International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 **ISSN: 2582-3930**

Convolutional Neural Network: an overview and application in Radiology

Abhishek Hiwarkar[1] Adnan Neyazi[2] Raj Shingare[3]

Asif Ansari[4] Saud Ahmed[5] Kunal Giradkar[6]

Prof. Imteyaz Shahzad[7]

Department of Computer Science and Engineering

Anjuman College of Engineering and Technology

Nagpur, India

ABSTRACT

Convolution neural network, a group of artificial neural networks that has become main feature in variety of applications computer vision tasks, is attracting interest on a variety of domains, including radiology. CNN is made to automatically and adaptive modify the spatial areas of features through back propagation by using number of building blocks, such as convolution layers, pooling layers, and fully connected layers. This research article offers a vision on the basics of CNN and its application to various radiological tasks, and discusses its difficulties for implementation and future scope in the field of radiology. Two challenges in using CNN to radio logical task, small database, will be get covered in this paper, also techniques to reduce them being introduced with the concepts and advantages, as well as limitations, of CNN is essential to leverage its potential in research radiology, with the aim of augmenting the features of radiologists and improving medical care.

INTRODUCTION

Convolutional neural networks (CNNs) have revolutionized the field of image analysis, and have shown great promise in various fields, including radiology. CNNs are a type of artificial neural network that are designed to automatically and adaptively learn spatial hierarchies of features from input images, without the need for handcrafted features. This has made CNNs particularly effective in tasks such as image segmentation, classification, and detection, and has led to the development of state-of-the-art techniques in these areas.

In radiology, CNNs have been applied to a variety of tasks, including the detection of lung nodules, the segmentation of brain tumors, and the classification of breast cancer. CNNs have shown great promise in these tasks, achieving high levels of accuracy and demonstrating the potential for improving diagnosis accuracy, reducing interpretation time, and enhancing patient outcomes.

This paper provides an overview of CNNs, including their architecture, training, and application in radiology. It also discusses some of the challenges associated with the use of CNNs, such as the need for larger and more diverse datasets, the development of explainable models, and the mitigation of bias in the training data. Finally, the paper discusses future directions for the use of CNNs in radiology and medical imaging, highlighting the potential for new techniques to improve diagnosis accuracy and enhance patient outcomes.. The rest of this paper is as organized as follows. The remaining sections of the paper provide a comprehensive overview of the proposed approach, its literature review including its technical details and experimental results, research methodology, result and discussions, conclusions, discuss its potential implications for future research with its references.

LITERATURE REVIEW

Convolutional Neural Networks (CNNs) have revolutionized many areas of image analysis, including radiology. CNNs have been used in a variety of applications in radiology, including image classification, segmentation, and detection of abnormalities. In recent years, Neural Style Transfer (NST) using CNNs has emerged as a promising technique for generating stylized medical images that can aid in diagnosis and visualization.

In a study by Anirudh et al. (2019), a deep learning-based approach using CNNs was proposed for the segmentation of bone and soft tissue in CT images. The authors used a pre-trained CNN as an encoder to extract features from the CT images, and then used a decoder to reconstruct the segmented images. The results showed that the proposed method achieved high accuracy in segmenting the bone and soft tissue in the CT images.

In another study by Nie et al. (2020), a deep learning-based method using CNNs was proposed for the detection of pulmonary nodules in CT images. The authors used a CNN to extract features from the CT images, and then used a classification model to detect the pulmonary nodules. The results showed that the proposed method achieved high sensitivity and specificity in detecting the pulmonary nodules.

Similarly, in a study by Wang et al. (2018), a deep learning-based method using CNNs was proposed for the classification of breast masses in mammograms. The authors used a pre-trained CNN to extract features from the mammograms, and then used a classification model to classify the breast masses. The results showed that the proposed method achieved high accuracy in classifying the breast masses.

In the context of NST, a study by Choy et al. (2018) proposed a method for generating stylized medical images using CNNs. The authors used a pre-trained CNN to extract features from both the content image (a medical image) and the style image (a medical illustration), and then used an optimization-based approach to generate the stylized medical image. The results showed that the proposed method was able to generate stylized medical images that were visually appealing and informative.

Overall, CNNs have shown great promise in a variety of applications in radiology, including image classification, segmentation, detection of abnormalities, and NST. These studies demonstrate the potential of CNNs to aid in the diagnosis and visualization of medical images, and highlight the importance of continued research in this area.

METHODOLOGY

The objective of this study is to evaluate the performance of different neural style transfer techniques and compare their results. The research methodology includes the following steps:

1) Image Preprocessing: The content and style images are preprocessed before being fed into the CNN. This involves resizing the images, normalizing the pixel values, and converting them to tensors.

2) CNN Architecture Selection: A pre-trained CNN architecture is chosen, such as VGG-19, and the weights are imported. The choice of architecture may vary depending on the specific problem and the trade-off between performance and computational complexity.

3) Feature Extraction: The pre-trained CNN is used to extract feature maps from both the content and style images. This is done by passing the images through the CNN and recording the activations of certain layers.

4) Loss Calculation: The content loss and style loss are calculated based on the extracted features. The content loss measures the difference between the activations of a certain layer in the generated image and the content image. The style loss measures the difference between the correlations of the activations in certain layers between the generated image and the style image.

5) Total Loss Calculation: The total loss is calculated by weighting the content loss and style loss based on hyperparameters. This total loss is then used to update the pixel values in the generated image.

6) Backpropagation: The gradients of the total loss with respect to the pixel values in the generated image are computed using backpropagation.

7) Image Generation: The pixel values in the generated image are updated based on the gradients computed in step 6. This process is repeated for multiple iterations until the generated image converges.

8) Postprocessing: The final generated image is postprocessed by converting the tensor back to an image, denormalizing the pixel values, and adjusting the brightness and contrast.

It is important to note that there are many variations and improvements to CNN methodology for NST that are being researched and developed, such as multi-scale processing, instance normalization, and adaptive loss functions

In summary, the research methodology involves collecting a dataset of images, pre-processing the images, training and evaluating four different neural style transfer techniques, and comparing their results using several metrics and a human perception study. The methodology is designed to provide a comprehensive evaluation of the performance of different neural style transfer techniques and to guide the selection of the most appropriate method for different applications.

Additionally, there are many variations and improvements to CNN methodology that are being researched and developed, such as transfer learning, data augmentation, and adversarial training.

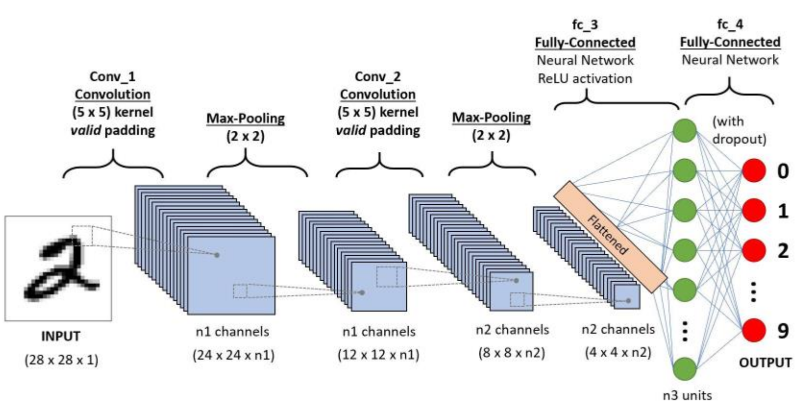
TERMINOLOGY

The mentioned terms are employed in this paper study, to avoid confusion. A Parameters in this paper stands for a variable that is automatically learned as paper study goes in deep. A hyper parameter is related to a variable which needs to be set before the training process starts for implementing CNN. A "kernel" related to the set of informative parameters used in convolution performances. A "weight" issued interchangeably used and have tried to employ this concept when referring to a parameter outside of convolution layers, which is, a kernel, for example used in fully connected layers in CNN.

