




Medical images classification using deep learning: a survey

Rakesh Kumar¹ · Pooja Kumbharkar¹ · Sandeep Vanam¹ · Sanjeev Sharma¹ 

Received: 26 April 2022 / Revised: 12 February 2023 / Accepted: 19 April 2023 /

Published online: 28 July 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Deep learning has made significant advancements in recent years. The technology is rapidly evolving and has been used in numerous automated applications with minimal loss. With these deep learning methods, medical image analysis for disease detection can be performed with minimal errors and losses. A survey of deep learning-based medical image classification is presented in this paper. As a result of their automatic feature representations, these methods have high accuracy and precision. This paper reviews various models like CNN, Transfer learning, Long short term memory, Generative adversarial networks, and Autoencoders and their combinations for various purposes in medical image classification. The total number of papers reviewed is 158. In the study, we discussed the advantages and limitations of the methods. A discussion is provided on the various applications of medical imaging, the available datasets for medical imaging, and the evaluation metrics. We also discuss the future trends in medical imaging using artificial intelligence.

Keywords Image classification · Deep Learning models · Performance · Features · Medical imaging

1 Introduction

Medical imaging [133] is the technique and process which takes images of the various interior parts of the human body for diagnostics and treatment purposes. It also has a database

✉ Sanjeev Sharma
sanjeevsharma@iiitp.ac.in

Rakesh Kumar
rakeshkumar@cse.iiitp.ac.in

Pooja Kumbharkar
poojakumbharkar20@ece.iiitp.ac.in

Sandeep Vanam
vanamsandeep20@ece.iiitp.ac.in

¹ Indian Institute of Information Technology Pune, Pune, Maharashtra, India

of normal humans to make it possible to identify abnormalities. Using medical imaging in health care allows doctors to have more accurate results and make better decisions.

Analyzing medical images involves image classification, object detection, and segmentation [89]. The purpose of image classification [87] is to determine a particular class out of many possible ones based on the image. The purpose of object detection [183] is to locate or find objects in an image. Image segmentation helps to change the representation of an image into something that is easier to analyze. Through image analysis, it is possible to detect many diseases, assess a patient's condition based on the medical images generated by a variety of methods [30]. Some of the uses of medical image analysis are calculating the volume of the human left ventricle (LV) for diagnosing cardiac diseases [58], the chest X-rays or CT scans are helpful for doctors in finding out the type of pulmonary nodules [116], Alzheimer's disease detection [10], Parkinson's disease detection, brain disorders [85] identification, covid detection [36, 104], and many more. By analyzing medical images, doctors can detect the disease, identify the problem, and determine the rate of tumor growth. The above points demonstrate the importance of image analysis in all medical fields.

Most of the medical image analyses are still done manually by the doctors or physicians is time-consuming and error-prone. In some of the applications like estimating the volume of the human left ventricle in the heart, to classify some of the retinal diseases (there are high similarities in various kinds of retinal disease) [63] are done manually and it is a tedious job and requires high expertise in those areas and error-prone. Traditional models cannot perform better in such applications due to their complexity. Deep learning models are system such that they can learn any complex function.

Deep learning [8, 101, 102] is a subset of machine learning [45]. There are many neural networks for both supervised and unsupervised learning in this system. The functionality of neural network [94] is very much similar to the functionality of the brain. It has 3 layers those are input, hidden, and output layer. A neural network consists of a set of nodes in every layer. The input layer takes the inputs. Hidden layers process the input data. Output layers produce output which could be classification, regression. Output functions for real values is a linear function and the output function for probabilities is softmax. The connection between the layers happens through weights and biases. The weights and biases are variables and will get updates when the network is trained. The neural network is trained with a back propagation algorithm. The back propagation algorithm runs until the loss function reduces to an acceptable range. The loss function for real valued outputs is squared error and the loss function for probabilities is cross entropy. Each node has pre-activation and post-activation. The activation could be linear or non-linear. Some of the examples of activation functions are sigmoid, logistic, linear, relu functions.

In supervised learning [156] the outputs are known for training data and we have to train the model to find the outputs for test data. In unsupervised learning [90] the outputs of the data are not known. The weights and biases of the network get updated from the update rule which involves a learning rate. There are different update rules, Adam and RMSprop have shown better results.

Autoencoders [88], Convolutional neural networks [40, 70], Recurrent neural networks [152], Long short term memory [114], Generative adversarial networks [33, 127, 158] are some of the examples of neural networks. Each of these has its own applications. Autoencoders are used for efficient data representations. Convolutional neural networks are widely used for image classification. Recurrent neural networks are used for algorithms for sequential data.

Applying deep learning models to the medical image processing improves the accuracy, time saving and automate the results which were manually done by the physicians and doctors [12]. Now a days Robot Assisted Surgery (RAS) [67] can be done, unlike traditional surgery, a RAS system includes a camera arm and few other mechanical arms with surgical instruments attached. Implementing an Inception network to detect and localize pulmonary nodules and predicting whether it is malignant or benign is an example of implementing a deep learning in the medical field [162]. In the medical field, we need much higher accuracies because it depends on the life of people.

Feature selection plays an important role in any algorithm or model to get higher accuracies. In traditional methods like SIFT [47] and HOG algorithms, manually selecting the features is more complicated and the obtained results are unstable [12, 58]. Even in machine learning algorithms, we will select the features and the accuracies will greatly depend on the selected features. These handcrafted features have limited representation capability and didn't get better accuracies and had high false-positive rates [174]. Deep learning has an advantage over traditional algorithms because it selects features based on their characteristics. Models that use deep learning automatically select features and are driven to select good features to reduce loss functions on their own. A deep learning model's accuracy can also be improved by conducting occlusion experiments.

Despite the fact that many researchers have worked on exploring medical imaging and image processing methods using deep learning, the main objective of this paper is to examine and study extensive models/methods for medical imaging. With COVID19, this work explores almost every type of disease in medical diagnosis. Most of the work we are highlighting is related to medical imaging and image processing.

The main objective of this paper is to explore and study extensive models/methods for medical imaging, even though many researchers were worked on exploring medical imaging and image processing method using Deep learning but novelty of this work to explored almost every type of diseases in medical diagnosis with recent work on COVID19. We are highlighting mostly recent work done on Medical imaging and image processing.

Detailed literature reviews of deep learning architectures for medical imaging are presented in this survey paper. As a result of this survey paper, the following important contributions can be made:

- The purpose of this paper is to explore the different deep learning architectures for different domains of medical imaging, which will help us understand which deep learning architectures are best for each type of image.
- Discusses the different types of medical imaging datasets.
- This paper discusses the different evaluation metrics used in medical imaging classification.
- Provides a conclusion and future directions in the field of medical image processing using deep learning.

This is the outline of the survey paper. In Section 2, medical image analysis is discussed in terms of its applications. Section 3 discusses the search criteria for the survey. Detailed descriptions of deep learning models for medical image processing are provided in Section 4. In Section 5, we discuss different datasets that can be used to classify medical images. Metrics for evaluating performance are discussed in Section 6. Conclusions and future trends are discussed in Section 7.

2 Applications of medical image analysis

A medical image analysis can be used to detect disease, locate a tumor, calculate the size or volume of various body parts.

Chest x-ray images can be used for pneumonia [137], lung cancer [130], emphysema, heart problems [43], Multiple fundus images detection [142] and covid detection [51]. X-ray images can also be used to find the fractures in various body parts. Pneumonia is an infection in that air sacs are filled with fluid or air and have difficulty in breathing. There are some white patches in the chest x-ray of a person affected with pneumonia. Pulmonary nodules are a kind of spot in the lungs. They can be malignant (cancerous) or benign (non-cancerous). Pulmonary nodules appear as white spots on x-ray images. Pulmonary nodules can also be detected from low-dose computed tomography. Covid-19 detection also happens through x-ray images. There are several deep learning models to detect covid -19 and achieve higher accuracies [139]. The x-ray image of the brain helps for early detection of Alzheimer's disease. A head MRI scan can be used for tumors, swelling, blood vessel issues, Parkinson's disease, and Alzheimer's disease. The MRI images of the heart are helpful in calculating the volume of the human left ventricle. The volume of the left ventricle is important for diagnosing cardiac diseases. Ultrasound images help to visualize body structures like tendons, muscles, liver, gall bladder. Various deep learning approaches have been proposed for breast cancer detection using ultrasound images [76]. CT scans can detect complex bone fractures, tumors. They also show internal injuries, bleeding. OCT imaging is used for ophthalmology [4] (retina and tissue), vulnerable cardiovascular plaques, skin lesions. CT scans can be used for the detection of hemorrhage (flow of blood from a blood vessel). CT scans are also used for covid 19 pattern identification [2, 31]. Microscopic images can be used for the detection of white blood cell cancer. Finger vein is the emerging hand based bio metrics. [119] describes highly accurate finger vein detection using depth wise separable CNN (Table 1).

3 Search criteria for survey

For conducting this survey, we have explored the following databases:

- Google Scholar
- IEEE Xplore Digital Library
- Science Direct
- Wiley Online Library
- ACM Digital Library

The selection of papers was done using the search criteria as Medical Imaging, Medical imaging using deep learning, Medical Imaging, and Classification, and Deep learning or convolutional neural network or CNN or neural network for Medical Image, and Transfer Learning. We reviewed over 240 recent research papers and shortlisted 133 research papers that focus on Medical Image classification using Deep Learning techniques. We have explored and studied research papers published within the years 1998 to 2022. The distribution of articles for Deep Learning in Medical Imaging is as shown in the Table 2.

Figure 1 shows the graph for distribution of paper over the years.

Table 1 Applications of Medical Image Analysis

Sr.No	Applications	Research paper
1.	Covid-19 screening from x-ray images	[36, 112, 164, 170, 175]
2.	Heartbeat classification from 2-lead electrocardiogram	[83]
3.	Brain tumor classification using Brain MRI images	[6, 9, 92, 113, 141, 160]
4.	Classification of vulnerable cardiovascular plaques from OCT images	[63, 155]
5.	Breast lesion classification in Ultrasound images	[179]
6.	Classification of pulmonary lung nodules	[115, 116, 143, 162]
7.	Dendrobium classification	[163]
8.	Classification of blood cell WBC	[64, 86]
9.	Detection of White Blood Cancer from Bone Marrow Microscopic Images	[80]
10.	Classification of Liver Tumor from MRI	[149]
11.	Liver fibrosis classification from Ultrasound images	[99, 128]
12.	Classification of Cervical Cells	[66, 176, 180]
13.	Detection of Alzheimer's disease from MRI	[73, 117]
14.	Classification of Common Thorax Diseases from chest x-ray images	[7]
15.	Segmentation and tracking of heart left ventricle from Ultrasound images	[58, 109, 110]
16.	Multi-Lesion Classification in Chest X-Ray Images	[131, 172]
17.	Lung Nodule Classification -Benign-Malignant	[126, 143, 169]
18.	Breast Cancer Detection	[28, 29, 35, 42, 140, 166]

4 Deep learning models/methods for medical image classification

Medical Image classification is the task of identifying the class or category of the image. It is one of the critical use cases in digital image analysis performed using deep neural networks [75]. The class of the image is predicted using the neural network that extracts the features within the image. The convolutional neural network forms the base for image classification. Also, various transfer learning models using CNN [157] are developed for classification. Unsupervised techniques like GAN [184] are used for data generation and classification. The Fig. 2 below shows some of the methods for Medical Image classification covered within this survey.

4.1 Convolutional Neural Network (CNN)

Convolution Neural Network (CNN) is a class of deep neural networks. The CNN has its applications in image processing, image segmentation, video processing and recognition, image classification, medical image analysis, natural language processing, recommender systems, etc.

It is a multi-layered neural network with one or more hidden layers [50]. The basic structure of a CNN model is as shown in the Fig. 3. The CNN makes use of convolution operation in minimum one of its layers. It requires a fixed-size input image for processing. The layers

Table 2 Count of No. of reviewed Article

S. No.	Articles in journals/book chapters	Count
1	IEEE Aaccess	38
2	arXiv	22
3	Multimedia Tools and Applications	7
4	IEEE Journal of Biomedical and Health Informatics	7
5	IEEE Transactions on Medical imaging	6
6	Knowledge-Based Systems	5
7	Neurocomputing	5
8	IET Image Processing	4
9	Neural Networks	3
10	Applied Intelligence springer	4
11	Artificial Intelligence Review	3
12	Information Fusion	3
13	Expert Systems with Applications	6
14	Pattern Recognition	3
15	IEEE Sensors Letters	2
16	Big Data	1
17	IET Computer Vision	1
18	Academic Press	1
19	Machine Learning with Applications	1
20	Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning Workshop	1
21	Autoencoders	1
22	Optik	1
23	IEEE Transactions on Cybernetics	1
24	IEEE Transactions on Pattern Analysis and Machine Intelligence	1
25	IEEE Transactions on Image Processing	5
26	Heart Dieses Data Set	1
27	Tissue and Cell	1
28	Procedia Computer Science	2
29	Proceedings of the 25th International Conference on Neural Informa- tion Processing Systems	1
30	IEEE Signal Processing Letters	1
31	Proceedings of the IEEE	1
32	Journal of Electrocardiology	1
33	Complexity 2021	1
34	SLAS Technology	1
35	Information Fusion	1
36	Physica D: Nonlinear Phenomena	1
37	Journal of King Saud University - Computer and Information Sciences	1
38	IRBM	1
39	IEEE Photonics Journal	1
40	Sensors and Actuators	2

Table 2 (continued)

S. No.	Articles in journals/book chapters	Count
41	Journal of Ambient Intelligence and Humanized Computing	1
42	Neural Computing and Applications	1
43	Applied Sciences	1
44	Expert Systems	1
45	Electronics	1
46	Journal of Healthcare Engineering	2
47	International Journal of Healthcare Information Systems and Informatics	1
58	International Conference on Advances in Information Communication Technology, Computing	1
	Total	158

included in this network are the input layer, convolution layer, pooling layer, fully connected layer, and drop out layer [13]. The functions of each layer are given as :

- The convolutional layer consists of attributes like input channels, output channels, user-defined filter or kernel size, and additional parameters like padding, strides, etc.
- The pixel values of the image form the input to the convolutional layer. It produces output based on the pixel values and the filter values by performing convolution operation pixel by pixel [69].
- The pooling layer also has a user-defined filter size. It performs dimensionality reduction of the input data. There are two types of pooling operations Max pooling and Average pooling.
- The fully connected layer or dense layer connects every neuron in one layer to other neurons in the next layer through the activation function. This layer is a densely connected layer which means that each neuron in the dense layer receives input from all

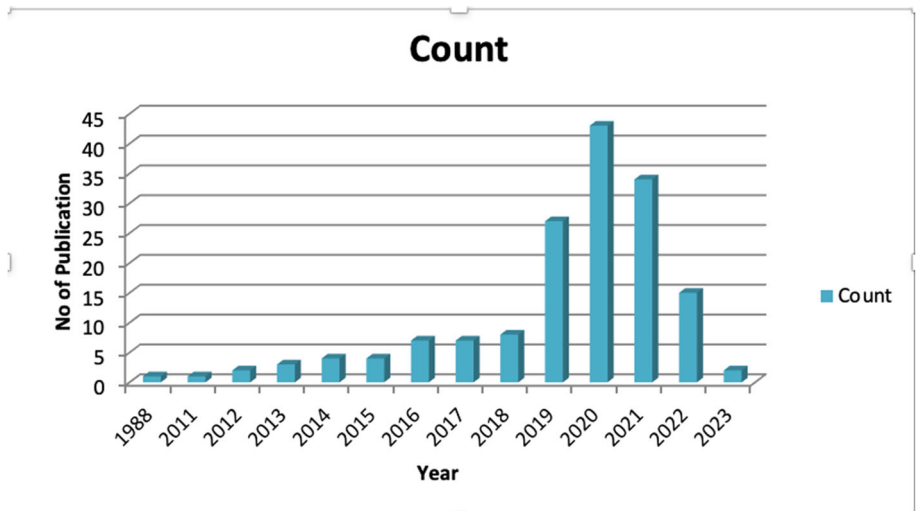


Fig. 1 Year-wise distribution of selected research papers

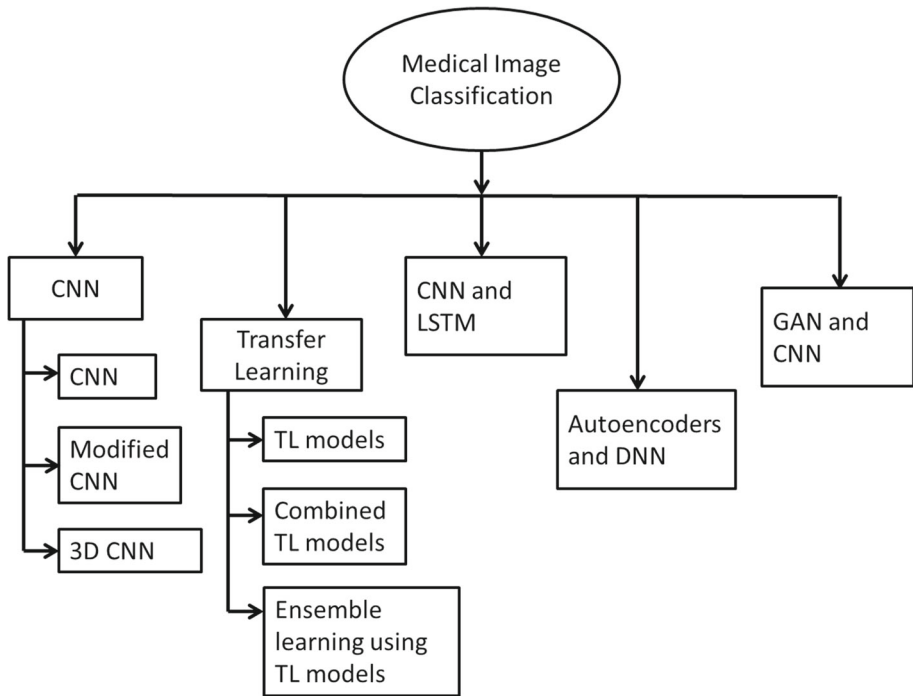


Fig. 2 Medical Image classification techniques

neurons from the previous layer. It also helps in changing the dimensionality of the vector.

