UNIVERSITÉ DE BRETAGNE OCCIDENTALE

DOCTORAL THESIS

Tomographic Image Reconstruction with Neural Networks

Author:
Venkata Sai Sundar KANDARPA

Supervisor:
Dr. Dimitris VISVIKIS

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

LATIM
Biologie Santé

September 6, 2021

Declaration of Authorship

I, Venkata Sai Sundar KANDARPA, declare that this thesis titled, "Tomographic Image Reconstruction with Neural Networks" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."

Dave Barry

UNIVERSITÉ DE BRETAGNE OCCIDENTALE

Abstract

Biologie Santé
Biologie Santé

Doctor of Philosophy

Tomographic Image Reconstruction with Neural Networks

by Venkata Sai Sundar KANDARPA

Neural Networks are extensively used in the field of medical imaging for biomedical image segmentation, cancer diagnosis, image analysis, etc. The advancements in computation power (GPUs) and efficient memory utilization have propelled the spread of deep neural networks into various domains. The main motivation behind the use of neural network approaches is faster prediction (compared to traditional methods) without compromising on the quality of the result. Tomographic image reconstruction has also benefited from the development of neural networks. Medical image reconstruction involves the task of mapping raw measurement data collected by the detector to images that are comprehensible to a radiologist. A medical image reconstruction algorithm essentially approximates this mapping to predict the best possible image. There are established analytical and iterative reconstruction algorithms which have over the years proven to be effective in producing the best image possible. Convolutional neural networks (CNN) specifically have proven to be exceptional in tasks related to images such as denoising, deblurring, and super-resolution. The use of neural networks in Positron Emission Tomography (PET) and Computed Tomography (CT) reconstruction has been explored in this thesis. Novel frameworks called DUG-RECON (Double U-Net Generator) for PET, CT image reconstruction, and LRR-CED (Low-Resolution Reconstruction aware Convolutional Encoder-Decoder) for Sparse-view CT image reconstruction and Total-Body PET image reconstruction are proposed in this manuscript. Quantitative analysis of the images reconstructed with the proposed methods indicated that image quality was either better or on par with standard reconstruction algorithms.

Acknowledgements

The acknowledgments and the people to thank go here, don't forget to include your project advisor...

Contents

List of Figures

List of Tables

List of Abbreviations

LAH	**L**ist **A**bbreviations **H**ere
WSF	**W**hat (it) **S**tands **F**or

Physical Constants

Speed of Light $c_0 = 2.997\,924\,58 \times 10^8\,\mathrm{m\,s^{-1}}$ (exact)

List of Symbols

a	distance	m
P	power	W ($\mathrm{J\,s^{-1}}$)
ω	angular frequency	rad

For/Dedicated to/To my…

Chapter 1

Introduction

1.1 Motivation

The use of deep learning in medical imaging has been on the rise over the last few years. It has widely been used in various tasks across medical imaging such as image segmentation (Ronneberger, Fischer, and Brox, 2015; Guo et al., 2019; Sinha and Dolz, 2019; Dolz et al., 2018; Hatt et al., 2018), image denoising (Kadimesetty et al., 2018; Li et al., 2020; Chen et al., 2017; Yang et al., 2018), image analysis (Litjens et al., 2017; Amyar et al., 2019; Cui et al., 2018). Deep learning based algorithms produce faster results along with best possible quality in accordance with existing state of the art methods (Leuschner et al., 2021). Medical Image reconstruction too has benefited hugely with the advancement of deep learning (Reader et al., 2020; Zhang and Dong, 2020). Medical Image reconstruction corresponds to the task of mapping raw projection data retrieved from the detector to image domain data. During the course of this thesis, the focus has been towards PET and computed tomography (CT) image reconstruction. Both these modalities present a unique of set of challenges for image reconstruction.