What is CNN? (Fig.01 is showing the architecture)

CNN refers to type of deep learning model for processing data that is grid pattern, such as images, that is emerged by the organization of animal visual cortex and designed to automatically learn spatial hierarchies on features, from low to high-level pattern. CNN is a mathematical concept that is mainly consist of three types of layers. convolution, pooling, fully connected layer two convolution and pooling layer parts perform feature extraction, whereas the third a fully connected layer part, measure the taken out features to final output, differentiation. a small grid of parameter called kernel, feature extractor this is applied at image positions which makes CNN highly workable to image processing, feature might form anywhere in the image. So as one-layer feeds its output into the next layer-extracted features gets hierarchically and progressively more complex. The process of arrange parameter sas kernels is called training which is performed to minimize difference between outputs and ground truth labels through an optimization algorithm called back propagation and gradient descent, compared to others.

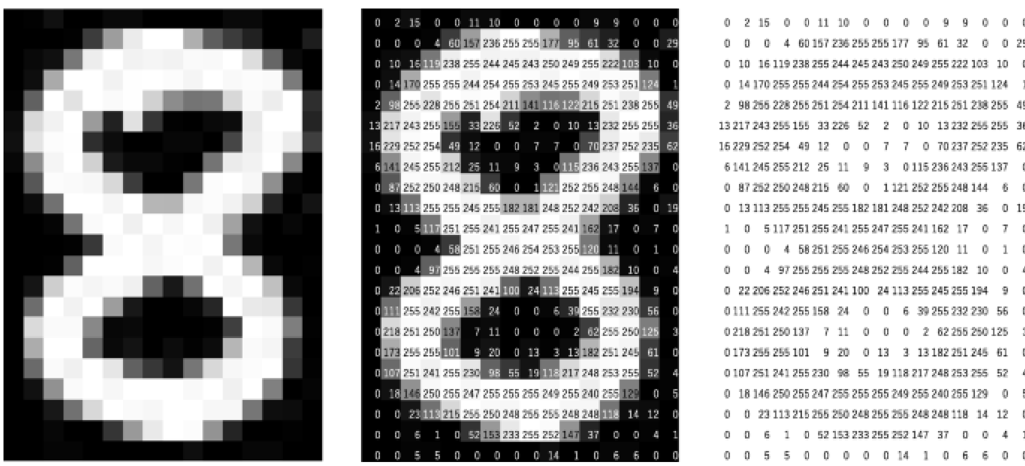
Fig. 1



International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 **ISSN: 2582-3930**

An overview of a convolutional neural network (CNN) architecture and the training process. A CNN is composed of a stacking of several building blocks: convolution layers, pooling layers (e.g. Max pooling), and fully connected (FC) layers. A model's performance under particular kernels and weights is calculated with a loss function through forward propagation on a training data set, and learn able parameters, i.e. kernels and weights, are updated according to the loss value through back propagation with gradient descent optimization algorithm. ReLU, rectified linear unit

Fig. 2



A computer records image as an array of numbers & the matrix onside the right contain numbers between 0 and 255, each of which related to the pixel brightness in the left image.

How is CNN different from other methods in radio mics?

Recent radio mics study use handmade craft feature extraction techniques, such as texture analysis follow by conventional machine learning classifiers in such as random forests and support vector machines. There are many differences to note in such method & CNN.

CNN does not require hand crafted featured extraction. CNN architecture do not necessarily require segmentation tumors or organs by human experts. CNN is way more data hungry because of its millions of parameters to perform & is more computationally expensive, resulting in requiring graphical processing units for model training.

Building blocks of CNN architecture

The CNN architecture consist of several building blocks, which are convolution layers, pooling layers & fully connected layers. An architecture contains of repetitions of stack of several convolution layers & pooling layer, connected by one or more fully connected layers. The step in which input data is transformed to output in these layers is called forward propagation which is shown in figure of convolution and pooling operations described in this paper are for 2D CNN, same operations also be performed on three-dimensional 3DCNN

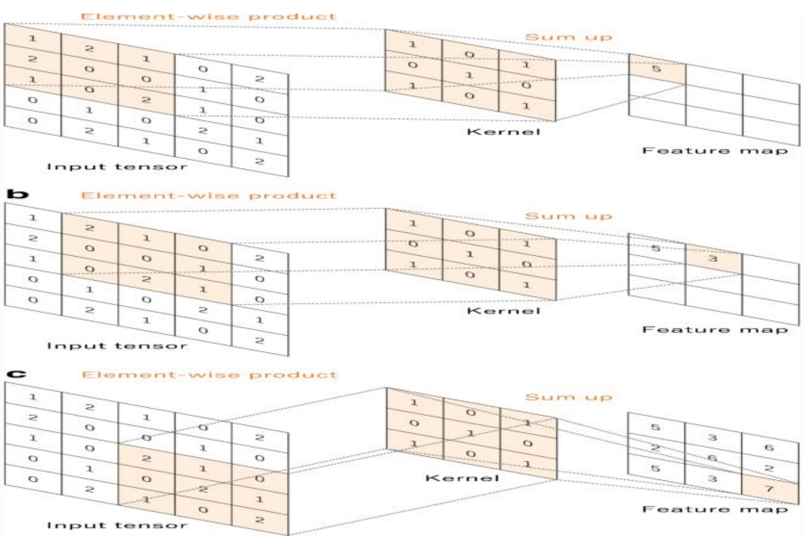
Convolution layer

A convolution layer is a basic part of the CNN architecture that performs extraction, which mainly consist of a combination of linear and nonlinear operation i.e. convolution operational function.

Convolution

Convolution is a specialized type of linear operation used for feature extraction. A small array of numbers, called a kernel, is applied across the input, which is an array of numbers called a "tensor". An element-wise product between each element of the kernel and the input tensor is calculated at each location of the tensor and mixed to obtain the output in the corresponding position of the output tensor, called a feature map (Fig. 03). This process is repeated by applying multiple kernels to form an arbitrary number of feature maps, which present differential characteristics of the input tensors. Different kernels can be considered as different feature extractors (Fig.03). Two hyperparameters define the convolution operation: size and number of kernels. The size is typically 3 × 3 but sometimes 5 × 5 or 7 × 7. This is arbitrary and determines the depth for output feature maps.