- Thus in image classification, an appropriate activation function is applied to the dense layer to produce output or the category of image classified.
- This neural network is trained using backpropagation for weight and bias updation based on the loss function.

4.1.1 Single CNN

The CNN architecture has its applications in medical image classification. A Multi-class Image-based Cell Classification for types THP1, MCF7, MB231, PBMC is proposed [100]. It is an 8 layered CNN model. The input cell images are preprocessed and the images having noise, blur are discarded. It can be seen that the model accuracy is better for the CNN model with 0.995 than machine learning models like SVM having 0.977 and kNN 0.808. A deep learning model using CNN is used to classify the brain tumor types using brain MRI [141]. The tumor is classified into four types menin, gioma, glioma, and pituitary tumor. Using another dataset the glioma is classified into three grades (Grade II, Grade III, and Grade IV). The model proposed consists of a 16 layered architecture having combinations of input, convolution, max pooling, dropout, relu, fully connected, normalization, and dense layers. The dropout layer is used for reducing the overfitting and the activation function used within hidden layers is Relu. A softmax layer is used at the output layer for the

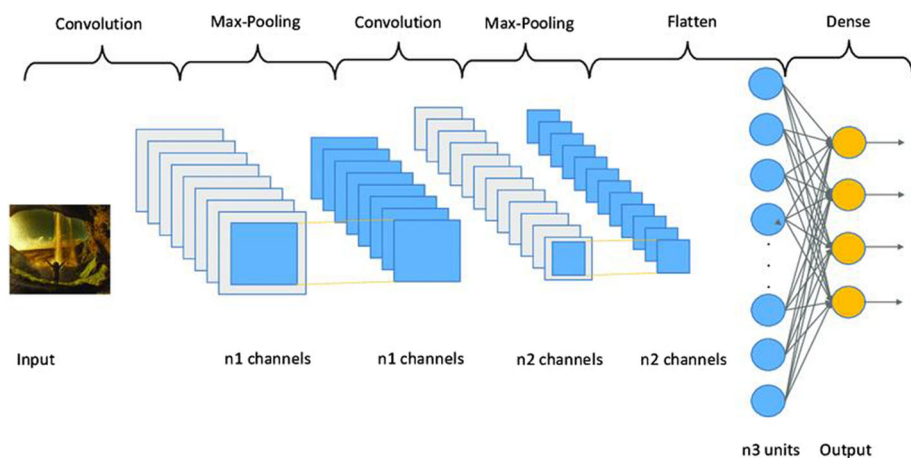


Fig. 3 CNN architecture for image classification [46]

prediction of the tumor type. The accuracy obtained is 97% and 100% respectively for both datasets. A 5 layered CNN is used for Covid-19 detection with data augmentation, it obtained accuracy of 92.6% [132]. Similarly a 6 layered CNN using data augmentation is used for detecting the Covid-19 disease using chest xray image dataset [123]. It obtained accuracy 99.2%. A fully connected 3 layer CNN with 20 CT scan cases to develop diagnosis is studied in [91]. It obtained a accuracy of 98% for classification task and dice score of 98% and 91% for lung and infection segmentation tasks. A modified U net model is used for covid-19 lung infection segmentation [134] with Medseg1 dataset it obtained dice score of 92.46%.

Adopt proper optimization function Various different types of optimizers can be used for the image classification problem. Studies are proposed for comparing the optimizer performance on the problem. One of these studies makes use of CNN architecture used for classifying the vulnerable cardiovascular plaques using OCT images [155]. The learning rate is set to 0.001 and various optimization functions like adam, SGD, RMSProp, Adadelata, Adagrad, Nadam, Maxadam are used for comparing their performance based on accuracy and overfitting ratio. It has been found that SGD has the highest accuracy 96.65%. Similarly, optimization function comparison is made for Breast lesion classification in ultrasound images by [179]. The classified types are benign cases and malignant. The model consisted of 4 convolutional layers and the activation function used is relu. For avoiding overfitting regularization, batch normalization and dropout are performed. Again it has been observed that SGD optimization has the highest AUC as compared to other functions. The RMSProp optimization function has some degree of overfitting due to the exponential nature of the function.

Combine image preprocessing and CNN The structure of the medical images is different than that of the natural images. Medical images consist of lot of quantitative data, intensity variation, scale variation, location of interest, etc. Thus medical image processing, before applying it to the input of neural network helps to improve the performance. The preprocessing of the medical images like histogram equalization, segmentation, denoising and

image augmentation are performed. A study for tumor stage classification of pulmonary lung nodules using deep convolutional neural networks is performed [143]. It is a 8-layered architecture. The nodule regions are extracted and segmented according to their malignancy levels by radiologists. These segmented images form input to the CNN and are classified as Benign, Malignant, and Non-Cancerous with an accuracy of 96%. Thus image preprocessing helps to improve the performance of CNN. Image preprocessing steps like denoising, thresholding, and morphological dilation are performed and these processed images form input to CNN model for classification of retinal OCT images [125].

Reduce model training time The time required to train a model is quite large. The training time required for CNN is large depending on the training parameters. To reduce the computational complexity of the DCNN, a method is used in which the weight updation of the convolution layers and the fully connected layers is modified for Abnormal Brain image classification [55]. The weight estimation is independent of the number of iterations and the model can be trained in a single iteration by setting the value of entropy to 0.01 and calculating the weight values in a single pass using the mathematical equations. The drawback of this method is it is trained for a very small dataset. To reduce the training time of the model a new method called Fast Convolutional Neural Network was developed for detection of hemorrhage [153]. It uses the technique of selective sampling where the higher weight is assigned to those pixels where the error obtained is high. The CNN model is trained and the score for the image is obtained using the pixel score calculated using the Gaussian filter. Thus the iterations required for training reduced from 170 to 60.

4.1.2 Variants of CNN

Modifications in the kernel sizes, weight initialization, and updation are performed to make the model better in terms of accuracy. The changes in kernel size were made in method optimized DCNN was proposed for dendrobium classification [163]. Instead of traditional CNN that uses 2D filters, here, 1 D filters are used for convolution and pooling layers. The convolution and pooling layer have a kernel of size 2×1 . There are 5 convolution layers and 3 pooling layers with a single fully connected layer and a softmax layer of 10 predictions. The accuracy obtained using optimized DCNN improved to 87% whereas DCNN has the accuracy of 81%. Changes in the activation function and adding a normalization layer also help to perform better and reduce overfitting. Multiple combinations of convolutional layers, activation layer, and batch normalization layers are performed for Brain image recognition. It is observed that using PRelu improved the accuracy than that of the Relu function. Also, a combination of CNN followed by Batch Normalization and then PRelu helps to improve the model performance. A 30 layered CNN with Batch Normalization is developed for Breast Cancer detection [148]. The preprocessing steps such as data augmentation, lesion detection using a Single Shot MultiBox Detector for object detection are performed. The accuracy for this model achieved is 89.47% for the classification of normal, benign cases and malignant cases using data provided by Mammographic Image Analysis Society (MIAS).

Fuse CNN kernels with other types of mathematical functions For the selection of the quantitative information and the variation in the frequencies or intensities of pixels, the kernel needs to be modified. For this, the kernel weights are combined with other functions for feature selection. This is performed in another approach proposed for the classification of blood cell WBC by combining the Gabor wavelet and CNN kernels that help in getting modulated kernels at different frequencies and orientation [64]. The images used are

hyperspectral images. The Gabor wavelet is a function of the product of Gaussian function and sinusoidal plane wave. The convolutional kernel and the Gabor wavelet kernel are multiplied to form a modulated kernel and the images are classified using this modified CNN structure as Basophil, Eosinophi, Lymphocyte, Monocyte, and Neutrophil.

Appropriate feature selection for model training is also an important task. In one method, the input features are selected based on a Chi squared test using CNN for Detection of White Blood Cancer from Bone Marrow Microscopic Images [80]. A mechanism for the classification of Acute Lymphoblastic Leukemia (ALL) and Multiple Myeloma (MM) is proposed that makes use of a 9 layered CNN model with Adam optimizer. If the value of Chi-square is high, then the feature is highly dependent on the output, hence it is selected for model training. Thus using feature selection the trained model obtained an accuracy of 97.25%. The limitation of this method is that it is database-oriented. It cannot perform that effectively for other data or image types.

4.1.3 Multi-CNN

For selecting the high-level features and low-level features within an image, different convolutional layers output is combined in the final feature vector. This combination of different convolutional neural networks layers and combining their results is used for the detection of Covid 19 that uses Multi-scale learning for classification [60]. The Fig. 4 shows the architecture of this approach. The structure consists of a total of 5 convolutional layers. The output of the last three convolutional layers is convolved separately using 1×1 filters and performed Global Max pooling. The score vectors of these layers are aggregated and applied to softmax activation for the prediction of the image for a positive or negative report.

Similarly multiple convolutional neural networks and their results are combined for the uncertainty of a model for detection of TB that is proposed by [5]. The developed model

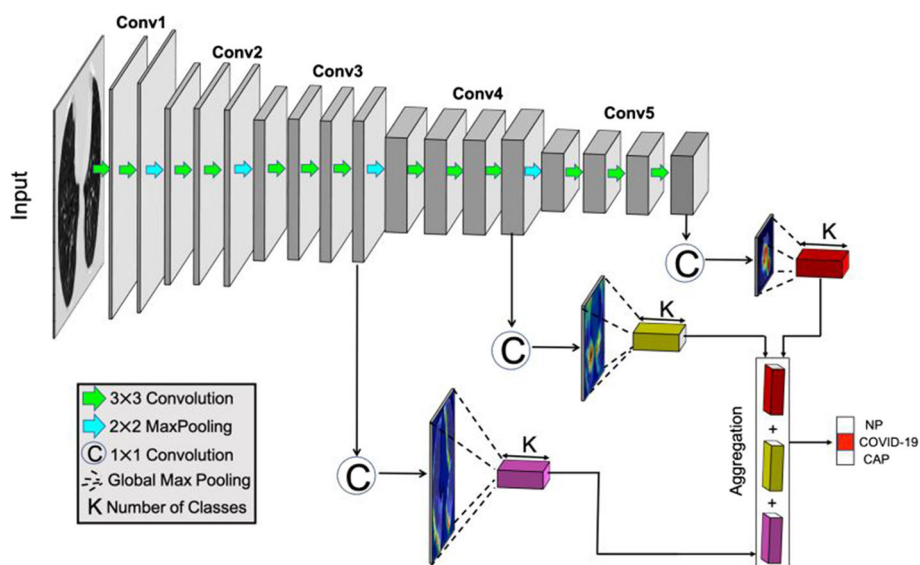


Fig. 4 Multi scale learning Network Architecture [60]

consists of multiple CNN layers. The output of every layer is fed to the Bayesian network and a dropout operation is performed. The magnitude of variation of probabilities defines that the model is predicting with high confidence or low confidence. If the magnitude is large it indicates higher uncertainty in prediction and low means that the model is confident while prediction. Another model proposed uses multiple CNN for Glaucoma detection [24]. It uses 2 CNN networks and the output is concatenated using fully connected layers and then the sigmoid activation function is used for binary classification and obtained an accuracy of 91%.

One model is proposed for abnormal kidney detection using multi fusion CNN using ultrasound images [167]. It makes use of multiple CNN for feature extraction. One CNN is used for extraction of ROI, second CNN for the image feature extraction. Both outputs of these CNN are fused together using convolutional layers and relu layers and provided as an input to another CNN network using transfer learning model DenseNet169 that classifies the abnormal images. This method obtained an accuracy of 99%. This model requires larger training data. For Covid 19 detection, a similar kind of CNN fusion-based model is proposed using ultrasound lung images [106]. It makes use of CNN blocks and the output of every block is concatenated to dense layers and then a softmax layer for prediction. It classifies for covid, pneumonia, and healthy lung images and obtained an accuracy of 92.5%. Implementation of multiple such CNN models is complex.

4.1.4 3D-CNN

All the above-discussed methods were for the 2D convolutional neural network. A medical image may be in forms such as 3D volume and 4D volume that change over time. For analysis of 3D images, 3D CNN is required. A 3D CNN [61] uses a kernel having 3 dimensions. Thus it learns more hierarchical data representations and can be applied to videos and volumetric data. A 3D CNN uses a kernel having 3 dimensions. Thus it learns more hierarchical data representations and can be applied to videos and volumetric data. A 3D convolutional neural network is proposed for classification for MRI Liver Tumor by [149]. This proposed structure consists of four consecutive 3D convolutional layers with kernel size $3 \times 3 \times 3$ followed by the Relu activation function layer and a fully connected layer. Softmax layer is used for the binary classification of the output as primary and metastatic and the input images are 3 dimensional. These 3D CNN network does not require any image preprocessing steps. The accuracy obtained is 83% and that of a 2D network is 62%. Thus 3D CNN helps to improve model performance. In a similar way, 3D CNN is used for 3D brain MRI image classification [160]. The experiment was performed on two datasets, one for detection of schizophrenia and other for classification of developing children (TDC) vs. children with ADHD. This model used a total of 6 convolutional layers and Max pooling and fully connected layers. The kernels of all the layers were 3D. From the results, it can be observed that the performance of the model is better than the transfer learning models.

Thus CNN is used for Medical Image Classification. The training time required for the model is large. The images are preprocessed for improving model performance. Multiple variations are made in the CNN model like changing kernel sizes, kernel, and other function operations, adding more filters are made depending on the application and input images. The CNN models proposed are dependent on the input data. So the performance will vary depending on the dataset considered.