PET imaging is a form of emission tomography wherein the image reconstruction task revolves around identifying the radio-tracer distribution emitted from the patient. A PET image gives functional information about the organs in a patient making it invaluable for oncology. Some of the challenges in PET image reconstruction are scatter, attenuation and difficulty in identifying the exact annihilation point of the electron-positron. Despite being the most sensitive emission tomography modality, the number of photons captured is low relative to the photons emitted contributing to further image degradation. These challenges result in very noisy images when reconstructed with analytical algorithms. These challenges are addressed by Iterative/Model-based approaches which take into account detector geometry, noise statistics

and approximate scatter and attenuation correction resulting in better image quality.

CT imaging on the other hand is an example of transmission tomography. The extent of attenuation undergone by X-Rays that pass through a patient are measured to obtain attenuation maps. In CT imaging research, there has been active interest in sparse-view and low-dose reconstruction scenarios. In both cases, severe artifacts are introduced in reconstructed images either due to incomplete projections or low counts. Many established model-based iterative methods account for the low-dose and sparse-view settings to remove artifacts and noise from the reconstruction (Nuyts et al., 1998; Elbakri and Fessler, 2002; Liu et al., 2013). However, these methods for require the knowledge of the noise and artifacts statistics and generally have longer reconstruction times (Kim, Ramani, and Fessler, 2014).

The main tasks involved in image reconstruction can be broadly categorized into three: sinogram correction, domain translation from sinogram to image, and image correction. Algorithms either tackle the three task individually or simultaneously account for them. One can relate to these tasks in the domain of Computer Vision wherein deep learning architectures have revolutionized the field by producing the state of the art results in most applications (Guo et al., 2016). For example, effective use of deep learning-based methods is seen in dealing with image denoising (Kadimesetty et al., 2018; Li et al., 2020; Chen et al., 2017; Yang et al., 2018), super resolution (Ledig et al., 2017; Lim et al., 2017) and image-to-image translation (Isola et al., 2017; Zhu et al., 2017) tasks. The continuous improvement in the availability of public data has further propelled interest in data-driven medical image reconstruction making it an active area of research. This thesis aims to explore novel deep learning approaches for PET and CT image reconstruction. Most common ways to introduce deep learning architectures in the image reconstruction pipeline are for pre-processing to correct raw projection data from the detector and post-processing to improve images reconstructed with existing methods. Another way is to embed the network into an iterative algorithm to enable faster convergence. The relatively less explored way called direct image reconstruction is to utilize neural networks alone for the entire reconstruction process. In this thesis CNN approaches are proposed for direct image reconstruction with neural networks.

1.2 Thesis Organization

This thesis is divided into six chapters with the first two chapters being introduction and literature review, followed by three chapters that focus on different deep learning methods explored during the thesis, and finally conclusion and perspectives. In the introduction various aspects of PET and CT image reconstruction are discussed along with the relevant background in deep learning background. The second chapter throws light on deep learning applied to medical image reconstruction and reviews the state of the art approaches in the scope of this thesis. In chapter 3 , we discuss reconstruction framework double U-Net generator (DUG) for PET and CT image reconstruction. A novel method for Sparse-view CT reconstruction called low-resolution reconstruction aware convolutional encoder decoder (LRRCED) is covered in chapter 4. A modified version of LRRCED for total body PET is discussed in chapter 5. Potential improvements and ideas for future work are presented in the final chapter.

1.3 Imaging Modalities and Reconstruction Models

1.3.1 PET

PET images provide functional information to the radiologist making them invaluable in image analysis. The application of PET imaging has been on the rise in oncology, cardiology and neuropsychiatry. The increased application lead to the development of many novel reconstruction approaches that lead to improved image quality. This section focuses on the standard analytical and iterative algorithms which are applicable in the context of this thesis.