Fig:03

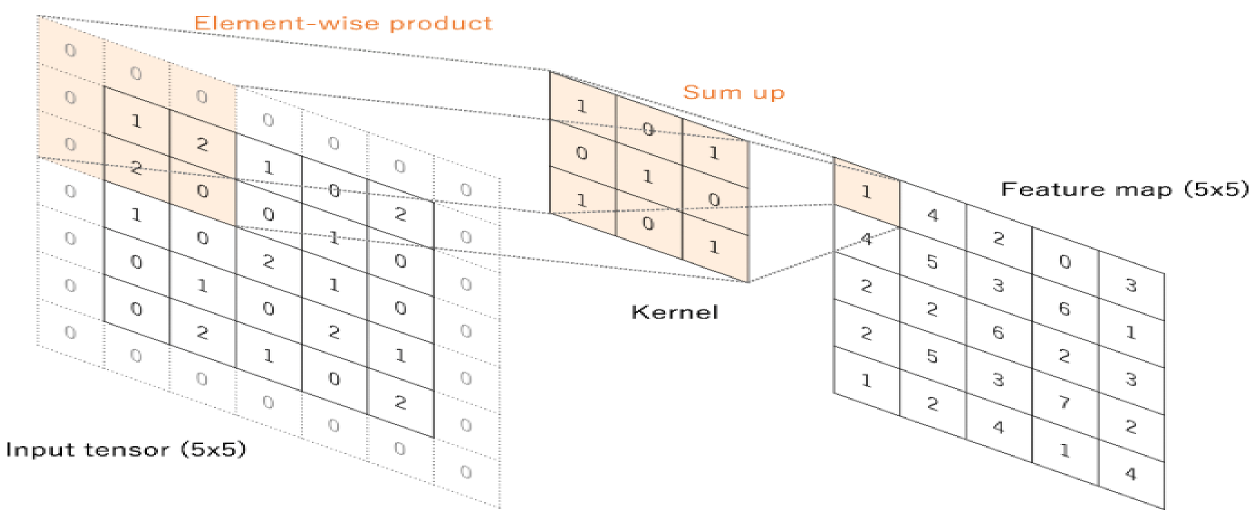


International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

An example of a convolution operation within a kernel size of 3 × 3, no padding, and a layer of 1 kernel is applied across the input tensor, and an element-wise product between each element of the kernel and the input tensor is calculated at each location and combined to obtain the output value in the corresponding position of the output tensor, called a feature map. Examples of how kernels in convolution layers are made of features from an input tensor are shown. Multiple kernels work as different feature extractors, such as a horizontal detector, a vertical edge detector (in the middle), and an outline detector (at the bottom). Note that the left image is an input, those in the middle are kernels, and those on the right are output feature maps.

The convolution operation described above does not allow the center of each kernel to overlap with the outermost element of the input tensor and reduces the height and width of the output feature map compared to the input tensor. To address this issue, a technique is used where rows and columns of zeroes are added on each side of the input tensor, so as to fit the center of a kernel on the outermost element and keep the same in-plane dimension through the convolution operation (Fig. 04). Modern CNN architectures mainly employ zero padding to retain in-plane dimensions to apply more layers. Without zero padding, each successive feature map would get smaller after the convolution operation.

Fig. 04



A convolution architecture with zero pads so as to retain in plane dimensions. Note that input dimension of 5 × 5 is kept in the output feature map. In this given example a kernel size and a stride are set as 3 × 3 and 1 respectively.

The distance between two kernel positions is called a stride, which is also defines the convolution operation. The common choice of a stride is 1. However, a stride larger than 1 is sometimes uses in order to achieve down sampling of the featured maps. optional technique to down sampling is a pooling operation as described in following.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

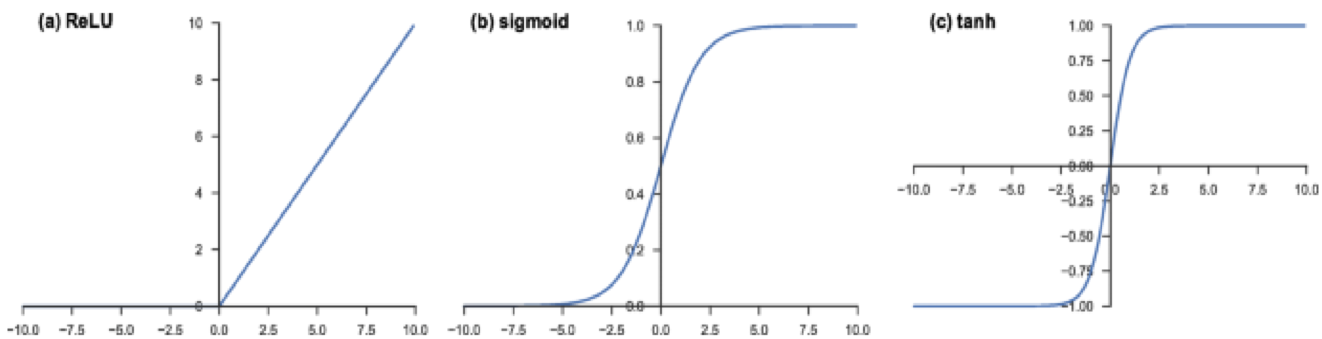
The main feature of a convolution operation is weight sharing, kernels are shared across all the image positions of the architecture. Weight sharing creates the following characteristics of convolution operations (1) letting the local feature patterns got by kernels translation invariant as kernels which travel across all the image positions and detect learned local patterns, (2) learning hierarchies of feature patterns by down sampling in conjunction with a pooling operation. Results in getting an increasingly larger field of view (3) increasing model efficiency by reducing the number of parameters to learn in comparison with fully connected neural networks over the period.

A list of parameters and hyper parameters in a convolution neural network

Nonlinear activation function

The outputs of a linear operation such as convolution are t passed through a nonlinear activation function. though smooth nonlinear function such as hyperbolic tangent function, were used previously as they are mathematical representations of a biological neuron behavior, the most common nonlinear activation function for CNN is used in present days is the rectified linear unit which simply summed the function $f(x) = \max(0, x)$.

Fig. 05



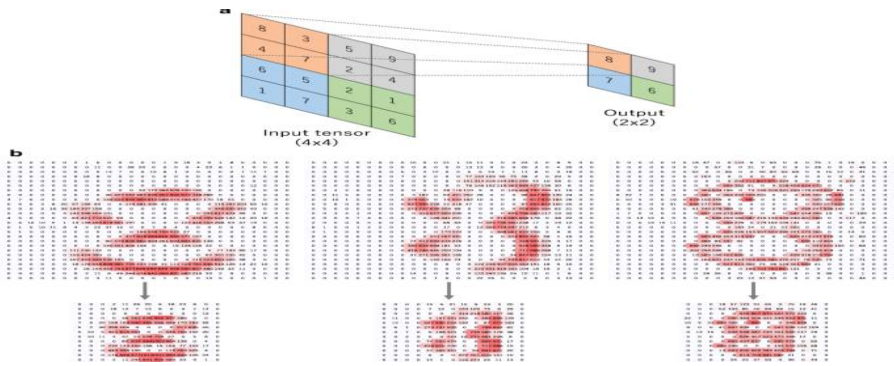
Activation functions are commonly applied to neural networks rectified linear unit and hyperbolic tangent.

Pooling layer

A pooling layer provided with a typical down sampling operations which reduces the in plane dimensional of the feature maps in order to introduce a translation invariant of all shifts and distortions, and decreases the number of subsequent learn able parameters in the frame. It is of note that there Is no search able parameters present in any of the pooling layers, where the filter size, stride, and padding are hyper parameters in pooling operations same in nature to convolution operation.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 **ISSN: 2582-3930**

Fig. 06



An example maximum pooling operation with a filter size of 2 × 2 with no padding, and a stride of 2, which extracts 2 × 2 patches which the input tensors, outputs the maximum value in each patch & discards all the other values, resulting in down sampling dimension of an input tensor by a factor of 2. example of the max pooling operation on the same images in Fig.03 Note that images in the upper row are down sampling a factor of 2 from 26 × 26 to 13 × 13.

Global average pooling

Another pooling operation believes noting is a global average pooling. A global average pooling performs a serious type of down sampling in which the feature map with size of height × width is down sampling to a 1 × 1 array by simply taking the average of all elements in each of the feature map, whereas the depth of feature maps is get back. This operation is typically applied only at one time before the fully connected layers. The advantages of applying global average pooling are as follows (1) reduce the number of parameters. (2) make it able to the CNN to accept inputs of variable size.

Fully connected layer

The output feature maps of the final convolution product i.e. transformed into a one dimensional array of numbers, and connected to one or more fully connected layers, also known as dense layers, in which every input is attached to every output by a learn weight. Once the features extracted by the convolution layers and down by the pooling layers are made, they are mapped by a subset of fully connected layers to the final outputs of the network, such as the possible forms for each class in classification task.

