 + e3 \times 3)$, so the squeeze layer helps to limit the number of input channels to the 3×3 filters.

In left side begins with s single layer followed by 8 Fire module (fire 2–9) ending with final layer (Convolutional 10). Both complex and simple bypass connection yield an accuracy improvement over the vanilla Squeeze Net architecture. With using the fire module, we can reduce the size of model while maintaining the prediction the accuracy.

6.15 U-Net

U-Net is a convolutional neural network that was introduced for image segmentation in the biomedical field. It was introduced by Olag Ronneberger. Its architecture consist upon encoder and decoder. The encoder in its architecture deals with covenant layers that are followed by the pooling operation. The Decoder part in U-Net architecture uses transport convolutions for localization [105].

According to the image this model architecture looks like u and its justify the name. this model architecture consists of three section

- Contraction
- Bottleneck
- Expansion section

Contraction section is made of many blocks and each block takes input applies \times convolutional layers followed by a 2×2 max pooling. We can say that main part or heart of this architecture lies in expansion section this section was similar to the contraction like make block and each block several layers etc. But each block of feature mas used by convolutional layer get half to maintain the symmetry. Loss Calculation in UNet so it is the energy function that is computed by a pixel wise soft max over the final feature map combined with the cross entropy loss function. Image segmentation is important so UNet contributed significantly in such manner to help and solve the problem.

6.16 V-Net

V-Net (Volumetric-Convolutional Network) is a Convolutional network that was introduced medical image segmentation of 3D images. Before Vnet most of the bio medical image segmentation was done using 2D images. Vnet CNN model is trained on 3D MRI scans to perform image segmentation [106].

Left part of the architecture consists of the compression path and right part decompress the signal until its original size reached.

- *Left Side* This side divided into different stages and every stage comprise one to three convolutional layers. Each stage residual function is learnt all layers connected and added to the last output convolutional layer of that stage in order to enable learning a residual function. Convolutional performed in each stage use volumetric kernels having size $5 \times 5 \times 5$ voxels. (Voxels represent the value on s regular grid in 3d space). Output of the pooling layers needed for back propagation. Down sampling also perform because it helps to increase the respective field. In this architecture PReLU is used as activation function.
- *Right Side* This side of network extract the features and expand the spatial of the lower resolution feature maps in degree to collect and generate the important information to output two channel V segmentation.

At each stage reverse of the convolutional layer perform deconvolution to increase the size of the input followed by one to three convolutional layers. Residual function learnt

is similar to the left part. Last convolutional layer computed two feature maps having $1 \times 1 \times 1$ kernel size and produce the outputs of the same size as input volume. These features maps are the probabilistic segmentation of foreground and background regions by applying soft max function voxel wise. Feature extracted from early stages of left part so that will have forwarded to the right part through this connection and this will help to provide location information to the right side so that's cause improve or better the quality of the final contour prediction.

6.17 SegNet

SegNet is a convolutional network that was introduced by the researchers of the University of Cambridge. SegNet architecture consist of encoder and decoder. SegNet uses encoder and decoder based architecture for multiclass pixel-wise segmentation. SegNet performs very well in indoor scene understanding [107].

In encoder convolutional layers with batch normalization and a ReLU non-linearity followed by non-overlapping max pooling and subsampling in other words we can say that down sampling. In this network there are 13 convolutional layers from VGG-16. During the 2×2 max pooling corresponding max pooling locations can be stored. In the decoder convolutional as well as up sampling are performed in the end softmax classifier is present for each pixel. In above image show during the up sampling max poling indices the corresponding encoder layer then it is called up sampling. So ate end soft max classifier is present which can predict class for each pixel.

6.18 Unet++

Unet++ is the modified version of Unet CNN. It provides more accuracy with a series of nested, dense and skip pathways. The changes that are made in Unet to modify it into a Unet++ are redesigned skip pathways, dense skip connections, deep supervision [108].

Unet++ is uses the dense block and its differ from Unet with three ways:

- CL (Convolutional layers) on skip pathways.
- Dense skip connections on skip pathways.
- Deep supervision which enables model pruning.

In first point or step which bridges the semantic gap between encoder and decoder and second point will tell improves the gradient flow third and last point tells achieve comparable performance to using only one loss layer. This network starts with encoder and decoder sub-network and backbone followed by a decoder sub network. There are re-designed

skip pathways green and blue that connect the two sub networks and use of deep supervision (red). Above figure shows that how the feature map travel through top skip pathway of unet++. Each layer in convolutional is preceded by a concatenation layer that fuses the output from previous layer of the same dense block with corresponding up sample output.

6.19 FCNN

Fully convolutional neural network model is a deep learning model based on simple and classic convolutional neural network. It can take input in any arbitrary size. In Fully convolutional neural network there are no fully connected convolutional layers. FCNN has learning filters everywhere. FCNN also has filters in decision making layers.

6.20 Fast-RCNN

Fast-RCNN was introduced in 2015 by [109]. Fast-RCNN is used for faster object detection. Its architecture is pretty much similar to the RCNN but in Fast-RCNN we direct give input image to convolutional neural network to generate a convolutional feature map. The reason that the fast-RCNN is faster than the RCNN is that there is no need to give input a lot of region proposals to CNN every time.

6.21 Mask R-CNN

Mask R-CNN is used for instance segmentation. Mask R-CNN is formed on the base of Faster R-CNN. Mask R-CNN uses anchor boxes to detect multiple objects and overlapping objects in an image. This method increases its speed [110].

6.22 DeepLab

DeepLab was introduced in 2016 by Google. It is one of the best semantic segmentation models present. DeepLab has been modified and improved over the years and now there are many promising versions of DeepLab present [111]. The following are the versions of DeepLab:

1. *DeepLab-v2* DeepLabv2 was built on DeepLab it contained an atrous spatial pooling scheme. It consist of parallel dilated convolutions with different rates applied in the input feature map.
2. *DeepLabv3* DeepLabv3 was an improvement made over DeepLabv2, in this version the problem of segmenting objects at multiple scale was solved by designing and

employing atrous in cascade to capture multiscale context by adopting multiple atrous rates.

3. *DeepLabv3+* DeepLabv3+ is an improvement over DeepLabv3+. In this version semantic labels are assigned to every pixel in the given input image.

6.23 RetinaNet

RetinaNet is used for one stage object detection models. RetinaNet provides good results in dense and small objects detection. RetinaNet is mostly used with aerial and satellite object imagery. RetinaNet is formed by making improvements in existing object detecting models which are Feature Pyramid networks and Focal Loss [112].

6.24 YOLO

YOLO (You only look once) is a network that is used for object detection in deep learning. YOLO is faster than its competition and is mostly used in real time object detection. YOLO does object detection by classifying different objects in the given frame [113]. Yolo has been modified year by year and has been converted into different versions:

1. *YOLOv4* YOLOv4 is an object detection network that can be operated on single GPU with a smaller mini batch size. YOLOv4 increases the speed of object detection for the systems with a single GPU.
2. *PP-YOLO* PP-YOLO (Paddle Paddle-YOLO) is an improvement over YOLOv4 it has improved inference speed and mAP score.

Channel (Input) exploitation based CNNs The lack of variance and the absence of distinct information within the input might have an effect on CNN's performance as an individual that's why the concept of channel boosting was proposed to boost the representation of the networks.

Attention is the method that is used to copy cognitive attentions. This effect will magnify the important parts of input data and will fade out the all-other data.

7 Deep Learning for Medical Image Analysis and CAD

CNN's beginnings may be traced back to the nonrecognition idea put out by [114] in the beginning 1980 s. In early 1990, [115] training a CNN using backpropagation to categories sequences of character recognition. In the

initial nineties, CNN was employed in several domains including image identification, character segmentation and facial detection. In 1993, [116] trained a CNN to recognize pulmonary nodules in chest radiographs, which was the first time CNN had been used to analyze healthcare images [117]. Chan et al. [118–121] used CNN for lung nodules and microcalcification recognition [122, 123] on radiography within the identical year. In 1994, [124] used a comparable shift-invariant neural network to find regions of microcalcifications. Despite the fact that these initial CNNs seemed not really sophisticated, they were able to recognize patterns in healthcare images. Profound CNN was made possible by a number of crucial neural network learning methods that have been evolved over time, such as layer-wise unsupervised representation learning accompanied by closely monitored fine [125–127], the use of rectified linear unit (ReLU) [128, 129] as an activation function in place of sigmoid operations, pooling to enhance functionality normalization and minimize granularity [130], dropout to control under- and over-fitting problem [131] and batch normalization [132]. Such methods enable the training of neural networks with increasing layers and parameters. In 2012, [133] presented a CNN with five convolution layer and three fully connected layers (named “AlexNet”) that contained more than 60 million weights. This CNN outperformed previous models in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [134], which categorized more than 1000 categories of common items on images. AlexNet exhibited the hierarchical structure’s several layers’ capacity for pattern classification. Since AlexNet, DCNNs have been built with higher density.

The 2015 ILSVRC was won by [94], who presented residual training and demonstrated that a residual network (ResNet) having 110 to 152 hidden layers might beat a number of previous DCNNs. According to [135], a DCNN’s learning ability grew with complexity, but it was only used if the training set were adequately substantial.

Unexpected hopes for deep learning in healthcare have been elevated by the achievement of AI and learning techniques in driving-less cars, social networks, board and GO games. Numerous medical image analysis tasks for CAD have used deep learning [136–138]. The categorization of illness and healthy trends, the classification of cancerous and mild tumors and the forecasting of high- and low-risk trends of acquiring tumors in the long term are among the major popular applications of deep learning in CAD. The division and categorization of various parts of the body and tumor classifications, as well as the categorization of alterations in tumor volume or appearance for the evaluation of medication reaction or the forecast of outcome or recurrence, were further uses. Numerous research for lung illnesses and prostate tumors were carried out utilizing the public data sets since there exist sizable public data sets for Computed

Tomography (CT Scan), thoracic CT and mammography. The diagnosis of eye disorders using retina images or optically Computed Tomography (CT Scan), as well as the categorization of cell types using histopathological images, have all been accomplished using deep learning-based imaging techniques [139, 140]. The majority of the experiments presented extremely encouraging findings, adding to the excitement around deep-learning-based CAD. Even though this latest range of CAD technologies is referred described as AI, they are nowhere near being ‘intelligent’, instead behaving like an extremely complicated numerical method that memorizes data in its billions of weights.

One of the earliest applications of deep learning to medical image processing is image or assessment categorization. Generally, one or more image graphs (an assessment) are used as the source and just one medical parameter is used as the outcome in exam categorization (e.g., disease present or not). Every screening test serves as a trial in this situation, and dataset volumes are often lower than in machine vision (e.g., hundreds or even thousands vs. millions of observations). Hence, it is not unexpected that ensemble learning is attractive for these purposes.

Attempt to circumvent the (considered) need for massive amounts of data for deep CNN training, ensemble learning basically involves using pre-trained networks (usually on real imagery). Using a pre-trained CNN model as a feature representation and fine-tuning a pre-trained CNN model on health data were other transfer learning methodologies that were discovered. The suggested technique also has the advantage of not having any deep CNN training, making it simple to integrate the derived features into current image processing workflows. These methods are well-liked and often used. Many writers, though, actually provide a detailed analysis of which approach produces the greatest results. Kim et al. [141] and [142], two studies that attempt, present inconsistent findings. In the study by [142], fine-tuning achieved 57.6% performance in inter class evaluation of knee osteoarthritis against 53.4% performance for extracting features. However, [141] demonstrated that the efficiency of cytopathology image categorization was improved by utilizing CNN as a feature representation (70.5% vs 69.1%). Several current articles, both issued in reputable journals, that optimized a pre-trained CNN variation of Google’s Inception v3 framework on healthcare records and attained (nearly) human expert efficiency are cited if any insight can be offered as to which approach could be the most effective [143, 144]. Such findings have not yet been obtained utilizing only pre-trained CNN models as feature representation, as far as the researchers are concerned.

A time-frame resembling machine vision is evident for the neural architectures that are frequently utilized in

test categorization. Traditionally, the medical imaging industry concentrated on network designs like SAEs and RBMs as well as unsupervised pre-training CNN models. The very earliest articles utilizing these methods for test categorization concentrated on neuroimaging and were published in 2013. Based on brain MRI, [145–148] used DBNs and SAEs to categorize individuals as suffering Alzheimer's syndrome. There has been a noticeable trend favoring CNNs presently. Among the 397 publications on test categorization that were released in 2018, 2019, 2020, 2021 and 2022, 313 use CNNs, 51 use AEs and 33 use RBMs. These techniques have a wide range of applications, including pulmonary CT Scans, retinal scanning, brain MRI, and computational pathology.

Instead of utilizing pre-trained CNN models, researchers of relatively latest researches using CNNs frequently develop their customized network designs from beginning. In tests evaluating fine-tuning of pre-trained CNN networks versus training from beginning, [149] shown that fine-tuning scored significantly given a limited collected data of roughly 1000 images of skin disease. But the size of these studies makes it impossible to extrapolate any broad generalizations.

Three studies had to utilize a layout that took advantage of the particular characteristics of healthcare records: two was using 3D convolutions rather than 2D to identify Alzheimer's patients [150–152] implemented a CNN-like structure to a brain network diagram from MRI Diffusion Tensor Imaging (DTI). In effort to accomplish this, they created three additional layers, known as node-to-graph, edge-to-edge and edge-to-node layers, which served as the framework for their structure. They demonstrated that their structure surpassed current approaches in evaluating intellectual and physical scores by using their system to anticipate brain growth.

Mahapatra et al. [153] present deep learning algorithms' dependability and integrity are highly coveted qualities in the healthcare profession. In this study, they present an interpret-ability-guided inductive bias method that requires learned features to produce saliency maps that are more distinguishable and regionally stable for various classifiers of classification model, resulting in enhanced simulation results. By including a category cost and a regularization cost term for spatial coherence, they are able to accomplish our goals. The suggested methodology surpasses existing approaches, according to practical findings for healthcare image categorization and differentiation challenges, and produces feature map that are more in line with the opinions of healthcare staff. They also demonstrate the usage of unmarked image data to further improve efficiency. In conclusion, the suggested method produces better acquisition rates, model resilience, and model understandability. It is also flexible, adaptable to current infrastructure designs utilized for medical imaging purposes [153].

Lee and Chung [154] proposed a Deep Neural Networks (DNNs) classify medical images using transfer learning from image features to enable computer-assisted medical assessment. Contrarian threats are anticipated to be confined even though learning datasets (healthcare illustrations), which are frequently needed for malicious examples, are typically unreachable in contexts of privacy and safety conservation and protection, despite the fact that the adversarial weakness of CNN architectures impedes real world applications due to the elevated shareholdings of prognosis. However, in this paper, they proved that, despite the absence of such clinical images, antagonistic assaults are still feasible using visual features for healthcare neural network models with domain adaptation. In specifically, they established that Universal Adversarial Perturbations (UAPs) may be produced from image features. Both non-targeted and targeted assaults can benefit from UAPs derived from real-world images. UAPs generated from real images performed noticeably better than randomized counterparts. Transfer learning usage creates a privacy gap, which reduces the accuracy and trustworthiness of machine illness identification. Although UAP effectiveness from real images was lowered by model development from random selection, UAP susceptibility was still present.

Liu et al. [155] highlighting two issues that must be addressed for Semi-Supervised Learning (SSL) in medical image analysis (MIA) to be effective. Both inter- and cross-label issues should be successfully handled by it. Anti-curriculum pseudo-labeling is a novel SSL technique that they suggest in this research. It presents brand-new methods for choosing insightful unlabeled samples. In this study, they recommend an innovative SSL method called Anti-Curriculum Pseudo-Labeling (ACPL), that also incorporates innovative strategies to can choose insightful unmarked samples, working to improve learning harmony and enabling the prototype to collaborate for both inter- and cross-label issues, as well as to approximate pseudo labels by a precise ensemble of classification models, in contrast to conventional methodologies that can choose self-assured pseudo label by criterion.

Xue et al. [156] proposed a novel training methodology for effective clinical image categorization from chaotic class labels using domestic and global learning algorithm. To effectively choose cleaner and chaotic data, they use the self-ensemble approach with a messy labeling filter. The proposed approach succeeds better than existing training from chaotic label approaches.

Apostolidis and Papakostas [157] proposed a technique for adversarial assaults have established themselves as a serious danger to machine vision. This study explores the hypothetical black-box challenges and respond aspect of computerized watermarking, which is a separate aspect of the technology. The widespread usage of watermarks for

safety purposes appears to offer major hazards to machine vision, which is why they emphasized a critical issue in this respect in addition to introducing a new class of adversarial assaults called watermarking attacks. The methodology with the greatest decrease was CT scans, while the framework with the greatest vulnerability was MobileNetV2.

Jin et al. [158] proposed a theoretical paradigm for combining the unrivalled advantages of superior human understanding (native intellectual ability) and machine intelligence to create a reliable, precise, and all-encompassing item detection technique for clinical image processing. The Automatic Anatomy Recognition (AAR-DL) method integrates deep learning image recognition systems, model-based object classification and advanced anatomy modeling technique. In each of the four major components that make up AAR-DL, past information is carefully considered. High precision and resistance to image distortions and aberrations have been shown with AAR-DL. AAR-DL exhibits extraordinary resilience and positioning precision within 1-2 grid cells, performing like a skilled human observer in object identification.

In this study, [159] efficiently utilize the entire clinical dataset for semi-supervised clinical image categorization, they present a new mutual consistency network (MC-Net+) in this research. One common encoder and a number of marginally unique decoders are present in the framework. Consequently, our suggested MC-Net+ architecture comprises of two fresh layouts. First, the architecture has a single common encoder and a number of marginally unique decoders (i.e. using various up-sampling approaches). The unidentified difficult sections are indicated by the statistically significant difference of the outcomes from numerous decoders, which is estimated to signify the architecture's ambiguity. Second, they add a unique reciprocal consistency requirement between the smooth pseudo labels of one parser and the probabilistic outcome of another parser. On three open datasets, they evaluated the categorization outcomes of our MC-Net+ system with five cutting-edge methods.

Wang et al. [160] proposed technique for Diabetic Retinopathy (DR). One of the most serious consequences of hyperglycemia is Diabetic Retinopathy. For an earlier detection of DR, precise separation of DR symptoms is crucial. Technically speaking, concurrent categorization of several DR tumors is difficult due to 1) the dearth of pixel-level labels and 2) the wide variety of DR tumors. They provide a brand-new Poisson-Blending Data Augmentation (PBDA) approach to create fake images that can simply be used to increase the training examples already available for tumor separation. The approach underwent significant validation using comparative and elimination tests on two open clinical

datasets. The outcomes showed that the suggested approach significantly improved than the most recent techniques.

Zhao et al. [161] proposed a scientific studies and healthcare diagnostics both heavily rely on Magnetic Resonance (MR) scanning. The location of the layer cluster has a significant impact on the utility of the recovery since MR scanning has a strong in-slice accuracy and a poor through-slice resolution. Conventional medical workflows rely on laborious hand adjustments that are difficult to duplicate. In comparison to our earlier research, a localization network was included, concluding the automated slice grouping positioning operation. To improve the categorization efficiency when processing full resolution images, they develop a multi-resolution architecture. Using variational Weighted Principal Component Analysis (WPCA) regularization they enhance the plane separation technique. They created the Performance Measurement Index, or PMI, that is utilized to show the client how confidence.

Wang et al. [162] proposed a novel methodology for dermoscopic images containing skin lesions. Professionals must put in a lot of research to better understand the features from medical and dermoscopic images in order to diagnose skin lesions accurately. When contrasted to solitary modality-based approaches, deep heterogeneous learning-based techniques can lower intra- and inter-reader heterogeneity and increase classification output. In order to accomplish heterogeneous skin lesion detection, this paper introduces a unique technique called Adversarial Multi-modal Fusion with Attention Mechanism (AMFAM). To be more precise, they utilize an adversarial learning-based discriminator to drive the feature extraction technique to deliberately acquire the linked data. By understanding the associated and supportive data, an unique heterogeneous fusion approach is suggested to conduct automatic skin lesion categorization utilizing medical and dermoscopic images. To direct the feature extraction technique in learning the connected data, a modality classifier is created. An image reconstruction method built on self-attention that autonomously directs the extracted features to focus on tumor regions.

Tilborghs et al. [163] explored the most advanced method for several brain tumor segmentation challenges, such as the separation of the myocardium in cardiac magnetic resonance imaging (MR) images, is semantic segmentation utilizing convolutional neural networks (CNNs). However, the anticipated division maps produced by such a typical CNN would not permit precise comparison of local form characteristics such local layer strength. They offer a CNN for immediate myocardial geometry and posture factor estimation. An appropriate mathematical geometry model relying on landmarks is connected to the parameters. To help with accurate feature estimation and form adaptation,

semantic segmentation is utilized. The error function enforces uniformity between estimated attributes and integrated segmentation. The adapted geometry model may be used to instantly determine regional myocardial characteristics.

Yang et al. [164] highlights a basic issue in computational histopathology is the retrieval of pictorial representations. Self-supervised training has evolved as an important method to derive useful pictorial representations from unprocessed histopathology images in light of deep learning's strong representational capabilities and the dearth of labels. While a few self-supervised techniques have been presented expressly for histopathology images, the majority of them include flaws which may limit their adaptability or capability for depiction. A brand-new multi-modal self-supervised training technique for pictorial imagery designed specifically for H & E stained histopathology images. Self-supervised methodologies that are discriminatory and creative can improve one another. using histopathology-specific area information to make wise choices. Excellent adaptability for various computational histopathology problems.

In this study, [165] discuss and analyze the tissue level semantic segmentation. In computational pathology, tissue-level semantic analysis is a key step. With extensive pixel-level labels, fully-supervised algorithms have indeed demonstrated remarkable achievement. Unfortunately, it is quite costly and time-consuming to create such markers on the giga-pixel entire slide images. With only texture categorization markers as input, they describe a cell semantic segmentation technique for histopathological images that significantly reduces the amount of time pathologists must spend annotating the scans. To close the knowledge gap across patch-level and pixel-level tags, multi-layer pseudo-supervision with continuous dropout attention is suggested. Additionally, a categorization gate function is included to lower the frequency of false positives. On two datasets, our design plan outperforms weakly-supervised semantic segmentation algorithms in terms of effectiveness while also performing on pace with the fully-supervised background. The first LUAD dataset for weakly-supervised tissue semantic segmentation has been made available.

Zhang et al. [166] proposed a novel methodology for Nuclear Cataract (NC) using Optical Coherence Tomography. A major source of vision loss and visual impairment worldwide is nuclear cataract (NC). Strategy for medically early treatment and cataract surgery requires precise and trustworthy NC assessment. The core area may be seen easily in Anterior Segment Optical Coherence Tomography (AS-OCT) images, which can also be used to quantify the transparency of NC. Recent medical studies have shown that there is a strong inter- and intra-class association and recurrence between the averaged nuclei volume on AS-OCT scans and the NC intensity of injury.

The medical approaches serve as the inspiration for this article's basic yet successful Region-based Integration-and-Re-calibration attention (RIR), which combines several Convolutional feature area depictions and softmax attention iteratively readjusts the values of individual part. The network may concentrate on area depictions with high contributions and inhibit those that are less valuable thanks to this region re-calibration technique. To dynamically forecast the threat standard of NC, they integrate the RIR block with the Residual convolution layer to create a Residual-RIR module. A series of Residual-RIR modules are then layered to build the region-based integration-and-re-calibration network (RIR-Net).