The aim of image reconstruction in PET is to predict the tracer distribution emitted from the patient. The emission is a result of positron emitting radionuclide injected into the patient which causes positron-electron annihilation. The annihilation results in the production of gamma photons that travel in opposite directions due to the law of conservation of momentum. The simultaneous detection of these photons (also called coincidence events) enables the estimation of tracer distribution in PET imaging. A PET scanner detects the coincidence events through a set of detectors arranged in a circular fashion. This design of the scanner facilitates detection of coincidence photons between a pair of detectors (d_p and d_q). The centers of two detectors

are connected by a straight line called LOR. Photon pairs that are not subject to scatter are a result of annihilation events that occur along a thin volume surrounding the LOR. In PET, $\boldsymbol{x} = \boldsymbol{\lambda}$ is the distribution of a radiotracer delivered to the patient by injection, and is measured through the detection of pairs of γ-rays emitted in opposite directions (indirectly from the positron-emitting radiotracer).

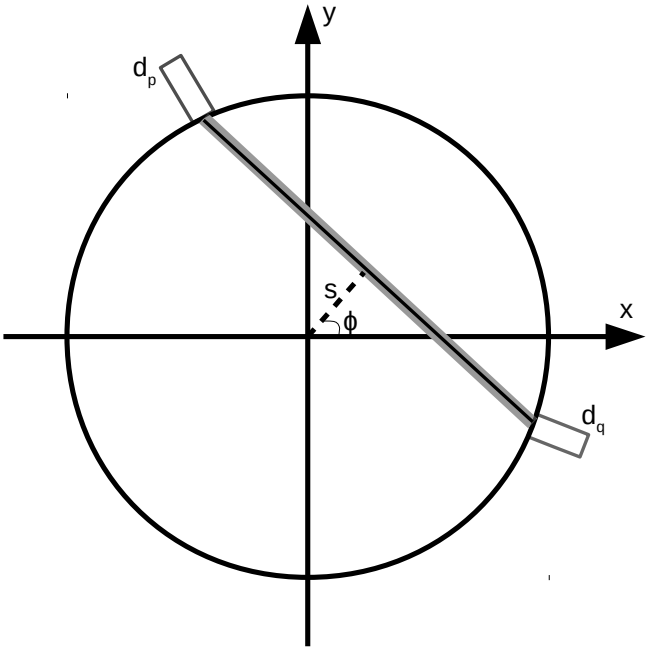

FIGURE 1.1: Depiction of a circular PET detector with detectors d_p and d_q connected with a LOR indicated in gray.

The measurement $\boldsymbol{y}$ is a random vector modeling the number of detection (photon counting) at each of the n detector bins, and follows a Poisson distribution with independent entries:

$$\boldsymbol{y} \sim \text{Poisson}(\bar{\boldsymbol{y}}(\boldsymbol{x})) \tag{1.1}$$

where $\bar{\boldsymbol{y}}(\boldsymbol{x}) \in \mathbb{R}^n$ is the expected number of counts (noiseless), which is a function of the image $\boldsymbol{x}$.

The expected number of counts is

$$\bar{\boldsymbol{y}}(\boldsymbol{\lambda}) = \boldsymbol{P}\boldsymbol{\lambda} \tag{1.2}$$

where $\boldsymbol{P} \in \mathbb{R}^{n \times m}$ is a system matrix such that each entry $[\boldsymbol{P}]_{i,j}$ represents the probability that a photon pair emitted from voxel j. Image reconstruction is

achieved by finding a suitable image $\hat{\boldsymbol{x}} = \hat{\boldsymbol{\lambda}}$ that approximately solves

$$\boldsymbol{y} = \bar{\boldsymbol{y}}(\boldsymbol{x})\,. \tag{1.3}$$

1.3.2 CT

Let an image be represented by $\boldsymbol{x} \in \mathbb{R}^m$ and the scanner measurement by $\boldsymbol{b} \in \mathbb{R}^n$ where m is the number of voxels and n is the number of measurements. In two-dimensional (2-D) CT imaging n depends on the number of detectors N_d and the number of angles N_a. The task of medical image reconstruction corresponds to finding a mapping from $\boldsymbol{b}$ to $\boldsymbol{x}$. The measurement $\boldsymbol{b}$ is a random vector modeling the number of detection (photon counting) at each of the n detector bins, and follows a Poisson distribution with independent entries, i.e.,