Last layer activation function

The activation function performed to the last fully connected layer is usually different from the other features. An activation function needs to be selected according to each task. An activation function applied to the multi class classification task is a soft max function which normalizes output real values from the last fully connected layer to target class probabilities, where each value ranges between 0 and 1 and all values sum to 1. Typical choices of the last layer activation function for various types of tasks are summarized in Table2.

list of commonly applied last layer activation functions for various tasks

Training a network

Training a network is a process of finding kernels in convolution layers & weights in fully connected layers which minimize difference between output predictions & given ground truth labels on a training data. Back propagation algorithm is the method commonly used for training neural networks where loss function and gradient descent optimization algorithm play essential role. A model performance under particular kernels & weights is calculated by a loss function through forward propagation on a training data set, learn able parameters.

Loss function

A loss function referred to as a cost function, measures the compatibility between output predictions of the network through forward propagation & given ground truth labels included. Commonly used loss function for multi class classification is cross entropy, whereas mean squared error is typically applied to regression to continuous values of data. A type of loss and need to be determined according to the given tasks.

Gradient descent

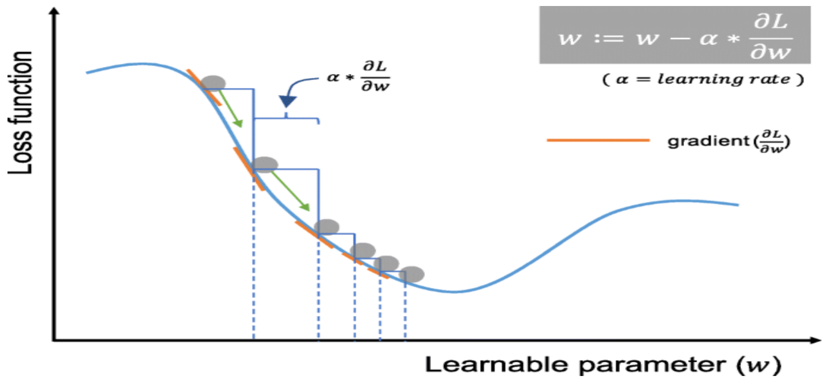
Gradient descent is commonly used in an optimization algorithm that iterative updates the learn able parameters, that is kernels and weights, of the network so as to minimize the loss. The gradient of the loss function provides us the direction in the function has the steepest rate of increase, and each learn able parameter is updated in the negative direction of the gradient with an arbitrary step size determined based on a hyper parameter called learning rate (Fig.07). The gradient is, mathematically, a partial derivative of the loss with respect to each learn able parameter, and a single update of a parameter is formulated as follows:

$$w := w - \alpha * \partial L \partial w.$$

where *w* stands for each learn able parameter, *α* stands for a learning rate, and *L* stands for a loss function. It is of note that, in practice, a learning rate is one of the most important hyper parameters to be set before the training starts. In practice, for reasons such as memory limitations, the gradients of the loss function with regard to the parameters are computed by using a subset of the training data set called mini-batch, and applied to the parameter updates. This method is called mini-batch gradient descent, also frequently referred to as stochastic gradient descent (SGD), and a mini-batch size is also a hyper parameter. In addition, many improvements on the gradient descent algorithm have been proposed and widely used, such as SGD with momentum, RMS prop, and Adam, though the details of these algorithms are beyond the scope of this article.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 **ISSN: 2582-3930**

Fig. 7



Gradient descent is an optimization algorithm that iterative updates the learn able parameters so as to minimize the loss, which measures the distance between an output prediction and a ground truth label as shown. The gradient of the loss function provides the direction in which the function has the steepest rate of increase and all parameters are updated in the negative direction of the gradient with a step size determined based on a learning rate.

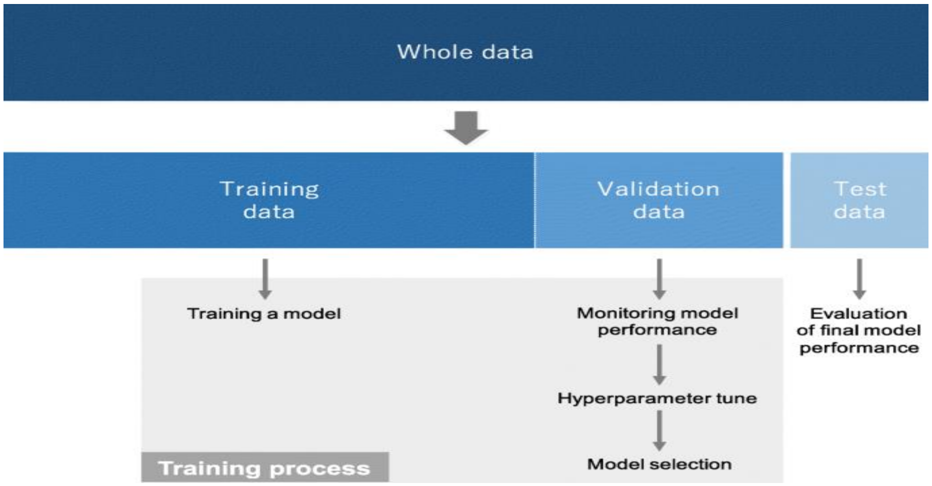
Data and ground truth labels

Data and ground truth labels are most important components in research applying deep learning or other machine learning methods. As famous proverb originating in computer science notes: “Garbage in, garbage out. careful collection of data and ground truth labels with which to train and test a model is mandatory for a successful deep learning project, but obtaining high-quality labeled data can be costly and time-consuming. While there may be multiple medical image data set open to the public, special attention should be paid in these cases to the quality of the ground truth labels.

Available data are typically split into three sets a training, a validation, and a test set (Fig.08) though there are some variants, such as cross validation. A training set is used to train a network where loss values are calculated via forward propagation and learn able parameters are updated through back propagation. A validation set is used to examine the model during the training process fine tune hyper parameters and perform model selection. A test set is ideally used only once at the very end of the project in order to evaluate the performance of the final model that was fine tune and selected on the training process with training and validation sets of given project.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Fig. 8



Available data are typically split into three sets: a training a validation, and a test set. A training set is used to train a network, where loss values are calculated via forward propagation and learn able parameters are updated through back propagation. A validation set is used to monitor the model performance during the training process fine tune hyper parameters & perform model selection. A test set is ideally used only once at the very end of the project in order to examine the performance of the final model that is fine tune and selected on the training process with training and validation sets related to project.

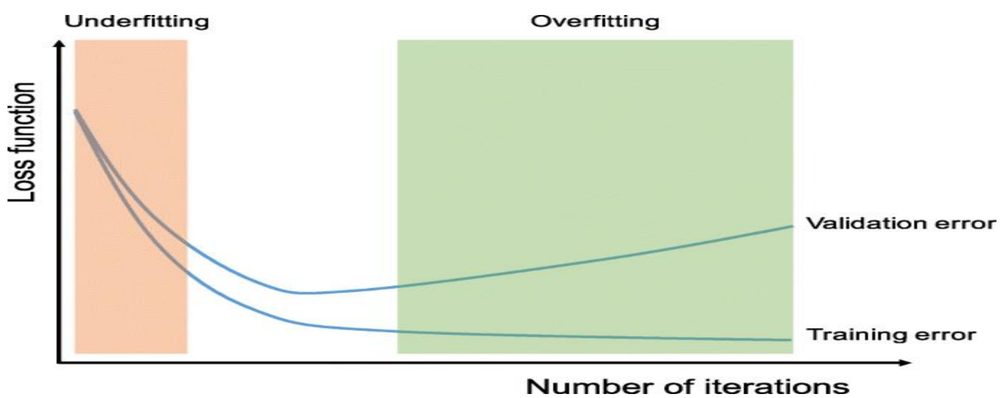
Separate validation and test sets are needed because training a model always involves fine using its hyper parameters and performing model selection. As this process is performed based on the performance on the validation set some information about this validation set leaks into the model itself, i.e. over fitting to the validation set, even though the model is never directly trained on it for the learn able parameters in project. For reason it is guaranteed that the model with fine-tuned hyper parameters on the validation set will perform well on this same validation set. Therefore, a completely unseen data set, i.e. a separate test set is necessary for the appropriate evaluation of the model performance, as what we care about is the model performance on never before seen data, i.e. generalizes.