Huang et al. [167] proposed concept for breast cancer detection using Ultrasound (US). In particular for women with thick mammary, US is essential for breast cancer prevention. Prior to making the assessment, it is standard procedure for a sonographer to identify the main clinical characteristics of a tumor and capture one or more sample images throughout the active imaging. They put forth a reinforcement learning-based approach that can autonomously retrieve distinguish frame from breast ultrasound movies of variable period in order to tackle these issues. It has a detection-based tumor screening component and a cutting-edge incentives that enable keyframe searches to include topographical and clinical aspects of the tumors.

Yi et al.[168] proposed methodology for nerve segmentation using ultrasonography for anesthesiologists. When administering anesthetic during peripheral nerve blockade, anesthesiologists frequently employ ultrasound-guided injection. Even for seasoned anesthesiologists, it can be a challenge to correctly detect nerve anatomy in ultrasonography. The proper implementation of PNB operations depends on the precise recognition of nerve in ultrasonography. In addition to being a clinical issue, precisely segmenting the nerve anatomy from ultrasonography is a difficult work for machine vision. To enhance the categorization efficiency of nerve anatomy by concurrently segmenting multiple structures in ultrasonography, an unique brachial plexus segmentation model (called MallesNet) combined with past healthcare expertise of nerve recognition is presented.

Huang et al. [169] explored concept of Graph Convolutional Network for tumor segmentation using ultrasound. Breast ultrasound (BUS) has shown to be a reliable method for finding breast malignancy in its initial stages. A tumor categorization is an essential first stage in making an adequate assessment since it identifies the target's boundaries, form, and position. An expert separation technique that can handle hazy or obscured boundaries in ultrasonography. The framework for discrete edge choosing that can concentrate on the borderline area autonomously. The edge rendering architecture that uses the Graph

Convolutional Network (GCN) and can use global contour data. a comprehensive architecture capable of concurrently segmenting and classifying data.

Li et al. [170] discovered a novel concept for Daibetics patients using Vision Transformer and Tongue Diagnosis Analysis System (TDAS). Diabetes is a long-term illness that is prevalent and highly prevalent Globally, especially in China. A serious global healthcare issue is diabetes. However, it is challenging to manage the progression of diabetes with the present diagnostic and therapeutic approaches. Due to its inexpensive price, potent therapeutic impact, and easy availability, traditional Chinese medicine has emerged as a viable choice for the prescription of diabetes. This work introduces a brand-new algorithm for automatically categorizing images of diabetic tongues. The individuals' tongue images are taken using the TFDA-1 tongue diagnostic tool. The TDAS characteristics of the tongue images are extracted using the Tongue Diagnosis Analysis System. Images of the tongue are used to collect distinguish characteristics using the Vector Quantized Variational Autoencoder (VQ-VAE). K-means clustering of tongue images using VQ-VAE characteristics. The disparities across regions are described using TDAS characteristics. The categorization findings are validated and positional monitoring feedback is calculated using Grad-weighted Class Activation Mapping (Grad-CAM) and Vision Transformer (ViT).

Oyelade et al. [171] proposed a novel concept for histopathological images using Whole-image based CNN (WCNN) and Region-based CNN (RNN). Interpretation and localization of anomalies in clinical images is thought to be a highly difficult process. To mitigate this problem, several computer-aided techniques have been used, and the popularity of deep learning network topologies is evidence of the remarkable contribution that has been documented in the research. In this study, they propose a dual subsidiary deep learning framework that makes use of two alternative Convolutional neural network architectural topologies to handle problems associated with the coupling of categorization with image localization and recognition. To categorize and locate anomalies in collections, whole-image based CNN (WCNN) and region-based CNN (RCNN) models are rigorously mixed. The technique does not need images that are reliant on labeling to classify anomalies into many categories or to pinpoint their location. Additionally, a flawless assurance and reasoning method is offered, allowing the results from WCNN and RCNN to be combined for additional study.

Sadik et al. [172] propose a new paradigm for COVID-19 detection using customized U-Net. Computer-aided diagnostic technologies are becoming increasingly necessary in the Coronavirus disease-2019 (COVID-19) outbreak for the quick and reliable detection of a significant amount of

individuals in addition to conventional approaches. In this study, an efficient, precise COVID-19 detection method based on deep convolutional neural networks (CNNs) is suggested using pulmonary computed tomography (CT) scans images. Initially, by adding more skip linkages to the U-Net structure to compensate for the information loss caused by dimension shifting, a customized CNN architecture called SKICU-Net is developed for the detection and segmentation of pulmonary areas in a chest CT scan. Following that, an agglomerative hierarchical algorithm is used to exclude the CT segments with insufficient data. Finally, a customised DenseNet structure called P-DenseCOVNet is configured for efficient feature mining and treatment of COVID-19 and pneumonia from fragmented chest slices. This structure adds concurrent convolution layer pathway on highest part of the standard DenseNet framework to improve effectiveness by conquering the cost of strategical reasoning.

Sun et al. [173] propose early detection mechanism for retinal disorder using Hybrid graph convolution. For the purpose of preventing permanent vision damage, timely identification and management of retinal problems are essential. The creation of multi-label fundus illness recognition systems that can check for numerous illnesses is better in accordance with practical demands considering that individuals in the healthcare environment may have several forms of visual sickness. This study provided a synthesis framework for cross-label fundus ailment diagnosis at the individual level relying on hybrid graph convolution. A foundation component, a hybrid graph convolution subsystem, and a classification subsystem made up the composite model. This study used graph convolution to construct the relationships across the categories, and it then used a self-attention method to create a composite graph convolution architecture. While the classification component produced cross-label data using LightGBM, the foundation component used EfficientNet-B4 for feature extraction. The input data of binocular images and the effect of tag association on the model's recognition rate were also examined in this study. On the readily viewable ODIR dataset, the design plan MCGL-Net surpassed all other cutting-edge techniques, with F1 achieving 91.60% on the testing dataset.

Liu et al. [174] advise a novel approach for neuroblastoma categorization using histopathological whole-slide images. The most frequent extra-cranial malignant tumor in young children is neuroblastoma. Physicians can utilize the International Neuroblastoma Pathology Categorization (INPC), a widely employed classification approach, as a resource for therapy stratification. To categorize individuals having neuroblastoma into several prognosis categories, an autonomous, extensive, and accurate categorization approach is required. In this research, 107 participants having

neuroblastoma who had surgical excision provided 563 Hematoxylin and Eosin-stained (H &E) histopathological whole-slide images. For nuclear separation, cell-level visual feature collection, and individual-level feature accumulation, they suggested a unique computational strategy. To categorize individuals with Favorable Histology (FH) and those with Unfavorable Histology (UH), a logistic regression classifier was created.

Saini and Susan [175] propose a unique model for multi-class imbalance dataset of Diabetic Retinopathy (DR). Diagnosing and monitoring for diabetic retinopathy is a well notable topic in the biomedical field. A component of Computer-Aided Diagnosis (CAD) that has advanced significantly over the previous several decades as a result of the development and effectiveness of deep learning is the application of medical imaging from a patient's vision to identify the injury inflicted to blood vessels. Comprehensive comparison of three standard datasets of varying dimensions, each containing images of diabetic retinopathy for DR scoring (identification), separation, tumor recognition, and optical disc, using several state-of-the-art transfer learning approaches. For optimum efficiency when dealing with various unbalanced circumstances of various dataset dimensions, several pre-trained CNN models used to fundus images together with rejected interpolation (randomized under-sampling at mini-batch level) approach.

Xin et al. [176] explore Vision Transformer (ViT) and apply on skin lesion classification using multi-scale and overlapping sliding windows. Current decades have seen significant advancements in the initial detection and medication of melanoma, whose prevalence is rising annually throughout the world and pose potential a serious risk to human health. These advancements have been made possible by the use of machine learning to recognize dermoscopic images. Three stage process is used to confirm its efficacy. In order to confirm the efficiency of SkinTrans in the categorization of melanoma, a ViT system is first constructed. The image is then serialized using cross-scale and overlapping sliding windows, and cross-scale patch implantation is performed with a focus on cross-scale characteristics. In order to create the encoding outcomes of diverse input source as distinct as feasible, contrastive learning is utilized to create the comparable facts of melanoma encode uniquely.

Shabani et al. [177] propose a novel strategy for COVID-19 segmentation using self-supervised learning. Segmenting medical images is an important first phase in many medical strategies. We provide an unique and new labeled-free process to do categorization of COVID-19 Computed tomography. Whereas the majority of autonomous categorization approaches are supervised and need a easily-labeled paired dataset. They suggested a self-supervised approach that may not necessitate any human categorization

labeling at the cell level. We created wholesome CT images from COVID-19 Computed tomography using GAN. By doing region-aware categorization, we were able to improve the findings by using a contrastive cost. Our approach behaves reasonably better than the state-of-the-art approaches for segmenting COVID-19 Tomography at this time.

Hayat et al. [178] discover Malaria infection using Genetic Algorithm (GA) and Computational Linguistic (CL) to classify the Plasmodium falciparum. Malaria is a contagious and lethal infection that is brought on by Plasmodium falciparum. Initially, a microscope was used to identify cells that were contaminated with malaria. Due to the sheer volume of data to analyse and the complexity of the timing, it could result in false diagnosis. Increasing time requirements and inaccurate diagnosis have led to a significant need for autonomous parasite diagnosis methods. For the purpose of identifying the malaria parasite, a smart mathematical framework is developed. As extracting features strategies, biochemical, physiochemical, and Computational Linguistic (CL) approaches are employed. The ensemble method is GA-based. The analysis of several identification methods. Support Vector Machine incorporates Bose Chaudhuri Hocquenghem (BCH) loss correcting algorithm to lower loss.

Hussain et al. [179] propose a novel concept for Breast tumor segmentation using embedded U-Net model. For the purposes of evaluating, recognizing, and detecting cancers, breast tumor separation in B-mode ultrasound imaging is crucial. The degree setup strategy is one that is quite frequently employed for breast separation, and it is continually being refined. In order to utilize thematically augmented characteristics for breast tumor separation, this research suggests a brand-new deep-feature integrated level set block. A UNet-based model is initially trained to retrieve various characteristics at various phases. Every level depicts distinct aspects differently. Furthermore, at the conclusion of every phase, an unique level-set mechanism is added to acquire better detailed and realistic feature maps. In the performance component of the level-set technique, a novel feature-discriminator is developed to improve the low confidence values at the borders. Finally, to strengthen the separation procedure even more, the outcomes of the level-set approach at multiple phases are merged into the resulting feature maps.

Sendra-Balcells et al. [180] adopt domain generalization concept for MRI scans. For non-contrast scanning throughout the past few decades, the domain generalization issue has been extensively researched in machine learning, while for contrast-enhanced scanning, it has gotten less interest. In this article, they give a thorough analysis of deep learning methods for contrast-enhanced image separation that may be generalized to unknown healthcare situations.

In order to accomplish this, a number of techniques including data preprocessing, domain combining, transfer learning, and domain adaptation are researched, optimized, and rigorously assessed. The techniques are assessed for ventricular separation in contrast-enhanced cardiac MRI scans to show the possibility of domain generalization for contrast-enhanced scanning.

Yalçın and Vural [181] explore a new concept for brain injury using customized U-Net CNN model. For doctors to concentrate on certain parts of the brain and provide individuals the appropriate therapy, precise brain injury assessment, categorization, and separation are crucial. Numerous artificial intelligence technologies have successfully included encoder-decoder deep learning techniques. The Computed Tomography (CT) scans from the medical dataset employed to detect if a brain injury has occurred are examined by the suggested approach. Once an injury has happened, it is possible to ascertain whether it was triggered by ischemia or bleeding. The suggested approach can also perfectly separate the current stroke and highlight the area that the physician has overlay.

Qin et al. [182] proposed a novel concept for Airway Tree Segmentation using convolutional neural network (CNN). It is essential for bronchial illness detection and endobronchial navigation that the airways are segmented on CT images. The challenging construction and varied morphology of the airway need laborious hand-crafted extraction attempts. Convolutional neural network (CNN) oriented techniques have lately emerged as the most advanced technique for autonomous airway segmentation. Nevertheless, it is still difficult for convolutional neural networks to understand the inter-connectedness of the airways and recognize the tree-like layout. In order to solve this problem, we provide AirwayNet, a voxel-connectivity informed technique for precise airway identification.

Lian et al. [183] proposed a novel concept for Computer Tomography (CT) scans for lungs tumor classification using Tumor-CNN. Appropriate patient lifespan forecasting can help with medication management and may even boost performance because pulmonary disease is the most common cancer-related reason for death. In this research, utilizing Computed Tomography (CT) scans from individuals with Non-Small Cell Lung Cancer (NSCLC), we developed an autonomous method competent of pulmonary separation and mortality forecasting. In this retrospective investigation, we constructed unique pulmonary graphs from the 10 chest Computed Tomography scans segments and trained a GCN algorithm to estimate 5-year mortality risk. The present Tumor Nodes Metastases (TNM) staging system was compared to a series of machine learning algorithms, a convolutional neural network based on tumors (Tumor-CNN), and a Cox proportional-hazard predictor.

Zheng et al. [184] propose a novel concept for pre-operative diagnostic and intra-operative positioning for lung treatment, autonomous airway separation is a requirement. This is hindered by a significant imbalanced data across backgrounds and foreground parts, which finds it difficult for CNN-based algorithms to interpret distal minor airways due to the tiny dimension and dispersed spatial patterns of peripheral bronchi. In this study, we show how the neighborhood voxels' gradient degradation and elongation cause this issue. The salient gradients may be degraded by their neighborhoods throughout back-propagation if the frontal slope to overlay slope proportion is low and the imbalanced data is regional.

Zhang et al. [185] use 3D CNN for airway segmentation. For airway segmentation, 3D convolutional neural networks (CNNs) are frequently used. The medical datasets has a significant impact on how well 3D CNNs work, however the available airway datasets are primarily pristine Computed Tomography (CT) images with limited labeling, making it challenging to generalize to noisier COVID-19 CT scans. In this study, they suggest a novel dual-stream system that employs pristine Medical images and a tiny handful of labeled noisy Computed tomography for airway segmentation to solve the heterogeneity between the clean region and noisy region. To differentiate the distinctive noisy characteristics from the shareable clean characteristics, they create two distinguish encoders, accompanied by two distinct decoders.

In conclusion, CNNs are the prevailing industry norm for testing categorization. CNNs in particular have produced remarkably good outcomes after being trained on genuine images, surpassing the competence of human specialists in several applications. The researchers have also demonstrated how CNNs may be modified to take advantage of the inherent organization of medical images.

8 Anatomical Domains of Medical Images

An summary of the achievements made by deep learning towards the different healthcare imaging technology fields is provided in this section. We analyse machine efficiency on big data sets and on challenge data sets, and we emphasize certain valuable discoveries. On the website <http://www.grand-challenge.org>, these problems are all mentioned.

8.1 Lungs or Chest

The most frequently discussed topic in pulmonary image processing from ct imaging and radiography is the identification, description, and categorization of tumors.

Numerous studies evaluate CNNs with traditional machine learning techniques employing handmade features or supplement current feature sets with deep network-derived features. With an unified technique, numerous organizations may identify various illnesses with a lung X-ray. Another well-liked study area in CT is the identification of topographical features suggestive of interstitial pulmonary disorders [186–188].

The most frequent radiological examination is a chest X-ray, and various studies have used a huge collection of images and textual descriptions to build algorithms that integrate CNNs for visual interpretation and RNNs for textual processing. In the foreseeable future, we anticipate seeing additional study in this area [189–191].

All of the highest approaches in a current competition for tumor identification in CT, LUNA16, utilized CNN frameworks. In comparison, ANODE09's chest tumor identification challenge included handmade characteristics to categorize potential nodules. The top systems in LUNA16 continue to employ nodule possibilities calculated by rule-based image analysis, while deep network candidacy identification techniques also delivered excellent results (e.g. U-net). The 2017 Kaggle Data Science Bowl, with \$1 million in awards and also more than 1,000 competing groups, has as its goal predicting the likelihood that a person has lung disease from a Computed Tomography Scans [192–197]. COVID-19 Lung CT Lesion Segmentation Challenge - 2020 organized by MICCAI and more than 1976 teams are participating to predict the COVID-19. Airway Tree Modeling (ATM22) Challenge 2022 organized by MICCAI2202 and more than 288 teams are participating to predict pulmonary disease diagnosis and endobronchial navigation.

8.2 Digital Pathology (Histopathological Image) and Microscopy

Deep learning algorithms are increasingly being used in the field of digital pathology and microscopy due to the increasing accessibility of massive gigapixel Whole Slide Imaging (WSI) of cell samples. The approaches that have been created and used in this arena concentrate on three primary problems: Segmenting large organs, identifying and defining the condition of relevance at the tumor- or WSI-level and finding malignant cells on those regions.

Images from histology have also been normalized using deep learning algorithms. An essential topic is Color Normalization in the research of histopathology image processing. A technique for stain normalization of Hematoxylin and Eosin (H &E) stained tissue samples utilizing deep sparse auto-encoders was proposed by [217]. Janowczyk and Madabhushi [140] briefly demonstrated the value of color

normalization for CNN-based tumor categorization in H &E colored images.

Digital pathology has faced some significant obstacles, which has aided in the growth of computerized digital pathology methods. AMIDA 2013 Mitosis detection challenges sponsored by MICCAI Grand Challenge ICPR 2012 Contest on Mitosis Detection, EM segmentation challenge (ISBI - 2012) for the 2D segmentation of neuronal processes, GLAS for gland segmentation CAMELYON16 and TUPAC for processing breast cancer tissue samples are among the challenges that assessed both established and novel methods for analyzing digital pathology images [218–220].

In clinical specimens from colon tumor patients, the challenge of gland instance segmentation was handled by GLAS. The highest rank was attained by [221] using three CNN architectures. Pixels are categorized as gland or non-gland in the initial CNN. The comprehensively layered edge approach, which creates an edge map via side convolutions, is used to retrieve edge detection out of each feature space of the initial CNN. Lastly, the complete segmentation is created by merging edge and gland patterns in a final CNN architecture [222].

The IDZIA research group surpassed rival methods with a CNN-based strategy in both the ICPR 2012 and the AMIDA13 contests on mitosis recognition by [223]. For the 2-dimensional segmentation of neuronal networks at EM 2012, the very identical team's approach [224] obtained the best performance. As part of their method, a CNN's outcome was lightly smoothed and threshold-ed to accomplish the goal of dividing the membranes of cells.

The inaugural contest to deliver competitors WSIs was CAMELYON16. In contrast to certain other medical image processing, the vast amount of labeled data that was available in this competition permitted for the development of extremely deep models, such as the ResNet-101, VGG-Net-16, GoogLeNet-22 utilized by [91] and [94]. Each of these architectures was utilized by the top five devices. The CAMELYON16 challenge's top-performing approach was described in [225] The approach is built on an ensemble of two GoogLeNet architecture, one learned using hard negative mining and the other without it. AUC of 99.35 was obtained for job 2 in the most recent response from this research group utilizing the WSI standardization technique by [226], outperforming the AUC of a clinician who individually evaluated the entire test set (AUC = 96.6).

The most current TUPAC competition focused on predicting tumor classification at the Whole Slide Imaging (WSI) level and detecting mitosis in breast malignant tissue. The best solution was developed by [227], and it excelled at every challenge. The technique consists of three basic parts: Identifying areas with high cell numbers, using a CNN to identify mitoses there, turning the mitosis identification

Table 2 Selected articles of brain MRI using deep learning

Author	Details
Ragab et al. [198]	Two-phased segmentation approaches are used for breast tumor segmentation
Mambou et al. [199]	Used Independent Component Analysis (ICA) with convolutional neural network to classify breast cancer
Selvathi and Poornila [200]	Used Sparse-autoencoder, Stacked Sparse-autoencoder with Convolutional Neural Network for mammogram classification of breast cancer
Mohamed et al. [201]	Used multi-fold based technique for breast cancer on biopsy dataset
Kavitha et al. [202]	For diagnosing digital mammogram of breast cancer, they used Optimal multi-level Thresholding based segmentation with capsule network
Chowdhury et al. [203]	Used transfer learning approach with customized CNN and ResNet101 for classification of breast cancer
Escorcia-Gutierrez et al. [204]	Used CNN model with ResNet34 with distinct preprocessing steps for classification of breast cancer
Jasti et al. [205]	Used various distinct approaches for feature extraction, selection, image processing and classification of mammograms
Jabeen et al. [206]	Used five-fold based deep learning approaches for classification of breast tumor classification from ultrasound
Naseem et al. [207]	Used various machine learning based ensemble algorithms to classify breast tumor
Singh et al. [208]	Proposed a hybrid approach comprises residual and inception block of CNN for breast cancer classification
Liu et al. [209]	Used a pre-trained CNN model AlexNet and fine-tuned on BreakHis, IDC and UCSB datasets
Wang et al. [210]	Developed a novel deep learning approach DeepGrad model comprises InceptionV3 blocks for Histopathological image classification
Reshma et al. [211]	Used Fourier Transform based Segmentation for Histopathological image (Biopsy) classification of breast cancer
Ragab et al. [212]	Proposed Ensembled based deep learning approach containing multi-level thresholding based segmentation for breast tumor
Ahmad et al. [213]	Used Gated Recurrent Unit with pretrained CNN model (AlexNet) for classification of Lymph Node of breast tumor
Maqsood et al. [214]	Used multi-phase approach comprises contrast enhancement, Transferable texture using pretrained CNN models
Ibrokhimov and Kang [215]	Proposed two-stage CNN network to extract local patches from breast cancer and locate Region of Interest (ROI)
Mohamed et al. [216]	Used two-step approach, comprises of U-Net (CNN) for extraction of breast from the whole body and second step is to classify into binary classes

findings into feature vectors for each Whole Slide Imaging (WSI), and employing a Support Vector Machine (SVM) classifier model to calculate malignant growth and molecular information ratings are the next steps.

8.3 Breast

Breast tomography was one of the initial areas where [121] applied DNN. Subsequently, research has resurfaced, leading to considerable advancements over the state-of-the-art and obtaining ROI effectiveness comparable to that of sentient readers [248]. Although typical breast imaging approaches are two directional, strategies that work well with real images can be simply applied to artificial imaging.

Including one exemption, there are only three objectives that are acknowledged: (1) the identification and categorization of mass-like tumors, (2) the identification and categorization of microcalcifications, and (3) the image-based breast tumor probability assessment. The modality that receives the maximum utilization and, as a result, the maximum interest is mammography. There is presently little research on tomosynthesis, US, and tidal pulse elastography, and there is just one publication that used deep learning to evaluate

breast MRI. These additional modalities will certainly get increasing emphasis in the next decades. The literature's key points are summarized in Table 2.

While several nations have breast tumor monitoring programs, there ought to be tonne of data accessible, particularly for mammography, and consequently sufficient room for deep CNN models to thrive. Outdated digitized screen-film data sets are still being used because huge public digitized libraries are regrettably inaccessible. Competition like the just-launched MITosis DOmain Generalization Challenge 2022 (MIDOG 2022), Tumor Infiltrating lymphocytes in breast cancer (TIGER) and Breast Cancer Immunohistochemical Image Generation Challenge (BCI) challenge haven't yet achieved the results expected.