$$\boldsymbol{b} \sim \mathrm{Poisson}(\bar{\boldsymbol{b}}(\boldsymbol{x})) \tag{1.4}$$

where, $\boldsymbol{b} = [b_1(\boldsymbol{x}), \ldots, b_n(\boldsymbol{x})]^\top \in \mathbb{R}^n$ and $\bar{\boldsymbol{b}}(\boldsymbol{x}) = [\bar{b}_1(\boldsymbol{x}), \ldots, \bar{b}_n(\boldsymbol{x})]^\top \in \mathbb{R}^n$ is the expected number of counts (noiseless), which is a function of the image $\boldsymbol{x}$.

The image $\boldsymbol{x} \in \mathbb{R}^m$ is a vectorized input image (also referred to as attenuation) representing the measure of X-rays absorbed or scattered as they pass through the patient. In a monochromatic setting, the expected number of counts $\bar{\boldsymbol{b}}(\boldsymbol{x})$ is given by the Beer-Lambert law, i.e.,

$$\bar{b}_i(\boldsymbol{x}) = I \cdot \exp(-[\boldsymbol{P}\boldsymbol{x}]_i) \quad \forall i = 1, \ldots, n \tag{1.5}$$

where, I is the intensity and $\boldsymbol{P} \in \mathbb{R}^{n \times m}$ is a system matrix such that each entry $[\boldsymbol{P}]_{i,j}$ represents the contribution of the j-th image voxel to the i-th detector. Given the raw projections $\bar{\boldsymbol{b}}$, we take the logarithm as follows

$$y_i = \log\left(\frac{I}{b_i}\right) \quad \forall i = 1, \ldots, n \tag{1.6}$$

where we assumed that the intensity I is sufficiently high so that $b_i > 0$ for all i. Image reconstruction is based on finding a suitable image $\hat{\boldsymbol{x}}$ that approximately solves

$$\boldsymbol{y} = \boldsymbol{P}\hat{\boldsymbol{x}} \tag{1.7}$$

where $\boldsymbol{y} = [y_1, \ldots, y_n]^\top \in \mathbb{R}^m$. The reconstruction can also be achieved with more sophisticated iterative techniques that account for the stochastic

properties of the measurement (1.1) Nuyts et al., 1998; Elbakri and Fessler, 2002.

In a sparse-view setting, the number of rotation angles of the detector is decreased in order to reduce the radiation passing through the patient. This implies a reduction in the number of projection angles in the measurement $\boldsymbol{y}$.

1.3.3 Analytical Reconstruction Algorithms

Analytical algorithms can efficiently solve (1.3) and (1.7) and have been the cornerstone of tomographic image reconstruction. One of the most famous reconstruction algorithms for both PET and CT is the filtered-backprojection (FBP). The projections ($\boldsymbol{y}$) are first filtered (typically with a ramp filter) and then back-projected to get an image. The discrete implementation of the FBP can be written as follows:

$$\boldsymbol{x}(i,j) = \frac{\pi}{N_\phi} \sum_{l=0}^{N_\phi - 1} \boldsymbol{y}_f(s = i\cos\phi_l + j\sin\phi_l, \phi_l) \tag{1.8}$$

where x is the image for a set of pixels (i,j), $\boldsymbol{y}_f$ are the filtered projections obtained by filtering the projections, expressed in terms of radial variable s and projection angle ϕ, and N_ϕ number of projection angles. The above equation is the approximation of backprojection by a discrete quadrature.

Analytical methods are faster to implement and practical in a clinical setting but they are vulnerable to noise. The assumptions made in analytical formations are that the measurements are continuous and the solutions are of integral formulation. Sampling is done to the data a posteriori. They are also highly susceptible to system geometry. Since the 80's, model-based iterative reconstruction (MBIR) techniques Shepp and Vardi, 1982; Fessler, Sonka, and Fitzpatrick, 2000 became the standard approach. They consist in iteratively approximating a solution $\hat{\boldsymbol{x}}$ such that $\bar{\boldsymbol{y}}(\hat{\boldsymbol{x}})$ maximizes the likelihood of the measurement $\boldsymbol{y}$. As they model the stochasticity of the system, they are more robust to noise as compared with FBP, and can be completed with a penalty term for additional control over the noise De Pierro, 1995.