It is worthy of mention that the term validation is used differently in the medical field and the machine learning field. As described above research done in machine learning, the term validation usually refers to a step to fine tune and select models during the training process. On the other hand, in medicine, validation generally usually stands for the process of verifying the performance of a prediction model, which is analogous to the term test in machine learning. In order to avoid this confusion, the word “development set” is sometimes used as a substitute for validation set”.

Over fitting

Over fitting refers to a situation where a model learns statistical regularities specific to the training set, i.e. ends up memorizing the irrelevant noise instead of learning the signal, therefore, performs less well on a subsequent new data set for project. This is one of the main challenges in machine learning present till the date, as an over fitted model is not generalize able to never seen before data. In that sense a test set play a pivotal role in the proper performance evaluation of machine learning model as discussed in the previous section in given research. A routine check for recognizing over fitting to the training data is to monitor the loss and accuracy on the training and validation set (Fig.09). If the model performs well on the training set compared to the validation set, then the model has likely been over fit to the training data. There have been several methods proposed to minimize over fitting (Table 03). The best solution for reducing over fitting is to obtain on given training data. A model trained on a larger data set typically generalizes better, though that is not always attainable in medical imaging. The other solutions include regularization with dropout or weight decay, batch normalization, and data augmentation as well as reducing architectural complexity. Dropout is a recently introduced regularization technique where randomly selected activities are set to 0 during the training, so that the model becomes less sensitive to specific weights in the network. Weight decay, also referred to as L2 regularization, reduces over fitting by penalizing the models weights so that the weights take only small values. Batch normalization is a type of supplemental layer which adaptive normalizes the input values of the following layer, mitigating the risk of over fitting, and also improving gradient flow through the network, allowing higher learning rates, and reducing the dependence on initialization. Data augmentation is an also effective for the reduction of over fitting, which is a process of modifying the training data through random transformations, such as flipping, translation, cropping, rotating, and random erasing, so that the model will not see exactly the same inputs during the training iterations. In spite of these efforts, there is still a concern of over fitting to the validation set rather than to the training set because of information leakage during the hyper parameter fine-tuning and model selection process. Therefore, reporting the performance of the final model on a separate (unseen) test set, and ideally on external validation data sets if applicable, is crucial for verifying the model generalize ability.

Fig:08



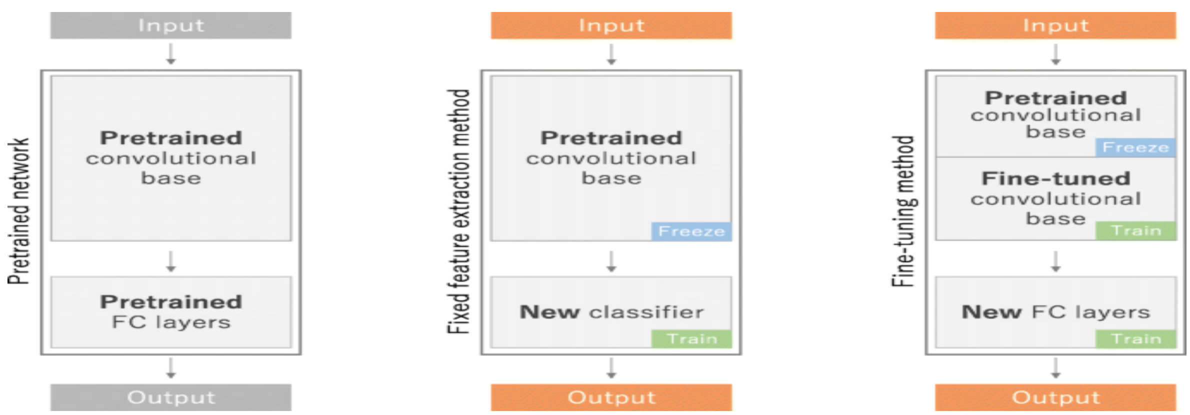
A routine check for recognizing overfitting is to monitor and check the loss on the training and validation sets during the training iteration of architecture. If the model performs well on the training set compared to the validation set, then the model has been overfit to the training data for the given situation. If the model performs poorly on both the training and validation sets, then the model is being under fit to the data for the given situation. Though the longer a network is trained, the better it performs on the training set, at some point, the network gets fit too well to the training data and loses its capability for generalization.

A list of common methods to mitigate over fitting.

Training on a small data set

An abundance of well-labeled data in medical imaging is desirable but rarely available due to the cost and necessary workload of radiology experts out there. There are a couple of techniques available to train a model efficiently on a smaller data set. Data augmentation and transfer learning as concern. As data, augmentation was briefly covered in the previous section focuses on transfer learning. Transfer learning is a common and effective strategy to train a network on a small data set where a network is depend on an extremely large data set such as Image Net, which contains 1.4 million images with 1000 classes, then reused and applied to the given task of interest which is required. The underlying assumption of transfer learning is that generic features learned on a large enough data set could be shared among all seemingly disparate data sets.

Fig. 1



Transfer learning is a common and effective strategy to train a network on a small data set, where a network is already trained on an extremely large data set, such as Image Net, then reused and applied to the given task of interest. A fixed feature extraction method is a process to remove FC layers from a already trained network and while maintaining the remaining network, which consists of a series of convolution and pooling layers, referred to as the convolution base, as a fixed feature extractor. In this scenario, any machine learning classifier, such as random forests and support vector machines, as well as the usual FC layers, can be added on top of the fixed feature in the extractor, resulting in training limited to the added classifier on a given data set of interest. A fine-tuning method, which is more often applied to radiology research, is to not only replace FC layers of the already model with a new set of FC layers to retrain them on a given data set, but to fine-tune all or part of the kernels in the already convolution base by means of back propagation. FC fully connected

A fixed feature extraction method is a process to remove fully connected layers from a network already trained on Image Net and while maintaining the remaining network, which consists of a series of convolution and pooling layers, referred to as the convolution base, as a fixed feature extractor. In this scenario, any machine learning classifier, such as random forests and support vector machines, as well as the usual fully connected layers in CNN, can be added on top of the fixed feature extractor, resulting in training limited to the added classifier on a given data set of interest. This approach is not common in deep learning research on medical images because of the dissimilarity between Image Net and given medical images.

A fine-tuning method, which is more often applied to radiology research is to not only replace fully connected layers of the already trained model with a new set of fully connected layers to retrain on a given data set, but to fine-tune all or part of the kernels in the already trained convolution base by means of back propagation. all the layer included in the convolution] base can be fine-] tuned alternatively, some earlier layers can be fixed while fine tuning the rest of the deeper layers. This is motivated by the observation that the early layer features appear more generic, including features such as edges applicable to a variety of data sets and tasks whereas later features progressively become more specific to a particular data set or task.