Numerous research examined limited data sets as a response, which had varying results. This problem has been explored in a number of studies by investigating transfer learning [249, 250], semi-supervised learning [251–253] and weakly supervised learning [254–256]. Some other approach integrates customized features with cnn networks, which has been demonstrated to be complimentary even for extremely large amounts of data [257–259]. Modern methods for classifying mass-like tumors often use a multiple workflow with

Table 3 Selected articles of Diabetic Retinopathy (retinal fundus images) using deep learning

Author	Details
Li et al. [228]	Used deep ensemble algorithm for to detect Diabetic Retinopathy (DR) using retinal fundus images
Pinedo-Diaz et al. [229]	Proposed a unique lightweight Convolutional Neural Network for detection of diabetic retinopathy using retinal fundus images
Saranya et al. [230]	Used pretrained CNN model (VGG-16) for classification of neovascularization using color fundus images
Boreiko et al. [231]	Proposed an adversarially robust ensemble model containing visual counterfactual explanations, Saliency maps (Integrated Gradients (IG) and Guided Backprop (GBP)) using retinal fundus images
Gunasekaran et al. [232]	Identified a novel model using deep recurrent neural network for classification of diabetic retinopathy
Saranya et al. [233]	Used a pretrained CNN block (DenseNet) for classification of daibetic mellitus
Mikram et al. [234]	Discovered a hybrid approach using distributed and non-distributed technique deep feature extraction and classification
Gao et al. [235]	Used a pretrained CNN models such as VGG16, RestNet50, and DenseNet for screening fundus fluorescein angiography (FFA) images
Abbood et al. [236]	Used a novel two-stage approach comprises of cropping, removing noise (gaussian blur), contrast enhancement and ResNet50 for classification of diabetes mellitus
Nneji et al. [237]	Used Contrast-Limited Adaptive Histogram Equalization (CLAHE) and Contrast-Enhanced Canny Edge Detection (CECED), to enhance the low quality images into high quality
Dayana and Emmanuel [238]	Used various preprocessing techniques for removal of noise, contrast enhancement and deep attention based Fusion Network to classify DR stages
Zhang et al. [239]	Developed a novel CNN model containing InceptionV3 characteristics for diagnosing severe Diabetic retinopathy
Canayaz [240]	Acquired 512 features using EfficientNet and DenseNet models and classify those features via Support Vector Machine
Venkaiappalaswamy et al. [241]	Used pretrained CNN models namely AlexNet and GoogleNet for classification of retinal fundus images
Gupta et al. [242]	Used guided and median filter in processing stage and employed U-Net for infection segmentation
Granty Regina Elwin et al. [243]	Proposed a novel Autoregressive-Henry Gas Sailfish Optimization (Ar-HGSO) based deep learning model for Macular Edema classification
Nderitu et al.[244]	For retinal and gradability classification using customized deep learning approach
Desika et al. [245]	Used Wild Geese Algorithm with customized Region-based Convolution Neural Network (R-CNN)
Islam et al. [246]	Used CLAHE and Xception based CNN model is used for detection of diabetic retinopathy
Babenko et al. [247]	Proposed a novel deep learning model for classification of diabetic macular oedema and poor blood glucose control using retinal fundus images

a candidate sensor; this architecture condenses the image to a list of possibly cancerous tumors that are given to a deep CNN [260–264].

To avoid the cascaded method by using a region-based proposal network (R-CNN) [265–267]. Excellent outcomes can be attained when there are massive data sets accessible. A research scientist from a top breast imaging CAD organization described to a crowded meeting hall at the SPIE Medical Imaging conference in 2016 how just a several times of experimental studies with a basic deep CNN architecture (AlexNet) prepared on the business's patented technology database produced effectiveness that was better than what decades of designing manual feature structures had produced [268].

8.4 Eye

Although ophthalmic scanning has advanced significantly in previous decade, deep learning techniques have just

gradually been used to comprehend eye images. The majority of studies use straightforward CNNs to analyze Color Fundus Images (CFI), as seen in Table 3. Techniques for segmenting anatomical features, detecting and segmenting retinal anomalies, diagnosing eye disorders, and evaluating image resolution are all covered.

In order to create systems to determine the extent of the infection in 53,000 trial images, over 35,000 colour fundus scans were supplied in a diabetic retinopathy detection challenge conducted by Kaggle in 2015. The bulk of the 661 groups that participated in the challenge used deep learning, and four, all of which used end-to-end CNN model, outperformed individuals in terms of outcomes. A Google Inception v3 model just underwent a rigorous investigation by [144] in order to identify diabetes mellitus, and the results showed accuracy on par with a board of seven board-certified ophthalmologists.

The top-performing groups to date have all employed CNNs in the 2022 and 2021 challenges such as Diabetic

Table 4 Selected articles of brain MRI using deep learning

Author	Details
Havaei et al. [272]	Tumor segmentation; two-path way CNN with different receptive fields
Moeskops et al. [273]	Tissue segmentation; CNN trained on multiple patch sizes
Ghafoorian et al. [274]	Lacune detection; FCN for candidate segmentation then a multi-scale 3D CNN with anatomical features as false positive reduction CNN
van der Burgh et al. [275]	Survival prediction; DBN on MRI and fusing it with clinical characteristics and structural connectivity data
Kamnitsas et al. [276]	Tumor segmentation; 3D multi-scale fully convolutional network with CRF for label consistency
Kleesiek et al. [277]	Brain extraction; 3D fully convolutional CNN on multi-modal input
Ghafoorian et al. [278]	Lesion segmentation; CNN trained on non-uniformly sampled patch to integrate a larger context with a foviation effect
Havaei et al. [279]	Tumor segmentation; CNN handling missing modalities with abstraction layer that transforms feature maps to their statistics
Dou et al. [280]	Microbleed detection; 3D FCN for candidate segmentation followed by a 3D CNN as false positive reduction
Shi et al. [281]	AD/MCI/HC classification; multi-modal stacked deep polynomial networks with an SVM classifier on top using MRI and PET
Bashir et al. [282]	Using Stationary wavelet transform (SWT) and Principal Components Analysis (PCA) for Brain image fusion
Panigrahy et al. [283]	Using a Weighted Parameter Adaptive Dual Channel PCNN
Yazdan et al. [284]	Using multi-scale CNN for robust classification of brain tumor and eliminate the affect of Rician noise
Wahlang [285]	Utilizing pretrained CNN model for feature extraction and use traditional classification algorithm SVM to classify Magnetic Resonance Imaging (MRI) for brain tumor classification
Alanazi [286]	Use 22-layer custom CNN with transfer learning for multi-classification of brain tumor
Bangare [287]	Use Modified Region Growing, Gabor filtering are used and Fuzzy Min-Max Neural Network is used for classification
Siddiqi et al. [288]	Use Stepwise Linear Discriminant Analysis (SLDA) with Support Vector Machine
Ouchicha et al. [289]	Use customized dense block with pretrained CNN model for Alzheimer's Disease Classification Using Brain MRI
Haq et al. [290]	Three step methodology is proposed (1) initially extract pertinent features of brain MRI space. (2) for localizing tumor area use a faster region-based CNN (3) conventional and traditional CNN algorithm is used to categorize and segment the brain tumor
Saurav et al. [291]	Constitutes a novel lightweight attention-guided convolutional neural network (AG-CNN) with skip connection classification of brain tumor using MRI images
Nayak et al. [292]	Propose a parametric based metaheuristics methodology is used with CNN model to classify brain tumor using MRI

Retinopathy Analysis Challenge (DRAC-2022) sponsored by MICCAI 2022, RAVIR: A Dataset and Methodology for the Semantic Segmentation and Quantitative Analysis of Retinal Arteries and Veins in Infrared Reflectance Imaging 2022 [269], AIROGS: Artificial Intelligence for ROBust Glaucoma Screening Challenge sponsored by IEEE ISBI 2022 [270, 271].

8.5 Brain

Neuroscience image processing using DNNs has been widely employed in a variety of professional sectors Table 4. Numerous investigations categorise Alzheimer's dementia and divide up the brain's cell and structural elements (e.g. the hippocampus). The identification and segmentation of diseases are additional crucial regions (e.g. tumors, white matter lesions, lacunes, micro-bleeds).

The majority of approaches train projections from regional portions to representations and then from

abstractions to categories, with the exception of those that seek for a scan-level categorization (such as Alzheimer assessment). But for applications where anatomical knowledge is crucial, the regional portions could not have the contextual details needed (e.g. white matter lesion segmentation). Ghafoorian et al. [274] employed non-uniformly observed patches to address this issue by significantly reducing sampling frequency on patch ends to encompass a wider context. Multi-scale investigation and the synthesis of images in a fully-connected layer are alternate methods that are employed by numerous entities. Despite the fact that all of the investigated research used 3-dimensional volumes of brain imaging, the majority of approaches only analyze the 3-dimensional volumes slice-by-slice. The thin slices compared to in-plane sharpness in specific data collections or the decreased computational resources are frequently the driving forces behind this. Additionally, more published studies had used 3D networks.

Table 5 Selected articles of Abdomen (CT/MRI/Colonoscopy) using deep learning

Author	Details
Lu et al. [297]	Competing on SILVER07 using 3D CNN and produce plausible outcomes
Ravishankar et al. [298]	Using transfer learning, tuning of existing model and hybrid approach to acquire credible findings
Roth et al. [299]	Using random forests and two CNNs, we can distinguish the pancreas' inside and exterior
Näppi et al. [300]	In our approach, digital endoluminal imaging of the potential polyps were used to recognize polyps in response to a computer-aided detection (CADe) methodology. The pretrained CNN models are used with millions of natural images
Ren et al. [301]	The unique Computed Tomographic Colonography based CAD technique that is proposed in this study uses three-dimensional radiomic characteristics to provide excellent identification specificity at low FP rate. To provide prospects for polyps, we apply our prior shape-based techniques
Uemura et al. [302]	They built a 3-dimensional Convolutional Neural Network oriented on a flow-based generative model. Using our medical Computed Tomographic Colonography case data, the 3-dimensional Glow was programmed to generate artificial Volumes of Interest of polyps
Yasuda et al. [303]	A simulation experiment was conducted before this one in order to determine the best procedure for ultra-low-dose Computed Tomography Colonography (CTC). 206 individuals were enlisted who had a significant or medium risk of developing colorectal cancer. After completing complete bowel preparation, the individuals were examined in the upright and supine postures using an adaptive incremental reconstruction technique with the Computed Tomography settings set to 120kV, standard error 45 to 50
Younas et al. [304]	In contrast to transfer learning, this research examines six previous pretrained Convolutional neural networks models and chose the best outperforming architecture exclusively for traditional ensemble structures
Tanwar et al. [305]	The colonoscopy images are primarily enhanced and filtered using dynamic histogram equalization and directed image filtering techniques. Then, colorectal polyps in colonoscopy images are accurately detected and classified using Single Shot MultiBox Detector. Lastly, the polyp categories are categorized using fully linked layers with dropouts
Wesp et al. [306]	To estimate the polyp category, two deep learning networks, Segmentation (SEG) and noSEG, were learned using 3D CT colonography image sub-volumes. System SEG was also learned using polyp segmentation markers
Biffi et al. [307]	In order to perform in real-time using a traditional white light (WL) endoscopic video stream without the use of simulated chromoendoscopy, researchers created a revolutionary sophisticated medical equipment (blue light, BL). On a systematically obtained dataset of unmodified colonoscopy recordings, researchers investigated the stand-alone effectiveness of this computer-aided diagnosis tool (CADx) in this research
Sánchez-Peralta et al. [308]	The initial step is to describe the components needed for implementing supervised DL techniques for Colorectal Polyps in colonoscopy. The system design, collection, Error function, data enrichment, and indicators are highlighted. While covering techniques for polyp identification, positioning, separation, and categorization, the publically accessible datasets that may be beneficial for future research are discussed
Kusters et al. [309]	Confidence calibration directly mentions the trustworthiness of a Convolutional Neural Network (CNN) for identification of Colorectal Polyps in this research. Well-calibrated algorithms generate categorization ratings that accurately represent the model probability of being accurate, enabling accurate forecasts with dependable and illuminating confidence ratings
Byeon et al. [310]	Every diseased disc was digitally captured in order to fine-tune two Convolutional neural networks that had already been developed
Nisha et al. [311]	Using a DP-CNN to categorize polyp and non-polyp scan from the colonoscopy images
Cao et al. [312]	To extract polyps descriptors with the help of the gradient operator and Hessian operator

Several issues in brain image processing have been totally replaced by Cnn architectures. The top-performing groups to date have all employed CNNs in the 2014 and 2015 tumor segmentation challenges (BRATS), 2015 longitudinal multiple sclerosis lesion segmentation challenge, 2022 Ischemic Stroke Lesion Segmentation Challenge (ISLES'22), CuRIOUS 2022 Segmentation Challenge (comprising two major task (1) brain tumor

segmentation using intra-operative US (2) cavity segmentation using intra-operative US), automatic segmentation of Head and Neck (H & N) for primary and lymph nodes (HECKTOR challenge in 2020, 2021 and 2022), Surface Learning for Clinical Neuroimaging (SLCN 2022) sponsored by MICCAI2022 [293–295], Anatomical Tracings of Lesions After Stroke (ATLAS R2.0) [296] and 2013 MR brain image segmentation challenge (MRBrains). The

Table 6 Selected articles of Musculoskeletal (CT and MRI) using deep learning

Author	Details
Harini et al. [313]	Using pre-trained CNN models such as Xception, Inception, VGG-19, DenseNet, and MobileNet to analyze Musculoskeletal Radiographs (MURA) data set
Kamiya [314]	Analysis of Whole body Computed Tomography (CT) scans using distinct image features such as area, shape, volume and texture
Kijowski et al. [315]	Using MRI and CT scans to detect lesion and forecasting of musculoskeletal illness
Mall and Singh [316]	Proposed a customized CNN (ChampNet) with multi-phase training to classify musculoskeletal using X-ray images
Singh et al. [317]	Discovered a hybrid CNN ComDNet-512 containing three steps training for finger binary classification
Cheng et al. [318]	Identified an eminent resolution based Generative Adversarial Network to generate and enrich the network training of Musculoskeletal Ultrasound (MSUS)
Zhu et al. [319]	To segment inter- and intra-slice features of Cerebral Palsy (CP) using end to end attention based hybrid approach
Shin et al. [320]	To diagnosing musculoskeletal disorders such as meniscus using Convolutional Neural Network
He et al. [321]	Discovered Musculoskeletal model (MSM) to acquired from Electromyography (EMG) to predict joint angle of human body
Malik et al. [322]	Proposed a three-phase hybrid CNN model for classification of elbow fracture
Gitto et al. [323]	Atypical Cartilaginous tumour (ACT) and chondrosarcoma (CS ₂) of long bones classification using MRI + Radiomics classification
Luong et al. [324]	Used transfer learning technique for binary classification of MUculoskeletal Radiographs (MURA) dataset
Gao et al. [325]	Used Modic Changes (MCs) for diagnosing MRI with T ₁ - and T ₂ -weighted using V-Net
Fabry et al. [326]	Used MRI for discriminating Facioscapulo-Humeral Dystrophy (FSHD1) and myositis for visual grading

Table 7 Selected articles of Cardiac (CT/MRI/US) using deep learning

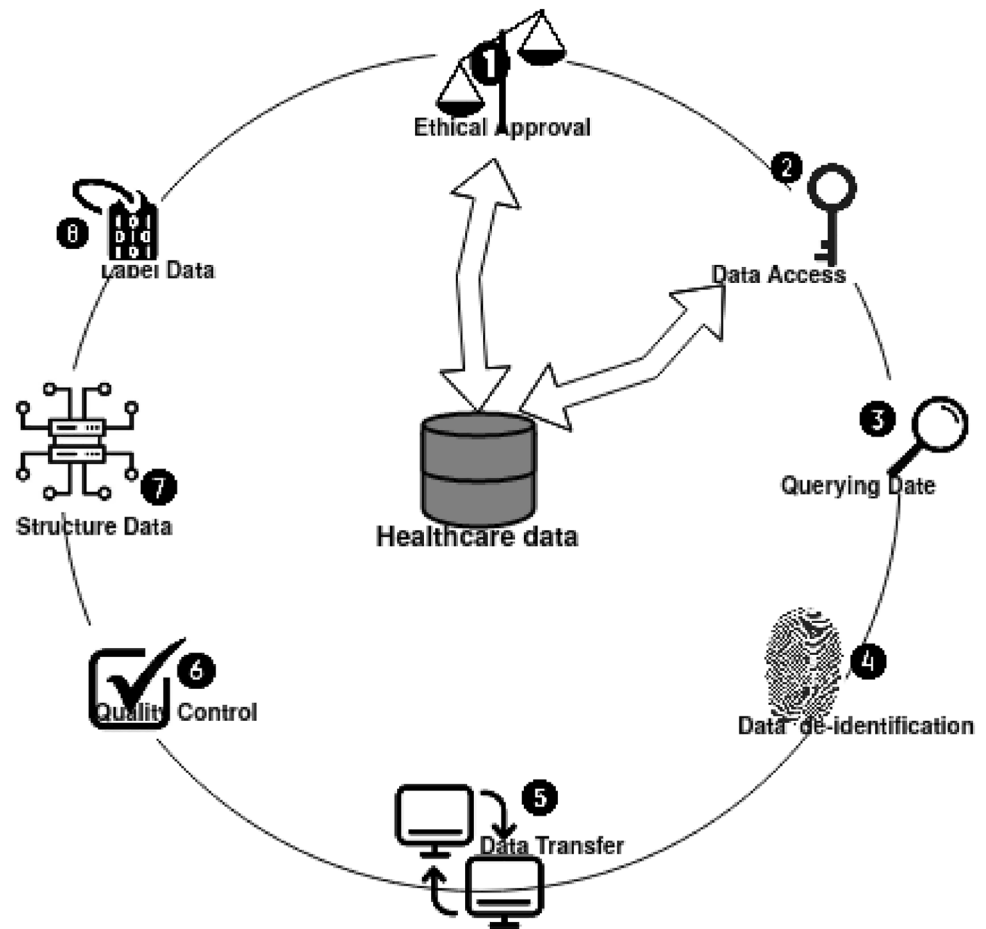
Author	Details
Lu et al. [330]	In this paper, SparseVoxNet, a brand-new and highly efficient U-Net-based 3-dimensional sparse convolutional model, is proposed. Any two layers that have the identical feature size have deep links in this network, which decreases the quantity of interconnections
Song et al. [331]	Utilizing different deep learning techniques such as Recurrent Neural Network, Long-Short Term Memory (LSTM), Boltzmann Machines and U-Net for cardiac MRI segmentation
Wang et al. [332]	A Multi-scale Multi-skip connection Network (MMNet) is suggested to completely outsource the left cardiac ventricular separation of cardiac magnetic resonance imaging (MRI) data, influenced by the potential of Convolutional Neural Network
Sandooghdar and Yaghmaee [333]	Using U-Net for segmentation process and random forest classifier for classification
Ahmad et al. [334]	Using ensemble method 2-dimensional residual network with Atrous Spatial Pyramid Pooling (ASPP) for cardiac MRI segmentation
Alabed et al. [335]	This research uses machine learning to evaluate the predictive cardiac magnetic resonance characteristics in Pulmonary arterial hypertension
Fernández-Llaneza et al. [336]	Using Customized U-Net model for evaluating systole and diastole phases (2MSA) and another model trained for all phases(1MSA)
Popescu et al. [337]	3-stage neural network is trained to identify and segment ROI of left ventricle (LV)
da Silva et al. [338]	This solution employs a feedforward strategy that combines FCNs and image analysis methods
Arian et al. [339]	using traditional machine learning and radiomics algorithms to predict myocardial function in cardiac MR (LGE-CMR)
Corrado et al. [340]	The separation of 4D flow velocity data was then applied using the bSSFP images that had been mapped to the appropriate 4D flow images
Lalande et al. [341]	Using inverted Residual Blocks with U-Net for assessment of myocardial infarction

majority of the proposed techniques focus on brain MRI imaging. We anticipate that deep learning evaluation will also be beneficial for other brain scans modalities like CT and US.

8.6 Musculoskeletal

Deep learning techniques have recently successfully employed to dissect and identify bone, ligament, and accompanying tendon anomalies in musculoskeletal

Fig. 24 Management of health care data cycle



images from different imaging techniques. Table 6 provides a summary of the publications. There are a remarkable number of finished implementations with encouraging outcomes; one that sticks out is [327] approach, which they learned with 12K discs and promised to execute nearly as well as humans on four separate radiological assessment challenges.

8.7 Abdomen

The majority of abdominal publications focused on localizing and segmenting various parts, primarily the liver, kidneys, bladder and pancreas Table 5. Partition of liver tumors is covered in two studies. For prostate examination, MRI is the primary modality, while CT is used for all other parts. Only in the colon are different techniques discussed, but they were generally done simply: a CNN was employed

to extract features, and these characteristics were then used to classify data [328].

It is important to highlight that relatively conventional image processing techniques predominated up to 2016 in two segmentation challenges, SLIVER07 for the liver and PROMISE12 for the prostate. The present second and third-place autonomous algorithms in PROMISE12 used active appearance algorithms. For approximately five years, IMorphics' approach held the top spot (now ranked second). Meanwhile, [329] a 3D fCNN that is comparable to U-net just overtook the top spot. This study employs a novel method by combining the ResNet and U-net architectures by using a sum procedure in place of the composition process employed in U-net. Convolutional networks have also begun to compete at the top of the rankings in SLIVER07, a 10-year-old liver segmentation competition,

replacing previously dominating approaches centered on structure and aspect modeling in 2016.

8.8 Cardiac

The research is compiled in Table 7 and shows how deep learning has been employed in several parts of cardiac image processing. Although left ventricle categorization represents the most frequently performed challenge in MRI and the methodology that has received the most investigations, there are many other implementations that can be used, including segmentation, monitoring, slice categorization, visual performance evaluation, automated calcium grading, coronary alignment monitoring, and super-resolution.

The majority of research employed straightforward 2-dimensional Convolutional neural and examined the 3-dimensional and frequently 4-dimensional data scan by scan; [342] develop an exception framework using 3D CNNs. Four publications use Deep Belief Networks, however they all come from the exact same author team. The Deep Belief Networks are incorporated into complicated segmentation architectures and are only employed for extracting features. Due to their integration of Convolutional Neural Networks and Recurrent Neural Networks, the following two publications stand out: In order to separate the left ventricle scans by scans and retain what characteristics to recall from the earlier scans when separating the subsequent one, [343] developed a recurrent connection inside the U-net framework. Kong et al.[344] performed sequential regression to detect particular images and a cardiac pattern using an architectures containing a two - dimensional CNN and a Long short - term memory. Readily viewable data are used in numerous articles. The 2015 Kaggle Data Science Bowl was the biggest challenge in this area, with the objective of autonomously measuring end-systolic and end-diastolic contents in cardiac MRI. The top group in a competition with 192 groups for a winnings of \$20,000 each used deep learning, specifically fCNN or U-net categorization techniques.

9 Workflow

9.1 Data Preparation

Several procedures must be followed prior healthcare data can be utilized to create an deep learning based algorithm or application. Before using healthcare data to construct a scientific or industrial artificial intelligence based system or application, the appropriate regional ethical committee permission is generally necessary. An organizational regulatory committee must assess the potential and client rewards of the trial. Available data are frequently utilised, necessitating a retrospective analysis. Formal written

permission is typically omitted considering the individuals in this kind of research are not required to perform any further treatments. In medical studies, each lead researcher may be required to give consent before sharing patient data. When research data are collected continuously, as in a controlled trial, written permission is required. After receiving ethical permission, it is necessary to acquire, query, correctly de-identify, and safely archive pertinent data. Both the Digital Imaging and Communications in Medicine (DICOM) metadata and the images must be cleared of any sensitive healthcare facts [345, 346].