1.3.4 Iterative Reconstruction Algorithms

These algorithms are modeled on discrete representation of measurement and the image. They also incorporate corrections for scatter and are independent of detector geometry. Two different MBIR methodologies, namely

MLEM and WLS for PET and CT image reconstruction respectively, are discussed in this section.

MLEM

The key components of an iterative method are the data model, cost function and optimization. The cost function for MLEM is based on Poisson likelihood given as follows:

$$\Pr\{\bar{\boldsymbol{y}} \mid \boldsymbol{x}\} = \prod_{j=1}^{N_{LoR}} \exp\left(-\langle y_j \rangle\right) \langle y_j \rangle^{y_j} / y_j! \tag{1.9}$$

Putting 1.2 in 1.9, taking log and dropping terms that do not depend on unknown image x we get the cost function for MLEM,

$$Q(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) = \sum_{j=1}^{N_{LOR}} \left\{ -\sum_{i=1}^{m} P_{j,i} x_i + y_j \log\left(\sum_{i=1}^{m} P_{j,i} x_i\right) \right\} \tag{1.10}$$

where Q is the cost function and the definition of other variables is consistent from above. As long as the matrix P is singular, the above cost function remains convex and results in a unique image. The update step to map from the current estimate x^n to the next estimate x^{n+1} can be written as follows:

$$\bar{\boldsymbol{x}}_i^{n+1} = x_i^n \frac{1}{\sum_{j'=1}^{N_{LoR}} P_{j',i}} \sum_{j=1}^{N_{LOR}} P_{j,i} \frac{y_j}{\sum_{i'=1}^{m} P_{j,i'} x_{i'}^n} \quad i = 1, \ldots, m \tag{1.11}$$

The initial estimate x_i^1, $i = 1, \ldots, m$ typically follows a uniform distribution. The denominator with sum over index i' is the forward projection operation. Hence it estimates the measured data for the current image estimate. The numerator with sum over index j is the back projection over the ratio of measured and estimated data. The MLEM algorithm does not include a prior and it converges to the image that best fit the data. This estimate has inherent instabilities as the fitting is done closely to the noisy measured data. Various strategies like adding a Bayesian prior to the cost function, filtering the measured data or image post reconstruction, are employed to improve image quality.

WLS

One of the most common iterative techniques for CT image reconstruction is the WLS method, the image $\hat{x}$ is estimated by minimizing the following:

$$\hat{x} = \arg\min_{x \succeq 0} \frac{1}{2} \|y - Ax\|_W^2 \tag{1.12}$$

where $W = \text{diag}\{w_i\}$ is the diagonal weighting matrix that constitutes for the variance of each ray and $\|z\|_W^2 = z'Wz$. Despite the statistical weighting, due to the noise in measurements and ill-conditioned problem of image reconstruction the image estimate will still be noisy.

An improvement to the above can be brought by finding a balance between the desired a priori characteristics of the image and the data fitting. This balance is realized through a regularized WLS cost function:

$$\hat{x} = \arg\min_{x \succeq 0} \Psi(x), \quad \Psi(x) \triangleq \frac{1}{2} \|y - Ax\|_W^2 + \beta R(x) \tag{1.13}$$

where $R(x)$ is the regularization term, that enforces piece-wise smoothness on the image and β the regularization parameter. The usual choice for R is an edge preserving regularizer that penalized the differences between neighboring voxels:

$$R(x) = \sum_{j=1}^{n_p} \sum_{k \in \mathcal{N}_j} \psi_{jk} \left(x_j - x_k \right) \tag{1.14}$$

where N_j are the set of neighboring indices of the j^{th} voxel. ψ is a potential function that controls the penalization of differences in the neighboring voxels.