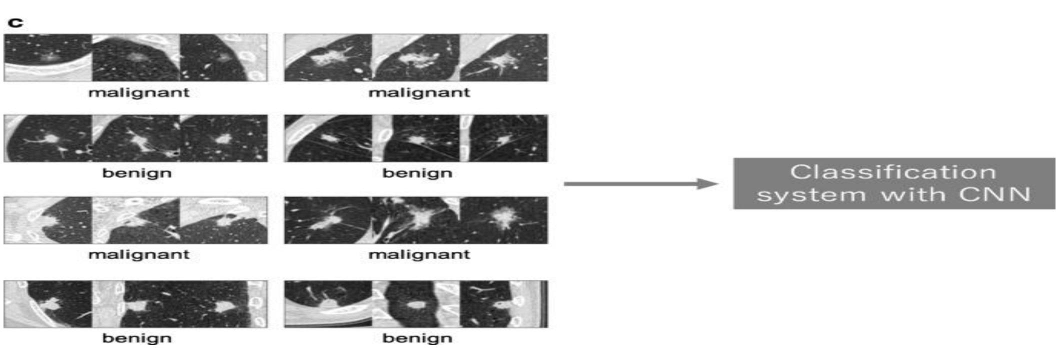
Applications in radiology

This section introduces recent applications within radiology, which are divided into the sub parts the following categories: classification, segmentation, detection, and others.

Classification

In medical image analysis, classification with deep learning generally utilizes target lesions depicted in medical images and these lesions are classified into two or more classes. For example, deep learning is frequently used for the classification of lung nodules on computed tomography (CT) images as benign or malignant (Fig. 11). As shown above it is necessary to prepare a large number of training data with corresponding labels for efficient classification using CNN. For lung nodule classification, CT images of lung nodules and their labels (i.e., benign or cancerous) are used as training data. Figure 11b, c show two examples of training data of lung nodule classification between benign lung nodule and primary lung cancer; Fig. 11(b)shows the training data where each datum includes an axial image and its label, and Fig. 11c shows the training data where each datum includes three images (axial, coronal, and images of a lung nodule) and their labels. After training CNN. the target lesions of medical images can be specified in the deployment phase by medical doctors or computer-aided detection (CAD) systems.

Fig:11



A schematic illustration of a classification system with CNN and representative examples of its training data. Classification system with CNN in the deployment phase. Training data used in training phase

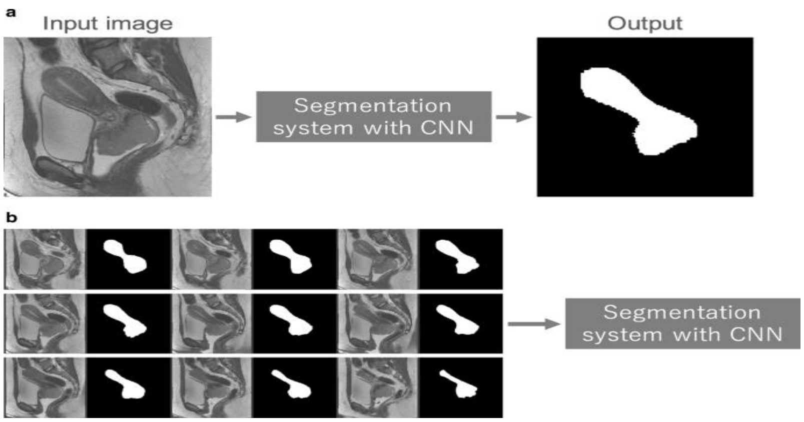
Because 2D images are frequently utilized in computer vision, deep learning networks developed for the 2D images (2D-CNN) are not directly applied to 3D images obtained in radiology [thin-slice CT or 3D-magnetic resonance imaging (MRI) images]. To apply deep learning to 3D radio logical images, different approaches such as custom architectures are used. For example, Setio et al used a multi stream CNN to classify nodule candidates of chest CT images between nodules or non-nodules in the databases of the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI). ANODE09, and the Danish Lung Cancer Screening Trial. They extracted differently oriented 2D image patches based on multi planar reconstruction from one nodule candidate (one or nine patches per candidate), and these patches were used in separate streams and merged in the fully connected layers to obtain the final classification output. One previous study used 3D-CNN for fully capturing the spatial 3D context information of lung nodules. Their 3D-CNN performed binary classification (benign or malignant nodules) and ternary classification (benign lung nodule, and malignant primary and secondary lung cancers) using the LIDCIDRI database. They used a multi view strategy in 3D-CNN, whose inputs were obtained by cropping three 3D patches of a lung nodule in different sizes and then resizing them into the same size. They also used the 3D Inception model in their 3D-CNN, where the network path was divided into multiple branches with different convolution and pooling operators.

Time series data are frequently obtained in radio logical examinations such as dynamic contrast-enhanced CT/MRI or dynamic radio isotope (RI)/positron emission tomography (PET). One previous study used CT image sets of liver masses over three phases (non-enhanced CT, and enhanced CT in arterial and delayed phases) for the classification of liver masses with 2D-CNN. To utilize time series data, the study used triphasic CT images as 2D images with three channels, which corresponds to the RGB color channels in computer vision, for 2D-CNN. The study showed that 2D-CNN using triphasic CT images was superior to that using biphasic or monophasic CT images.

Segmentation

Segmentation of organs or anatomical structures is a fundamental image processing technique for medical image analysis, such as quantitative evaluation of clinical parameters (organ volume and shape) and computer-aided diagnosis (CAD) system. In the previous section, classification depends on the segmentation of lesions of interest. Segmentation can be performed manually by radiologists or dedicated personnel, a time-consuming process. However, one can also apply CNN to this task as well. Figure 12a shows a representative example of segmentation of the uterus with a malignant tumor on MRI. In most cases, a segmentation system directly receives an entire image and outputs its segmentation result. In contrast to classification, because an entire image is inputted to the segmentation system, it is necessary for the system to capture the global spatial context of the entire image for efficient segmentation.

Fig. 12



A schematic illustration of the system for segmenting a uterus with a malignant tumor and representative examples of its training data. Segmentation system with CNN in deployment phase. Training data used in the training phase. Note that original images and corresponding manual segment are arranged next to each other.

One way to perform segmentation is to use a CNN classifier for calculating the possibility of an organ or anatomical structure which is to be examine. In this part the segmentation process is divided into two steps the first step is construction of the probability map of the organ or anatomical structure using CNN and image parts and the second is a refinement step where the global context of images and the probability map are utilized so far. One previous study used a 3D CNN classification for segmentation of the liver on 3D CT images. The input of 3D CNN were 3D image patches collected from entire 3D CT images, and the 3D CNN calculated probability cut was used for refinement of liver segmentation, based on the probability map of the liver. In this method, the local context of CT images was evaluated by 3D CNN and the global context was evaluated by the graph cut algorithm included in the process.

Though segmentation based on image patch was successfully performed in deep learning used. U net of Ronneberger outperformed the image patch based method on the ISB IEEE (The Institute of Electrical and Electronics Engineers) International Symposium on Biomedical Imaging challenge for segmentation of neuronal structures in electron microscopic images. The architecture of U net consists of a contracting path to capture anatomical context and a symmetric expanding path that enables precise localization. Though it was difficult to capture global context and local context at the same time by using the image patch Based method, U net enabled the segmentation process to incorporate a multi scale spatial context. As a result, U net could be trained end to end from a limited number of training data.