Particularly when building industrial methods, Intelligence scientists frequently do not even have immediate connection to healthcare imaging data via the Picture Archiving and Communication System (PACS) because they are not generally housed inside a radiology clinic or hospital. Only authorized experts such as medical practitioners, doctors, technologists, PACS managers, and medical scientists are allowed accessibility to PACS systems. Providing healthcare data available to deep learning scientists are complex and takes numerous stages, particularly de-identification of data. The optimal strategy is interaction between physicians and deep learning programmers or developers, either in-house or via collaborative academic agreements [347].

Even though the U.S. does not usually require individuals to provide formal documented permission, both prospectively and retroactively collected data must be properly de-identified in order to comply with the Health Insurance Portability and Accountability Act (HIPAA), or and the European General Data Protection Regulation. Among other things, confidential material may contain your name, your medical record number, and your birth date. Personal data is frequently contained in the DICOM metadata (header), and there are numerous methods accessible to effectively eliminate this material [348].

The last method for anonymizing health data is k-anonymity, which modifies an initial information source including private healthcare data in order to keep prospective hackers from discovering the identify of the individual. The DICOM information is frequently totally erased or transformed to some other standard, such as NIFTI (Neuroimaging Informatics Technology Initiative), which simply keeps track of voxel size and participant position, when sharing radiology images in open-source scientific projects. While eliminating all DICOM metadata from open-source scientific projects addresses privacy concerns, doing so lowers the quality of the information because metadata is essential for the creation of AI algorithms [349].

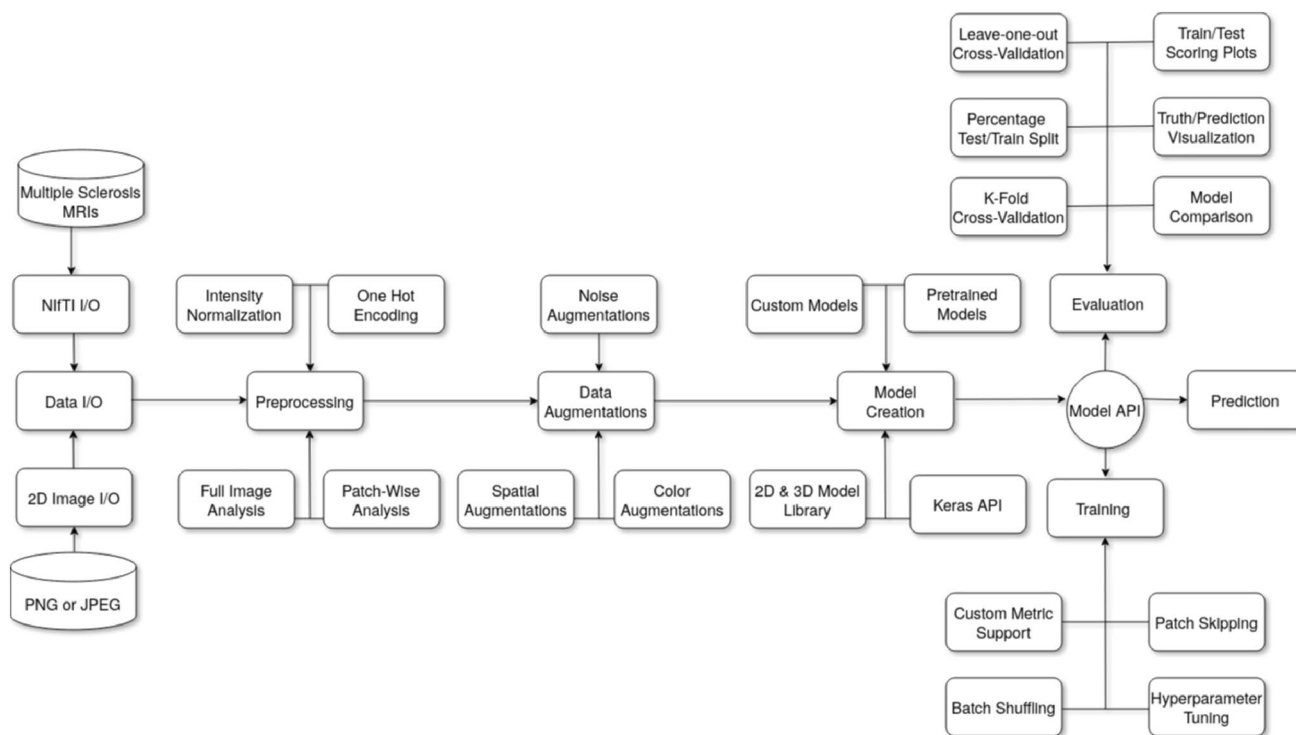


Fig. 25 Flowchart showing the data inputs from different modalities at the beginning and a training pretrained or custom deep learning model at the conclusion

9.2 Format

One of the most challenging aspects of processing healthcare images is understanding the image data format. The main file formats now employed in medical imaging are Analyze, Neuroimaging Informatics Technology Initiative (Nifti), Minc, and Digital Imaging and Communications in Medicine (DICOM) [350].

The amount of nibbles utilized to represent the data for every pixel is known as pixel depth. Each image is retained in a file and saved in a system memory as a collection of bytes. The minimal amount that can be saved in a system memory is a byte, which is a collection of 8 binary bits [351, 352].

A healthcare image is an organized collection of visual components called pixels or voxels that depicts the actual shape or activity of an anatomical area. It is a categorical formulation that links quantitative value to spatial positions as a consequence of a sampling/reconstruction operation. The amount of pixels employed to characterize a particular collecting modality's field-of-view is an indication of the level of complexity that can be captured in the structure or operation [353, 354].

9.3 Metadata

Metadata are details that the image uses to define itself. It may sound unusual, but in every storage format, the images has additional information beyond the image pixels. The file header containing the image dimensions with all coordinates, the depth of pixel, the spatial resolution and the photometric interpretation are contained in this data, which is known as metadata, which is normally put at the start of the document as a preamble [355, 356]. According to the characteristics of medical datasets, metadata play a larger significance in this situation. Images obtained from diagnostic modalities frequently contain details about the image's creation. For instance, the pulse sequence used to create a magnetic resonance image will contain characteristics such as time accusation, flip point of view, amount of acquisitions, etc. The radio-pharmaceutical that was administered and the patient's weight are both included in nuclear medicine images like PET images [357, 358].

9.4 Model Building

The construction of an acceptable deep learning based CNN model for a particular issue in medical image analysis is challenging and necessitates extensive analysis and testing. Before selecting a CNN model that best achieves their goals,

professionals must become knowledgeable in a variety of models and evaluate their efficiency traits [31, 359]. Figure 25 depicts the building of a deep learning model comprises all necessary step that we follow.

For instance, there is currently no universal way to evaluate a predetermined the ideal configuration of neurons and layers for an CNN model provided a scenario statements. A popular strategy begins with an educated assumption based on past knowledge about networks used to solve associated issues. This assumption might be based on first-hand knowledge from a person or second or third-hand knowledge from a training deep learning program, weblog, or academic publication. While choosing a course of action, the analyst may then experiment with other options and thoroughly evaluate the classification results [360–362]. Several hyper-parameters communicate with the dimensions and height of deep learning neural networks, thus altering one hyper-parameter can have an impact on the others. A straightforward step-by-step strategy is to build a CNN network with multi-level hidden layers that are the same or variable size as the source:

- Try adjusting the CNN network's height and breadth with multiple hyper-parameters [363].
- Try various methods (such as dropout, pooling, kernel/filter, regularization, momentum, decay, learning rate, activation functions, optimization algorithms, batch size and loss functions, etc.) as well as eliminating certain cells [364, 365].
- Choose a model that is more accurate altogether after making a few tweaks.

There will constantly be improved algorithms, thus users should really not waste their time fine-tuning deep learning models. Forming a realistic assumption of precision is aided by investigating the data. To set goals to beat, start with straightforward linear tactics. Various deep learning algorithms could be more useful, quick, and efficient than your initial choice [366–368].

9.5 Training, Testing and Validation

For the purpose of developing reliable machine learning algorithms, the source data must be divided into training, testing, and validation sets. Beginning with a training dataset is necessary for algorithms for machine learning. In order to lessen this disparity, it measures the discrepancy between both the results that were anticipated and those that really occurred throughout each repetition. The generated technique is specially adapted to the training set of data. The finished method and its learned characteristics are evaluated on a different evaluation data set to see how generalisable they are. Training, validating, and evaluation samples

are often required for deep learning based CNN models. In real-world situations, the entire population (dataset) is initially divided into a training and a testing datasets. A training and a validation datasets are created from the learning materials (training datasets). To adapt the model, learning samples are utilized. A authentic assessment of the algorithm during learning is provided via validation samples. Testing datasets allow for an immediate assessment of the finished product. The ideal distribution of the three different samples is not predetermined [369, 370]. K-fold cross-validation is the process of repeating the division K time periods and averaging the correctness to make sure the model is impartial. The biggest problem that could occur is the system becoming over-fit to a particular training dataset. Whenever this happens, the system will be very efficient on the training data but extremely wrong on the testing set of data (low false negative rate among expected and real outcomes). Another distinct validation samples is applied to get around this.

9.6 Sensitivity Analysis

[371] made the initial suggestion for sensitivity analysis in the context of misdiagnosis brought on by weight perturbations as a result of chaotic intake and system inaccuracy [371]. Since then, the phrase “sensitivity analysis” has become saturated with connotations that are tied to one another. The subject of neural network susceptibility to parameter error has indeed been extensively studied and documented [372]. Sensitivity analysis is defined here as the investigation of the impact of input modifications on system projections. The method that employs technique sensitivity analysis for cell segmentation is the one that is most similar to the one that is provided here [373]. The overall concept is the same in this study, but it varies in a number of ways, including the use of automated measurement discovery and the emphasis on computing efficiency [374]. Deep learning unpredictability is measured using a variety of techniques called Sensitivity assessment. It examines the significance of each source variable's characteristic for every outcome. Comparing the system outcomes with all factors involved and the model with one element eliminated, or maintaining the quantities of all other factors and simply adjusting the strength of one input element, allows for the measurement of the effect of each input variable [375]. Sensitivity and Specificity analysis is essential for the actual uses of deep learning applications in the real life; it clearly and unambiguously demonstrates the extent to which CNN model output depends on each factor and gives medical practitioners and specialists more authority, particularly when the recent discoveries are catastrophic events that may represent additional anomalies that surpasses the predictive power of CNN models [376,

377]. Categorization method sensitivity assessment can be qualitative or quantitative. In the qualitative scenario, both the initial data and its altered (for example, flipped) variants are used in the categorization process. The researcher is shown the classification filters as overlays over the modified photos for review. A ground truth separation of the source image must be available for quantitative sensitivity evaluation. By computing a comparison measure between the predicted and reality, the categorization efficiency of the algorithm may then be assessed. According on the conversion (such as color and brightness modifications), the ground truth mask either stays unchanged or also has to be altered (e.g., flip, crop and zoom).

9.7 Training Optimization

Optimization approaches, as a turbocharger in neural network techniques, perform a significant role in this subject. Due to their quick convergence, adaptive optimizers like ADAGrad [378] and ADAM [379] have revolutionized deep learning for the the past several years. The fundamental principle of these optimizer is to take gradient history into account while changing variables in the phase cycle.

The majority of optimization techniques have been developed, evaluated, and assessed in the research for classification techniques where the prediction range is generally constrained to a few dozen examples. Nevertheless, since each pixel is projected to have a tag in healthcare image classification issues, the forecast of state space is substantially greater; as a result, the usual optimizer of selection may not always be the best option [380].

For many newcomers, tweaking Artificial intelligence systems is a necessary but hard step. In order to reduce the objective functions, hyperparameters are adjusted. Optimization techniques are programs that modify characteristics like batch size, epochs, weights and learning rates to reduce losses. Adaptive Moment Estimation, Nesterov's Accelerated Gradient (NAG), Gradient Descent (GD), AdaDelta, and other frequently implemented optimizers are examples [381]. The majority of minimum locations discovered by gradient-based optimizers are local minima, which presents a frequent difficulty. Since the gradient gets less as training progresses and the learning rate is too high to approach the correct response, it is challenging to find the global minima. The genetic algorithm, which integrates the idea of development to ML, seems to be another method. Only the finest models emerge at the conclusion of the procedure, which is performed several occasions [382]. Every optimization technique has weaknesses. A one-size-fits-all approach cannot adjust to every dataset and accelerate learning to arrive to minima more quickly. The learning must swiftly converge to the goal with the least amount of loss in the fewest possible

repetitions or epochs. It is important to properly solve the gradient disappearing concern (the gradient is too tiny to adjust the weight in the following loop) [383, 384].

9.7.1 Parallel Computing

Parallel computing or multitasking is a useful technique in medical image analysis since it increases the effectiveness of deep learning based medical image segmentation, classification, registration and other operations. The first concerns the continuously growing size of healthcare data especially images that is now available as a result of improvements in both research and development methods and mathematical simulation or visualization of healthcare data [380, 385].

The second argument is that AI algorithm intricacy is rising. The complexity of ML models, particularly DL models, is increasing as predictive performance is sought for. As an illustration, the Microsoft Turing natural language generation model comprises 17 billion parameters. Due to these two factors, training a sophisticated AI algorithm without concurrently computation may take many days or even months [386, 387].

Numerous initiatives have been made to investigate ways to enable concurrent ML from various angles [388, 389]. The following is a summary of three concurrent ML potential, the first of which is applicable to all ML jobs and the second and third of which are exclusive to medical image analysis using deep learning. The first potential is the need to create a unified system that combines simultaneous deep learning algorithm with parallel hyper-parameter adjustment [390, 391].

Nowadays, these two activities are frequently carried out using various technologies, such as Spark for concurrent hyper-parameter tweaking and Tensorflow for simultaneous deep learning. Still required is a framework or method that supports both effectively but is more integrated. The capability for concurrent training on top of datasets for the system based on medical image analysis using deep learning. Current collaborative initiatives on quickly acquiring and analyzing DICOM and Nifti datasets. However, it is currently unclear how to incorporate these methods with machine learning. The final chance is concurrent machine learning capability for Histopathological, MRI, CT and X-ray datasets containing images with clinical details, which are typical of the medical image analysis. When splitting clinical data and medical images, as opposed to conventional independent and identically distributed (IID) datasets, the clinical/medical dependency and correlation would be broken. Therefore, parallel deep learning using clinical data and medical images has to be given specific consideration [392–394].

The training issue, the throughput constraints, and the accessible budget all play a role in selecting the best infrastructure for DL. The growth of Graphics Processing Units (GPUs) and advances in parallel computing have considerably boosted the performance of Deep Neural Networks for clinical applications. Due to their efficient massively parallel processing architecture for both training and interpretation techniques, Graphic Processing Units are crucial in deep learning. An integrated GPU chip on the computer system is needed for the usual operation. There are hundreds of Arithmetic Logic Units (ALUs) for each GPU core. The very identical directions will be applied to a significant number of neurons in CNN through each layer [395]. It's fairly usual to parallelize processing operations over numerous GPUs because a single GPU can't handle large-scale deep learning delegates. By utilizing additional dispersed resources, distributed computing is an effective concurrent method to improve DL efficiency [396, 397].

9.7.2 Explainable AI

The term smart healthcare describes the application of techniques like cloud computing, the Internet of Things (IoT) and artificial intelligence AI to create a healthcare system that is effective, accessible, and individualized [1]. These platform enables real-time medical surveillance via medical apps on smartphones or wearables, enabling people to take charge of their own healthcare. In addition to being communicated with doctors for additional assessment, user-level health records may also be utilised in conjunction with AI for illness monitoring, early diagnosis of diseases, and specific medication assessment. There is a demand for Algorithms that can be described in the medical field due to the ethical dilemma of openness around AI and the absence of confidence in the black-box functioning of AI systems [398]. Explainable AI (XAI) approaches are the AI approaches that are used to illustrate AI models and their outcomes [399].

Complicated deep learning models (e.g., U-NET, DenseNet, ResNet, Inception and other GAN based models) cannot offer a self-explainable explanation for their outcomes in comparison to basic or tree-structured Machine learning (e.g., linear regression, DT, Bayesian, RF). Machine learning techniques should include description, according to several medical image analysis scientists, to make them easier to grasp and increase user confidence. The initial "black-box" model may be examined using explainable AI (XAI) tools, which offer "explanations" that give a subjective knowledge of the interactions among model elements and projections.

This procedure provides solutions to inquiries concerning the model, including which aspects are most crucial and the reasons why some aspects influence outcomes more

than others. It also offers information that enables the algorithms to be changed in a significant way. Molnar et al. [400] provide an assessment of frequent comprehensible procedures. The necessity for model-agnostic or model-specific procedures, the level of description necessary, and spatiotemporal or computational limits can all be key considerations when choosing XAI approaches.

The inability of existing XAI techniques to identify issues in the training sample as well as their emphasis on RGB images and user-friendliness for higher dimensional space visuals are among its drawbacks [401]. Given the issues, XAI offers possibilities for enhancing healthcare based CNN models. XAI might identify artifacts that lead to inaccuracies in numerical algorithms.

9.7.3 Generalization

The traditional objective of generalization is to improve the performance of learned AI models on testing samples. It gets challenging, though, because the training sample is a tiny fraction of the enormous medical images sample. Maintaining the ideal balance among over-fitting and under-fitting is no more sufficient for deep learning based medical image analysis since algorithms developed in one location at one time could not be applicable in another location at a later date. Conventional statistically learning algorithms, meanwhile, do not generalize effectively on fresh data with various distributions, which is a major contributor to typical AI errors. A significant barrier to using AI in medical image analysis research is identifying a sound generalization approach to make model performed outside of the training samples [402, 403].

Generalization philosophy, which is still in its infancy, could be able to offer solutions to these issues. The generalization of AI has been extensively researched. The Ockham's Razor principle described by Ariew [404], it demonstrates that a positive estimated results is more likely to be due to factors other than the idiosyncrasies of the selected examples the simpler the theory is. The boundaries between inadequate and excessive training sample learning are hazy. The division of data into a training sub-samples and an assessment subset is one of the traditional techniques for identifying under-fitting or over-fitting [405]. The software will execute the classification classifier on the testing subset at each iteration process in order to determine the accuracy rate on data that are not in the initial training pool. The algorithm is overfitting if the precision of testing samples begins to steadily decline. On the other hand, it indicates that the system is still under-fitting if the evaluate the efficiency hasn't achieved its optimum [210].

Finding a balance among bias (under-fitting) and variance (over-fitting) requires a generalize and proper technique. Cross-validation is a popular method for ensuring that

there is no accidental learning bias. Another method used to improve how well a generalize learning method is regularization. It concentrates on minimizing the effects of noise data, which are arbitrary mistakes and coincidences rather than the true properties of the collection [406, 407].

In order to lower the likelihood of inaccurate generalization, it avoids training increasingly complex models. A recently suggested technique for neural networks called “dropout” involves arbitrarily removing units to make the following levels rely entirely on the connections to the different layers above them. However, no technique can escape continual, extensive modification to improve the CNN model’s generalization [408, 409].

An appealing aspect of artificial intelligence is that as a machine is supplied with greater samples, productivity will increase. It will ultimately, nonetheless, come up against some restrictions imposed by the training ability of the system. Many DL algorithms have excessive parameters and are susceptible to bias as they learn additional noise data. In a long-term operational run, solving the generalization issue will enable AI systems of the medical image analysis’s system far more robust and noise-proof. An automated method that self-adjusts in absorbing data by evaluating their reliability would be a strategy and deliver. Data that might potentially cause the network to become unstable should immediately be assigned less weight in the propagation, and their effects on subsequent updates must be minimized [410, 411].

9.7.4 Uncertainties in CNN Based Medical Models

Primarily, deep learning frameworks are procedures made up of a collection of instructions that use optimization and random number synthesis to establish parameters of the model. As a result, deep learning models created using the same information are practically never identical. Data and knowledge-based ambiguities combine to form the uncertainty of machine learning applications. Aleatory uncertainty, which is not created by the framework but is irreducible, is the term for the uncertainty connected to the intrinsic disturbance of the actual figures. Epistemic uncertainty, which is frequently the outcome of the discrepancy among the facts in system training and forecasting, is the ambiguity brought on by insufficient information and data [412].

We must evaluate the level of uncertainty in each data source used as input to ML models and comprehend how uncertainty spreads inside the model in order to be able to measure aleatory uncertainty. The standardized production intricacy can make this difficult for deep learning models. A deep learning model’s ultimate outcomes can be significantly altered by a tiny variation in the raw data. The problem of generalization is connected to the epistemic

uncertainty. Since the majority of machine learning models are built using a particular set of facts, it may be difficult to generalize the algorithm to situations that are not included in the entire sample.

It can be exceedingly difficult to precisely measure the uncertainty associated with generalization due to the lack of information in the initial data collection. For deep learning models to be more user-friendly and to build trustworthiness, precise uncertainty assessment is crucial. There are several mathematical and statistical approaches that have been suggested to deal with uncertainty quantification.

The best often used techniques may be divided into two categories: ensemble uncertainty quantification and Bayesian uncertainty quantification. In Bayesian uncertainty quantification strategies, the learning sample is used to approximate the posterior probability distribution [412]. Ensemble uncertainty quantification entails training many models, computing the synthesized estimate (for example, mean), and computing the variance to represent uncertainty. Subsequently, many Monte Carlo model [413] modifications for uncertainty quantification, as Monte Carlo Dropout [414] have been developed to more effectively quantify forecasting uncertainty.

9.8 Model Development

Massive hyper-parameter ranges and lengthy training duration are characteristics of contemporary deep learning. These characteristics drive the requirement for the development of sophisticated hyperparameter optimization capability in computing environments environments, along with the growth of distributed programming and the increasing desire to productionize deep learning applications [415, 416]. The procedure of selecting an appropriate framework or tailoring a linked framework for one or more training samples is known as model development. Candidates for off-the-shelf models comprise ensemble models like Random Forest, XGBoost, Support Vector Machine, Decision Tree and the majority of Deep learning as well as individual models like AlexNet, Inception with different flavors, DenseNet, U-Net, V-Net, U-Net⁺⁺ and SegNet [417].

There is a high need for Auto Machine Learning (AutoML) that does not involve specialized expertise or human tweaking since discovering ideal models or connecting new models takes time and may never be sufficient. For instance, considering that the efficiency of an deep learning and machine learning methods are data-dependent such as Optuna, OptiML, AutoScikit-learn, and AutoWeka employ Bayesian optimization techniques (Exploration and Exploitation), Evolutionary method (NSGAIISampler (Non-dominated sorting genetic algorithm [418])) to forecast the model’s effectiveness on a particular dataset [415, 419].

For example, to minimize resources, OptiML can develop a regression model to forecast the efficiency of more models that haven't yet been evaluated. Bayesian optimization, meta-learning, and ensemble building are also utilized in Auto-Sklearn's hyperparameter tweaking. Significant issues, however, stay unanswered. Initially, distinct models should be chosen based on distinct use cases using distinct optimal measurement parameters. Furthermore, huge data training does not benefit from the cross-validation approach. Third, consistency, dependability, computational cost, and generalizability are all highly essential and sometimes ignored when looking for remedies. Achievement on precision shouldn't be the sole consideration [420].