4.2 Transfer learning models

Various transfer learning models are developed for image classification using convolutional neural networks. Transfer learning is using the model trained for one problem for another similar problem. It allows the use of these pre-trained models for feature extraction, fine-tuning, and integrating it for the new models. ImageNet is one of the popular computer vision completion that tasks the researchers to develop a model that accurately classifies the images in the ImageNet dataset which contains images belonging to 1000 categories. Thus many image classification models were developed. The Table 3 shows a few of the transfer models developed for image classification and the year released.

Single Transfer learning model The comparison of various transfer learning models has been tested based on their performance. The detection of Covid has been compared using different types of transfer learning models using the x-ray images [112]. The different models used are VGG16, DenseNet121, Xception, NASNet, and EfficientNet. It is observed that the efficiency of EfficientNet is highest followed by DenseNet121 and VGG16 with the lowest accuracy. Similarly for skin disease classification, it is observed that Xception Net has the highest accuracy. Another study for TB detection shows that without image segmentation ChexNet model has the highest accuracy and with segmentation DenseNet 121 has high accuracy [120]. This paper [79] describes feature extraction both from deep learning features and traditional hand crafted features.

The Table 4 below shows some of the methods used for the classification using transfer learning models.

Combining multiple transfer learning models The feature vectors of multiple transfer learning models are combined and concatenated to extract the vectors using a fully connected layer. Then a softmax or any other activation function is used for the prediction of

Table 3 Transfer Learning Models and year released

Sr No	Model	Year	Reference
1	LeNet	1998	[82]
2	AlexNet	2012	[77]
3	ZFNet	2013	[178]
4	VGG 16	2014	[136]
5	VGG 19	2014	[136]
6	GoogleNet/Inception V1	2014	[144]
7	ResNet 50	2015	[53]
8	Wide ResNet	2016	[177]
9	Inception V3	2015	[145]
10	SqueezeNet	2016	[59]
11	Xception V1	2016	[27]
12	Xception V3	2017	[27]
13	Inception V4	2016	[146]
14	MobileNet	2016	[57]
15	DenseNet264	2018	[62]

Table 4 A summary of implementation using Transfer Learning Models

S. No.	Title	Reference	Application	Model Used	Dataset	Accuracy
1.	TLCoV: An automated Covid-19 screening model using Transfer Learning from chest X-ray images	[36]	Covid detection	VGG 16	Covid-19 radiography	97.67%
2.	Heartbeat classification using Deep Residual Convolutional Neural Network from 2-lead electrocardiogram	[83]	Heartbeat classification	Modified ResNet-31	MIT-BIH arrhythmia	99.38%
3.	COVID-19 Detection System Using Chest CT Images and MultipleKernels-Extreme Learning Machine Based on Deep Neural Network	[151]	Covid detection	Multiple Kernels-121 ELM-based DCNN using DenseNet	COVID-CT dataset	98.36%
4.	Pulmonary Image Classification Based on Inception-v3 Transfer Learning Model	[161]	Pulmonary disease	InceptionNet v3	standard public CT digital image database JSRT	94.71%
5.	On OCT Image Classification via Deep Learning	[159]	Diabetic macular edema (DME) and age-related macular degeneration (AMD) detection	CliqueNet	Spectralis SD OCT imaging ,Spectralis SD-OCT imaging	99%
6.	Deep Convolutional Approaches for the Analysis of COVID-19 Using Chest X-Ray Images From Portable Devices	[37]	Covid 19 detection	DenseNet 161	CHUAC-X-ray images:normal/pathological,COVID-19	90%

Table 4 (continued)

S. No.	Title	Reference	Application	Model Used	Dataset	Accuracy
7.	Liver fibrosis classification based on transfer learning and FCNet for ultrasound images	[99]	liver fibrosis classification	VGGNet and FCNet	Raw liver fibrosis ultrasound image	96.06%
8.	Automatic Classification of Cervical Cells Using Deep Learning Method	[176]	Cervical cancer detection	CNN + Inception + SPP layer	Papanicolaou dataset,Herlev dataset	98.28%
9.	Transfer Learning With Intelligent Training Data Selection for Prediction of Alzheimer's Disease	[73]	Alzheimer's Disease detection	VGG 19 - Training and freezing blocks of CNN layer	Alzheimer's Disease Neuroimaging Initiative (ADNI)	99.20%
10.	Transfer Learning-Based DCE-MRI Method for Identifying Differentiation Between Benign and Malignant Breast Tumors	[185]	Breast Tumors- Benign and Malignant	InceptionNet v3 , VGG 19	patients underwent imaging with a 1.5 T scanner	InceptionNet - 0.90 VGG - 0.91
11.	A Novel Approach for Multi-Label Chest X-Ray Classification of Common Thorax Diseases	[7]	Common Thorax Diseases	DenseNet 121	ChestX-ray14 and ChestXpert	ChestX-ray14- 0.87,CheXpert -0.812
12.	DeepPap: Deep Convolutional Networks for Cervical Cell Classification	[180]	Cervical Cell Classification	ConvNet	Herlev and HEMLBC	98.3%
13	Deep Transfer Learning for Brain Magnetic Resonance Image Multi-class Classification	[20]	brain tumor multi-classification	ResNet50	NINS,Harvard Brain Atlas , Biomedical School Brain MRI	97.05%

Table 4 (continued)

S. No.	Title	Reference	Application	Model Used	Dataset	Accuracy
14	Classification of COVID-19 in CT Scans using Multi-Source Transfer Learning	[95]	COVID-19 detection	DenseNet169, ResNet101V2	COVID-19 CT image - UCSD	DenseNet169-90.4%, ResNet101V2-89.3%
15	Predictive Analysis of Diabetic Retinopathy with Transfer Learning	[81]	Diabetes Mellitus Retinopathy detection	ResNet50 V2, VGG16, EfficientNet diabetic retinopathy detection dataset	ResNet50 V2-89%, VGG16-95%, EfficientNet-96%	
16	Optimal Transfer Learning Model for Binary Classification of Fundus Images Through Simple Heuristics	[71]	diabetic retinopathy detection	EfficientNet	EYEPACS diabetic retinopathy dataset, ORIGA-650 dataset	90%
17	MANet: A two-stage deep learning method for classification of COVID-19 from Chest X-ray images	[170]	chest COVID-19, TB, BP and VP	Preprocessing-segmentation ResUNet, ResNet50 model	CXR data- Montgomery County, Germany, GitHub	97%
18	Neural networks model based on an auto-mated multi-scale method for mammogram classification	[168]	Breast Cancer detection	Uses 3 feature map for local lesion classification- DenseNet121/MobileNet	INbreast dataset	DenseNet- 94.51%, MobileNet - 92.68%
19	Transfer learning for establishment of recognition of COVID-19 on CT imaging using small-sized training datasets	[84]	COVID-19 detection	CheXNet 121 layered	COVID-CT-Dataset	87%
20	Deep Transfer Convolutional Neural Network and Extreme Learning Machine for lung nodule diagnosis on CT images	[65]	lung nodules benign/malignant	Xception ELM	LIDC-IDRI dataset	92.39%

Table 4 (continued)

S. No.	Title	Reference	Application	Model Used	Dataset	Accuracy
21	Transfer Learning to Detect COVID-19 Automatically from X-Ray Images Using Convolutional Neural Networks	[27]	Covid-19 detection	VGG16, MobileNet	Kaggle COVID-19 Radiography Database	VGG16 - 98.71% MobileNet- 98.39%
22	COVID-19 detection in X-ray images using convolutional neural networks	[11]	Covid-19 detection	VGG19 with image pre-processing and lung segmentation	BIMCV-COVID19+, BIMCV-COVID, BIMCV-Padchest, Spain Pre-COVID era, Montgomery, JSTR, NIH dataset, Data Collection by Cohen and by Daniel Kermany	99.06%
23	Malignancy Detection in Lung and Colon Histopathology Images Using Transfer Learning With Class Selective Image Processing	[97]	lung and colon cancers detection	AlexNet with histogram equalization	LC25000 dataset of lung and colon tissues	98.4%
24	Skin Lesion Classification by Multi-View Filtered Transfer Learning	[16]	Skin Lesion Classification- Malignant Melanoma (MM), Seborrheic Keratosis (SK), and Benign Nevi (BN)	ResNet-50	ISIC2017, HAM10000 dataset	91.8%
25	Automatic Thyroid Ultrasound Image Classification Using Feature Fusion Network	[182]	thyroid nodules - benign and malignant	Image processing, ROI extraction, ResNet-50	Data from Department of Ultrasound of Chinese PLA General Hospital	88.3%

Table 4 (continued)

S. No.	Title	Reference	Application	Model Used	Dataset	Accuracy
26	Multi-Class Skin Lesion Detection and Classification via Tele dermatology	[74]	Skin lesion - benign and melanoma	CNN for lesion segmentation, used DenseNet for lesion classification, ELM classifier for assigning labels	ISBI2016, ISIC2017, PH2, and ISBI2018	ISBI2016-92.28%, ISIC2017-91.52%, PH2-96.6%, and ISBI2018-92.25%
27	Deep Learning Based Staging of Bone Lesions From Computed Tomography Scans	[96]	prostate cancer - benign or malignant	2D ResNeXt-50, 3D ResNet-18, 3D ResNet-50, ensemble of 2D and 3D ResNet model	collected 2,880 annotated bone lesions from CT scans of 114 prostate cancer patients	2D ResNeXt-50 - 86.4%, 3D ResNet-18 - 79.4%, 3D ResNet-50 -77.2%, ensemble of 2D and 3D ResNet model - 87.0%
28	Deep Transfer Learning Based Parkinson's Disease Detection Using Optimized Feature Selection	[3]	Early detection of Parkinson's disease (PD)	ResNet50 Inception-V3	NewHandPD	90%

a class of the image. The architecture of this approach is as shown in Fig. 5 used for the Fundus image Quality Assessment [121]. It is observed that 4 models are combined for the feature vector and passed to the fully connected layer. Thus the output is predicted using the softmax layer. Also, transfer learning models and Ensemble learning models are combined for the prediction of the class of the image.

The Table 5 below gives some of the combinations of models and ensemble learning methods used for the classification by various authors.

In transfer learning, the model is fine-tuned for the weight adjustment using backpropagation. These pre-trained models are easy to implement and require less time for training. The combination of ensemble learning and transfer learning models helps to get new characteristics from various models and improves the prediction. Implementing an Ensemble learning model using these models is a complex task and requires more computational time. The transfer learning models are trained for natural image classification. Hence while implementing it for medical images; it might not be able to capture all the significant features. While using the transfer learning models, it is required to perform the image processing to capture the required image textures and features for better results. Thus modifications in the pre-trained model like changing the activation function of a few layers, modifications in filter operation, and combining it with functions are required for improving the performance. In transfer learning models, as the number of layers increases the requirement of the training data also increases. Hence if a sufficient amount of data is not available, then it may affect the performance of the model.

4.3 CNN and RNN

Recurrent Neural Networks are also known as RNNs are a class of Artificial Neural networks. In RNN the output from the previous step is fed as an input to the current state [135]. RNN is able to process the sequential data. For producing output, RNN considers the output of the previous state and the input of the current state. These networks make use of the internal memory or state known as a hidden state for processing the input or sequences.

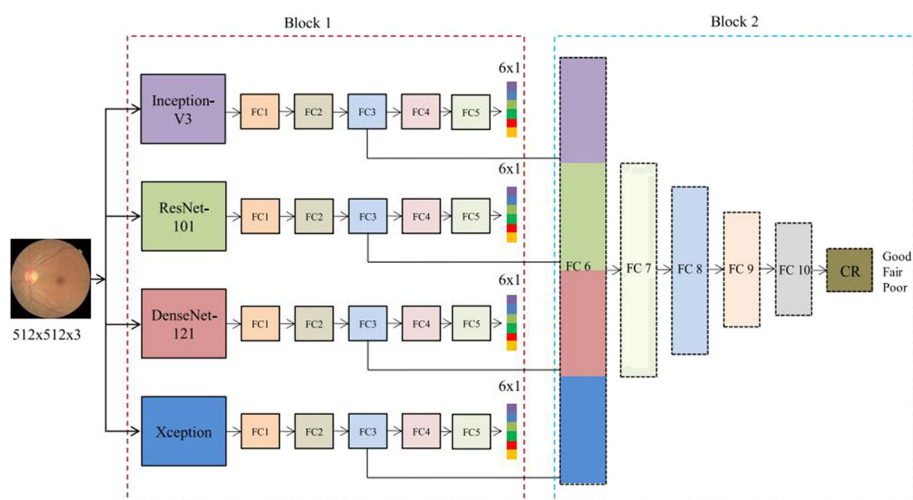


Fig. 5 Combined TL models for Fundus image QA [121]

Table 5 A summary of implementation using a combination of Transfer Learning Models

S. No.	Title	Reference	Application	Combination of Models	Dataset	Accuracy
1.	Multi-Label Classification of Fundus Images With EfficientNet	[21]	Fundus disease detection	Two EfficientNet models	ophthalmic data collected by Shangong Medical Technology Co. Ltd	89%
2.	A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network	[66]	cervical cancer diagnostic prediction	Resnet-50, Resnet-101 and Googlenet models	LBC dataset	97%
3.	Multivariate Regression-Based Convolutional Neural Network Model for Fundus Image Quality Assessment	[121]	fundus image quality assessment	Inception V3, DenseNet 121, ResNet 101 and Xception	kaggle - EyePACS	95.66%
4.	COVIDetection-Net: A tailored COVID-19 detection from chest radiography images using deep learning	[41]	Covid 19 detection	ShuffleNet and SqueezeNet and use SVM for classification	CRIs from the Github and kaggle	94.44%
5.	Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs	[122]	TB detection	Ensemble method - InceptionResNet-V2, Inception-V3, and DenseNet-121	Pediatric pneumonia dataset, RSNA dataset, Indiana dataset	94.10%

Table 5 (continued)

S. No.	Title	Reference	Application	Combination of Models	Dataset	Accuracy
6.	A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor	[113]	Brain Tumor detection	1. Concatenate Inception V3 model and classify using softmax 2. Concatenate DenseNet 201 model and classify using softmax	T1-weighted MR images-meningioma, glioma.and pituitary	DenseNet201-99.51% InceptionNet V3 -99.34%
7.	Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network	[111]	Breast cancer	4 CNN of Modified VGGNet - changing convolutional layers to 7 and using Relu and LeakyRelu activation function	CBIS-DDSM and mini-MIAS	malignant/ benign-77.66%, mass/calcification-92.2%,normal/abnormal-93.73%
8	DenResCov-19: A deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays	[93]	COVID-19, pneumonia, TB detection	Concatenate ResNet 50 and DenseNet 121 models	IEEE COVID-19 CXRs, Tuberculosis CXRs from Shenzhen Hospital, Pediatric CXRs	75.75%
9	Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT	[169]	Benign-Malignant Lung Nodule Classification	Extracting the OA, HVV and HS patches from x ray and ensemble learning for 3 ResNet 50 for class prediction	LIDC-IDRI database	92%
10	EMS-Net: Ensemble of Multiscale Convolutional Neural Networks for Classification of Breast Cancer Histology Images	[173]	Breast cancer detection-normal, benign, situ carcinoma and invasive carcinoma	Ensemble learning-Two DenseNet161 and ResNet152	BACH dataset	91.75%

Table 5 (continued)

S. No.	Title	Reference	Application	Combination of Models	Dataset	Accuracy
11	TRk-CNN: Transferable Ranking-CNN for image classification of glaucoma, glaucoma suspect, and normal eyes	[72]	Glaucoma detection	Image processing, augmentation, ROI extraction and ensemble for 6 DenseNet121 model	glaucoma image dataset-Korea University Medical Center	88.94%
12	An efficient deep Convolutional Neural Network based detection and classification of Acute Lymphoblastic	[34]	Acute Lymphoblastic Leukemia Blood cancer	MobileNetV2 and ResNet	ALLIDB1 and ALLIDB2	97.92% and 96.00%

The LSTM – Long Short Term Memory is an RNN architecture designed to address the vanishing gradient problem [49].