Detection

A common task for radiologists is to detect abnormalities within medical images. Abnormalities can be rare and they must be detected among many normal cases. One previous study investigated the usefulness of 2D-CNN for detecting tuberculosis on chest radio graphs. The study utilized two different types of 2D-CNN, Alex Net and Google Net, to detect pulmonary tuberculosis on chest radio graphs. To develop the detection system and evaluate its performance, the data set of 1007 chest radio graphs was used. According to the

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 **ISSN: 2582-3930**

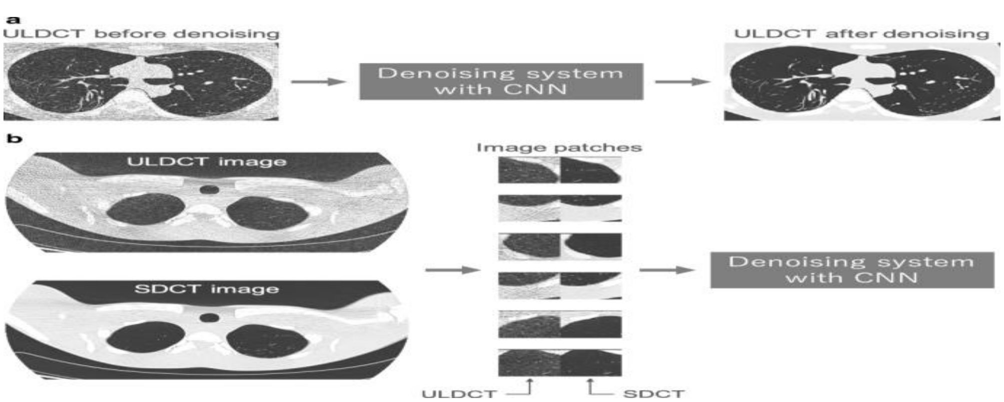
results, the best area under the curve of receiver operating characteristic curves for detecting pulmonary tuberculosis from healthy cases was 0.99, which was obtained by ensemble of the Alex Net and Google Net 2D-CNNs.

Nearly 40 million mammography examinations are performed in the USA every year. These examinations are mainly performed for screening programs aimed at detecting breast cancer at an early stage. A comparison between a CNN-based CAD system and a reference CAD system relying on hand-crafted imaging features was performed previously. Both systems were trained on a large data set of around 45,000 images. The two systems shared the candidate detection system. The CNN-based CAD system classified the candidate based on its region of interest, and the reference CAD system classified it based on the hand-crafted imaging features obtained from the results of a traditional segmentation algorithm. The results show that the CNN-based CAD system outperformed the reference CAD system at low sensitivity and achieved comparable performance at high sensitivity.

Others

Low-dose CT has been increasingly used in clinical situations. For example, low-dose CT was shown to be useful for lung cancer screening. Because noisy images of low-dose CT hindered the reliable evaluation of CT images, many techniques of image processing were used for denoising low-dose CT images. Two previous studies showed that low-dose and ultra-low-dose CT images could be effectively denoised using deep learning. Their systems divided the noisy CT image into image patches, denoised the image patches, then reconstructed a new CT image from the denoised image patches. Deep learning with encoder–decoder architecture was used for their systems to denoise image patches. data for the denoising systems consisted of pairs of image patches, which are obtained from standard-dose CT and low-dose CT. Figure13 shows a representative example of the training data of the systems.

Fig. 13



A schematic diagram of the system for denoising an ultra-low dose CT (ULDCT). image of phantom and representative examples of its training data included. Designing system with CNN in deployment phase.

One previous study on U-net to solve the inverse problem in imaging for obtaining a noiseless CT image reconstructed from a sub sampled sinogram. To train U-net for reconstructing a noiseless CT image from the sub sampled sinogram, the training data of U-net consist of (i) noisy CT images obtained from sub sampled sinogram by filtered back projection (FBP) and (ii) noiseless CT images obtained from the original sinogram. The study suggested that, while it would be possible to train U-net for reconstructing CT images directly from the sinogram, performing the FBP first greatly simplified the training. As a refinement of the original U-net, the study added a skip connection between the input and output for residual learning. Their study showed that U-net could effectively produce noiseless CT images from the sub sampled sinogram.

Although deep learning requires a large number of training data, building such a large-scale training data of radio logical images is a challenging problem. One main challenge is the cost of annotation (labeling); the annotation cost for a radio logical image is much larger than a general image because radiologist expertise is required for annotation. To tackle this problem, one previous study utilized radiologists' annotations which are routinely added to radiologists' reports (such as circle, arrow, and square). The study obtained 33,688 bounding boxes of lesions from the annotation of radiologists' reports. Then, unsupervised lesion categorization was performed to speculate labels of the lesions in the bounding box. To perform unsupervised categorization, the following three steps were iteratively performed: (i) feature extraction using pretrained VGG16 model from the lesions in the bounding box, (ii) clustering of the features, and (iii) fine-tuning of VGG16 based on the results of the clustering. The study named the labels obtained from the results of clustering as pseudo-category labels. The study also suggested that the detection system was built using the Faster R-CNN method, the lesions in the bounding box, and their corresponding pseudo-category. The results demonstrate that detection accuracy could be significantly improved by incorporating pseudo-category labels.

Radiologists routinely produce their reports as results of interpreting medical images. Because they summarize the medical images as text data in the reports, it might be possible to collect useful information about disease diagnosis effectively by analyzing the radiologists' reports. One previous study. evaluated the performance of a CNN model, compared with a traditional natural language processing model, in extracting pulmonary embolism findings from chest CT. By using word embedding, words in the radio logical reports can be converted to meaningful vectors. For example, the following equation holds by using vector representation with word embedding: king – man + woman = queen. In the previous study, word embedding enabled the radio logical reports to be converted to a matrix (or image) of size 300×300. By using this representation, 2D-CNN could be used to classify the reports as pulmonary embolism or not. Their results showed that the performance of the CNN model was equivalent to or beyond that of the traditional model.

RESULTS AND DISCUSSIONS

Results:

The study provides an overview of convolutional neural networks (CNNs) and their application in radiology. The authors discuss the architecture of CNNs, including the convolutional, pooling, and fully connected layers. They also explore various types of CNNs, such as AlexNet, VGG, GoogLeNet, and ResNet.

The study demonstrates the application of CNNs in radiology, particularly in medical image analysis. The authors discuss the use of CNNs for image classification, segmentation, and detection in various medical imaging modalities, including X-ray, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET).

The results of the study show that CNNs can improve the accuracy of medical image analysis tasks and reduce the workload of radiologists. For instance, the authors discuss how CNNs can detect and classify lung nodules in CT images with a high degree of accuracy. They also demonstrate the use of CNNs for brain tumor segmentation in MRI images.

Discussions:

The study highlights the potential of CNNs in radiology and medical image analysis. The authors note that CNNs can assist radiologists in making accurate diagnoses, particularly for complex

cases. Moreover, CNNs can reduce the workload of radiologists by automating repetitive and time-consuming tasks, such as image segmentation and classification.However, the authors also highlight some challenges associated with the use of CNNs in radiology. One of the primary challenges is the need for large annotated datasets to train CNNs effectively. Moreover, CNNs are vulnerable to overfitting, which can result in reduced generalization performance.

Future research can explore the development of novel CNN architectures and training techniques to address these challenges. Additionally, the authors suggest that CNNs can be combined with other machine learning techniques, such as transfer learning and ensemble methods, to improve their performance further.

Overall, the study demonstrates the potential of CNNs in radiology and medical image analysis. With continued research and development, CNNs can provide innovative solutions to current challenges in radiology and improve patient outcomes.

CHALLENGES AND FUTURE DIRECTIONS

Challenges:

1) Annotated datasets: One of the primary challenges associated with the use of CNNs in radiology is the need for large annotated datasets to train the networks effectively. Annotating medical images is a time-consuming and laborious process, and it can be difficult to obtain sufficient training data, particularly for rare conditions.

2) Over fitting: CNNs are vulnerable to over fitting, which can result in reduced generalization performance. Over fitting occurs when a model becomes too complex and learns to memorize the training data instead of generalizing to new data.

3) Interpretability: CNNs can be difficult to interpret, particularly for radiologists who are not familiar with deep learning. This lack of interpretability can make it challenging to understand the underlying factors that contribute to a diagnosis or image classification

Future Directions

1) Integration with Electronic Health Records (EHRs): Future research can explore the integration of CNNs with EHRs to enable more personalized and efficient diagnosis and treatment. EHRs can provide additional patient data, such as medical history and laboratory results, which can be incorporated into CNNs to improve their accuracy and effectiveness.