In addition to predictive efficiency, a strong AutoML system should autonomously generate a model that addresses all issues with huge data, scenario adaptation, and exhaustive measurements. Although there is a recognized need for more deep learning specialists in business and academia, these professionals are challenging to by and difficult to teach. AutoML may fill that vacuum and lead to several new career possibilities in the deep learning field, including those in medical image analysis. Model selection would be rapid and simple with AutoML and the difficulty of comparing ML algorithms would be significantly lower. Deep learning-powered value-added models and applications wouldn't just be available to the biggest technology and software companies [421, 422].

Small clusters will also be capable of rapidly construct reliable programs to imitate the actual world, collect useful information, and direct the development of climate change and environmental policies. For the upcoming version of deep learning based medical analysis applications, new avenues will be made available [423, 424].

9.9 Provenance, Reproducibility, Replicability, & Reusability

The foundation of the scientific methodology is reproducibility. Its broadest and most typical interpretation relates to the capacity to replicate the results of a certain exploratory studies. This is a prerequisite (but not the only one) for a research claim to be acknowledged as incoming knowledge [425, 426]. The phrase "producing and creating and then executing novel software based on the description of a computational model or procedure supplied in the original release, and producing outcomes that are close enough" is used to complete the definition of "replicating a documented finding" [427, 428].

For deep learning based medical image analysis research, there are four major issues that are connected:

- **Provenance:** What changes had the dataset go through before the results were published? Where did the training set, AI algorithm, software, and hardware come from?
- **Reproducibility:** Utilizing the identical facts and methodologies, can a third party independently verify the accurate AI procedure and outcomes identified?
- **Replicability:** How could an alternative party achieve the equivalent findings using equivalent (but not exact) machine learning (ML) analysis on similar (but not always the very similar) statistics?
- **Re-usability:** refers to how readily previously trained AI models may be deployed to fresh data or to different circumstances.

9.10 Workflow Automation of Healthcare Application

Deep Learning designing is a broad field that encompasses numerous technologies, methodologies, algorithms, techniques and libraries. Its product procedures are made up of a number of connections between hardware and software such as a raw data provider (source) and the sharing of useful information, online services (Web services), APIs and other utility software. It would be impractical to manage every piece manually [429]. Making deep learning based healthcare application usable in real-world circumstances requires optimizing every procedure. Workflow automation for deep learning based healthcare application is still being developed. The Computer Aided Diagnosis society requires a more efficient method for managing the lifespan of established deep learning models in order to sustain artificial intelligence acceptance and scalability [430, 431]. The technique of integrating an investigational machine learning model into a live online system is known as MLOps (MLDevOps). Production-level ML models are deployed, monitored, managed, and governed by it. There are numerous chances for open-source software programmers to engage on this project in the future [432].

Implementing effective and profitable deep learning based healthcare models needs the cooperation of many different organizations and assets, as well as a variety of different technologies, algorithms, scripts, libraries, and tools for automating data preprocessing, digestion, training, validation, and production. To finish a challenging task, a workflow, or logically connected stream of several procedures, is needed [433–435].

Process coordination can be done in a variety of methods, such as by scripting a Shell script, a Python notebook, or by utilizing workflow management software like Cylc. Processes involves fundamental elements are identical. Atomic operations and linkages exist in every workflow. Process automation is the automated execution of all atomic operations without user intervention after a process

has been initiated [436]. Workflow management software (WfMS) such as Pegasus-WMS, Apache Airflow, Galaxy, Cylc and Geoweaver [437] and others are being developed to enable automation. These Workflow management software can increase the replicability and reproducibility of CAD discoveries by not only automating the procedure but also documenting the provenance [438].

9.11 AI Ethics

With an unforeseen tremendous potential of predicting the healthcare future and handling biological risks and materials in preparation to protect lives and safeguard the humanity, healthcare AI is built to defend humankind. The ability has a limitation, though, and it can't save everybody in an incident like a COVID-19 or catastrophic one. What about if healthcare AI makes a mistake, overlooks a region or population, underestimates the impact, and causes more deaths or even more widespread injury? Healthcare AI is a sophisticated but still non-existent system that lacks legal status. However, it acts with a certain amount of self-will and its choices have an influence on society [439, 440].

The ethical issues raised by artificial intelligence as it is deployed are the subject of a multitude of study [441]. Researchers have looked at the connection between the oppression of racial minorities by computer aided diagnosis systems and how it reinforces preexisting prejudice, as well as the impact that cultural attitude performs in computational unfairness. Ultimately, there will likely be several rules and laws governing artificial intelligence morality on healthcare systems [442, 443].

In this article, we briefly discuss a few of the various routes approaching more moral AI in the computer aided diagnosis and healthcare sciences, including more accessible datasets and impartial algorithms. By collaborating with social scientists, medical practitioners and academics who are currently researching the societal consequences of deep learning in the fields of enforcement, legislation, economics and technologists should build computer aided diagnosis ethics-related logic. This involves creating guidelines for deep learning researchers to interact with ethics as both a theoretical and practical endeavor, as the gathering of data and the selection of one theory over another has an immediate influence on healthcare system and people [444–446].

Especially over the past but not least, we think that for there to be an ethical and equitable progression in artificial intelligence in healthcare, one must communicate their implementation of any machine learning or deep learning to the larger society it influences (for instance, if an automated method for developing COVID-19 detection will immediately affect on representations of healthcare system) [447, 448].

10 Tooling and Services

Strong computing is required due to the massive data structure of medical image analysis using deep learning and the great sensitivity of artificial intelligence algorithms. The common hardware and software for computer aided diagnosis are described in this section.

10.1 Computing Device

Graphics Processing Units (GPU), Central Processing Units (CPU), Field-Programmable Gate Arrays (FPGA) and specialized processors (such as TPU - Tensor Processing Unit) are often utilized deep learning devices. Due to their ability to accelerate matrix operations and convolution calculations, GPUs are dominating. In deep learning, the weights are changed after each cycle and kept in memory or a temporary cache so they may be used in subsequent iterations. Compared to Central Processing Units, Graphics Processing Units offer larger memory bandwidths and are better suited for demanding tasks and streaming memory models.

Additionally, researchers are currently looking at the subsequent AI intelligence revolution. Quantum computing, which was first hypothesized by [449], is thought to be the forthcoming major advancement because it can generate statistical formations that are computationally challenging for conventional computers to build [450–452]. Using the Internet of Things (such as edge devices, access points, wearable technology, mobile phones, sensors, etc.) with implanted artificial intelligence methodologies to analyze information locally without transferring much records is yet another option known as edge computing. This can decrease dependency on network systems and boost the artificial intelligence's adaptability and usefulness [453, 454].

10.2 Software

Operating systems developed from Linux that have ongoing, ongoing engineering assistance are advised. The artificial intelligence bellwether at the moment is Ubuntu, which has a lot of built-in components. Users may download and run GPU drivers like CUDA (a program that enables programming for NVIDIA GPUs) with ease, and Python package managers like Conda and Pip can make it easier to download and run packages. Jupyter server [455] (either of notebook, Lab, or Hub) is strongly advised for virtual machine interaction. It enables deep learning based healthcare scientists to perform their tasks, improve efficiency and facilitate cooperation by allowing them to generate and exchange their experimental studies scripts to complete result findings single document.

Python's vibrant, freely available, and supportive library ecosystem is primarily responsible for the language's prominence in the field of artificial intelligence. Several frequently used open-source libraries are Tensorflow, Pytorch, Keras, Theano, Channer, Caffe and Mxnet, OpenCV, AutoML, Scikit-Learn, Pandas, Matplotlib, Seaborn, Plotly, Numpy and etc.

11 Summary

This article provides a summary of the state-of-the-art technology and the advancement of deep learning research with a concentrate on application areas of healthcare analysis especially medical image analysis. Innovations in computer aided diagnosis frameworks and theory will advance healthcare into its next stage. The healthcare community needs to keep up with the explosion in empirical datasets and construct usable deep learning models fast, accurately and affordably. Healthcare AI study and development remain in their adolescence and all the major challenges—from data to models to operations—can lead to a wide range of opportunities across all fields, including industry, academia, Research & Development organizations and government.

Acknowledgments The authors would like to thank the National Key R&D Program of China for providing experimental facilities to conduct these experimentations.

Funding This study is supported by the National Key R & D Program of China with Project No. 2020YFB2104402.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Winsberg F, Elkin M, Jr Macy J, Bordaz V, Weymouth W (1967) Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. *Radiology* 89(2):211–215
2. Kimme C, O'Loughlin BJ, Sklansky J (1977) Automatic detection of suspicious abnormalities in breast radiographs. In: *Data structures, computer graphics, and pattern recognition*, pp 427–447. Elsevier
3. Spiesberger W (1979) Mammogram inspection by computer. *IEEE Trans Biomed Eng* 4:213–219
4. Ishida M, Kato H, Doi K, Frank PH (1982) Development of a new digital radiographic image processing system. In: *Application of optical instrumentation in medicine X*, vol 347, pp 42–48. SPIE
5. Chen CM, Chou YH, Tagawa N, Do Y (2013) Computer-aided detection and diagnosis in medical imaging
6. Chan H-P, Doi K, Galhotra S, Vyborny CJ, MacMahon H, Jokich PM (1987) Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography. *Med Phys* 14(4):538–548
7. Heang-Ping Chan E, Charles PM, Lam K, Wu Y, Macmahon H (1990) Improvement in radiologists? Detection of clustered microcalcifications on mammograms. *Arbor* 1001:48109–0326
8. Lin M, Hu H, Zhou S, Xu S (2022) Soft wearable devices for deep-tissue sensing. *Nat Rev Mater* 1(20):12
9. Azam MA, Khan KB, Salahuddin S, Rehman E, Khan SA, Khan MA, Kadry S, Gandomi AH (2022) A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. *Comput Biol Med* 144:105253
10. Haidekker MA (2013) X-ray projection imaging. In: *Medical imaging technology*, pp 13–35. Springer
11. Mikla VI, Mikla VV (2013) *Medical imaging technology*. Elsevier, Amsterdam
12. National Electrical Manufacturers Association Medical Imaging & Technology Alliance (2022) Medical image modalities URL <https://www.medicalimaging.org/about-mita/modalities>
13. James AP, Dasarathy BV (2014) Medical image fusion: A survey of the state of the art. *Inform Fus* 19:4–19
14. He Z, He Y, Cao W (2022) Deformable image registration with attention-guided fusion of multi-scale deformation fields. *Appl Intell* 8:1–15
15. El-Gamal FEIZA, Elmogy M, Atwan A (2016) Current trends in medical image registration and fusion. *Egypt Inform J* 17(1):99–124
16. Dietrich CF, Bolondi L, Duck F, Evans DH, Ewertsen C, Fraser AG, Gilja OH, Jenssen C, Merz E, Nolsoe C, Nårnberg D et al (2022) History of ultrasound in medicine from its birth to date, on occasion of the 50 years anniversary of EFSUMB. a publication of the European federation of societies for ultrasound in medicine and biology (EFSUMB), designed to record the historical development of medical ultrasound. *Medical Ultrasonography* 521(7553):436–444
17. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
18. Cheng Y-T, Patel A, Wen C, Bullock D, Habib A (2020) Intensity thresholding and deep learning based lane marking extraction and lane width estimation from mobile light detection and ranging (lidar) point clouds. *Remote Sens* 12(9):1379
19. Feng R, Fan C, Li Z, Chen X (2020) Mixed road user trajectory extraction from moving aerial videos based on convolution neural network detection. *IEEE Access* 8:43508–43519
20. Razzak MI, Naz S, Zaib A (2018) Deep learning for medical image processing: overview, challenges and the future. *Classif BioApps* 5:323–350
21. Pang S, Yang X (2016) Deep convolutional extreme learning machine and its application in handwritten digit classification. *Comput Intell Neurosci* 98:320–337
22. Lenc K, Vedaldi A (2015) Matconvnet: Convolutional neural networks for matlab. In: *Proceedings of the 23rd ACM international conference on multimedia*, pp 689–692

23. Guo Y, Ashour AS (2019) Neutrosophic sets in dermoscopic medical image segmentation. In: Neutrosophic set in medical image analysis, pp 229–243. Elsevier
24. Merjulah R, Chandra J (2019) Classification of myocardial ischemia in delayed contrast enhancement using machine learning. In: Intelligent data analysis for biomedical applications, pp 209–235. Elsevier
25. Wang J, Zhang M (2020) Deepflash: An efficient network for learning-based medical image registration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 4444–4452
26. Yabo F, Lei Y, Wang T, Curran WJ, Liu T, Yang X (2020) Deep learning in medical image registration: a review. *Phys Med Biol* 65(20):20TR01
27. Haskins G, Kruger U, Yan P (2020) Deep learning in medical image registration: a survey. *Mach Vis Appl* 31(1):1–18
28. Sharma H, Jain JS, Bansal P, Gupta S (2020) Feature extraction and classification of chest X-ray images using CNN to detect pneumonia. In: 2020 10th international conference on cloud computing, data science & engineering (Confluence), pp 227–231. IEEE
29. 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(2):854–864
30. Kowsari K, Sali R, Ehsan L, Adorno W, Ali A, Moore S, Amadi B, Kelly P, Syed S, Brown D (2020) Hmic: Hierarchical medical image classification, a deep learning approach. *Information* 11(6):318
31. Singh SP, Wang L, Gupta S, Goli H, Padmanabhan P, Gulyás B (2020) 3D deep learning on medical images: a review. *Sensors* 20(18):5097
32. Shen C, Nguyen D, Zhou Z, Jiang SB, Dong B, Jia X (2020) An introduction to deep learning in medical physics: advantages, potential, and challenges. *Phys Med Biol* 65(5):05TR01
33. Wenhao W, Gao L, Duan H, Huang G, Ye X, Nie S (2020) Segmentation of pulmonary nodules in CT images based on 3d-unet combined with three-dimensional conditional random field optimization. *Med Phys* 47(9):4054–4063
34. Huff DT, Weisman AJ, Jeraj R (2021) Interpretation and visualization techniques for deep learning models in medical imaging. *Phys Med Biol* 66(4):04TR01
35. Li L, Liu Z, Huang H, Lin M, Luo D (2019) Evaluating the performance of a deep learning-based computer-aided diagnosis (dl-cad) system for detecting and characterizing lung nodules: comparison with the performance of double reading by radiologists. *Thoracic Cancer* 10(2):183–192
36. Hsu H-H, Ko K-H, Chou Y-C, Wu Y-C, Chiu S-H, Chang C-K, Chang W-C (2021) Performance and reading time of lung nodule identification on multidetector CT with or without an artificial intelligence-powered computer-aided detection system. *Clin Radiol* 76(8):626–723
37. Zia M, Rehman U, Javaid M, Shah SIA, Gilani SO, Jamil M, Butt SI (2018) An appraisal of nodules detection techniques for lung cancer in CT images. *Biomed Signal Process Control* 41:140–151
38. Zhang J, Xia Y, Cui H, Zhang Y (2018) Pulmonary nodule detection in medical images: a survey. *Biomed Signal Process Control* 43:138–147
39. Zhang M, Kong Z, Zhu W, Yan F, Xie C (2021) Pulmonary nodule detection based on 3d feature pyramid network with incorporated squeeze-and-excitation-attention mechanism. *Pract Exp Concurr Comput*, p e6237
40. Monkam P, Qi S, Ma H, Gao W, Yao Y, Qian Wei (2019) Detection and classification of pulmonary nodules using convolutional neural networks: a survey. *IEEE Access* 7:78075–78091
41. Halder A, Dey D, Sadhu AK (2020) Lung nodule detection from feature engineering to deep learning in thoracic CT images: a comprehensive review. *J Digital Imaging* 33(3):655–677
42. Geirhos R, Jacobsen J, Michaelis C, Zemel R, Brendel W, Bethge M, Wichmann FA (2020) Shortcut learning in deep neural networks. *Nat Mach Intell* 2(11):665–673
43. Verma B (1997) Fast training of multilayer perceptrons. *IEEE Trans Neural Netw* 8(6):1314–1320
44. Delashmit WH, Manry MT et al (2005) Recent developments in multilayer perceptron neural networks. In: Proceedings of the seventh annual memphis area engineering and science conference, MAESC
45. Desai M, Shah M (2021) An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and convolutional neural network (CNN). *Clin eHealth* 4:1–11
46. Smits JRM, Melssen WJ, Buydens LMC, Kateman G (1994) Using artificial neural networks for solving chemical problems: part i. Multi-layer feed-forward networks. *Chemomet Intell Lab Syst* 22(2):165–189
47. Svozil D, Kvasnicka V, Pospichal J (1997) Introduction to multilayer feed-forward neural networks. *Chemomet Intell Lab Syst* 39(1):43–62
48. Ojha VK, Abraham A, Snášel V (2017) Metaheuristic design of feedforward neural networks: a review of two decades of research. *Eng Appl Artif Intell* 60:97–116
49. Grossberg S (2013) Recurrent neural networks. *Scholarpedia* 8(2):1888
50. Yong Y, Si X, Changhua H, Zhang J (2019) A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput* 31(7):1235–1270
51. Durstewitz D, Koppe G, Meyer-Lindenberg A (2019) Deep neural networks in psychiatry. *Mol Psychiatry* 24(11):1583–1598
52. Sen S, Raghunathan A (2018) Approximate computing for long short term memory (LSTM) neural networks. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 37(11):2266–2276
53. Van Houdt G, Mosquera C, Nápoles G (2020) A review on the long short-term memory model. *Artif Intell Rev* 53(8):5929–5955
54. Sherstinsky A (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D* 404:132306
55. Dey R, Salem FM (2017) Gate-variants of gated recurrent unit (GRU) neural networks. In: 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS), pp 1597–1600. IEEE
56. Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. [arXiv:1412.3555](https://arxiv.org/abs/1412.3555)
57. Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. *IEEE Trans Signal Process* 45(11):2673–2681
58. Berglund M, Raiko T, Honkala M, Kärkkäinen L, Vetek A, Karhunen JT (2015) Bidirectional recurrent neural networks as generative models. *Adv Neural Inform Process Syst* 28:1–89
59. Lv W, Xiong J, Shi J, Huang Y, Qin S (2021) A deep convolution generative adversarial networks based fuzzing framework for industry control protocols. *J Intell Manuf* 32(2):441–457
60. Meng Y, Shen J, Zhang C, Han J (2018) Weakly-supervised neural text classification. In: proceedings of the 27th ACM International Conference on information and knowledge management, pp 983–992
61. Jonschkowski R, Stone A, Barron JT, Gordon A, Konolige K, Angelova A (2020) What matters in unsupervised optical flow. In: European conference on computer vision, pp 557–572. Springer

62. Diederik PK, Max W et al (2019) An introduction to variational autoencoders. *Found Trends Mach Learn* 12(4):307–392
63. Casper KS, Tapani R, Lars M, Søren K, Ole W (2016) Ladder variational autoencoders. *Adv Neural Inform Process Syst* 29:88
64. Tanaka T (1998) Mean-field theory of Boltzmann machine learning. *Phys Rev E* 58(2):2302
65. Sutskever I, Hinton GE, Taylor GW (2008) The recurrent temporal restricted Boltzmann machine. *Adv Neural Inform Process Syst* 21:878
66. Hua Y, Guo J, Zhao H (2015) Deep belief networks and deep learning. In: *Proceedings of 2015 International conference on intelligent computing and internet of things*, pp 1–4. IEEE
67. Kuremoto T, Kimura S, Kobayashi K, Obayashi M (2014) Time series forecasting using a deep belief network with restricted Boltzmann machines. *Neurocomputing* 137:47–56
68. Xin Y, Ekta W, Paul B (2019) Generative adversarial network in medical imaging: a review. *Med Image Anal* 58:101552
69. Aggarwal A, Mittal M, Battineni G (2021) Generative adversarial network: an overview of theory and applications. *Int J Inform Manag Data Insights* 1(1):100004
70. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK (2018) Medical image analysis using convolutional neural networks: a review. *J Med Syst* 42(11):1–13
71. Gobert L, Hiroshi F (2020) *Deep learning in medical image analysis: challenges and applications*, vol 1213. Springer, New York
72. Jiuxiang G, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J et al (2018) Recent advances in convolutional neural networks. *Pattern Recognit* 77:354–377
73. Riad SM (1986) The deconvolution problem: an overview. *Proc IEEE* 74(1):82–85
74. Shen-Orr SS, Gaujoux R (2013) Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol* 25(5):571–578
75. Liu X, Deng Z, Yang Y (2019) Recent progress in semantic image segmentation. *Artif Intell Rev* 52(2):1089–1106
76. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 9(4):611–629
77. Murphy J (2016) An overview of convolutional neural network architectures for deep learning. *Microway Inc*, pp 1–22
78. Zafar A, Aamir M, Nawi NM, Arshad A, Riaz S, Alruban A, Dutta AK, Almotairi S (2022) A comparison of pooling methods for convolutional neural networks. *Appl Sci* 12(17):8643
79. Yu D, Wang H, Chen P, Wei Z (2014) Mixed pooling for convolutional neural networks. In: *International conference on rough sets and knowledge technology*, pp 364–375. Springer
80. Shabbeer Basha SH, Dubey SR, Pulabaigari V, Mukherjee S (2020) Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing* 378:112–119
81. Dubey SR, Singh SK, Chaudhuri BB (2022) Activation functions in deep learning: a comprehensive survey and benchmark. *Neurocomputing*
82. Santurkar S, Tsipras D, Andrew I, Madry A (2018) How does batch normalization help optimization? *Adv Neural Inform Process Syst* 31:45
83. Li Z, Gong B, Yang T (2016) Improved dropout for shallow and deep learning. *Advan Neural Inform Process Syst* 29:189
84. Garbin C, Zhu X, Marques O (2020) Dropout vs. batch normalization: an empirical study of their impact to deep learning. *Multimedia Tools Appl* 79(19):12777–12815
85. Yan X, Yang X, Xing X, Cui B, Guo Z, Guo T (2020) Trustworthy network anomaly detection based on an adaptive learning rate and momentum in IIOT. *IEEE Trans Indust Inform* 16(9):6182–6192
86. Xie Z, Wang X, Zhang H, Sato I, Sugiyama M (2022) Adaptive inertia: Disentangling the effects of adaptive learning rate and momentum. In: *International conference on machine learning*, pp 24430–24459. PMLR
87. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
88. Yu N, Jiao P, Zheng Y (2015) Handwritten digits recognition base on improved lenet5. In: *The 27th Chinese control and decision conference (2015 CCDC)*, pp 4871–4875. IEEE
89. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60(6):84–90
90. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. In: *European conference on computer vision*, pp 818–833. Springer
91. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
92. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1–9
93. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L (2009) Imagenet: a large-scale hierarchical image database. In: *2009 IEEE conference on computer vision and pattern recognition*, pp 248–255. IEEE
94. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
95. Srivastava RK, Greff K, Schmidhuber J (2015) Highway networks. *arXiv preprint* [arXiv:1505.00387](https://arxiv.org/abs/1505.00387)
96. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017a) Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 4700–4708
97. Zagoruyko S, Komodakis N (2016) Wide residual networks. [arXiv:1605.07146](https://arxiv.org/abs/1605.07146)
98. Han D, Kim J, Kim J (2017) Deep pyramidal residual networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 5927–5935
99. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 2818–2826
100. Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, Fischer I, Wojna Z, Song Y, Guadarrama S et al (2017b) Speed/accuracy trade-offs for modern convolutional object detectors. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 7310–7311
101. Xia X, Xu C, Nan B (2017) Inception-v3 for flower classification. In: *2017 2nd international conference on image, vision and computing (ICIVC)*, pp 783–787. IEEE
102. Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1251–1258
103. Xie S, Girshick R, Dollár P, Tu Z, He K (2017) Aggregated residual transformations for deep neural networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1492–1500
104. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) Squeezenet: Alexnet-level accuracy with 50x fewer parameters and < 0.5 mb model size. [arXiv:1602.07360](https://arxiv.org/abs/1602.07360)

105. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, pp 234–241. Springer
106. Abdollahi A, Pradhan B, Alamri A (2020) Vnet: an end-to-end fully convolutional neural network for road extraction from high-resolution remote sensing data. *IEEE Access* 8:179424–179436
107. Badrinarayanan V, Kendall A, Cipolla R (2017) Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 39(12):2481–2495
108. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J (2018) Unet++: a nested u-net architecture for medical image segmentation. In: Deep learning in medical image analysis and multimodal learning for clinical decision support, pp 3–11. Springer
109. Girshick R (2015) Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 1440–1448
110. He K, Gkioxari G, Dollár P, Girshick R (2017) Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 2961–2969
111. Liang-Chieh C, George P, Iasonas K, Kevin M, Alan LY (2017) DeepLab: semantic image segmentation with deep convolutional nets, Atrous convolution, and fully connected CRFS. *IEEE Trans Pattern Anal Mach Intell* 40(4):834–848
112. Lin T-Y, Goyal P, Girshick R, He K, Dollár P (2017) Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision, pp 2980–2988
113. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788
114. Fukushima K, Miyake S (1982) Neocognitron: a new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognit* 15(6):455–469
115. LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L (1989) Handwritten digit recognition with a back-propagation network. *Adv Neural Inform Process Syst* 2:568
116. Lo SC, Lin JS, Freedman MT, Mun SK (1993) Computer-assisted diagnosis of lung nodule detection using artificial convolution neural network. In: Medical imaging 1993: image processing, vol 1898, pp 859–869. SPIE
117. Lo SCB, Chan H-P, Lin J-S, Li H, Freedman MT, Mun SK (1995) Artificial convolution neural network for medical image pattern recognition. *Neural Netw* 8(7–8):1201–1214
118. Chan HP, Sahiner B, Lo SC, Helvie M, Petrick N, Adler DD, Goodsitt MM (1994) Computer-aided diagnosis in mammography: detection of masses by artificial neural network. *Med Phys* 21:875–6
119. Sahiner B, Chan H-P, Petrick N, Wei D, Helvie MA, Adler DD, Goodsitt MM (1995) Image classification using artificial neural networks. In: Medical imaging 1995: image processing, volume 2434, pp 838–845. SPIE
120. Wei D, Sahiner B, Chan H-P, Petrick N (1995) Detection of masses on mammograms using a convolution neural network. In: 1995 international conference on acoustics, speech, and signal processing, vol 5, pp 3483–3486. IEEE
121. Sahiner B, Chan H-P, 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
122. Chan HP, Lo SCB, Helvie MA, Goodsitt MM, Cheng SNC, Adler DD (1993) Recognition of mammographic microcalcifications with artificial neural network. *Radiology* 189(P):318
123. Chan H-P, Lo SCB, Sahiner B, Lam KL, Helvie MA (1995) Computer-aided detection of mammographic microcalcifications: pattern recognition with an artificial neural network. *Med Phys* 22(10):1555–1567
124. Zhang W, Doi K, Giger ML, Yuzheng W, Nishikawa RM, Schmidt RA (1994) Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. *Med Phys* 21(4):517–524
125. Hinton GE, Osindero S, Teh Y-W (2006) A fast learning algorithm for deep belief nets. *Neural Comput* 18(7):1527–1554
126. Bengio Y, Lamblin P, Popovici D, Larochelle H (2006) Greedy layer-wise training of deep networks. *Adv Neural Inform Process Syst* 19:895
127. Erhan D, Courville A, Bengio Y, Vincent P (2010) Why does unsupervised pre-training help deep learning? In: Proceedings of the thirteenth international conference on artificial intelligence and statistics, pp 201–208. JMLR workshop and conference proceedings
128. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. In: *Icml*
129. Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp 315–323. JMLR workshop and conference proceedings
130. Ranzato M, Huang FJ, Boureau YL, LeCun Y (2007) Unsupervised learning of invariant feature hierarchies with applications to object recognition. In: 2007 IEEE conference on computer vision and pattern recognition, pp 1–8. IEEE
131. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 15(1):1929–1958
132. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning, pp 448–456. PMLR
133. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Process Syst* 4:25
134. Russakovsky O, Deng J, Hao S, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) Imagenet large scale visual recognition challenge. *Int J Comput Vis* 115(3):211–252
135. Sun C, Shrivastava A, Singh S, Gupta A (2017a) Revisiting unreasonable effectiveness of data in deep learning era. In: Proceedings of the IEEE international conference on computer vision, pp 843–852
136. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Laak JAVD, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42:60–88
137. Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, Summers RM, Giger ML (2019) Deep learning in medical imaging and radiation therapy. *Med Phys* 46(1):e1–e36
138. Mazurowski MA, Buda M, Saha A, Bashir MR (2019) Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI. *J Magnetic Reson Imaging* 49(4):939–954
139. De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, Askham H, Glorot X, O'Donoghue B, Visentin D (2018) Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 24(9):1342–1350
140. Janowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J Pathol Inform* 7(1):29
141. Kim E, Corte-Real M, Baloch Z (2016) A deep semantic mobile application for thyroid cytopathology. In: Medical imaging 2016: PACS and imaging informatics: next generation and innovations, volume 9789, pp 36–44. SPIE

142. Antony J, McGuinness K, O'Connor NE, Moran K (2016) Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In: 2016 23rd international conference on pattern recognition (ICPR), pp 1195–1200. IEEE
143. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542(7639):115–118
144. Gulshan V, Peng L, Coram M, Stumpe MC, Derek W, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316(22):2402–2410
145. Brosch T, Tam R, et al (2013) Initiative Alzheimer? Disease neuroimaging manifold learning of brain MRIS by deep learning. In: International conference on medical image computing and computer-assisted intervention, pp 633–640. Springer
146. 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
147. Suk H-I, Shen D (2013) Deep learning-based feature representation for ad/mci classification. In: International conference on medical image computing and computer-assisted intervention, pp 583–590. Springer
148. Suk H-II, Lee S-W, Shen D et al (2014) Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. *NeuroImage* 101:569–582
149. Menegola A, Fornaciali M, Pires R, Avila S, Valle E (2016) Towards automated melanoma screening: exploring transfer learning schemes. arXiv preprint [arXiv:1609.01228](https://arxiv.org/abs/1609.01228)
150. Hosseini-Asl E, Gimel'farb G, El-Baz A (2016) Alzheimer's disease diagnostics by a deeply supervised adaptable 3d convolutional network. arXiv preprint [arXiv:1607.00556](https://arxiv.org/abs/1607.00556)
151. Payan A, Montana G (2015) Predicting Alzheimer's disease: a neuroimaging study with 3d convolutional neural networks. arXiv preprint [arXiv:1502.02506](https://arxiv.org/abs/1502.02506)
152. Kawahara J, Brown CJ, Miller SP, Booth BG, Chau V, Grunau RE, Zwicker JG, Hamarneh G (2017) Brainnetcnn: convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage* 146:1038–1049
153. Mahapatra D, Poellinger A, Reyes M (2022) Interpretability-guided inductive bias for deep learning based medical image. *Med Image Anal* 81:102551
154. Lee J, Chung SW (2022) Deep learning for orthopedic disease based on medical image analysis: present and future. *Appl Sci* 12(2):681
155. Liu F, Tian Y, Chen Y, Liu Y, Belagiannis V, Carneiro G (2022) ACPL: Anti-curriculum pseudo-labelling for semi-supervised medical image classification. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 20697–20706
156. Xue C, Lequan Y, Chen P, Dou Q, Heng P-A (2022) Robust medical image classification from noisy labeled data with global and local representation guided co-training. *IEEE Trans Med Imaging* 41(6):1371–1382
157. Apostolidis KD, Papakostas GA (2022) Digital watermarking as an adversarial attack on medical image analysis with deep learning. *J Imaging* 8(6):155
158. Jin C, Udupa JK, Zhao L, Tong Y, Odhner D, Pednekar G, Nag S, Lewis S, Poole N, Mannikeri S et al (2022) Object recognition in medical images via anatomy-guided deep learning. *Med Image Anal* 81:102527
159. Yicheng W, Ge Z, Zhang D, Minfeng X, Zhang L, Xia Y, Cai J (2022) Mutual consistency learning for semi-supervised medical image segmentation. *Med Image Anal* 81:102530
160. Wang H, Zhou Y, Zhang J, Lei J, Sun D, Feng X, Xiayu X (2022) Anomaly segmentation in retinal images with poisson-blending data augmentation. *Med Image Anal* 81:102534
161. Zhao Y, Zeng K, Zhao Y, Bhatia P, Ranganath M, Kozhikkavil Muhammed Labeeb, Li Chen, Hermsillo Gerardo (2022) Deep learning solution for medical image localization and orientation detection. *Med Image Anal* 81:102529
162. Wang Y, Feng Y, Zhang L, Zhou JT, Liu Y, Goh RSM, Zhen L (2022) Adversarial multimodal fusion with attention mechanism for skin lesion classification using clinical and dermoscopic images. *Med Image Anal* 81:102535
163. Tilborghs S, Bogaert J, Maes F (2022) Shape constrained CNN for segmentation guided prediction of myocardial shape and pose parameters in cardiac MRI. *Med Image Anal* 81:102533
164. Yang P, Yin X, Haiming L, Zhongliang H, Zhang X, Jiang R, Lv H (2022) Cs-co: a hybrid self-supervised visual representation learning method for h & e-stained histopathological images. *Med Image Anal* 81:102539
165. Han C, Lin J, Mai J, Wang Y, Zhang Q, Zhao B, Chen X, Pan X, Shi Z, Xu Z et al (2022) Multi-layer pseudo-supervision for histopathology tissue semantic segmentation using patch-level classification labels. *Medical Image Anal*, pp 102487
166. Zhang X, Xiao X, Fu H, Yuan J, Xu Y, Higashita R, Liu J (2022) Attention to region: region-based integration-and-recalibration networks for nuclear cataract classification using as-oct images. *Med Image Anal* 80:102499
167. Huang R, Ying Q, Lin Z, Zheng Z, Tan L, Tang G, Zhang Q, Luo M, Yi X, Liu P et al (2022) Extracting keyframes of breast ultrasound video using deep reinforcement learning. *Med Image Anal* 80:102490
168. Yi D, Member QY, Wang Y, Chen D, Qin Z, Zhang J (2022) Malleynet a multi-object assistance based network for brachial plexus segmentation in ultrasound images. *Med Image Anal* 80:102511
169. Huang R, Lin M, Dou H, Lin Z, Ying Q, Jia X, Xu W, Mei Z, Yang X, Dong T et al (2022) Boundary-rendering network for breast lesion segmentation in ultrasound images. *Med Image Anal* 89:102478
170. Li J, Huang J, Jiang T, Tu L, Cui L, Cui J, Ma X, Yao X, Shi Y, Wang S, Wang Y et al (2022) A multi-step approach for tongue image classification in patients with diabetes. *Comput Biol Med* 89:105935
171. Oyelade ON, Ezugwu AE, Venter HS, Mirjalili S, Gandomi AH (2022) Abnormality classification and localization using dual-branch whole-region-based cnn model with histopathological images. *Comput Biol Med* 943:105
172. Sadik F, Dastider AG, Subah MR, Mahmud T, Fattah SA (2022) A dual-stage deep convolutional neural network for automatic diagnosis of Covid-19 and pneumonia from chest CT images. *Comput Biol Med* 149:105806
173. Sun K, He M, Xu Y, Wu Q, He Z, Li W, Liu H, Pi X (2022) Multi-label classification of fundus images with graph convolutional network and lightgbm. *Comput Biol Med* 909:105
174. Liu Y, Jia Y, Hou C, Li N, Zhang N, Yan X, Yang L, Guo Y, Chen H, Li J, Hao Y (2022) Pathological prognosis classification of patients with neuroblastoma using computational pathology analysis. *Comput Biol Med* 980:105
175. Saini M, Susan S (2022) Diabetic retinopathy screening using deep learning for multi-class imbalanced datasets. *Comput Biol Med* 989:105
176. Xin C, Liu Z, Zhao K, Miao L, Ma Y, Zhu X, Zhou Q, Wang S, Li L, Yang F, Xu S (2022) An improved transformer network for skin cancer classification. *Comput Biol Med* 939:105
177. Shabani S, Homayounfar M, Vardhanabhuti V, Mahani MAN, Koohi-Moghadam M (2022) Self-supervised region-aware

- segmentation of Covid-19 CT images using 3d GAN and contrastive learning. *Comput Biol Med* 149:106033
178. Hayat M, Tahir M, Alarfaj FK, Alturki R, Gazzawe F (2022) Nlp-bch-ens: Nlp-based intelligent computational model for discrimination of malaria parasite. *Comput Biol Med* 962:105
 179. Hussain S, Xi X, Ullah I, Inam SA, Naz F, Shaheed K, Ali SA, Tian CA (2022) A discriminative level set method with deep supervision for breast tumor segmentation. *Comput Biol Med* 105995:105995
 180. Sendra-Balcells C, Campello VM, Martín-Isla C, Medel DV, Descalzo M, Guala A, Palomares JF, Lekadir K (2022) Domain generalization in deep learning for contrast-enhanced imaging. *Comput Biol Med* 149:106052
 181. Yalçın S, Vural H (2022) Brain stroke classification and segmentation using encoder-decoder based deep convolutional neural networks. *Comput Biol Med* 149:105941
 182. Qin Y, Chen M, Zheng H, Gu Y, Shen M, Yang J, Huang X, Zhu YM, Yang GZ (2019) Airwaynet: a voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks. In: *International conference on medical image computing and computer-assisted intervention*, pp 212–220. Springer
 183. Lian J, Long Y, Huang F, Ng KS, Lee FM, Lam DC, Fang BX, Dou Q, Vardhanabhuti V (2022) Imaging-based deep graph neural networks for survival analysis in early stage lung cancer using ct: A multicenter study. *Front Oncol* 12:15
 184. Zheng H, Qin Y, Yun G, Xie F, Yang J, Sun J, Yang G-Z (2021) Alleviating class-wise gradient imbalance for pulmonary airway segmentation. *IEEE Trans Med Imaging* 40(9):2452–2462
 185. Zhang M, Yu X, Zhang H, Zheng H, Yu W, Pan H, Cai X, Gu Y (2021) Fda: Feature decomposition and aggregation for robust airway segmentation. In: *Domain adaptation and representation transfer, and affordable healthcare and AI for Resource Diverse Global Health*, pp 25–34. Springer
 186. Wang J, Zhu H, Wang S-H, Zhang Y-D (2021) A review of deep learning on medical image analysis. *Mobile Netw Appl* 26(1):351–380
 187. Dayanandhan K, Shaby SM (2020) Classification of CT image lung cancer disease using hybrid clustering and deep learning techniques. *Eur J Mol Clin Med* 3:7
 188. Tiwari L, Raja R, Awasthi V, Miri R, Sinha GR, Alkinani MH, Polat K (2021) Detection of lung nodule and cancer using novel mask-3 fcm and twedlnn algorithms. *Measurement* 172:108882
 189. Miró-Nicolau M, Moyà-Alcover G, Capó AJ (2022) Evaluating explainable artificial intelligence for x-ray image analysis. *Appl Sci* 12(9):4459
 190. Suganyadevi S, Seethalakshmi V (2022) Cvd-hnet: Classifying pneumonia and Covid-19 in chest x-ray images using deep network. *Wirel Personal Commun* 58:1–25
 191. Kogilavani SV, Prabhu J, Sandhiya R, Sandeep Kumar M, UmaShankar Subramaniam, Alagar Karthick, Muhibbullah M, Imam SBS (2022) Covid-19 detection based on lung CT scan using deep learning techniques. *Comput Math Methods Med*
 192. Cao W, Rui W, Cao G, He Z (2020) A comprehensive review of computer-aided diagnosis of pulmonary nodules based on computed tomography scans. *IEEE Access* 8:154007–154023
 193. Yu G, Chi J, Liu J, Yang L, Zhang B, Dahua Y, Zhao Y, Xiaoqi L (2021) A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning. *Comput Biol Med* 137:104806
 194. Chenyang L, Chan S-C (2020) A joint detection and recognition approach to lung cancer diagnosis from CT images with label uncertainty. *IEEE Access* 8:228905–228921
 195. Liu W, Liu X, Li H, Li M, Zhao X, Zhu Z (2021) Integrating lung parenchyma segmentation and nodule detection with deep multi-task learning. *IEEE J Biomed Health Inform* 25(8):3073–3081
 196. Painuli D, Bhardwaj S et al (2022) Recent advancement in cancer diagnosis using machine learning and deep learning techniques: a comprehensive review. *Comput Biol Med* 89:105580
 197. Guo Z, Zhao L, Yuan J, Hengyong Y (2021) Msanet: multiscale aggregation network integrating spatial and channel information for lung nodule detection. *IEEE J Biomed Health Inform* 26(6):2547–2558
 198. Ragab DA, Sharkas M, Marshall S, Ren J (2019) Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ* 5:e6201
 199. Mambou SJ, Maresova P, Krejcar O, Selamat A, Kuca K (2018) Breast cancer detection using infrared thermal imaging and a deep learning model. *Sensors* 18(9):2799
 200. Selvathi D, Poornila AA (2018) Deep learning techniques for breast cancer detection using medical image analysis. In: *Biologically rationalized computing techniques for image processing applications*, pp 159–186. Springer
 201. Mohamed A, Amer E, Eldin N, Hossam M, Elmasry N, Adnan GT et al (2022) The impact of data processing and ensemble on breast cancer detection using deep learning. *J Comput Commun* 1(1):27–37
 202. Kavitha T, Mathai PP, Karthikeyan C, Ashok M, Kohar R, Avanija J, Neelakandan S (2022) Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images. *Interdisc Sci* 14(1):113–129
 203. Chowdhury D, Das A, Dey A, Sarkar S, Dwivedi AD, Mukkamala RR, Murmu L (2022) Abcandroid: a cloud integrated android app for noninvasive early breast cancer detection using transfer learning. *Sensors* 22(3):832
 204. Escorcia-Gutierrez J, Mansour RF, Beleño K, Jiménez-Cabas J, Pérez M, Madera N, Velasquez K (2022) Automated deep learning empowered breast cancer diagnosis using biomedical mammogram images. *Comput Mater Continua* 71:3–4221
 205. Jasti V, Zamani AS, Arumugam K, Naved M, Pallathadka H, Sammy F, Raghuvanshi A, Kaliyaperumal K (2022) Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. *Secur Commun Netw*, 2022
 206. Jabeen K, Khan MA, Alhaisoni M, Tariq U, Zhang Y-D, Hamza A, Mickus A, Damaševičius R (2022) Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion. *Sensors* 22(3):807
 207. Naseem U, Rashid J, Ali L, Kim J, Haq QEU, Awan MJ, Imran M (2022) An automatic detection of breast cancer diagnosis and prognosis based on machine learning using ensemble of classifiers. *IEEE Access* 10:78242–78252
 208. Singh S, Kumar R (2022) Breast cancer detection from histopathology images with deep inception and residual blocks. *Multi-media Tools Appl* 81(4):5849–5865
 209. Liu M, Hu L, Tang Y, Wang C, He Y, Zeng C, Lin K, He Z, Huo W (2022) A deep learning method for breast cancer classification in the pathology images. *IEEE J Biomed Health Inform* 26(10):5025–5032
 210. Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wählby C, Hartman J, Rantalainen M (2022) Improved breast cancer histological grading using deep learning. *Ann Oncol* 33(1):89–98
 211. Reshma VK, Arya N, Ahmad SS, Wattar I, Mekala S, Joshi S, Krah D (2022) Detection of breast cancer using histopathological image classification dataset with deep learning techniques. *BioMed Res Int*
 212. Ragab M, Albukhari A, Alyami J, Mansour RF (2022) Ensemble deep-learning-enabled clinical decision support system for breast