The Fig. 6 shows the structure of an RNN.

The CNN and RNN architectures are combined for image classification problems. The detection and classification of the types of blood cells make use of a combination of CNN and RNN [86]. The architecture is as shown in the Fig. 7. The BCCD dataset is used and the image preprocessing steps like rotation are applied. These images form input to the CNN and the BiLSTM layers. The obtained feature vectors are merged and given as input to the fully connected layer and then the softmax activation function is used for class prediction. Initially, the CNN layers are trained using the Xception transfer learning model by freezing the BiLSTM network. Then the LSTM layers are trained by freezing the CNN layers. Then the combined model is trained and the result is obtained by the combination of both these networks. The BiLSTM helps to consider the spatiotemporal characteristics of the information contained in the image. Thus it can learn the structural characteristics of blood cell images. Thus the Xception-LSTM model achieves better accuracy of 90.79% as compared to the Xception model of 88.70%. The drawback of this method is that the training time and the classification time required are high. Also, it is more data specifically for single-cell images, cannot classify images with overlapping nuclei and cell cluster images.

Another study also makes use of CNN and LSTM layers for Covid detection using X-rays [38]. The images are classified for covid positive, normal, and pneumonia. The images are preprocessed for gradient magnitude adjustment, performing watershed transformations, and resizing. These preprocessed images form input to the CNN layers. The feature vector produced by CNN is fed as an input to the LSTM layers. The output of these layers is given to the fully connected layers and then followed by Relu, dropout and softmax layers for output prediction. This model achieved performance accuracy upto 100%. Again there is a limitation of its data-specific approach. A study is proposed to detect the Covid-19 using the chest xray images making use of CNN and LSTM [105]. The images form the input to the 5 layered CNN model and its output is fed to the 3 layered LSTM network. The output of the LSTM is given to the 3 fully connected layers which classifies the image. This method obtained accuracy of 91%.

Thus LSTM and CNN layers can be combined for the image classification. LSTM helps to identify the structural dependent features, spatial information, positional attributes within the image that CNN may fail to identify. The implementation of the LSTM and CNN model

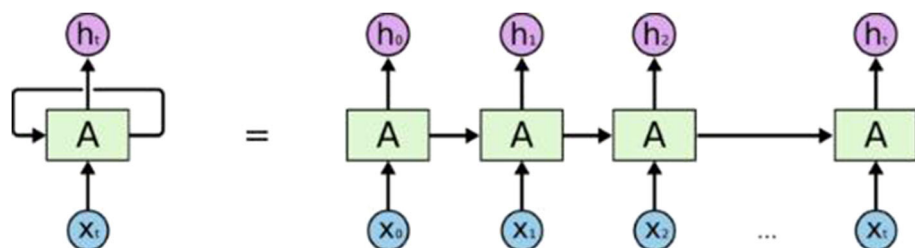


Fig. 6 Structure of RNN

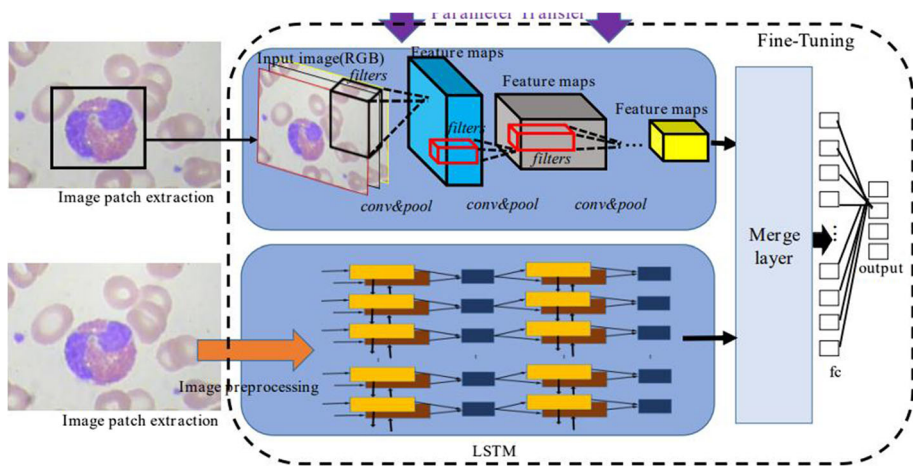


Fig. 7 Architecture proposed by combining CNN and LSTM [86]

is complex. These models require higher training time and are dependent on the input dataset.

4.4 DNN and auto encoders

Autoencoders is a neural network that falls under the category of unsupervised learning for the task of representation learning [14]. It learns to copy its input to the output. These networks are used for data compression. An autoencoder is composed of the encoder and a decoder. The encoder compresses the data and the decoder is used to reconstruct the original data using this compressed encoded data. It uses back propagation for generating the output value that is almost close to the input value [15]. Thus the encoder is able to extract the vital features that are required for the reconstruction of the input at the decoder. Hence after training the encoder model is saved as it performs feature extraction and data compression. The types of autoencoders are sparse autoencoders, denoising autoencoders, contractive autoencoder, concrete autoencoders, and variational autoencoders [25]. The Fig. 8 shows the basic structure of an autoencoder.

The autoencoders can be trained to reconstruct the original input, then the decoder part can be discarded and the encoder output can be used as a feature vector in supervised learning models [118]. For encoding an image, the autoencoder requires a series of convolutions, thus input image is encoded into a condensed vector. A Sparse autoencoder is a type of autoencoder that has only one hidden layer that is connected to the input vector with a weight matrix forming an encoding layer. The hidden layer outputs to the reconstruction vector along with the weight matrix that forms the decoder. This sparse autoencoder is used for the Nuclei Detection on Breast Cancer Histopathology Images [171]. It uses two sparse autoencoders forming stacked sparse autoencoder where the encoder network represents the pixel intensities and the decoder reconstructs the pixel intensities as low dimensional features. The SSAE is used for learning the high-level features from regions containing the nuclei. After the high-level feature learning, the encoder output is fed to the softmax layer that uses sigmoid activation for classifying it as patches containing the presence of nuclei or not. This method obtained F1 score of 84.49%. The training time required for this method is higher and it is specific to cell nuclei detection.

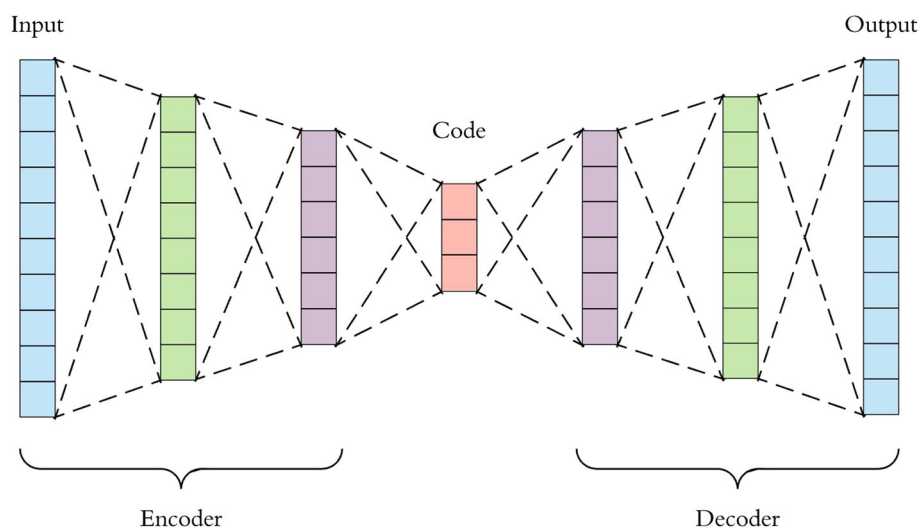


Fig. 8 Basic structure of an Autoencoder

A similar use of autoencoders combining it with wavelet transforms made for the Brain MRI Image Classification for Cancer Detection [92]. It uses a deep wavelet autoencoder-based deep neural network for detection. The autoencoder is trained for generating the output images as input images using the Discrete Wavelet Transform (DWT) operation for generating the high feature vectors. For this, the input images are encoded and processed by passing through the low pass filter and high pass filter and performing the operation of the Discrete Wavelet Transform (DWT) on it. The images are decoded by performing inverse wavelet transform. The autoencoder is trained for generating these feature vectors. These vectors form the input to the deep neural network using a sigmoid activation function that classifies the images for cancer detection. It is applicable for brain MRI images. The CNN and Autoencoder model is complex, it requires more training time. It is required to have a large input image dataset for this approach.

4.5 CNN and GAN

GAN is Generative Adaptive Network used for generative modeling using deep neural networks such as a convolutional neural network. These generative models create new data instances that resemble the training data. Thus GAN can create images that look like the training images [48]. The GAN model consists of a generator and a discriminator. The generator generates the fake data and the discriminator learns to distinguish the true data and false data. During the initial phase of training, the generator produces data and the discriminator is easily able to distinguish between the real and fake data. But as the generator performs well the discriminator is unable to distinguish between the real and fake data. Both these neural networks produce generator loss and discriminator loss and are trained using back propagation algorithm for updating weights. The training of GAN is stopped when the discriminator prediction probability reaches 0.5. The Fig. 9 shows the structure of a GAN model.

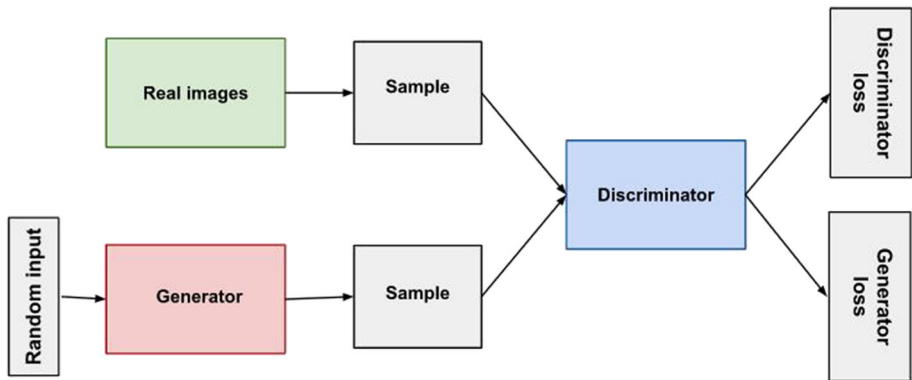


Fig. 9 Structure of GAN

GAN can be used for image classification for image augmentation and when the available training data is less. The neural network used in GAN for image classification is convolutional neural networks. The generator produces the images and the discriminator predicts the class of the image as well as distinguishes between the real and fake images. The discriminator model is modified by adding flatten and softmax layers to predict image class. The Fig. 10 shows the GAN model used for image classification of OCT retinal disease [33].

This similar kind of structure is used for detection and diagnosis of other infections or applications for medical image classification. The Table 6 below summarizes the different models for classification using GAN and CNN models.

Thus GAN using CNN can be used for image classification applications. The drawback of GAN and CNN is their computational complexity and the training time. Also, these

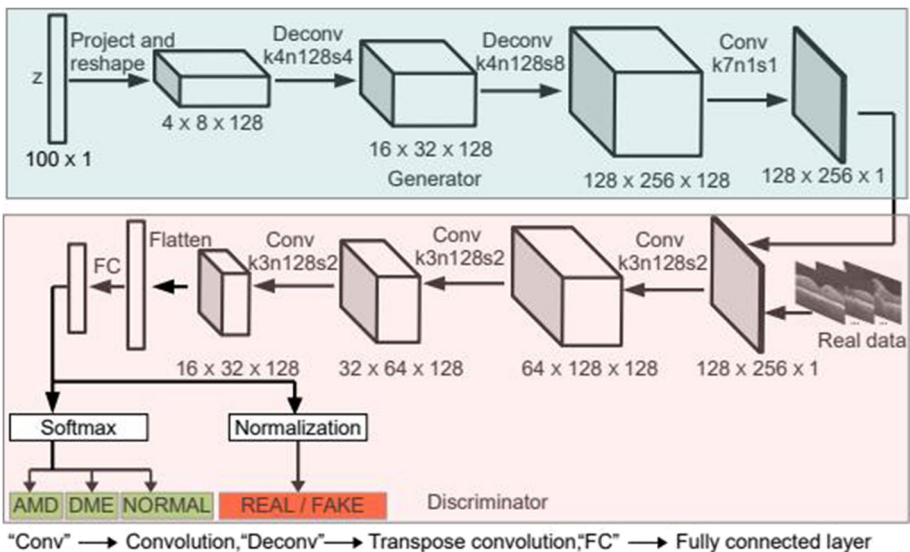


Fig. 10 Block diagram of the semi-supervised GAN for OCT classification [33]

Table 6 A summary of implementation using GAN and Convolutional Neural Network

S. No.	Title	Reference	Application	Combination of Models	Dataset	Accuracy
1.	DL-CRC: Deep Learning-Based Chest Radiograph Classification for COVID-19 Detection: A Novel Approach	[129]	COVID-19 Detection	GAN using Data Augmentation of Radiograph Images and CNN. Discriminator predicts image class	Covid Xray -Covid , Pneumonia and normal	95.25%
2.	A Data-Efficient Approach for Automated Classification of OCT Images using Generative Adversarial Network	[33]	Retinal OCT classification for- AMD,DME and normal	GAN network CNN. Generator generates images and discriminator predicts image class.	NEH OCT dataset	97.43%
3.	CovidGAN: Data Augmentation using Auxiliary Classifier GAN for Improved Covid-19 Detection	[154]	Covid-19 Detection	Modified Auxiliary Classifier GAN (ACGAN) - GAN generates images and Discriminator predicts the image class. Both networks are CNN.	IEEE Covid Chest X-ray dataset, COVID-19 Radiography Database, COVID-19 Chest X-ray Dataset Initiative	95%
4.	Synthesizing Chest X-Ray Pathology for Training Deep Convolutional Neural Networks	[131]	Chest X-Ray Pathology - Normal, Cardiomegaly, Pulmonary Edema, Pleuand Pneumothorax, Pleural Effusion	Deep convolutional GAN used for generating images and transfer learning models are used for classification. AlexNet, ResNet-50 and GoogLeNet achieved good accuracy.	adult posterior-anterior chest X-rays	AlexNet - 93.03% , GoogLeNet - 95.78% , ResNet - 96.14%
5.	Unsupervised Multi-discriminator Generative Adversarial Network for Lung Nodule Malignancy Classification	[78]	Lung Nodule Malignancy Classification-Benign, Uncertain, Malignant	Multi Discriminator GAN and encoder - GAN and encoder trained using multiple discriminators. Then encoder and GAN generates images and discriminator discriminates the images based on the anomaly score. Above threshold-malignant and below threshold - Benign	LIDC-IDRI dataset	95.32%

Table 6 (continued)

S. No.	Title	Reference	Application	Combination of Models	Dataset	Accuracy
6	Ultrasound Image Classification using ACGAN with Small Training Dataset	[128]	breast ultrasound image-benign /malignant	ACGAN using Data Augmentation and CNN. Discriminator predicts image class	Breast Ultrasound Image	98.80%
7	RANDGAN: Randomized Generative Adversarial Network for Detection of COVID-19 in Chest X-ray	[103]	Covid-19 Detection	GAN- Inception and Residual block layers. Discriminator - CNN for prediction	Covid-chestxray dataset	77%
8	GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification	[44]	Liver classification-cysts, metastases, hemangiomas	GAN and CNN. Generator generates images and discriminator predicts image class.	liver lesions- Sheba Medical center	85%
9	Retinal Optical Coherence Tomography Image Classification with Label Smoothing Generative Adversarial Network	[52]	OCT - AMD, CNV, DME and normal	Generator generates images and a CNN classifier used for image classification	UCSD, HUCM dataset	UCSD - 87.25% HUCM -68.36%
10	Dermoscopy Image Classification Based on StyleGAN and DenseNet201	[181]	OCT - AMD, CNV, DME and normal	Generator generates images and DenseNet121 used for image classification	ISIC2019 dataset	93.6%

methods are dependent on the database. For applying it for other image types, modifications in the CNN are required. As GAN generates the images, the correct labeling of data is very important as incorrect labeling would adversely affect the model performance.