2) Multi-modal fusion: CNNs can be trained on data from multiple modalities, such as CT and MRI, to provide a more comprehensive understanding of a patient's condition. Future research can explore how to effectively integrate and fuse data from multiple modalities to improve diagnostic accuracy and reduce false positives.

3) Real-time processing: CNNs can be computationally intensive, which can limit their use in real-time processing. Future research can explore the development of more efficient CNN architectures and training techniques that can enable real-time processing of medical images.

4) Automated reporting: CNNs can be used to generate automated reports based on medical images, which can reduce the workload of radiologists and enable faster diagnoses. Future research can explore the development of natural language processing techniques that can convert the output of CNNs into meaningful reports.

5) Clinical decision support: CNNs can be used as a clinical decision support tool to assist radiologists in making accurate diagnoses. Future research can explore how to effectively integrate CNNs into clinical decision support systems to provide real-time feedback and improve the overall quality of care.

Overall, the future directions of CNNs in radiology and medical image analysis are promising, with potential for further development and integration into clinical practice. Continued research and innovation can lead to improved patient outcomes and more efficient healthcare delivery.

Here's a comparing the state-of-the-art neural style transfer techniques based on their key features and performance:

Neural Style Transfer (NST)

Key Features-Basic technique that applies the style of one image onto the content of another image using a pre-trained CNN.

Performance-Fast but limited in flexibility and artistic control. Can result in "blurry" or "patchy" images.

Fast Neural Style Transfer (FST)

Key Feautures-Improves the speed of NST by replacing the Gram matrix computation with a more efficient method.

Performance-Faster than NST but limited in artistic control. Can produce "smoother" images.

Adaptive Instance Normalization (AdaIN)

Key Features-Introduces adaptive normalization to allow for more flexible and precise style transfer.

Performance-Provides greater artistic control and flexibility compared to NST and FST.Can result in high-quality images with precise stylization.

StyleGAN and StyleGAN2

Key Features-GAN-based techniques that allow for more advanced and diverse stylization.

Performance-Provides advanced control over style and content, resulting in high-quality images. Can generate diverse and complex styles, but may require more computational resources.

Overall, the state-of-the-art neural style transfer techniques continue to advance and offer greater artistic control, flexibility, and quality of results. The choice of technique depends on the specific application and requirements, such as speed, computational resources, and desired level of artistic control.

CONCLUSION

In conclusion, convolutional neural networks have become a powerful tool for image analysis and have shown great promise in various fields, including radiology. They have proven to be effective in tasks such as image segmentation, classification, and detection, and have achieved state-of-the-art results on several benchmarks. Furthermore, the development of transfer learning techniques has allowed for the application of pre-trained CNNs to new tasks with limited labeled data.

Despite the success of CNNs, there are still challenges that need to be addressed, such as the need for larger and more diverse datasets, the development of explainable models, and the mitigation of bias in the training data. Furthermore, there is a need for more research into the development of CNNs that can handle multi-modal data and generate high-quality images with limited data.

In the future, the use of CNNs in radiology is expected to grow, with the potential for the development of new techniques that can improve diagnosis accuracy, reduce interpretation time, and enhance patient outcomes. Moreover, CNNs can also be applied to other medical fields such as pathology and ophthalmology. Overall the use of convolutional neural networks has revolutionized the field of image analysis and holds great potential for further advancements in the field of radiology and medical imaging.

FUTURE SCOPE

1) Development of hybrid CNN models for radiology: Hybrid CNN models combine the strengths of different CNN architectures to improve their accuracy and efficiency. Future research can explore the potential of hybrid CNN models in radiology and investigate the effectiveness of different combinations of CNN architectures.

2) Application of CNNs in advanced imaging techniques: CNNs have shown great potential in analyzing various types of medical images, including MRI, CT, and PET. Future research can focus on applying CNNs in advanced imaging techniques, such as functional MRI and diffusion tensor imaging, to improve the accuracy of diagnosis and treatment planning.

3) Development of CNN-based systems for image segmentation: Image segmentation is a crucial step in many radiological applications, such as tumor detection and measurement of organ volumes. CNNs have shown promising results in automated image segmentation, and future research can focus on developing more accurate and efficient CNN-based systems for image segmentation.

4) Integration of CNNs with other AI techniques for radiology: CNNs can be combined with other AI techniques, such as reinforcement learning and transfer learning, to improve their accuracy and efficiency. Future research can investigate the potential of integrating CNNs with other AI techniques for radiology and explore the effectiveness of different integration strategies.

5) Development of explainable CNN models for radiology: Explainability is a critical aspect of AI models in medical applications, as it helps to build trust and ensure patient safety. Future research can focus on developing explainable CNN models for radiology and investigating the interpretability of these models.

6) Multi-modal CNNs for radiology: Radiological images are often accompanied by other types of data, such as clinical data and genomic data. Multi-modal CNNs can integrate these different types of data to improve the accuracy of diagnosis and treatment planning. Future research can explore the potential of multi-modal CNNs in radiology and investigate the best ways to integrate different types of data.

7) CNNs for personalized medicine in radiology: Personalized medicine is an emerging field that aims to tailor medical treatments to individual patients based on their unique characteristics. CNNs can play a key role in personalized medicine by analyzing radiological images and other patient data to predict treatment outcomes and guide treatment planning. Future research can focus on developing CNN-based systems for personalized medicine in radiology.

ABBREVATIONS

1D: One-dimensional**.**

2D: Two-dimensional**.**

3D: Three-dimensional**.**

CAD: Computer-aided diagnosis.

CADe: Computer-aided detection**.**

CAM: Class activation map**.**

CNN: Convolutional neural network**.**

CT: Computed tomography.

FBP: Filtered backprojection.

GAN: Generative adversarial network.

GPU: Graphical processing unit**.**

IEEE: The Institute of Electrical and Electronics Engineers**.**

INLSVRC: ImageNet Large Scale Visual Recognition Competition.

ISBI: IEEE International Symposium on Biomedical Imaging.

LIDC-IMDRI: Lung Image Database Consortium and Image Database Resource Initiative**.**

MRI: Magnetic resonance imagine.

PET: Positron emission tomography**.**

RELU: Rectified linear unit**.**

SDC: **Stochastic gradient descent.**

REFERENCES

1)Published on: 22nov,2020, convolutional neural networks: an overview and application in radiology, Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do & Kaori Togashi.

2) Published on: 11march, 2019, Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach: Sakshi Indolia, Anil Kumar Goswami, S.P. Mishra, Pooja Asopa.

3)Published on: IEEE ,14feb, 2018, A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects.

4) Published on: 11may,2018, Recent advances in convolutional neural networks. Author Jiuxian Gu, ZhenhuaWang, JasonKuen, LianyangMa, Amir Shahroudy, BingShuai, TingLiu, Xingxing Wang, Gang Wang, Jianfei Cai, Tsuhan Chenga.

5) Published on: researchGate,2017,3feb, Convolutional Neural Network.

6)Published on: 2020 4march, PA Review of Convolutional Neural Network, Arohan Ajit; Koustav Acharya; Abhishek Samanta.

7) **Introduction to Convolutional Neural Networks Jianxin Wu LAMDA Group National Key Lab for Novel Software Technology Nanjing University, China wujx.**

8) Published on: 2019,15may, A Convolutional Neural Network for Modelling Sentences, Nal Kalchbrenner, Edward Grefenstette, Phil Blunsom.

9) Published on: 2020,3july, Sparse Convolutional Neural Network, Proceedings of the IEEE Conference on Computer Vision and Pattern Reorganization.