- cancer diagnosis and classification on ultrasound images. *Biology* 11(3):439
213. Ahmad S, Ullah T, Ahmad I, Al-Sharabi A, Ullah K, Khan RA, Rasheed S, Ullah I, Uddin M, Ali M et al (2022a) A novel hybrid deep learning model for metastatic cancer detection. *Comput Intell Neurosci* 2022
 214. Maqsood S, Damaševičius R, Maskeliūnas R (2022) Ttcnn: a breast cancer detection and classification towards computer-aided diagnosis using digital mammography in early stages. *Appl Sci* 12(7):3273
 215. Ibrokhimov B, Kang J-Y (2022) Two-stage deep learning method for breast cancer detection using high-resolution mammogram images. *Appl Sci* 12(9):4616
 216. Mohamed EA, Rashed EA, Gaber T, Karam O (2022) Deep learning model for fully automated breast cancer detection system from thermograms. *PLoS ONE* 17(1):e0262349
 217. Janowczyk A, Basavanthally A, Madabhushi A (2017) Stain normalization using sparse autoencoders (stanosa): application to digital pathology. *Comput Med Imaging Graph* 57:50–61
 218. Srinidhi CL, Ciga O, Martel AL (2021) Deep neural network models for computational histopathology: a survey. *Med Image Anal* 67:101813
 219. Zhou X, Chen Li M, Rahaman M, Yao Y, Ai S, Sun C, Wang Q, Zhang Y, Li M, Li X et al (2020) A comprehensive review for breast histopathology image analysis using classical and deep neural networks. *IEEE Access* 8:90931–90956
 220. Sercan SG, Kusetogullari H, Tokat F, Bozaba E, Karakaya S, Ithem LO, Tekin E, Özsoy Gülşah A et al (2022) Mitnet: a novel dataset and a two-stage deep learning approach for mitosis recognition in whole slide images of breast cancer tissue. *Neural Comput Appl* 85:1–15
 221. Xu Y, Li Y, Liu M, Wang Y, Lai M, Chang E I et al (2016) Gland instance segmentation by deep multichannel side supervision. In: *International conference on medical image computing and computer-assisted intervention*, pp 496–504. Springer
 222. Iman R, Pinar D (2022) Multi-task learning for gland segmentation. *Signal Image Video Process*, pp 1–9
 223. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J (2013) Mitosis detection in breast cancer histology images with deep neural networks. In: *International conference on medical image computing and computer-assisted intervention*, pp 411–418. Springer
 224. Cireşan D, Giusti A, Gambardella L, Schmidhuber J (2012) Deep neural networks segment neuronal membranes in electron microscopy images. *Adv Neural Inform Process Syst*. vol 25
 225. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH (2016) Deep learning for identifying metastatic breast cancer. [arXiv:1606.05718](https://arxiv.org/abs/1606.05718)
 226. Bejnordi BE, Litjens G, Timofeeva N, Otte-Höller I, Homeyer A, Karsseneijer N, Van Der Laak JAWM (2015) Stain specific standardization of whole-slide histopathological images. *IEEE Trans Med Imaging* 35(2):404–415
 227. Paeng K, Hwang S, Park S, Kim M (2017) A unified framework for tumor proliferation score prediction in breast histopathology. In: *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pp 231–239. Springer
 228. Li F, Wang Y, Tianyi X, Dong L, Yan L, Jiang M, Zhang X, Jiang H, Zhizheng W, Zou H (2022) Deep learning-based automated detection for diabetic retinopathy and diabetic macular Oedema in retinal fundus photographs. *Eye* 36(7):1433–1441
 229. Pinedo-Diaz G, Ortega-Cisneros S, Moya-Sanchez EU, Rivera J, Mejia-Alvarez P, Rodriguez-Navarrete FJ, Sanchez A (2022) Suitability classification of retinal fundus images for diabetic retinopathy using deep learning. *Electronics* 11(16):2564
 230. Saranya P, Prabakaran S, Kumar R, Das E (2022) Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning. *Visual Comput* 38(3):977–992
 231. Boreiko V, Ilanchezian I, Ayhan MS, Müller S, Koch LM, Faber H, Berens P, Hein M (2022) Visual explanations for the detection of diabetic retinopathy from retinal fundus images. In: *International conference on medical image computing and computer-assisted intervention*, pp 539–549. Springer
 232. Gunasekaran K, Pitchai R, Chaitanya GK, Selvaraj D, Sheryl SA, Almoallim HS, Alharbi SA, Raghavan SS, Tesemma BG (2022) A deep learning framework for earlier prediction of diabetic retinopathy from fundus photographs. *BioMed Res Int* 2022
 233. Saranya P, Devi SK, Bharanidharan B (2022b) Detection of diabetic retinopathy in retinal fundus images using densenet based deep learning model. In: *2022 international mobile and embedded technology conference (MECON)*, pp 268–272. IEEE
 234. Mikram M, Moujahdi C, Rhanoui M, Meddad M, Khallout A (2022) Hybrid deep learning models for diabetic retinopathy classification. In: *International conference on big data and internet of things*, pp 167–178. Springer
 235. Gao Z, Jin K, Yan Y, Liu X, Shi Y, Ge Y, Pan X, Yifei L, Jian W, Wang Y et al (2022) End-to-end diabetic retinopathy grading based on fundus fluorescein angiography images using deep learning. *Graefes Arch Clin Exp Ophthalmol* 260(5):1663–1673
 236. Abbood SH, Hamed HNA, Rahim MSM, Rehman A, Saba T, Bahaj SA (2022) Hybrid retinal image enhancement algorithm for diabetic retinopathy diagnostic using deep learning model. *IEEE Access* 10:73079–73086
 237. Nneji GU, Cai J, Deng J, Monday HN, Hossin MA, Nahar S (2022) Identification of diabetic retinopathy using weighted fusion deep learning based on dual-channel fundus scans. *Diagnostics* 12(2):540
 238. Dayana AM, Emmanuel WR (2022) Deep learning enabled optimized feature selection and classification for grading diabetic retinopathy severity in the fundus image. *Neural Comput Appl* 34(21):1–21
 239. Zhang G, Lin J-W, Wang J, Ji J, Cen L-P, Chen W, Xie P, Zheng Y, Xiong Y, Hanfu W et al (2022) Automated multidimensional deep learning platform for referable diabetic retinopathy detection: a multicentre, retrospective study. *BMJ Open* 12(7):e060155
 240. Canayaz M (2022) Classification of diabetic retinopathy with feature selection over deep features using nature-inspired wrapper methods. *Appl Soft Comput* 128:109462
 241. Venkaiahppalawamy B, Prasad Reddy PVGD, Batha S (2023) Hybrid deep learning approaches for the detection of diabetic retinopathy using optimized wavelet based model. *Biomed Signal Process Control* 79:104146
 242. Gupta IK, Choubey A, Choubey S (2022) Mayfly optimization with deep learning enabled retinal fundus image classification model. *Comput Electr Eng* 102:108176
 243. Granty Regina Elwin J, Mandala J, Maram B, Ramesh Kumar R (2022) Ar-hgso: autoregressive-henry gas sailfish optimization enabled deep learning model for diabetic retinopathy detection and severity level classification. *Biomed Signal Process Control* 77:103712
 244. Nderitu P, Joan M, do Rio N, Webster ML, Mann SS, Hopkins D, Cardoso MJ, Modat M, Bergeles C, Jackson TL (2022) Automated image curation in diabetic retinopathy screening using deep learning. *Sci Rep* 12(1):1–12
 245. Desika Vinayaki V, Kalaiselvi R (2022) Multithreshold image segmentation technique using remora optimization algorithm for diabetic retinopathy detection from fundus images. *Neural Process Lett* 54(3):2363–2384

246. Islam MR, Abdulrazak LF, Nahiduzzaman M, Goni MOF, Anower MS, Ahsan M, Haider J, Kowalsk M (2022) Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images. *Comput Biol Med* 146:105602
247. Babenko B, Mitani A, Traynis I, Kitade N, Singh P, Maa AY, Cuadros J, Corrado GS, Peng L, Webster DR et al (2022) Detection of signs of disease in external photographs of the eyes via deep learning. *Nature biomedical engineering*. *Nat Biomed Eng* 48:1–14
248. Kooi T, Litjens G, Van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, den Heeten A, Karssemeijer N (2017) Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal* 35:303–312
249. Kooi T, van Ginneken B, Karssemeijer N, den Heeten A (2017) Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network. *Med Phys* 44(3):1017–1027
250. Samala RK, Chan H-P, 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):6654–6666
251. Arora R, Rai PK, Raman B (2020) Deep feature-based automatic classification of mammograms. *Med Biol Eng Comput* 58(6):1199–1211
252. Zhang J, Saha A, Zhu Z, Mazurowski MA (2018) Hierarchical convolutional neural networks for segmentation of breast tumors in MRI with application to radiogenomics. *IEEE Trans Med Imaging* 38(2):435–447
253. Sun W, Tseng TLB, Zhang J, Qian W (2017) Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. *Comput Med Imaging Graph* 57:4–9
254. Hwang SH, Kim HE (2021) Machine learning method and apparatus based on weakly supervised learning, December 14 US Patent 11,200,483
255. Charan S, Khan MJ, Khurshid K (2018) Breast cancer detection in mammograms using convolutional neural network. In: 2018 international conference on computing, mathematics and engineering technologies (iCoMET), pp 1–5. IEEE
256. Hwang S, Kim H-E (2016) Self-transfer learning for fully weakly supervised object localization. [arXiv:1602.01625](https://arxiv.org/abs/1602.01625)
257. Muduli D, Dash R, Majhi B (2022) Automated diagnosis of breast cancer using multi-modal datasets: a deep convolution neural network based approach. *Biomed Signal Process Control* 71:102825
258. Dhungel N, Carneiro G, Bradley AP (2016) The automated learning of deep features for breast mass classification from mammograms. In: International conference on medical image computing and computer-assisted intervention, pp 106–114. Springer
259. Zhou J, Luo LY, Dou Q, Chen H, Chen C, Li GJ, Jiang ZF, Heng PA (2019) Weakly supervised 3d deep learning for breast cancer classification and localization of the lesions in MR images. *J Magn Resonan Imaging* 50(4):1144–1151
260. Abed AH (2022) Deep learning techniques for improving breast cancer detection and diagnosis. *Int J Adv Netw Appl* 13(06):5197–5214
261. Fotin SV, Yin Y, Haldankar H, Hoffmeister JW, Periaswamy S (2016) Detection of soft tissue densities from digital breast tomosynthesis: comparison of conventional and deep learning approaches. In: *Medical Imaging 2016: Computer-Aided Diagnosis*, vol 9785, pp 228–233. SPIE
262. Gao M, Fessler JA, Chan H-P (2022b) Deep convolutional neural network regularized digital breast tomosynthesis reconstruction with detector blur and correlated noise modeling. In: *Medical Imaging 2022: physics of medical imaging*, vol 12031. SPIE
263. Mota AM, Clarkson MJ, Almeida P, Matela N (2022) Automatic classification of simulated breast tomosynthesis whole images for the presence of microcalcification clusters using deep cnns. *J Imaging* 8(9):231
264. Jayasankar S, Bajhaiya D, Unni SN (2022) Deep learning-enabled soft tissue tumor localization using spatially offset Raman spectral analysis: in-silico investigations. *J Phys D* 55(39):394006
265. Akselrod-Ballin A, Karlinsky L, Alpert S, Hasoul S, Ben-Ari R, Barkan E (2016) A region based convolutional network for tumor detection and classification in breast mammography. In: *Deep learning and data labeling for medical applications*, pp 197–205. Springer
266. Frueh M, Kuestner T, Nachbar M, Thorwarth D, Schilling A, Gatidis S (2022) Self-supervised learning for automated anatomical tracking in medical image data with minimal human labeling effort. *Comput Methods Programs Biomed* 45:107085
267. Wei T, Aviles-Rivero AI, Wang S, Huang Y, Gilbert FJ, Schönlieb CB, Chen CW (2022) Beyond fine-tuning: classifying high resolution mammograms using function-preserving transformations. *Med Image Anal* 865:10268
268. Wang AS, Cunningham IA, Danielsson M, Fahrigr R, Flohr T, Hoeschen C, Noo F, Sabol JM, Siewerdsen JH, Tingberg A et al (2022) Science and practice of imaging physics through 50 years of SPIE medical imaging conferences. *J Med Imaging* 9(S1):012205
269. Hatamizadeh A, Hosseini H, Patel N, Choi J, Pole CC, Hoefelrin CM, Schwartz SD, Terzopoulos D (2022) Ravir: a dataset and methodology for the semantic segmentation and quantitative analysis of retinal arteries and veins in infrared reflectance imaging. *IEEE J Biomed Health Inform* 26(7):3272–3283
270. Khader F, Haarbuerger C, Kirr J-C, Menke M, Kather JN, Stegmaier J, Kuhl C, Nebelung S, Truhn D (2022) Elevating fundoscopic evaluation to expert level-automatic glaucoma detection using data from the airogs challenge. In: 2022 IEEE International symposium on biomedical imaging challenges (ISBIC), pp 1–4. IEEE
271. Camara J, Rezende R, Pires IM, Cunha A (2022) Retinal glaucoma public datasets: what do we have and what is missing? *J Clin Med* 11(13):3850
272. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin PM, Larochelle H (2017) Brain tumor segmentation with deep neural networks. *Med Image Anal* 35:18–31
273. Moeskops P, Viergever MA, Mendrik AM, De Vries LS, Benders MNL, Išgum I (2016) Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Trans Med Imaging* 35(5):1252–1261
274. Ghafoorian M, Karssemeijer N, Heskes T, Bergkamp M, Wissink J, Obels J, Keizer K, de Leeuw FE, van Ginneken B, Marchiori E et al (2017) Deep multi-scale location-aware 3d convolutional neural networks for automated detection of Lacunes of presumed vascular origin. *NeuroImage: Clin* 14:391–399
275. van der Burgh HK, Schmidt R, Westeneng H-J, de Reus MA, van den Berg LH, van den Heuvel MP (2017) Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis. *NeuroImage: Clin* 13:361–369
276. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B (2017) Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 36:61–78
277. Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bendszus M, Biller A (2016) Deep MRI brain extraction: a 3d convolutional neural network for skull stripping. *NeuroImage* 129:460–469
278. Ghafoorian M, Karssemeijer N, Heskes T, Van Uder IWM, de Leeuw F-E, Marchiori E, van Ginneken B, Platel B (2016)

- Non-uniform patch sampling with deep convolutional neural networks for white matter hyperintensity segmentation. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI), pp 1414–1417. IEEE
279. Havaei M, Guizard N, Chapados N, Bengio Y (2016) Hemis: hetero-modal image segmentation. In: international conference on medical image computing and computer-assisted intervention, pp 469–477. Springer
280. Dou Q, Chen H, Lequan Y, Zhao L, Qin J, Wang D, Mok VCT, Shi L, Heng P-A (2016) Automatic detection of cerebral microbleeds from MR images via 3d convolutional neural networks. *IEEE Trans Med Imaging* 35(5):1182–1195
281. Shi J, Zheng X, Li Y, Zhang Q, Ying S (2017) Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of alzheimer's disease. *IEEE J Biomed Health Inform* 22(1):173–183
282. Bashir R, Junejo R, Qadri NN, Fleury M, Qadri MY (2019) SWT and PCA image fusion methods for multi-modal imagery. *Multimedia Tools Appl* 78(2):1235–1263
283. Panigrahy C, Seal A, Mahato NK (2020) MRI and spect image fusion using a weighted parameter adaptive dual channel PCNN. *IEEE Signal Process Lett* 27:690–694
284. Yazdan SA, Ahmad R, Iqbal N, Rizwan A, Khan AN, Kim D-H (2022) An efficient multi-scale convolutional neural network based multi-class brain MRI classification for samd. *Tomography* 8(4):1905–1927
285. Wahlang I, Maji AK, Saha G, Chakrabarti P, Jasinski M, Leonowicz Z, Jasinska E (2022) Brain magnetic resonance imaging classification using deep learning architectures with gender and age. *Sensors* 22(5):1766
286. Alanazi MF, Ali MU, Hussain SJ, Zafar A, Mohatram M, Irfan M, AlRuwaiti R, Alruwaiti M, Ali NH, Albarrak AM (2022) Brain tumor/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model. *Sensors* 22(1):372
287. Bangare SL (2022) Classification of optimal brain tissue using dynamic region growing and fuzzy min-max neural network in brain magnetic resonance images. *Neurosci Inform* 2(3):100019
288. Siddiqi MH, Alsayat A, Alhwaiti Y, Azad M, Alruwaiti M, Alanazi S, Kamruzzaman MM, Khan A (2022) A precise medical imaging approach for brain MRI image classification. *Comput Intell Neurosci*
289. Ouchicha C, Ammor O, Meknassi M (2022) A novel deep convolutional neural network model for Alzheimer? Disease classification using brain MRI. *Automatic Control Comput Sci* 56(3):261–271
290. Haq EU, Jianjun H, Huarong X, Li K, Weng L (2022a) A hybrid approach based on deep cnn and machine learning classifiers for the tumor segmentation and classification in brain MRI. *Comput Math Method Medi*, 2022
291. Saurav S, Sharma A, Saini R, Singh S (2022) An attention-guided convolutional neural network for automated classification of brain tumor from MRI. *Neural Comput Appl* 56:1–20
292. Nayak DR, Padhy N, Mallick PK, Bagal DK, Kumar S (2022) Brain tumour classification using noble deep learning approach with parametric optimization through metaheuristics approaches. *Computers* 11(1):10
293. Fawaz A, Williams LZJ, Alansary A, Bass C, Gopinath K, da Silva M, Dahan S, Adamson C, Alexander B, Thompson D et al (2021) Benchmarking geometric deep learning for cortical segmentation and neurodevelopmental phenotype prediction. *bioRxiv*
294. Dahan S, Fawaz A, Williams LZJ, Yang C, Coalson TS, Glasser MF, Edwards AD, Rueckert D, Robinson EC (2022) Surface vision transformers: Attention-based modelling applied to cortical analysis. [arXiv:2203.16414](https://arxiv.org/abs/2203.16414)
295. Kong R, Li J, Orban C, Sabuncu MR, Liu H, Schaefer A, Sun N, Zuo X-N, Holmes AJ, Eickhoff SB et al (2019) Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cereb Cortex* 29(6):2533–2551
296. Liew S-L, Lo BP, Donnelly MR, Zavaliangos-Petropulu A, Jeong JN, Barisano G, Hutton A, Simon JP, Juliano JM, Suri A et al (2022) A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation algorithms. *Sci Data* 9(1):1–12
297. Lu Fang W, Peijun FH, Zhiyi P, Kong D (2016) Automatic 3d liver location and segmentation via convolutional neural network and graph cut. *Int J Comput Assist Radiol Surg* 12(2):171–182
298. Ravishankar H, Sudhakar P, Venkataramani R, Thiruvankadam S, Annangi P, Babu N, Vaidya V (2016) Understanding the mechanisms of deep transfer learning for medical images. In: *Deep learning and data labeling for medical applications*, pp 188–196. Springer
299. Roth HR, Lu L, Farag A, Sohn A, Summers RM (2016) Spatial aggregation of holistically-nested networks for automated pancreas segmentation in international conference on medical image computing and computer-assisted intervention
300. Näppi J J, Hironaka T, Regge D, Yoshida H (2016) Deep transfer learning of virtual endoluminal views for the detection of polyps in ct colonography. In: *Medical imaging 2016: computer-aided diagnosis*, vol 9785, pp 590–597. SPIE
301. Ren Y, Chen Y, Lin L, Zhao J (2018) Novel 3d radiomic features for computer-aided polyp detection in CT colonography. *IEEE Access* 6:74506–74520
302. Uemura T, Näppi JJ, Ryu Y, Watari C, Kamiya T, Yoshida H (2021) A generative flow-based model for volumetric data augmentation in 3d deep learning for computed tomographic colonography. *Int J Comput Assist Radiol Surg* 16(1):81–89
303. Yasuda T, Honda T, Utano K, Kato T, Togashi K, Yamaguchi S, Yasaka T (2022) Diagnostic accuracy of ultra-low-dose CT colonography for the detection of colorectal polyps: a feasibility study. *Jpn J Radiol* 18:1–9
304. Younas F, Usman M, Yan WQ (2022) A deep ensemble learning method for colorectal polyp classification with optimized network parameters. *Appl Intell* 74:1–24
305. Tanwar S, Vijayalakshmi S, Sabharwal M, Kaur M, AlZubi AA, Lee H-N (2022) Detection and classification of colorectal polyp using deep learning. *BioMed Res Int* 2022
306. Wesp P, Grosu S, Graser A, Maurus S, Schulz C, Knösel T, Fabritius MP, Schachtner B, Yeh BM, Cyran CC et al (2022) Deep learning in CT colonography: differentiating premalignant from benign colorectal polyps. *Eur Radiol*, pp 1–11
307. Biffi C, Salvagnini P, Dinh NN, Hassan C, Sharma P, Cherubini A (2022) A novel AI device for real-time optical characterization of colorectal polyps. *NPJ Digital Med* 5(1):1–8
308. Sánchez-Peralta LF, Pagador JB, Sánchez-Margallo FM (2022) Artificial intelligence for colorectal polyps in colonoscopy. In: *Artificial intelligence in medicine*, pp 967–981. Springer
309. Kusters KC, Scheeve T, Dehghani N, van der Zander QEW, Schreuder QEW, Masclee AAM, Schoon EJ, van der Sommen F et al (2022) Colorectal polyp classification using confidence-calibrated convolutional neural networks. In: *Medical imaging 2022: computer-aided diagnosis*, vol 12033, pp 442–454. SPIE
310. Byeon S, Park J, Cho YA, Cho BJ (2022) Automated histological classification for digital pathology images of colonoscopy specimen via deep learning. *Scie Rep* 12(1):1–8
311. Nisha JS, Gopi VP, Palanisamy P (2022) Automated colorectal polyp detection based on image enhancement and dual-path CNN architecture. *Biomed Signal Process Control* 73:103465
312. Cao W, Pomeroy MJ, Liang Z, Abbasi AF, Pickhardt P, Hongbing L (2022) Vector textures derived from higher order derivative domains for classification of colorectal polyps. *Vis Comput Ind Biomed Art* 5(1):1–14

313. Harini N, Ramji B, Sriram S, Sowmya V, Soman KP (2020) Musculoskeletal radiographs classification using deep learning. In: *Deep learning for data analytics*, pp 79–98. Elsevier,
314. Kamiya N (2020) Deep learning technique for musculoskeletal analysis. *Deep Learn Med Image Anal* 586:165–176
315. Kijowski R, Liu F, Caliva F, Pedoia V (2020) Deep learning for lesion detection, progression, and prediction of musculoskeletal disease. *J Magn Resonance Imaging* 52(6):1607–1619
316. Mall PK, Singh PK (2022) Boostnet: a method to enhance the performance of deep learning model on musculoskeletal radiographs x-ray images. *Int J Syst Assuran Eng Manag* 13(1):658–672
317. Singh G, Anand D, Cho W, Joshi GP, Son KC (2022) Hybrid deep learning approach for automatic detection in musculoskeletal radiographs. *Biology* 11(5):665
318. Cheng Y, Jin Z, Zhou X, Zhang W, Zhao D, Tao C, Yuan J (2022) Diagnosis of metacarpophalangeal synovitis with musculoskeletal ultrasound images. *Ultrasound Med Biol* 48(3):488–496
319. Zhu J, Bolsterlee B, Chow BVY, Song Y, Meijering E (2022a) Hybrid attentive unet for segmentation of lower leg muscles and bones from MRI scans for musculoskeletal research. In: *2022 IEEE 19th international symposium on biomedical imaging (ISBI)*, pp 1–5. IEEE
320. Shin H, Choi GS, Shon OJ, Kim GB, Chang MC (2022) Development of convolutional neural network model for diagnosing meniscus tear using magnetic resonance image. *BMC Musculoskelet Disord* 23(1):1–9
321. He Z, Qin Z, Koike Y (2022) Continuous estimation of finger and wrist joint angles using a muscle synergy based musculoskeletal model. *Appl Sci* 12(8):3772
322. Malik S, Amin J, Sharif M, Yasmin M, Kadry S, Anjum S (2022) Fractured elbow classification using hand-crafted and deep feature fusion and selection based on whale optimization approach. *Mathematics* 10(18):3291
323. Gitto S, Cuocolo R, van Langevelde K, van de Sande MAJ, Parafioriti A, Luzzati A, Imbriaco M, Sconfienza LM, Bloem JL (2022) MRI radiomics-based machine learning classification of atypical cartilaginous tumour and grade ii chondrosarcoma of long bones. *EBioMedicine* 75:103757
324. Luong HH, Le Lan TT, Nguyen HT, Hua VQ, Nguyen KV, Bach TNP, Nguyen TNA, Nguyen HTQ (2022) Transfer learning with fine-tuning on mobilenet and grad-cam for bones abnormalities diagnosis. In: *Computational intelligence in security for information systems conference*, pp 171–179. Springer
325. Gao KT, Tibrewala R, Hess M, Bharadwaj UU, Inamdar G, Link TM, Chin CT, Pedoia V, Majumdar S (2022c) Automatic detection and voxel-wise mapping of lumbar spine modic changes with deep learning. *JOR Spine*, pp e1204
326. Fabry V, Mamalet F, Laforet A, Capelle M, Acket B, Sengenès C, Cintas P, Faruch-Bilfeld M (2022) A deep learning tool without muscle-by-muscle grading to differentiate myositis from facio-scapulo-humeral dystrophy using MRI. *Diagnostic and Interventional Imaging*
327. Jamaludin A, Kadir T, Zisserman A (2017) Spinenet: automated classification and evidence visualization in spinal MRIS. *Med Image Anal* 41:63–73
328. Nogueira-Rodríguez A, Domínguez-Carbajales R, López-Fernández H, Iglesias Á, Cubiella J, Fdez-Riverola F, Reboiro-Jato M, Glez-Peña D (2021) Deep neural networks approaches for detecting and classifying colorectal polyps. *Neurocomputing* 423:721–734
329. Yu L, Yang X, Chen H, Jing Q, Pheng AH (2017) Volumetric convnets with mixed residual connections for automated prostate segmentation from 3d MR images. In: *Thirty-first AAAI conference on artificial intelligence*
330. Lu Y, Zhao Yan, Chen Xing, Guo Xiaoxin (2022) A novel u-net based deep learning method for 3d cardiovascular MRI segmentation. *Comput Intell Neurosci* 2022
331. Song Y, Ren S, Lu Y, Fu X, Wong KL (2022) Deep learning-based automatic segmentation of images in cardiac radiography: a promising challenge. *Comput Methods Programs Biomed* 59:106821
332. Wang Z, Peng Y, Li D, Guo Y, Zhang B (2022) Mmnet: a multi-scale deep learning network for the left ventricular segmentation of cardiac MRI images. *Appl Intell* 52(5):5225–5240
333. Sandooghdar A, Yaghmaee F (2022) Deep learning approach for cardiac MRI images. *J Inform Syst Telecommun (JIST)* 1(37):61
334. Ahmad I, Qayyum A, Gupta BB, Alassafi MO, AlGhamdi RA (2022) Ensemble of 2d residual neural networks integrated with atrous spatial pyramid pooling module for myocardium segmentation of left ventricle cardiac MRI. *Mathematics* 10(4):627
335. Alabed S, Uthoff J, Zhou S, Garg P, Dwivedi K, Alandejani F, Gosling R, Schobs L, Brook M, Shahin Y et al (2022) Machine learning cardiac-MRI features predict mortality in newly diagnosed pulmonary arterial hypertension. *Eur Heart J-Digital Health* 3(2):265–275
336. Fernández-Llaneza D, Gondová A, Vince H, Patra A, Zurek M, Konings P, Kagelid P, Hultin L (2022) Towards fully automated segmentation of rat cardiac MRI by leveraging deep learning frameworks. *Sci Rep* 12(1):1–13
337. Popescu DM, Abramson HG, Rebecca Y, Lai C, Shade JK, Wu KC, Maggioni M, Trayanova NA (2022) Anatomically informed deep learning on contrast-enhanced cardiac magnetic resonance imaging for scar segmentation and clinical feature extraction. *Cardiovasc Digital Health J* 3(1):2–13
338. da Silva IFS, Silva AC, de Paiva AC, Gattass M (2022) A cascade approach for automatic segmentation of cardiac structures in short-axis cine-mr images using deep neural networks. *Expert Syst Appl* 197:116704
339. Arian F, Amini M, Mostafaei S, Kalantari KR, Avval AH, Shahbazi Z, Kasani K, Rajabi AB, Chatterjee S, Oveisi M et al (2022) Myocardial function prediction after coronary artery bypass grafting using MRI radiomic features and machine learning algorithms. *J Digital Imaging* 89:1–11
340. Corrado PA, Wentland AL, Starekova J, Dhyani A, Goss KN, Wieben O (2022) Fully automated intracardiac 4d flow MRI post-processing using deep learning for biventricular segmentation. *Eur Radiol* 89:1–10
341. Lalande A, Chen Z, Pommier T, Decourselle T, Qayyum A, Salomon M, Ginhac D, Skandarani Y, Boucher A, Brahim K et al (2022) Deep learning methods for automatic evaluation of delayed enhancement-MRI. the results of the emidec challenge. *Med Image Anal* 79:102428
342. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Išgum I (2016) Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Med Image Anal* 34:123–136
343. Poudel RPK, Lamata P, Montana G (2016) Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. In: *Reconstruction, segmentation, and analysis of medical images*, pp 83–94. Springer
344. Kong B, Zhan Y, Shin M, Denny T, Zhang S (2016) Recognizing end-diastole and end-systole frames via deep temporal regression network. In: *International conference on medical image computing and computer-assisted intervention*, pp 264–272. Springer
345. Bhardwaj R, Nambiar AR, Dutta D (2017) A study of machine learning in healthcare. In: *2017 IEEE 41st annual computer software and applications conference (COMPSAC)*, vol 2, pp 236–241. IEEE