5 Dataset for medical image classification

In recent research findings, medical imaging can be used to diagnose COVID-19, Breast Cancer, TB, Brain Tumor, Blood cell classification, and many more. Furthermore, some studies suggest that radiological imaging may be a better way to identify medical diseases. Convolutional Neural Network is used to classify computer tomography pictures of medical disease cases automatically. So when we working on any deep learning methods or algorithm we need a huge amount of data to train and test our model. Dataset is the collection of instances, where a single row of data is called an instance. In Medical imaging, we need different types of datasets for respective methods. In this section, we will discuss some publicly available datasets on different data providers like Kaggle, UCI Machine repository. For Medical Imaging Dataset could be in image format for multiple classes or it could be in a compiled dataset in CSV format. In the dataset, we have different labeled classes for each class we have multiple images available for the training model. Some dataset we explored in a study in journals and some others we found on the dataset repository is mentioned below. Some of Medical imaging dataset and description shown in Table 7.

The following types of images are used in medical imaging and their uses in the medical field.

- X-ray radiography [175]: It is used to identify bone fractures, pathological changes in lungs.
- Magnetic Resonance Imaging scanning [1]: It is used to produce images of the body and brain.
- Fluroscopy: It is used to produce images of various internal parts and structures of the body.
- Ultrasonography: It is used to produce images of abdominal organs, heart, muscles, arteries, and veins.
- Elastography: It is used to map the elastic properties and stiffness of the soft tissue.
- Thermography: It uses an infrared camera to detect heat patterns and blood flow in body tissues. It is also used to diagnose breast cancer.
- Tomography [115, 126]: It is used to produce images by sections or sectioning through the use of any kind of penetrating wave.

The following are some of the techniques that don't produce images. They produce data that are represented in graphs or maps.

- Electro Cardiography (ECG): It gives the detailed structure of the heart, including chamber size, heart function, the valves of the heart.
- Electroencephalography (EEG): It is a monitoring method to record electrical activity on the scalp to represent the activities of the surface layer of the brain.
- **COVID19 Radiography Dataset:** It [22] is a dataset of Chest X-ray of COVID-19 positive cases. Team of researchers from Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh along with their collaborators from Pakistan and Malaysia in collaboration with medical doctors. Dataset has four classes COVID-19, Normal and

Table 7 Medical Imaging dataset and their description

Sr No	Name of the data set	Description	Reference
1	Covid 19 Radiography dataset	Chest X-ray of covid-19 positive cases dataset has four classes	[22]
2	Wisconsin Breast Cancer Database	Breast cancer database with two classes	[165]
3	Chest X-Ray Images (Pneumonia)	It has 5856 images of chest x-ray of pneumonia cases with 3 classes	[26]
4	Blood Cell Images	It has 12,500 images of blood cells with 4 classes	[17]
5	Brain MRI Images for Brain Tumor Detection	It has 253 grey scale MRI images with 2 labels	[18]
6	Retinal OCT Images (optical coherence tomography)	It has 84,495 x-ray retina images with 4 labels	[124]
7	NIH Chest X-ray Dataset	It has 1,12,120 chest x-ray images with disease labels with 15 labels	[107]
8	Tuberculosis (TB) Chest X-ray Database	It has 7,000 images of TB chest x-ray images with 2 labels	[150]
9	Diabetic Retinopathy Detection	It has large set of retina images with 5 labels	[39]
10	Skin Cancer MNIST	It has 10,015 dermatoscopic images	[138]
11	NIH DeepLesion dataset	It has 10,594 CT scan images from 4427 individuals	[108]
12	CT Images in COVID-19	It has Unenhanced chest CT's. Data is collected from 661 patients.	[23]
13	Brain-Tumor-Progression	The dataset is collected from 20 patients' images are taken before and after treatment	[19]
14	Heart Disease Data Set	It has 4 databases concerning heart disease diagnosis	[54]
15	Histology Image Collection Library	It has 3870 historical images from a variety of disease	[56]
16	The Cavy dataset	It comprises of wide range of situations captured by a stationary camera, this dataset is used in many disciplines like computer vision	[147]
17	COPD Machine Learning Datasets	It is used in the diagnosis of chronic obstructive pulmonary disease	[21]
18	DRIVE: Digital Retinal Images for Vessel Extraction	It is collected from 400 individuals ranging in age from 25 to 90 years old.	[32]
19	Indian Diabetic Retinopathy Image Dataset (IDRID)	This dataset is divided into 3 parts namely segmentation, disease grading and original colour of fundus images	[68]
20	Melanoma Cancer Cell Dataset	This database contains 69 image sequences of control melanoma cells and 69 images of hydroxyurea treated melanoma cells	[98]

- Viral Pneumonia, Lung Opacity images. It has COVID-19 class to 3616 CXR, 10,192 Normal, 6012 Lung Opacity (Non-COVID lung infection), and 1345 Viral Pneumonia images. All the images are in PNG file format and the resolution is 299*299 pixels.
- **Wisconsin Breast Cancer Database:** This breast cancer database [165] was obtained from the University of Wisconsin Hospitals, Madison from Dr. William H. Wolberg. This dataset has compiled test and trained radiologic images in CSV sheet and it has two labeled classes Benign and Malignant and contained 458 (65.5%) and 241 (34.5%) images respectively. The dataset contains 16 missing values and images categorized into 8 groups as images are added periodically.
 - **Chest X-Ray Images (Pneumonia):** The dataset [26] contains 5856 images of Chest X-RAY of Pneumonia Cases. It has three classes Normal and Bacterial pneumonia and Viral Pneumonia. Dataset is well categorized into train test and validation. After X-Ray, the diagnoses for the photographs were rated by two experts before being approved for use in the AI system. The assessment set was additionally verified by a third expert to account for any grading problems.
 - **Blood Cell Images:** This dataset [17] contains 12,500 augmented images of blood cells in JPEG and accompanying cell type labels in CSV. It has four classes Eosinophil, Lymphocyte, Monocyte, Neutrophil, and each group contained approx 3000 images. A supplementary dataset contains the original 410 images (pre-augmentation), as well as two extra subtype labels (WBC vs WBC) and bounding boxes for each cell in each of these 410 images (JPEG + XML metadata). The 'dataset-master' folder provides 410 images of blood cells with subtype labels and bounding boxes (JPEG + XML) and the 'dataset2-master' folder provides 2,500 updated pictures with four additional subtype labels (JPEG + CSV).
 - **Brain MRI Images for Brain Tumor Detection:** The dataset [18] has 253 greyscale MRI images. It has only two labeled classes either Yes or No. Class yes contains an MRI Scan image of those who have a Brain tumor and in No does have a tumor. No class contain 98 images while Yes contained 155 images.
 - **Retinal OCT Images (optical coherence tomography):** The imaging technique of retinal optical coherence tomography (OCT) is used to acquire high-resolution cross-sections of the retinas of alive individuals. Optical coherence tomography (OCT) images [124] (Spectralis OCT, Heidelberg Engineering, Germany) were selected from retrospective cohorts of adult patients at the University of California San Diego's Shiley Eye Institute, the California Retinal Research Foundation, Medical Center Ophthalmology Associates, Shanghai First People's Hospital, and Beijing Tongren Eye Center between July 1, 2013, and June 30, 2014. The dataset contains 84,495 X-Ray images (JPEG) in 4 categories (NORMAL, CNV, DME, DRUSEN). The dataset is organized into 3 folders (train, test, val) and contains subfolders for each class (NORMAL, CNV, DME, DRUSEN).
 - **NIH Chest X-ray Dataset:** This NIH Chest X-ray Dataset [107] contained 112,120 of size 1024 x 1024 X-ray images with disease labels from 30,805 unique patients. The authors applied Natural Language Processing to text-mine disease classifications

from the related radiological reports to construct these labels. There are 15 classes (14 diseases, and one for "No findings"). Images can be classified as "No findings" or one or more disease classes. The Classes are Atelectasis, Consolidation, Infiltration, Pneumothorax, Edema, Emphysema, Fibrosis, Effusion, Pneumonia, Pleural thickening, Cardiomegaly, Nodule Mass and Hernia. The labels are expected to be more than 90% accurate and suitable for weakly-supervised learning.

- **Tuberculosis (TB) Chest X-ray Database:** A group of researchers from Qatar University in Doha, Qatar, and the University of Dhaka in Bangladesh, along with Malaysian collaborators, have created a database [150] of chest X-ray images for Tuberculosis (TB) positive cases and normal images in collaboration with medical doctors from Hamad Medical Corporation in Bangladesh. The dataset contains 3500 TB images and 3500 normal images in the current release. In total images, 2800 CXR images are collected from the NIAID TB portal program dataset, 306 CXR images from the BELARUS TB portal program dataset, and 394 CXR images from NLM. It has two classes Normal and Tuberculosis. All images are in PNG file format and size of 512*512 pixels.
- **Diabetic Retinopathy Detection:** The dataset [39] contain large set of high- resolution retina images taken under a variety of imaging conditions. Images are labeled with a subject id as well as either left or right for example 2_left.jpeg is the left eye of patient id 2. The clinician has classified the presence of diabetic retinopathy in each image in 5 classes and those classes are No DR, Mild, Moderate, Severe, Proliferative DR. The data set contain train.zip which have the training set (5 files total),test.zip which have the test set (7 files total),sample.zip which a small set of images to preview the full dataset,sampleSubmission.csv contains sample submission file in the correct format,trainLabels.csv contains the scores for the training set.
- **Skin Cancer MNIST:** HAM10000 Dataset [138] contained dermatoscopic images from different populations, acquired and stored by different modalities. Dataset consists of 10015 dermatoscopic images which can use as a training set for academic machine learning purposes. The cases include a diverse range of pigmented lesions, including actinic keratoses and intraepithelial carcinoma / Bowen's disease (apiece), basal cell carcinoma (bcc), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, bkl), dermatofibroma (df), melanom (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, vasc).
- **NIH DeepLesion dataset:** The dataset [108] contain 32,120 axial computed tomography (CT) slices from 10,594 CT scans (studies) having 4,427 different individuals. Each image contains 1–3 lesions, as well as bounding boxes and size information, for a total of 32,735 lesions. In the data set, there is a folder called "Imagespng," which contains png image files. Each slice was given the name "patient index study index series index slice index.png," with the last underscore indicating sub-folders as / or. The photos are saved in 16-bit unsigned format. To get the original Hounsfield unit (HU) values, subtract 32768 from the pixel intensity.
- **CT Images in COVID-19:** These retrospective NIfTI image datasets [23] consists of unenhanced chest CTs. It consists of a set of two datasets, the First dataset is from

632 patients with COVID-19 infections at the initial point of care, and the second dataset - is the second set of 121 CTs from 29 patients with COVID-19 infections with serial/sequential CTs. Both datasets were collected at the point of care in an outbreak context from individuals who had SARS-CoV-2 confirmed by reverse transcription-polymerase chain reaction (RT-PCR). Data were collected from a total of 661 patients and have 771 series of data.

- **Brain-Tumor-Progression:** The datasets [19] collection come from 20 people who had primary freshly diagnosed glioblastoma and were treated with surgery and routine concomitant chemo-radiation therapy (CRT) before receiving adjuvant chemotherapy. Each patient has two MRI exams: one after CRT completion and one after progression (decided clinically and based on a combination of clinical performance and/or imaging data, punctuated by a change in treatment or intervention). T1w (before and post-contrast agent), FLAIR, T2w, ADC, normalized cerebral blood flow, normalized relative cerebral blood volume, standardized relative cerebral blood volume, and binary tumor masks are included in all DICOM picture sets (generated using T1w images).
- **Heart Dieses Data Set :** This dataset [54] contains 4 databases concerning heart disease diagnosis. All attributes in this dataset are numeric-valued. The data was collected from Cleveland Clinic Foundation (Cleveland. data), Hungarian Institute of Cardiology, Budapest (Hungarian. data) V.A. Medical Center, Long Beach, CA (long-beach-va. data), University Hospital, Zurich, Switzerland (Switzerland. data). Dataset has four classes Cleveland, Hungarian, Switzerland, Long Beach VA and
- **Histology Image Collection Library:** The HICL is a collection [56] of 3870 histological images from a variety of diseases, including brain cancer, breast cancer, and HPV-Cervical cancer (Human Papilloma Virus). There are 116 breast cancer instances in this data set. A total of 872 pictures were created, 414 of which were stained with IHC (x40) and 458 with H&E. (227 at x20 and 231 x40). For digitalization, two light microscope imaging systems were used. There are 93 cases of brain cancer in the data. At x20 and x40, a total of 2548 H&E stained pictures of various grades were obtained, with 840 from low-grade cases, 1608 from high-grade cases, and 100 from ambiguous (low to high grade) instances. At x20 and x40, 450 P63 stained images of various grades were created, with 175 coming from grade I cases, 147 from grade II cases, and 128 from grade III instances.
- **The Cavy dataset:** This dataset [147] comprises a wide range of situations captured by a stationary camera. Sequences are captured from several viewpoints with varying light and at various times. It contains 16 640 480 quality sequences captured at 7.5 frames per second (fps), totaling roughly 31621506 frames (272 GB). The sequences are captured in ppm format and stored non-synchronously. The Cavy dataset can be useful in many disciplines, in addition to computer vision, since the dataset is taken at various times, it may help the biologists to study and monitor the cavy behavior in specific periods.
- **COPD Machine Learning Datasets:** The dataset [21] can be used in the diagnosis of chronic obstructive pulmonary disease (COPD). The dataset contains derived features (320-dimensional feature vectors) extracted from CT images of patients and controls

scanned at two separate facilities using different scanners and scanning conditions. Each image is made up of 50 feature vectors, each of which defines a volumetric ROI of size 41 x 41 x 41 voxels that were extracted at random positions inside the lung mask. Dataset has two types of features 1) `_gss.txt` Gaussian scale-space features, or histograms of intensity values in the ROI after filtering the image. Here we use eight filters (smoothed image, gradient magnitude, Laplacian of Gaussian, three eigenvalues of the Hessian, Gaussian curvature, and eigen magnitude), four scales (0.6, 1.2, 2.4, and 4.8 mm), and histograms of ten bins. There are 320 features in total. The file feature `_names.txt` specifies the ordering of the filters and the scales. 2) `_kdei.txt` Kernel density estimation features or a histogram of intensity values between -1100 and -600 Hounsfield units in the ROI. It has total of 256 features.

- **DRIVE: Digital Retinal Images for Vessel Extraction:** A diabetic retinopathy screening program in the Netherlands provided the pictures for the DRIVE database citeDB18. A total of 400 diabetic individuals ranging in age from 25 to 90 years old were screened. Thirty photos were chosen at random, 33 of which show no evidence of diabetic retinopathy and seven of which exhibit evidence of mild early diabetic retinopathy.
 25_training: pigment epithelium changes, probably butterfly maculopathy with a pigmented scar in the fovea, or choroidopathy, no diabetic retinopathy or other vascular abnormalities.
 26_training: background diabetic retinopathy, pigmentary epithelial atrophy, atrophy around optic disk.
 32_training: background diabetic retinopathy
 03_test: background diabetic retinopathy
 08_test: pigment epithelium changes, a pigmented scar in the fovea, or choroidopathy, no diabetic retinopathy or other vascular abnormalities
 14_test: background diabetic retinopathy
 17_test: background diabetic retinopathy
- **Indian Diabetic Retinopathy Image Dataset (IDRID):** The dataset [68] is divided into three parts: 1. Segmentation consists of original color fundus images. It has 81 images in total divided into train and test sets. Images are in JPG format. Groundtruth images for the Lesions (Microaneurysms, Haemorrhages, Hard Exudates, and Soft Exudates divided into train and test set - TIF Files) and Optic Disc (divided into train and test set - TIF Files) 2. Disease Grading consists of Original color fundus images (516 images divided into train set (413 images) and test set (103 images) - JPG Files). Groundtruth Labels for Diabetic Retinopathy and Diabetic Macular Edema Severity Grade (Divided into train and test set - CSV File). 3. Original color fundus images (516 images divided into train set (413 images) and test set (103 images JPG Files). Groundtruth Labels for Optic Disc Center Location (Divided into train and test set - CSV File). Groundtruth Labels for Fovea Center Location (Divided into train and test set - CSV File).
- **Melanoma Cancer Cell Dataset:** The melanoma cancer cell database [98] contains 69 image sequences of control melanoma cells and 69 picture sequences of hydroxyurea-treated melanoma cells. In this dataset, it is easy to see how the number of cells increases without any treatment. Long-term culture of metastatic murine melanoma

B16F10 cells in Roswell Park Memorial Institute (RPMI) medium (supplemented with 10% Fetal Bovine Serum (FBS), Streptomycin 10 mg/mL, and Penicillin 10,000 Units/mL) is shown in this dataset. First, B16F10 cells were plated (5x10⁴ cells/mL) in a 35mm polystyrene dish and treated to hydroxyurea (30mM) or medium for 24 hours (control group). The cells were then placed in a Nikon BioStation IM-Q inverted microscope, and images from 69 fields were recorded over the course of 24 hours using a high sensitivity cooled charge-coupled device (CCD) camera (40x objective). After all The final database contained 69 image sequences totaling 95 frames with a spatial resolution of 640x480 pixels and a one-minute length.