346. Ahsan MM, Luna SA, Siddique Z (2022) Machine-learning-based disease diagnosis: a comprehensive review. In: *Healthcare*, vol 10, p 541. MDPI
347. Dash S, Shakyawar SK, Sharma M, Kaushik S (2019) Big data in healthcare: management, analysis and future prospects. *J Big Data* 6(1):1–25
348. Jwa AS, Poldrack RA (2022) The spectrum of data sharing policies in neuroimaging data repositories. *Hum Brain Mapp* 43(8):2707–2721
349. Rasoo RU, Ahmad HF, Rafique W, Qayyum A, Qadir J (2022) Security and privacy of internet of medical things: A contemporary review in the age of surveillance, botnets, and adversarial ml. *J Netw Comput Appl* 45:103332
350. Pires F, Costa C, Dias P (2021) On the use of virtual reality for medical imaging visualization. *J Digital Imaging* 34(4):1034–1048
351. Oakley J (2003) *Digital imaging: a primer for radiographers, radiologists and health care professionals*. Cambridge University Press, Cambridge
352. Dorobanțiu A, Brad R (2019) Improving lossless image compression with contextual memory. *Appl Sci* 9(13):2681
353. Siegle JH, Jia X, Durand S, Gale S, Bennett C, Graddis N, Heller G, Ramirez TK, Choi H, Luviano JA et al (2021) Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature* 592(7852):86–92
354. Allen EJ, St-Yves G, Yihan W, Breedlove JL, Prince JS, Dowdle LT, Nau M, Caron B, Pestilli F, Charest I et al (2022) A massive 7t fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nat Neurosci* 25(1):116–126
355. Larobina M, Murino L (2014) Medical image file formats. *J Digital Imaging* 27(2):200–206
356. Besser H (2003) *Introduction to imaging*. Getty Publications, Los Angeles
357. Varma DR (2012) Managing dicom images: tips and tricks for the radiologist. *Indian J Radiol Imaging* 22(01):4–13
358. Gambino O, Rundo L, Cannella V, Vitabile S, Pirrone R (2018) A framework for data-driven adaptive GUI generation based on dicom. *J Biomed Inform* 88:37–52
359. Wong KKL, Fortino G, Abbott D (2020) Deep learning-based cardiovascular image diagnosis: a promising challenge. *Fut Gener Comput Syst* 110:802–811
360. Han T, Liu C, Yang W, Jiang D (2019) Learning transferable features in deep convolutional neural networks for diagnosing unseen machine conditions. *ISA Trans* 93:341–353
361. Greener JG, Kandathil SM, Moffat L, Jones DT (2022) A guide to machine learning for biologists. *Nat Rev Mol Cell Biol* 23(1):40–55
362. Sharma V, Mir RN (2020) A comprehensive and systematic look up into deep learning based object detection techniques: a review. *Comput Sci Rev* 38:100301
363. Osipyan H, Bosede IE, Adrian DC (2022) *Deep neural network applications*. CRC Press, Boca Raton
364. Nweke HF, Teh YW, Al-Garadi MA, Alo UR (2018) Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: state of the art and research challenges. *Expert Syst Appl* 105:233–261
365. Kampouraki V (2021) Patch-level classification of brain tumor tissue in digital histopathology slides with deep learning
366. Khan S, Yairi T (2018) A review on the application of deep learning in system health management. *Mech Syst Signal Process* 107:241–265
367. Voulodimos A, Protopapadakis E, Katsamenis I, Doulamis A, Doulamis N (2021) A few-shot u-net deep learning model for Covid-19 infected area segmentation in CT images. *Sensors* 21(6):2215
368. Shen D, Guorong W, Suk H-I (2017) Deep learning in medical image analysis. *Ann Rev Biomed Eng* 19:221
369. Rouzrokh P, Khosravi B, Faghani S, Moassefi M, Vera DV, Garcia YS, Zhang K, Conte GM, Erickson BJ (2022) Mitigating bias in radiology machine learning: 1. data handling. *Radiol Artif Intell* 4(5):e210290
370. Gurung B, Liu P, Harris PDR, Sagi A, Field RE, Sochart DH, Tucker K, Asopa V (2022) Artificial intelligence for image analysis in total hip and total knee arthroplasty: a scoping review. *Bone Joint J* 104(8):929–937
371. Widrow B, Hoff ME (1960) Adaptive switching circuits. Technical report, Stanford Univ Ca Stanford Electronics Labs
372. Shu H, Zhu H (2019) Sensitivity analysis of deep neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence* 33:4943–4950
373. Teodoro G, Kurç TM, Taveira LFR, Melo ACMA, Gao Y, Kong J, Saltz H (2017) Algorithm sensitivity analysis and parameter tuning for tissue image segmentation pipelines. *Bioinformatics* 33(7):1064–1072
374. Xiang L, Zeng X, Shengli W, Liu Y, Yuan B (2021) Computation of CNN? Sensitivity to input perturbation. *Neural Process Lett* 53(1):535–560
375. Herb J, Perin Y, Yum S, Mylonakis A, Demaziere C, Vinai P, Yu M, Wingate J, Hursin M (2022) Sensitivity analysis in core diagnostics. *Ann Nucl Energy* 178:109350
376. Roy Y, Banville H, Albuquerque I, Gramfort A, Falk TH, Faubert J (2019) Deep learning-based electroencephalography analysis: a systematic review. *J Neural Eng* 16(5):051001
377. Rjoob K, Bond R, Finlay D, McGilligan V, Leslie SJ, Rababah A, Iftikhar A, Guldenring D, Knoery C (2022) Machine learning and the electrocardiogram over two decades: time series and meta-analysis of the algorithms, evaluation metrics and applications. *Artif Intell Med* 89:102381
378. Duchi J, Hazan E, Singer Y (2011) Adaptive subgradient methods for online learning and stochastic optimization. *J Mach Learn Res* 7:12
379. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980)
380. Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K, Mannel RS, Liu H, Zheng B, Qiu Y (2022) Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal* 7:102444
381. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, García S, Gil-López S, Molina D, Benjamins R et al (2020) Explainable artificial intelligence (xai): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inform Fus* 58:82–115
382. Howard J, Gugger S (2020) Fastai: a layered API for deep learning. *Information* 11(2):108
383. Bischl B, Binder M, Lang M, Pielok T, Richter J, Coors S, Thomas J, Ullmann T, Becker M, Boulesteix A-L et al (2021) Hyperparameter optimization: foundations, algorithms, best practices and open challenges. [arXiv:2107.05847](https://arxiv.org/abs/2107.05847)
384. Patterson J, Gibson A (2017) *Deep learning: a practitioner's approach*. O'Reilly Media, Inc., Sebastopol
385. Wang S, Summers RM (2012) Machine learning and radiology. *Med Image Anal* 16(5):933–951
386. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, Liu Y, Topol E, Dean J, Socher R (2021) Deep learning-enabled medical computer vision. *NPJ Digital Med* 4(1):1–9
387. Kim J, Hong J, Park H (2018) Prospects of deep learning for medical imaging. *Precis Future Med* 2(2):37–52
388. Wang M, Fu W, He W, Hao S, Wu X (2020) A survey on large-scale machine learning. *IEEE Transactions on Knowledge and Data Engineering*

389. Verbraeken J, Wolting M, Katzy J, Kloppenburg J, Verbelen T, Rellermeyer JS (2020) A survey on distributed machine learning. *ACM Comput Surv (CSUR)* 53(2):1–33
390. Alcaín E, Fernández PR, Nieto R, Montemayor AS, Vilas J, Galiana-Bordera A, Martínez-Girones PM, Lastra CP, Rodríguez-Vila B, Bonet M et al (2021) Hardware architectures for real-time medical imaging. *Electronics* 10(24):3118
391. Xing F, Xie Y, Hai S, Liu F, Yang L (2017) Deep learning in microscopy image analysis: a survey. *IEEE Trans Neural Netw Learn Syst* 29(10):4550–4568
392. Kumar R, Wang WY, Kumar J, Yang T, Khan A, Ali W, Ali I (2021) An integration of blockchain and AI for secure data sharing and detection of CT images for the hospitals. *Comput Med Imaging Graph* 87:101812
393. Ali Hasan MD, Mahir LMD, Labib TH, Hossain M et al (2021) Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: an extensive study. *Inform Med Unlock* 24:100582
394. Kumar Y, Mahajan M (2020) 5. recent advancement of machine learning and deep learning in the field of healthcare system. *Comput Intell Mach Learn Healthcare Inform* 1:77
395. Kaoru O, Minh SD, Vasileios M, Francesco GB, DE Natale (2017) Deep learning for mobile multimedia: a survey. *ACM Trans Multimedia Comput Commun Appl* 13(3):1–22
396. Choi W, Duraisamy K, Kim RG, Doppa JR, Pande PP, Marculescu R, Marculescu D (2016) Hybrid network-on-chip architectures for accelerating deep learning kernels on heterogeneous manycore platforms. In: *Proceedings of the international conference on compilers, architectures and synthesis for embedded systems*, pp 1–10
397. Poorheravi PA (2022) Software and fpga-based hardware to accelerate machine learning classifiers. Master's thesis, University of Waterloo
398. Graziani M, Dutkiewicz L, Calvaresi D, Amorim JP, Yordanova K, Vered M, Nair R, Abreu PH, Blanke T, Pulignano V et al (2022) A global taxonomy of interpretable AI: unifying the terminology for the technical and social sciences. *Artif Intell Rev* 15:1–32
399. Ahmed M, Zubair S (2022) Explainable artificial intelligence in sustainable smart healthcare. In: *Explainable artificial intelligence for cyber security*, pp 265–280. Springer
400. Molnar C, Casalicchio G, Bischl B (2020) Interpretable machine learning—a brief history, state-of-the-art and challenges. In: *Joint European conference on machine learning and knowledge discovery in databases*, pp 417–431. Springer
401. Krishnan M (2020) Against interpretability: a critical examination of the interpretability problem in machine learning. *Philos Technol* 33(3):487–502
402. Faghani S, Khosravi B, Zhang K, Moassefi M, Nugen F, Vahdati S, Kuanar SP, Rassoulnejad-Mousavi SM, Singh Y, Singh Y et al (2022) Mitigating bias in radiology machine learning: 3. performance metrics. *Radiology* 4(5):e220061
403. Kumar Y, Gupta S (2022) Deep transfer learning approaches to predict glaucoma, cataract, choroidal neovascularization, diabetic macular edema, Drusen and healthy eyes: an experimental review. *Archiv Comput Methods Eng* 10:1–21
404. Ariew R (1976) Ockham's razor: a historical and philosophical analysis of Ockham's principle of parsimony University of Illinois at Urbana-Champaign
405. Mehta O, Liao Z, Jenkinson M, Carneiro G, Verjans J (2022) Machine learning in medical imaging—clinical applications and challenges in computer vision. *Artif Intell Med* 8:79–99
406. Rasheed J (2022) Analyzing the effect of filtering and feature-extraction techniques in a machine learning model for identification of infectious disease using radiography imaging. *Symmetry* 14(7):1398
407. Mamandra E (2022) Diabetes diagnosis using machine learning. Master's thesis, Πανεπιστήμιο Πειραιώς
408. Zhu Z, Dong Z, Liu Y (2022b) Detecting corrupted labels without training a model to predict. In: *International conference on machine learning*, pp 27412–27427. PMLR
409. Teney AE, Lucey S, van den Hengel A (2022) Evading the simplicity bias: training a diverse set of models discovers solutions with superior OOD generalization. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp 16761–16772
410. Kenny D (2022) Human and machine translation. *Mac Transl Everyone* 18:23
411. Vidal XC, Maroñas LD, Suárez Á (2022) Modern machine learning: Applications and methods. In: *Machine learning and artificial intelligence with industrial applications*, pp 19–61. Springer
412. Hüllermeier E, Waegeman W (2021) Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach Learn* 110(3):457–506
413. Ferrenberg AM, Swendsen RH (1989) Optimized monte Carlo data analysis. *Comput Phys* 3(5):101–104
414. Gal Y et al (2016) Uncertainty in deep learning
415. Li L, Jamieson K, Rostamizadeh A, Gonina E, Ben-Tzur J, Hardt M, Recht B, Talwalkar A (2020) A system for massively parallel hyperparameter tuning. *Proc Mach Learn Syst* 2:230–246
416. Wang Y, Zhang H, Zhang G (2019) CPSO-CNN: an efficient PSO-based algorithm for fine-tuning hyper-parameters of convolutional neural networks. *Swarm Evolut Comput* 49:114–123
417. Elaziz MA, Dahou A, Abualigah L, Liyang Y, Alshinwan M, Khasawneh AM, Songfeng L (2021) Advanced metaheuristic optimization techniques in applications of deep neural networks: a review. *Neural Comput Appl* 33(21):14079–14099
418. Deb K, Pratap A, Agarwal S, Meyarivan TAMT (2002) A fast and elitist multiobjective genetic algorithm: Nsga-II. *IEEE Trans Evolut Comput* 6(2):182–197
419. Bergstra J, Bardenet R, Bengio Y, Kégl B (2011) Algorithms for hyper-parameter optimization. *Adv Neural Inform Process Syst* 24:5
420. Haq A, Wilk S, Abelló A (2022b) Comparison of models built using automl and data fusion. In: *European Conference on advances in databases and information systems*, pp 301–314. Springer
421. Trizna D et al (2022) Adversarial robustness of hybrid machine learning architecture for malware classification
422. Leili MS, Reza S (2022) Machine learning: theory to applications. CRC Press, Boca Raton
423. Ogundokun RO, Misra S, Douglas M, Damaševičius R, Maskeliūnas R (2022) Medical internet-of-things based breast cancer diagnosis using hyperparameter-optimized neural networks. *Fut Internet* 14(5):153
424. Ma Y, Xie Y (2022) Evolutionary neural networks for deep learning: a review. *Int J Mach Learn Cybernet* 4:1–18
425. Klapwijk ET, van den Bos W, Tamnes CK, Raschle NM, Mills KL (2021) Opportunities for increased reproducibility and replicability of developmental neuroimaging. *Dev Cognit Neurosci* 47:100902
426. Glenn Begley C, Ioannidis JPA (2015) Reproducibility in science: improving the standard for basic and preclinical research. *Circ Res* 116(1):116–126
427. Rougier NP, Hinsén K, Alexandre F, Arildsen T, Barba LA, Benureau FCY, Titus Brown C, De Buyl P, Caglayan O, Davison AP et al (2017) Sustainable computational science: the resilience initiative. *PeerJ Comput Sci* 3:e142
428. Lopes PH, Oliveir BC, de Souza AA, Blanco W (2022) Mitigating computer limitations in replicating numerical simulations of

- a neural network model with hodgkin-huxley-type neurons. *Front Neuroinform* 16:8
429. Kenthapadi K, Lakkaraju H, Natarajan P, Sameki M (2022) Model monitoring in practice: lessons learned and open challenges. In: *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, pp 4800–4801
 430. Meel V, Bodepudi A (2022) Melatect: a machine learning approach for identifying malignant melanoma in skin growths. In: *Fourteenth international conference on machine vision (ICMV 2021)*, vol 12084, pp 433–441. SPIE
 431. Bangert P, Moon H, Woo JO, Didari S, Hao H (2022) Medical image labeling via active learning is 90% effective. In: *Future of information and communication conference*, pp 291–310. Springer
 432. Liang W, Tadesse GA, Ho D, Li FF, Zaharia M, Zhang C, Zou J (2022) Advances, challenges and opportunities in creating data for trustworthy AI. *Nat Mach Intell* 12:1–9
 433. Dave D, Naik H, Singha S, Dwivedi R, Patel P (2022) Explainable AI for ICT: System and software architecture. *Recent Adv ICT Infrast Appl* 7:198
 434. Zhang J, Budhdeo S, William W, Cerrato P, Shuaib H, Sood H, Ashrafian H, Halamka J, Teo JT (2022) Moving towards vertically integrated artificial intelligence development. *NPJ Digit Med* 5(1):1–9
 435. Fergus P, Chalmers C (2022) Deploying and hosting machine learning models. In: *Applied deep learning*, pp 299–317. Springer
 436. Kleftakis S, Mavrogiorgou A, Mavrogiorgos K, Kiourtis A, Kyriazis D (2022) Digital twin in healthcare through the eyes of the Vitruvian man. In: *Innovation in Medicine and Healthcare*, pp 75–85. Springer
 437. Sun Z, Di L, Burgess A, Tullis JA, Magill AB (2020) Geoweaver: advanced cyberinfrastructure for managing hybrid geoscientific AI workflows. *ISPRS Int J Geo-Inform* 9(2):119
 438. Reddy M, Dattaprasad B, Kammath S, Kn S, Manokaran S, Be R (2022) Application of mlops in prediction of lifestyle diseases. *ECS Trans* 107(1):1191
 439. Salahuddin Z, Woodruff HC, Chatterjee Ak, Lambin Pe (2022) Transparency of deep neural networks for medical image analysis: a review of interpretability methods. *Comput Biol Med* 140:105111
 440. Rajpurkar P, Chen E, Banerjee O, Topol EJ (2022) Ai in health and medicine. *Nat Med* 28(1):31–38
 441. Jin W, Li X, Fatehi M, Hamarneh G (2022b) Guidelines and evaluation for clinical explainable ai on medical image analysis. [arXiv:2202.10553](https://arxiv.org/abs/2202.10553)
 442. Singh A, Sengupta S, Lakshminarayanan V (2020) Explainable deep learning models in medical image analysis. *J Imaging* 6(6):52
 443. Ho CWL, Soon D, Caals K, Kapur J (2019) Governance of automated image analysis and artificial intelligence analytics in healthcare. *Clin Radiol* 74(5):329–337
 444. Castiglioni I, Rundo L, Codari M, Di Leo G, Salvatore C, Interlenghi M, Gallivanone F, Cozzi A, D'Amico NC, Sardanelli F (2021) Ai applications to medical images: from machine learning to deep learning. *Phys Medica* 83:9–24
 445. Canadian Association of Radiologists (CAR) (2019) Artificial Intelligence Working Group. Canadian association of radiologists white paper on ethical and legal issues related to artificial intelligence in radiology. *Can Assoc Radiol J*, 70(2):107–118
 446. Currie G, Hawk KE, Rohren EM (2020) Ethical principles for the application of artificial intelligence (AI) in nuclear medicine
 447. Abels E, Pantanowitz L, Aeffner F, Zarella MD, van der Laak J, Bui MM, Vemuri VNP, Parwani AV, Gibbs J, Agosto-Arroyo E et al (2019) Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the digital pathology association. *J Pathol* 249(3):286–294
 448. D'Antonoli TA (2020) Ethical considerations for artificial intelligence: an overview of the current radiology landscape. *Diagnostic Interv Radiol* 26(5):504
 449. Feynman RP (2018) *Simulating physics with computers*. In: *Feynman and computation*. CRC Press, pp 133–153
 450. Rodrigues PS, Ferreira MJ, Monteiro JL (2010) Quantum computation perspectives in medical image processing. In: *Handbook of research on developments in E-Health and telemedicine: technological and social perspectives*, pp 125–155. IGI Global
 451. Kavitha SS (2022) Quantum machine learning for support vector machine classification. *Evolut Intell* 78:1–10
 452. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S (2017) Quantum machine learning. *Nature* 549(7671):195–202
 453. Dey N, Ashour AS (2018) Computing in medical image analysis. In: *Soft computing based medical image analysis*, pp 3–11. Elsevier
 454. Elaraby A (2022) *Quantum medical images processing foundations and applications*. IET Quant Commun
 455. Kluyver T, Ragan-Kelley B, Pérez F, Granger BE, Bussonnier BE, Frederic J, Kelley K, Hamrick JB, Grout J, Corlay S et al (2016) *Jupyter Notebooks—a publishing format for reproducible computational workflows.*, vol 2016

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.