6 Evaluation metrics

It is always important to know how well our model performs and try to improve the model to higher accuracies. Before going to evaluation metrics there are some definitions we should know.

- **True positive:** When we predict a data point to a particular class and actually it belongs to that class.
- **True negative:** When we predict a data point doesn’t belong to a particular class and actually it doesn’t belong to that class.
- **False positive:** When we predict a data point belongs to a class and actually it doesn’t belong to that class.
- **False negative:** when we predict a data point doesn’t belong to that class and actually it belongs to that class.

These are the some well know performance metric

- **Confusion Matrix:** The below Table 8 shows the confusion matrix. For a high accuracy model, most of the non-diagonal elements are zero and the diagonal elements are non-zero integers. The confusion matrix is a square matrix and its size is the number of classes in the classification. A confusion matrix is really important and from this we can also calculate other evaluation metrics like accuracy, precision, and recall.
- **Precision:** It is the ratio of true positives to all the positives predicted by the model. For high precision, we need the false positives to be very much less. Low precision means the model predicts more false positives.

$$Precision = \frac{TP}{TP + FP}$$

Table 8 Confusion Matrix

		Actual Value	
		Positive	Negative
Predicted Value	Positive	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)

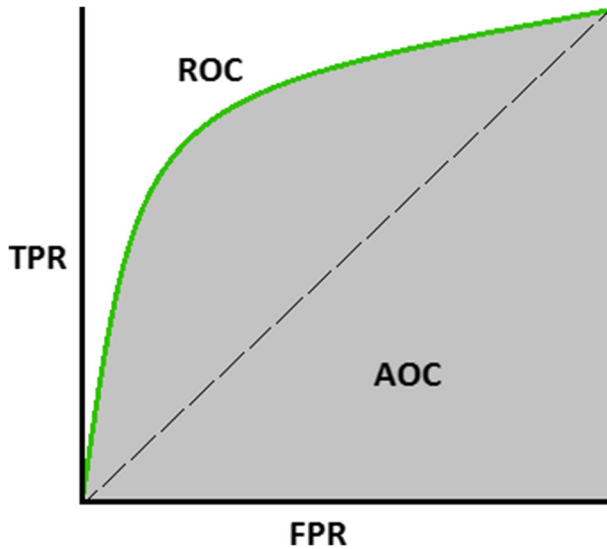


Fig. 11 AUC and ROC Curve

- **Recall/ Sensitivity:** It is the ratio of true positives to all the actual positives in the dataset. There will always be a trade-off between precision and recall.

$$Recall = \frac{TP}{TP + FN}$$

- **Specificity:** It is the ratio of true negative to the sum of a true negative and true positive. This can also be represented in the form of a false negative rate.

$$Specificity = \frac{TN}{TN + FP}$$

- **F1-Score:** If we need to find a balance between Precision and Recall AND there is an unequal class distribution, the F1 Score would be a preferable evaluation technique to use (for a large number of Actual Negatives).

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

- **AUC and ROC Curve:** AUC represents the degree or measure of separability, whereas ROC is a probability curve. It indicates how well the model can distinguish between classes. The AUC indicates how well the model predicts False as False and True as True. Higher the AUC, the model will perform better. In Fig. 11 TPR(True Positive Rate) or Recall or Sensitivity and FPR is 1-Specificity.

7 Discussion

As demonstrated by the 182 papers reviewed in this survey, medical image classification can assist in the automatic classification of many diseases. Within the last few years, it happened very quickly. The medical image classification problem was solved using a wide variety of deep architectures. These problems were successfully solved using pre-trained CNNs

as feature extractors. It was made possible to use these pre-trained networks by simply downloading them and directly applying them to any medical image. Our survey found that the majority of papers used this approach, so we can confidently state this is the current standard approach for medical image classification.

Various types of CNNs and multi CNNs can also be successfully applied to medical image classification problems. Some problems requiring sequence inputs can also be solved by combining CNNs and LSTMs.

Deep learning algorithms present several unique challenges when applied to medical image classification. One of the main obstacles is the lack of large training data sets.

Thus, image data availability is not the main challenge, but the correct and sufficient data are equally important. There is also the issue of class imbalance related to data. Depending on the task, it might be difficult to find images of the abnormal class in medical imaging. Together, AI scientists and medical science organizations can generate data that may be useful for deep learning model training.

8 Conclusion and future trends

This study provides a review of some of the work made on Medical Imaging using deep learning. We started with an introduction to Medical Imaging, the need for Medical Imaging, and have seen its applications and why is deep learning required for Medical Imaging. In this study, we focused on medical image classification using deep learning and have studied 156 research papers from various journals. We also reviewed some of the methods used for medical image classification they are CNN, Transfer learning models, CNN and LSTM, Autoencoders, and GAN using CNN. Also, we have discussed the strengths and limitations of these methods. The CNN models are easy to implement and can be combined with different image processing methods. We have also seen how CNN filters are modified with other mathematical function operations depending on the application and image type to improve the performance. We can also conclude that the transfer learning models require less training time as compared to other approaches and can be combined with other DNN for better accuracy. The transfer learning models not always give the expected results as these models are trained on the natural images. Thus it is required to perform image pre-processing, changes in the model functions for achieving better results. The LSTM, autoencoders and GAN implementation is complex, these methods require a higher training time and large image dataset. For GAN models, correct labeling of the data is an important task as it would affect the performance of the model. For medical image classification, combination of image processing for texture extraction and CNN or transfer learning model will be the promising method as it requires less training time and is less complex. We have also explored the available Medical Imaging datasets. A sufficient amount of training data is required for the deep learning models to perform well. The different evaluation metrics for image classification are discussed. Medical Imaging using deep learning is a vast field and a good research area. We have tried to cover all the possible methods for medical image classification. The future scope of this study can cover more details of the different image processing methods like segmentation, localization and using these processed images with CNN or any other deep learning model for classification.

At present deep learning and Artificial Intelligence has shown considerable progress in the field of health care. Deep learning plays a significant role in Medical Image Analysis

and is very beneficial for the detection and diagnosis of various diseases. Deep learning can be applied in various fields of medical imaging like radiology, ophthalmology, neuroimaging, and ultrasound. Nuclear imaging helps to predict thyroid, cancer, heart conditions, Alzheimer's disease, etc. Amyloid PET imaging helps to predict the progression of Alzheimer's disease. The amount of medical imaging data is huge and is in different formats. The use of this medical data for training different deep learning models can help to reduce the burden on the medical professionals and speed up their reading time. It has the ability to go through volumes of scans and provide insights into the data. It will help medical professionals for more precise and quick decision-making. The deep learning approaches are not only limited to the detection of a disease but can also find patterns and classify the lesion areas and provide more information about its diagnosis. It can perform tasks such as detection, localization, disease grading, data characterization, identification of the mutations of a specific disease, and quantification. Deep Learning can recommend the treatment required for the diagnosis. A good example of this is Google's DeepMind. It can read 3D retinal OCT scans and diagnose 50 different ophthalmic conditions and also detect indicators of eye disease. The AI Rad Companion developed by Siemens is used for the interpretation of medical images like MRI, CT scans, and X-ray machines. It provides a summary of the findings within an image thereby reducing the burden of basic repetitive tasks.

Also, the medical scanning devices are being improved to produce quality images and enhance the patient experience. Thus with the development and advancements of deep learning approaches for Medical Imaging, it will be very beneficial for health care and the medical field.

Funding No funding was received to assist with the preparation of this manuscript.

Data Availability Data sharing not applicable to this article as no datasets were generated during the current study. The datasets which used in this study their references are given in Table 7.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Abadeh MS, Shahamat H (2020) Brain MRI analysis using a deep learning based evolutionary approach. *Neural Netw* 126:218–234. <https://doi.org/10.1016/j.neunet.2020.03.017>
2. Abdulkareem K et al (2022) Automated system for identifying COVID-19 Infections in computed tomography images using deep learning models. In: *Journal of healthcare engineering* 2022. <https://doi.org/10.1155/2022/5329014>
3. Abdullah SM et al (2023) Deep transfer learning based parkinson's disease detection using optimized feature selection. *IEEE Access* 11:3511–3524. <https://doi.org/10.1109/ACCESS.2023.3233969>
4. Abdulsahib A, Mahmoud M (2022) An Automated Image Segmentation and Useful Feature Extraction Algorithm for Retinal Blood Vessels in Fundus Images. *Electronics* 11:1295. <https://doi.org/10.3390/electronics11091295>
5. Abideen ZU, Ghafoor M, Munir K, Saqib M, Ullah A, Zia T, Tariq SA, Ahmed G, Zahra A (2020) Uncertainty assisted robust tuberculosis identification with bayesian convolutional neural networks. *IEEE Access* 8:22812–22825. *IEEE*
6. Al-Saffar ZA, Yildirim T (2020) A novel approach to improving brain image classification using mutual information-accelerated singular value decomposition. *IEEE Access* 8:52575–52587. <https://doi.org/10.1109/ACCESS.2020.2980728>

7. Allaouzi I, Ahmed BM (2019) A novel approach for multi-label chest x-ray classification of common thorax diseases. *IEEE Access* 7:64279–64288. <https://doi.org/10.1109/ACCESS.2019.2916849>
8. Alzubaidi LZ, Humaidi J (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 8.53. <https://doi.org/10.1186/s40537-021-00444-8>
9. Anitha V, Murugavalli S (2016) Brain tumour classification using two-tier classifier with adaptive segmentation technique. *IET Comput Vis* 10:9–17. <https://doi.org/10.1049/iet-cvi.2014.0193>
10. Ansingkar NP, Patil R, Deshmukh PD (2022) An efficient multi class Alzheimer detection using hybrid equilibrium optimizer with capsule auto encoder. *Multimedia Tools and Applications*, pp 1–32
11. Arias-Garzón D et al (2021) COVID-19 detection in X-ray images using convolutional neural networks. *Mach Learn Appl* 6:100138. ISSN: 2666-8270. <https://doi.org/10.1016/j.mlwa.2021.100138>, <https://www.sciencedirect.com/science/article/pii/S2666827021000694>
12. Ashraf R et al (2020) Deep convolution neural network for big data medical image classification. *IEEE Access* 8:105659–105670. <https://doi.org/10.1109/ACCESS.2020.2998808>
13. Asifullah K et al (2020) A survey of the recent architectures of deep convolutional neural networks. *Artif Intell Rev* 53:5455–5516. ISSN: 1573-7462. <https://doi.org/10.1007/s10462-020-09825-6>
14. Baldi P (2011) Autoencoders, unsupervised learning and deep architectures. In: *Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning Workshop - vol 27. UTLW'11*. Washington, USA: JMLR.org, pp 37–50
15. Bank D, Koenigstein N, Giryas R (2021) Autoencoders. *arXiv:2003.05991* [cs.LG.]
16. Bian J et al (2021) Skin lesion classification by multi-view filtered transfer learning. *IEEE Access* 9:66052–66061. <https://doi.org/10.1109/ACCESS.2021.3076533>
17. Blood Cell Images (2018) <https://www.kaggle.com/paultimothymooney/blood-cells>
18. Brain MRI Images for Brain Tumor Detection (2019) <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection>
19. Brain-Tumor-Progression (2021) <https://wiki.cancerimagingarchive.net/display/Public/Brain-Tumor-Progression#g339481190e2ccc0d07d7455ab87b3ebb625adf48>
20. Brima Y, Tushar MHK, Kabir U, Islam T (2021) Deep transfer learning for brain magnetic resonance image multi-class classification. *arXiv:2106.07333* [cs.CV]
21. COPD Machine Learning Datasets (2018) <http://bigr.nl/research/projects/copd>
22. COVID-19 Radiography Dataset (2020) <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>
23. CT Images in COVID-19 (2021) <https://wiki.cancerimagingarchive.net/display/Public/CT+Images+in+COVID-19>
24. Chai Y, Liu H, Xu J (2018) Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. *Knowl Based Syst* 161:147–156. ISSN:0950-7051. <https://doi.org/10.1016/j.knosys.2018.07.043>, <https://www.sciencedirect.com/science/article/pii/S0950705118303940>
25. Charle D, Charle F, Garca S, del Jesus MJ, Herrera F (2018) A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines. *Inf Fusion* 44:78–96. ISSN: 1566-2535. <https://doi.org/10.1016/j.inffus.2017.12.007>, <https://www.sciencedirect.com/science/article/pii/S1566253517307844>
26. Chest X-Ray Images (Pneumonia) (2018) <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>
27. Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. *arXiv:1610.02357* [cs.CV]
28. Chowdhary CL, Acharjya D (2016) A hybrid scheme for breast cancer detection using intuitionistic fuzzy rough set technique. *Int J Healthc Inf Syst Inform* 11.2:38–61. <https://doi.org/10.4018/IJHISI.2016040103>
29. Chowdhary CL, Acharjya D (2016) Breast cancer detection using intuitionistic fuzzy histogram hyperbolization and possibilistic fuzzy c-mean clustering algorithms with texture feature based classification on mammography images. In: *Proceedings of the international conference on advances in information communication technology & computing*. <https://doi.org/10.1145/2979779.2979800>
30. Chowdhary CL, Acharjya D (2020) Segmentation and feature extraction in medical imaging: A systematic review. *Procedia Comput Sci* 167:26–36. <https://doi.org/10.1016/j.procs.2020.03.179>
31. DA Zebari, DA Ibrahim, HJ Mohammed (2022) Effective hybrid deep learning model for COVID-19 patterns identification using CT images. *Expert Systems*. <https://doi.org/10.1111/exsy.13010>
32. DRIVE: Digital Retinal Images for Vessel Extraction (2012) <https://drive.grand-challenge.org/>
33. Das V, Dandapat S, Bora PK (2020) A data-efficient approach for automated classification of ocr images using generative adversarial network. *IEEE Sens Lett* 4(1):1–4. IEEE
34. Das PK, Meher S (2021) An efficient deep convolutional neural network based detection and classification of acute lymphoblastic leukemia. *Expert Systems with Applications*, pp 115311. ISSN:

- 0957-4174. <https://doi.org/10.1016/j.eswa.2021.115311>, <https://www.sciencedirect.com/science/article/pii/S0957417421007405>
35. Das K et al (2020) Detection of breast cancer from whole slide histopathological images using deep multiple instance CNN. *IEEE Access* 8:213502–213511. <https://doi.org/10.1109/ACCESS.2020.3040106>
 36. Das AK et al (2021) TLCoV- An automated Covid-19 screening model using Transfer Learning from chest X-ray images. *Chaos, Solitons Fractals* 144:110713. ISSN: 0960-0779. <https://doi.org/10.1016/j.chaos.2021.110713>, <https://www.sciencedirect.com/science/article/pii/S0960077921000667>
 37. De Moura J et al (2020) Deep convolutional approaches for the analysis of COVID-19 using chest X-Ray images from portable devices. *IEEE Access* 8:195594–195607. <https://doi.org/10.1109/ACCESS.2020.3033762>
 38. Demir F (2021) DeepCoroNet: A deep LSTM approach for automated detection of COVID-19 cases from chest X-ray images. *IEEE Access* 103:107160. <https://doi.org/10.1016/j.asoc.2021.107160>
 39. Diabetic Retinopathy Detection (2015) <https://www.kaggle.com/c/diabeticretinopathy-detection>
 40. Diakite J, Xiaping X (2021) Hyperspectral image classification using 3D 2D CNN. *IET Image Proc* 15:1083–1092. <https://doi.org/10.1049/ipr2.12087>
 41. Elkorany AS, Elsharkawy ZF (2021) COVIDetection-Net: A tailored COVID-19 detection from chest radiography images using deep learning. *Optik* 231:166405. ISSN: 0030-4026. <https://doi.org/10.1016/j.jijleo.2021.166405>, <https://www.sciencedirect.com/science/article/pii/S0030402621001388>
 42. Elmannai H, Hamdi M, AlGarni A (2021) Deep learning models combining for breast cancer histopathology image classification. *Int J Comput Intell Syst* 14(1):1003. Atlantis Press BV
 43. Fradi M, Khriji L, Machhout M (2022) Real-time arrhythmia heart disease detection system using CNN architecture based various optimizers-networks. *Multimed Tools Appl* 81.29:41711–41732
 44. Frid-Adar M et al (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 321:321–331. ISSN: 0925-2312. <https://doi.org/10.1016/j.neucom.2018.09.013>, <https://www.sciencedirect.com/science/article/pii/S0925231218310749>
 45. Gao Y, Wang R et al, Shi Y (2013) Transductive cost-sensitive lung cancer image classification. *Appl Intell* springer 38:16–28. <https://doi.org/10.1007/s10489-012-0354-z>
 46. García-Ordás MT et al (2020) Detecting respiratory pathologies using convolutional neural networks and variational autoencoders for unbalancing data. *Sensors* 20.4. ISSN: 1424-8220. <https://doi.org/10.3390/s20041214>, <https://www.mdpi.com/1424-8220/20/4/1214>
 47. Garg NK, Chhabra P, Kumar M (2018) Content-based image retrieval system using ORB and SIFT features. *Neural Comput Appl* 32:2725–2733
 48. Goodfellow IJ et al (2014) Generative adversarial networks. *arXiv:1406.2661 [stat.ML]*
 49. Greg VH, Carlos M, Gonzalo N (2020) A review on the long short-term memory model. *Artif Intell Lett* 53:5929–5955. ISSN: 1573-7462. <https://doi.org/10.1007/s10462-020-09838-1>
 50. Gu J, Wang Z, Kuen J, Ma L, Shahrudiy A, Shuai B, Liu T, Wang X, Wang L, Wang G, Cai J, Chen T (2017) Recent advances in convolutional neural networks. *arXiv:1512.07108 [cs.CV]*
 51. Hasan MM et al (2023) Review on the evaluation and development of artificial intelligence for COVID-19 containment. *Sensors* 23.1:527
 52. He X, Fang L, Rabbani H, Chen X, Liu Z (2020) Retinal optical coherence tomography image classification with label smoothing generative adversarial network. *Neurocomputing* 405:37–47. ISSN: 0925-2312. <https://doi.org/10.1016/j.neucom.2020.04.044>, <https://www.sciencedirect.com/science/article/pii/S0925231220306111>
 53. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. *arXiv:1512.03385 [cs.CV]*
 54. Heart Dieses Data Set (1988) <http://archive.ics.uci.edu/ml/datasets/Heart+Disease>
 55. Hemanth DJ et al (2019) A modified deep convolutional neural network for abnormal brain image classification. *IEEE Access* 7:4275–4283. <https://doi.org/10.1109/ACCESS.2018.2885639>
 56. Histology Image Collection Library (1988) <https://medisp.bme.uniwa.gr/hicl/index.html>
 57. Howard AG et al (2017) MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv:1704.04861 [cs.CV]*
 58. Liao F, Chen X, Hu X, Song S (2017) Estimation of the volume of the left ventricle from MRI images using deep neural networks. *IEEE Trans Cybern* 49(2):495–504. IEEE
 59. Hu J et al (2019) Squeeze-and-excitation networks. *arXiv:709.01507 [cs.CV]*
 60. Hu S et al (2020) Weakly supervised deep learning for COVID-19 infection detection and classification from CT images. *IEEE Access* 8:118869–118883. <https://doi.org/10.1109/ACCESS.2020.3005510>

61. Hu Z-P, Zhang R-X, Qiu Y, Zhao M-Y, Sun Z (2021) 3D convolutional networks with multi-layer-pooling selection fusion for video classification. *Multimed Tools Appl* 80:33179–33192. Springer
62. Huang G et al (2018) Densely connected convolutional networks. arXiv:1608.06993 [cs.CV]
63. Huang L, Fang L, Rabbani H, Chen X (2019) Automatic classification of retinal optical coherence tomography images with layer guided convolutional neural network. *IEEE Signal Process Lett* 26.7:1026–1030. <https://doi.org/10.1109/LSP.2019.2917779>
64. Huang Q et al (2020) Blood cell classification based on hyperspectral imaging with modulated Gabor and CNN. *IEEE J Biomed Health Inf* 24.1:160–170. <https://doi.org/10.1109/JBHI.2019.2905623>
65. Huang X et al (2020) Deep transfer convolutional neural network and extreme learning machine for lung nodule diagnosis on CT images. *Knowl-Based Syst* 204:106230. ISSN: 0950-7051. <https://doi.org/10.1016/j.knosys.2020.106230>, <https://www.sciencedirect.com/science/article/pii/S0950705120304378>
66. Hussain E et al (2020) A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network. *Tissue Cell* 65:101347. ISSN: 0040-8166. <https://doi.org/10.1016/j.tice.2020.101347>, <https://www.sciencedirect.com/science/article/pii/S0040816619304872>
67. Hussain SM et al (2022) Deep learning based image processing for robot assisted surgery: a systematic literature survey. *IEEE Access* 10:122627–122657. <https://doi.org/10.1109/ACCESS.2022.3223704>
68. Indian Diabetic Retinopathy Image Dataset (IDRID) (2019) <https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-datasetidrid>
69. Indolia S et al (2018) Conceptual understanding of convolutional neural network- a deep learning approach. *Procedia Comput Sci* 132:679–688. ISSN: 1877-0509. <https://doi.org/10.1016/j.procs.2018.05.069>, <https://www.sciencedirect.com/science/article/pii/S1877050918308019>
70. Inthiyaz S et al (2023) Skin disease detection using deep learning. *Adv Eng Softw* 175:103361
71. Jammula R, Tejus VR, Shankar S (2020) Optimal transfer learning model for binary classification of fundus images through simple heuristics. arXiv:2002.04189 [cs.LG]
72. Jun TJ et al (2021) TRK-CNN: Transferable ranking-CNN for image classification of glaucoma, glaucoma suspect, and normal eyes. *Exp Syst Appl* 182:115211. ISSN: 0957-4174. <https://doi.org/10.1016/j.eswa.2021.115211>, <https://www.sciencedirect.com/science/article/pii/S0957417421006448>
73. Khan NM, Abraham N, Hon M (2019) Transfer learning with intelligent training data selection for prediction of alzheimer's disease. *IEEE Access* 7:72726–72735. <https://doi.org/10.1109/ACCESS.2019.2920448>
74. Khan MA, Muhammad K, Sharif M, Akram T, de Albuquerque VHC (2021) Multi-class skin lesion detection and classification via teledermatology. *IEEE J Biomed Health Inf* 25(12):4267–4275. IEEE
75. Kim S-H, Koh HM, Lee B-D (2021) Classification of colorectal cancer in histological images using deep neural networks: An investigation. *Multimed Tools Appl* 80.28:35941–35953
76. Kozegar E, Soryani M, Behnam H, Salamati M, Tan T (2020) Computer aided detection in automated 3-D breast ultrasound images: a survey. *Artif Intell Rev* 53:1919–1941. Springer
77. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet Classification with deep convolutional neural networks. In: *Proceedings of the 25th International conference on neural information processing systems - vol 1*. NIPS'12, Lake Tahoe, Nevada: Curran Associates Inc., pp 1097–1105
78. Kuang Y, Lan T, Peng X, Selasi GE, Liu Q, Zhang J (2020) Unsupervised multi-discriminator generative adversarial network for lung nodule malignancy classification. *IEEE Access* 8:77725–77734. <https://doi.org/10.1109/ACCESS.2020.2987961>
79. Kumar M, Bansal M, Sachdeva M (2021) Transfer learning for image classification using VGG19: Caltech-101 image data set. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-021-03488-z>
80. Kumar D et al (2020) Automatic detection of white blood cancer from bone marrow microscopic images using convolutional neural networks. *IEEE Access* 8:142521–142531. <https://doi.org/10.1109/ACCESS.2020.3012292>
81. Labhsetwar SR et al (2020) Predictive analysis of diabetic retinopathy with transfer learning. arXiv:2011.04052 [cs.CV]
82. Lecun Y et al (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86.11:2278–2324. <https://doi.org/10.1109/5.726791>
83. Li Z, Zhou D, Wan L, Li J, Mou W (2020) Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram. *J Electrocardiol* 58:105–112. ISSN: 0022-0736. <https://doi.org/10.1016/j.jelectrocard.2019.11.046>, <https://www.sciencedirect.com/science/article/pii/S0022073619304170>

84. Li C et al (2021) Transfer learning for establishment of recognition of COVID-19 on CT imaging using small-sized training datasets. *Knowl-Based Syst* 218:106849. ISSN: 0950-7051. <https://doi.org/10.1016/j.knosys.2021.106849>, <https://www.sciencedirect.com/science/article/pii/S095070512100112X>
85. Liang D, Sun L, Ma W, Paisley J (2020) A 3D spatially weighted network for segmentation of brain tissue from MRI. *IEEE Trans Med Imaging* 39:898–909. <https://doi.org/10.1109/TMI.2019.2937271>
86. Liang G et al (2018) Combining convolutional neural network with recursive neural network for blood cell image classification. *IEEE Access* 6:36188–36197. <https://doi.org/10.1109/ACCESS.2018.2846685>
87. Liu H, Huang KK, Ren CX, Lai ZR (2021) Hyperspectral image classification via discriminative convolutional neural network with an improved triplet loss. *Pattern Recognition* 112. <https://doi.org/10.1016/j.patcog.2020.107744>
88. Liu Y, Wang W (2015) Simultaneous image fusion and denoising with adaptive sparse representation. *IET Image Proc* 9:347–357. <https://doi.org/10.1049/iet-ipr.2014.0311>
89. Liu X-J et al (2022) Few-shot learning for skin lesion image classification. *Multimedia Tools and Applications*, pp 1–12
90. Ma Y, Niu D, Zhang J et al (2021) Unsupervised deformable image registration network for 3D medical images. *Applied Intelligence* springer. <https://doi.org/10.1007/s10489-021-02196-7>
91. Mahmoudi R, Benameur N, Mabrouk R (2022) A Deep Learning-Based Diagnosis System for COVID-19 Detection and Pneumonia Screening Using CT Imaging. *Appl Sci* 12:4825. <https://doi.org/10.3390/app12104825>
92. Mallick PK et al (2019) Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network. *IEEE Access* 7:46278–46287. <https://doi.org/10.1109/ACCESS.2019.2902252>
93. Mamalakis M et al (2021) DenResCov-19: A deep transfer learning network for robust automatic classification of COVID-19, pneumonia, and tuberculosis from X-rays. *arXiv:2104.04006 [eess.IV]*
94. Martín EX, Velasco M, Angulo C et al (2014) LTI ODE-valued neural networks. *Appl Intell Springer* 41:594–605. <https://doi.org/10.1007/s10489-014-0548-7>
95. Martinez AR (2020) Classification of COVID-19 in CT scans using multi-source transfer learning
96. Masoudi S et al (2021) Deep Learning Based Staging of Bone Lesions From Computed Tomography Scans. *IEEE Access* 9:87531–87542. <https://doi.org/10.1109/ACCESS.2021.3074051>
97. Mehmood S et al (2022) Malignancy detection in lung and colon histopathology images using transfer learning with class selective image processing. *IEEE Access* 10:25657–25668. <https://doi.org/10.1109/ACCESS.2022.3150924>
98. Melanoma Cancer Cell Dataset (2020) <https://sites.google.com/view/virginiafernandes/datasets/melanoma-cancer-cell-dataset>.
99. Meng D et al (2017) Liver fibrosis classification based on transfer learning and FCNet for ultrasound images. *IEEE Access* 5:5804–5810. <https://doi.org/10.1109/ACCESS.2017.2689058>
100. Meng N et al (2019) Large-scale multi-class image-based cell classification with deep learning. *IEEE J Biomed Inform* 23.5:2091–2098. <https://doi.org/10.1109/JBHI.2018.2878878>
101. Mercioni M-A, Stavarache LL (2022) Disease diagnosis with medical imaging using deep learning. In: *Advances in information and communication: proceedings of the 2022 future of information and communication conference (FICC)*, vol 2. Springer, pp 198–208
102. Mijwil MM (2021) Skin cancer disease images classification using deep learning solutions. *Multimed Tools Appl* 80.17:26255–26271
103. Motamed S, Rogalla P, Khalvati F (2020) RANDGAN: Randomized generative adversarial network for detection of COVID-19 in chest X-ray. *arXiv:2010.06418 [eess.IV]*
104. Moslehi S, Mahjub H, Farhadian M, Soltanian AR, Mamani M (2022) Interpretable generalized neural additive models for mortality prediction of COVID-19 hospitalized patients in Hamadan, Iran. *BMC Med Res Methodol* 22(1):339. Springer
105. Mousavi Z et al (2022) COVID-19 detection using chest X-ray images based on a developed deep neural network. *SLAS Technology* 27.1:63–75. ISSN: 2472-6303. <https://doi.org/10.1016/j.slast.2021.10.011>, <https://www.sciencedirect.com/science/article/pii/S247263032100011X>
106. Muhammad G, Shamim Hossain M (2021) COVID-19 and non-COVID-19 classification using multi-layers fusion from lung ultrasound images. *Inf Fusion* 72:80–88. ISSN: 1566-2535. <https://doi.org/10.1016/j.inffus.2021.02.013>, <https://www.sciencedirect.com/science/article/pii/S1566253521000361>
107. NIH Chest X-ray Dataset (2018) <https://www.kaggle.com/nih-chest-xrays/data>
108. NIH DeepLesion dataset (2018) <https://www.kaggle.com/kmader/nih-deepleesion-subset>.

109. Nascimento JC, Carneiro G (2013) Combining Multiple Dynamic Models and Deep Learning Architectures for Tracking the Left Ventricle Endocardium in Ultrasound Data. *IEEE Trans Pattern Anal Mach Intell* 35.11:2592. <https://doi.org/10.1109/TPAMI.2013.96>
110. Nascimento JC, Carneiro G, Freitas A (2012) The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Trans Image Process* 21.3:968–982. <https://doi.org/10.1109/TIP.2011.2169273>
111. Nasir Khan H et al (2019) Multi-view feature fusion based four views model for mam-mogram classification using convolutional neural network. *IEEE Access* 7:165724–165733. <https://doi.org/10.1109/ACCESS.2019.2953318>
112. Nigam B et al (2021) COVID-19: Automatic detection from X-ray images by utilizing deep learning methods. *Exp Syst Appl* 176:114883. ISSN: 0957-4174. <https://doi.org/10.1016/j.eswa.2021.114883>, <https://www.sciencedirect.com/science/article/pii/S0957417421003249>
113. Noreen N et al (2020) A deep learning model based on concatenation approach for the diagnosis of brain tumor. *IEEE Access* 8:55135–55144. <https://doi.org/10.1109/ACCESS.2020.2978629>
114. Pantrigo JJ, Nunez JC, Cabido R, Montemayor AS (2018) Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition. *Pattern Recognit* 76:80–94. <https://doi.org/10.1016/j.patcog.2017.10.033>
115. Peng Y, Zhu H, Han G, Zhao H (2021) Functional-realistic CT image super-resolution for early-stage pulmonary nodule detection. *Future Gener Comput Syst* 115:475–485. <https://doi.org/10.1016/j.future.2020.09.020>
116. Petrick N, Pezeshk A, Hamidian S, Sahiner B (2019) 3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT. In: *IEEE Journal of biomedical and health informatics* 23, pp 2080–2090. <https://doi.org/10.1109/JBHI.2018.2879449>
117. Poloni KM et al (2021) Brain MR image classification for Alzheimer’s disease diagnosis using structural hippocampal asymmetrical attributes from directional 3-D log-Gabor filter responses. *Neurocomputing* 419:126–135. ISSN: 0925-2312. <https://doi.org/10.1016/j.neucom.2020.07.102>, <https://www.sciencedirect.com/science/article/pii/S0925231220312972>
118. Pulgar FJ, Charfe F, Rivera AJ, del Jesus MJ (2020) Choosing the proper autoencoder for feature fusion based on data complexity and classifiers: Analysis, tips and guidelines. *Inf Fusion* 54:44–60. ISSN: 1566-2535. <https://doi.org/10.1016/j.inffus.2019.07.004>, <https://www.sciencedirect.com/science/article/pii/S1566253519300880>
119. Qureshi I, Shaheed K, Mao A, Zhang X (2022) Finger-vein presentation attack detection using depthwise separable convolution neural network. *Expert Systems with Applications* 198. <https://doi.org/10.1016/j.eswa.2022.116786>
120. Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mazhar R, Hamid T, Islam MT, Kashem S, Mahbub ZB et al (2020) Reliable tuberculosis detection using chest X-Ray with deep learning, segmentation and visualization. *IEEE Access* 8:191586–191601. *IEEE*
121. Raj A, Shah NA, Tiwari AK, Martini MG (2020) Multivariate regression-based convolutional neural network model for fundus image quality assessment. *IEEE Access* 8:57810–57821. <https://doi.org/10.1109/ACCESS.2020.2982588>
122. Rajaraman S, Antani SK (2020) Modality-specific deep learning model ensembles toward improving tb detection in chest radiographs. *IEEE Access* 8:27318–27326. <https://doi.org/10.1109/ACCESS.2020.2971257>
123. Reshi AA, Rustam F, Mehmood A, Alhossan A, Alrabiah Z, Ahmad A, Alsuwailem H, Choi GS (2021) An efficient CNN model for COVID-19 disease detection based on X-ray image classification. *Complexity* 2021:1–12. Hindawi Limited
124. Retinal OCT Images (optical coherence tomography) (2018) <https://www.kaggle.com/paultimothy/mooney/kermany2018>
125. Rong Y et al (2019) surrogate-assisted retinal OCT image classification based on convolutional neural networks. *IEEE J Biomed Health Inf* 23.1:253–263. <https://doi.org/10.1109/JBHI.2018.2795545>
126. Russell RL, Ozdemir O, Berlin AA (2020) A 3D probabilistic deep learning system for detection and diagnosis of lung cancer using low-dose CT scans. *IEEE Trans Med Imaging* 39:1419–1429. <https://doi.org/10.1109/TMI.2019.2947595>
127. SJS Gardezi, Elazab A, Wang C, Bai H (2020) GP-GAN: Brain tumor growth prediction using stacked 3D generative adversarial networks from longitudinal MR Images. *Neural Netw* 132:321–332. <https://doi.org/10.1016/j.neunet.2020.09.004>
128. Saha S, Sheikh N (2021) Ultrasound image classification using ACGAN with small training dataset. *arXiv:2102.01539 [eess.IV]*

129. Sakib S et al, Fouda MM, Fadlullah ZM, Guizani M (2020) DL-CRC: Deep learning-based chest radiograph classification for COVID-19 detection: A novel approach. *IEEE Access* 8:171575–171589. <https://doi.org/10.1109/ACCESS.2020.3025010>
130. Salama WM, Shokry A, Aly MH (2022) A generalized framework for lung cancer classification based on deep generative models. *Multimed Tools Applic* 81(23):32705–32722. Springer
131. Salehinejad H et al (2019) Synthesizing chest X-Ray pathology for training deep convolutional neural networks. *IEEE Trans Med Imaging* 38.5:1197–1206. <https://doi.org/10.1109/TMI.2018.2881415>
132. Saxena A, Singh SP (2022) A deep learning approach for the detection of COVID-19 from chest X-Ray images using convolutional neural networks. <https://europepmc.org/article/PPR/PPR454232>
133. Schmid V, Meyer-Baeke A (2014) Pattern recognition and signal analysis in medical imaging, 2nd edn. Academic Press, Cambridge, pp 1–20. <https://doi.org/10.1016/B978-0-12-409545-8.00001-7>
134. Shamim S et al (2022) Automatic COVID-19 lung infection segmentation through modified Unet model. *J Healthcare Eng* 2022:6566982. <https://doi.org/10.1155/2022/6566982>
135. Sherstinsky A (2020) Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Phys D: Nonlinear Phenom* 404:132306. ISSN: 0167-2789. <https://doi.org/10.1016/j.physd.2019.132306>
136. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 [cs.CV]
137. Singh S, Tripathi B (2022) Pneumonia classification using quaternion deep learning. *Multimed Tools Appl* 81.2:1743–1764
138. Skin Cancer MNIST: HAM10000 (2018) <https://www.kaggle.com/kmader/skin-cancer-mnist-ham10000>
139. Soomro TA (2021) Artificial intelligence (AI) for medical imaging to combat coronavirus disease (COVID-19): a detailed review with direction for future research. *Artificial Intelligence Review*. ISSN: 1573-7462. <https://doi.org/10.1007/s10462-021-09985-z>
140. Sudharshan PJ et al (2019) Multiple instance learning for histopathological breast cancer image classification. *Exp Syst Appl* 117:103–111. ISSN: 0957-4174. <https://doi.org/10.1016/j.eswa.2018.09.049>, <https://www.sciencedirect.com/science/article/pii/S0957417418306262>
141. Sultan HH, Salem NM, Al-Atabany W (2019) Multi-classification of brain tumor images using deep neural network. *IEEE Access* 7:69215–69225. <https://doi.org/10.1109/ACCESS.2019.2919122>
142. Sun G, Wang X, Xu L, Li C, Wang W, Yi Z, Luo H, Su Y, Zheng J, Li Z et al (2023) Deep learning for the detection of multiple fundus diseases using ultra-widefield images. *Ophthalmol Therapy* 12(2):895–907. Springer
143. Suresh S, Mohan S (2019) NROI based feature learning for automated tumor stage classification of pulmonary lung nodules using deep convolutional neural networks. *Journal of King Saud University - Computer and Information Sciences*. ISSN: 1319-1578. <https://doi.org/10.1016/j.jksuci.2019.11.013>, <https://www.sciencedirect.com/science/article/pii/S131915781931420X>.
144. Szegedy C et al (2014) Going deeper with convolutions. arXiv:1409.4842 [cs.CV]
145. Szegedy C et al (2015) Rethinking the inception architecture for computer vision, arXiv:1512.00567 [cs.CV]
146. Szegedy C et al (2016) Inception-v4, inception-ResNet and the impact of residual connections on learning. arXiv:1602.07261 [cs.CV]
147. The Cavy dataset (2016) <http://www.inf-cv.uni-jena.de/Research/Datasets/Cavy+Dataset.html>.
148. Ting FF, Tan YJ, Sim KS (2019) Convolutional neural network improvement for breast cancer classification. *Exp Syst Appl* 120:103–115. ISSN: 0957-4174. <https://doi.org/10.1016/j.eswa.2018.11.008>, <https://www.sciencedirect.com/science/article/pii/S0957417418307280>
149. Trivizakis E et al (2019) Extending 2-D convolutional neural networks to 3-D for advancing deep learning cancer classification with application to MRI liver tumor differentiation. *IEEE J Biomed Health Inf* 23.3:923–930. <https://doi.org/10.1109/JBHI.2018.2886276>
150. Tuberculosis (TB) Chest X-ray Database. 8.2 (2021) <https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>
151. Turkoglu M (2021) COVID-19 detection system using chest CT images and multiple kernels-extreme learning machine based on deep neural network. *IRBM*. ISSN: 1959-0318. <https://doi.org/10.1016/j.irbm.2021.01.004>, <https://www.sciencedirect.com/science/article/pii/S1959031821000051>
152. Vairamuthu S, Navaneethakrishnan M, Parthasarathy G (2021) Atom search-Jaya-based deep recurrent neural network for liver cancer detection. *IET Image Proc* 15:337–349. <https://doi.org/10.1049/ipr.2.12019>

153. van Grinsven MJJP, van Ginneken B et al (2016) Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. *IEEE Trans Med Imaging* 35(5):1273–1284. <https://doi.org/10.1109/TMI.2016.2526689>
154. Waheed A, Goyal M, Gupta D, Khanna A, Al-Turjman F, Pinheiro PR (2020) Covidgan: data augmentation using auxiliary classifier gan for improved covid-19 detection. *IEEE Access* 8:91916–91923. *IEEE*
155. Wang J (2020) OCT image recognition of cardiovascular vulnerable plaque based on CNN. *IEEE Access* 8:140767–140776. <https://doi.org/10.1109/ACCESS.2020.3007599>
156. Wang SW, Guo B, Y et al (2020) Twin labeled LDA: a supervised topic model for document classification. *Appl Intell Springer* 50:4602–4615. <https://doi.org/10.1007/s10489-020-01798-x>
157. Wang Q, Li Y, Wang Y, Ren J (2022) An automatic algorithm for software vulnerability classification based on CNN and GRU. *Multimedia Tools and Applications*, pp 1–22
158. Wang M, Jiang M (2019) Deep residual refining based pseudo-multi-frame network for effective single image super-resolution. *IET Image Process* 13:591–599. <https://doi.org/10.1049/iet-ipr.2018.6057>
159. Wang D, Wang L (2019) On OCT Image Classification via Deep Learning. *IEEE Photonics J* 11.5:1–14. <https://doi.org/10.1109/JPHOT.2019.2934484>
160. Wang Z et al (2019) Dilated 3D Convolutional neural networks for brain MRI data classification. *IEEE Access* 7:134388–134398. <https://doi.org/10.1109/ACCESS.2019.2941912>
161. Wang C et al (2019) Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access* 7:146533–146541. <https://doi.org/10.1109/ACCESS.2019.2946000>
162. Wang C et al (2019) Pulmonary image classification based on inception-v3 transfer learning model. *IEEE Access* 7:146533–146541. <https://doi.org/10.1109/ACCESS.2019.2946000>
163. Wang Y et al (2020) An optimized deep convolutional neural network for dendrobium classification based on electronic nose. *Sens Actuator A Phys* 307:111874. ISSN: 0924-4247. <https://doi.org/10.1016/j.sna.2020.111874>, <https://www.sciencedirect.com/science/article/pii/S0924424719303954>
164. Wang S-H et al (2021) Covid-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network. *Inf Fusion* 67:208–229. ISSN:1566-2535. <https://doi.org/10.1016/j.inffus.2020.10.004>, <https://www.sciencedirect.com/science/article/pii/S1566253520303705>
165. Wisconsin Breast Cancer Database (1992) <https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+%28original%29>
166. Wu J, Li Y, Wu Q (2019) Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning. *IEEE Access* 7:21400–21408. <https://doi.org/10.1109/ACCESS.2019.2898044>
167. Wu Y, Yi Z (2020) Automated detection of kidney abnormalities using multi-feature fusion convolutional neural networks. *Knowl-Based Syst* 200:105873. ISSN: 0950-7051. <https://doi.org/10.1016/j.knosys.2020.105873>, <https://www.sciencedirect.com/science/article/pii/S0950705120302306>
168. Xie L, Zhang L, Hu T, Huang H, Yi Z (2020) Neural networks model based on an automated multi-scale method for mammogram classification. *Knowl-Based Syst* 208:106465. Elsevier
169. Xie Y et al (2019) Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT. *IEEE Trans Med Imaging* 38.4:991–1004. <https://doi.org/10.1109/TMI.2018.2876510>
170. Xu Y, Lam H-K, Jia G (2021) MANet: A two-stage deep learning method for classification of COVID-19 from Chest X-ray images. *Neurocomputing* 443:96–105. ISSN: 0925-2312. <https://doi.org/10.1016/j.neucom.2021.03.034>, <https://www.sciencedirect.com/science/article/pii/S0925231221004021>
171. Xu J et al (2016) Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Trans Med Imaging* 35.1:119–130. <https://doi.org/10.1109/TMI.2015.2458702>
172. Xu S et al (2020) Cxnet-M3: A Deep quintuplet network for multi-lesion classification in chest X-Ray images via multi-label supervision. *IEEE Access* 8:98693–98704. <https://doi.org/10.1109/ACCESS.2020.2996217>
173. Yang Z et al (2019) EMS-Net: Ensemble of multiscale convolutional neural networks for classification of breast cancer histology images. *Neurocomputing* 366:46–53. ISSN: 0925-2312. <https://doi.org/10.1016/j.neucom.2019.07.080>, <https://www.sciencedirect.com/science/article/pii/S0925231219310872>
174. Yang X et al (2019) A two-stage convolutional neural network for pulmonary embolism detection from CTPA images. *IEEE Access* 7:84849–84857. <https://doi.org/10.1109/ACCESS.2019.2925210>
175. Yao R, Fan Y, Liu J, Yuan X (2021) COVID-19 detection from X-ray images using multi-kernel-size spatial-channel attention network. *Pattern Recognit* 119. <https://doi.org/10.1016/j.patcog.2021.108055>

176. Yu S et al (2021) Automatic classification of cervical cells using deep learning method. *IEEE Access* 9:32559–32568. <https://doi.org/10.1109/ACCESS.2021.3060447>
177. Zagoruyko S, Komodakis N (2017) Wide residual networks. arXiv:1605.07146 [cs.CV]
178. Zeiler MD, Fergus R (2013) Visualizing and Understanding Convolutional Networks. arXiv:1311.2901 [cs.CV]
179. Zeimarani B et al (2020) Breast lesion classification in ultrasound images using deep convolutional neural network. *IEEE Access* 8:133349–133359. <https://doi.org/10.1109/ACCESS.2020.3010863>
180. Zhang L et al (2017) DeepPap: Deep convolutional networks for cervical cell classification. *IEEE J Biomed Health Inf* 21.6:1633–1643. <https://doi.org/10.1109/JBHI.2017.2705583>
181. Zhao C et al (2021) Dermoscopy image classification based on StyleGAN and DenseNet201. *IEEE Access* 9:8659–8679. <https://doi.org/10.1109/ACCESS.2021.3049600>
182. Zhao X et al (2022) Automatic thyroid ultrasound image classification using feature fusion network. *IEEE Access* 10:27917–27924. <https://doi.org/10.1109/ACCESS.2022.3156096>
183. Zhou L, Gu X (2020) Embedding topological features into convolutional neural network salient object detection. *Neural Netw* 121:308–318. <https://doi.org/10.1016/j.neunet.2019.09.009>
184. Zhou Q, Zhang J, Han G, Ruan Z, Wei Y (2022) Enhanced self-supervised GANs with blend ratio classification. *Multimedia Tools and Applications*, pp 1–17
185. Zhou L et al (2020) Transfer learning-based DCE-MRI method for identifying differentiation between benign and malignant breast tumors. *IEEE Access* 8:17527–17534. <https://doi.org/10.1109/ACCESS.2020.2967820>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.