Sparse-view CT image reconstruction problem has been explored during the course of this thesis. Sparse-view CT image reconstruction is an underdetermined problem due to the limited number of projection data available for reconstruction. In such a scenario stronger forms of regularization like total variation (TV) are utilized. Consider a 2D digital image $x[p, q]$, discrete form of TV can be written as:

$$\sum_p \sum_p |x[p, q] - x[p-1, q]| + |x[p, q] - x[p, q-1]| \tag{1.15}$$

This is an an-isotropic version of TV regularization. Images reconstructed with TV, sometimes have inexplicable artifacts due to the fact that the absolute value potential is not differentiable at 0.

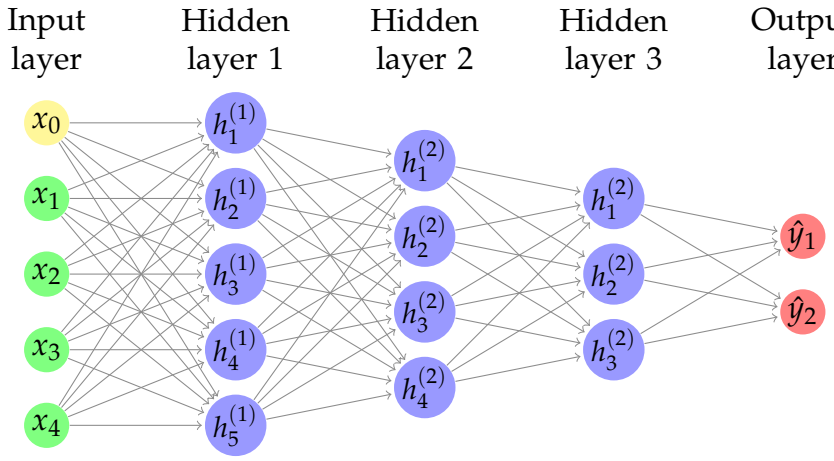

FIGURE 1.2: Depiction of a neural network with an input layer, three hidden layers and an output layer

1.4 Neural networks

Neural networks also known as artificial neural networks (ANN) are machine learning algorithms that form the basis of deep learning. They inherit name and structure from neurons in the brain. Biological neurons transmit signals to one another through complex networks. This interconnected networking is realized though various architectures by ANNs. Sets of artificial neurons are stacked on top of each other to form a layer. A typical neural network consists of many such layers that are connected to each other. The first layer is called input layer, the final layer is termed output layer and the layers in-between are called hidden layers. A neural network with three hidden layers is depicted in Fig 1.2. The transmission of data across the nodes or artificial neurons happens through the connections. Each and every node has a specific weight and threshold associated. The output from a node is passed through the connection only if the value is above the threshold. Neural network approaches are data-driven. Their performance improves as they learn through training on a dataset.

To further understand the working of a neural network, we can imagine each node to be solving the problem of linear regression. For example consider a node with four inputs ($x_i, i = 1, 2, \ldots 4$), four weights ($w_i, i = 1, 2, \ldots 4$) and a bias:

$$\sum_{i=1}^{m} w_i x_i + \text{ bias } = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 + \text{ bias } \tag{1.16}$$

The output of the node is the above summation after going through an activation function a:

$$\text{output } = a(x) = \begin{cases} 1 \text{ if } \sum w_1 x_1 + b \geq 0 \\ 0 \text{ if } \sum w_1 x_1 + b < 0 \end{cases} \tag{1.17}$$

In the above example, the given activation function of this node propagates the value 1 only when the weighted sum of it's inputs is non-negative. When the condition of an activation function are met the output of this node becomes an input to the node to which it's connected. Due to the process of forwarding values through a network, an ANN is also called feed-forward network. Complex networks with multiple layers of these nodes are used in practical tasks. An important category of machine learning task is supervised learning. It involves training a neural network on labeled datasets. The goal of training a neural network is to minimize a cost function that enforces the closeness of predicted and real output labels. During the training the network reorganizes it's weights based on the loss function. This process of updating weights is called optimization. Each update is aimed at reaching a minimum of the loss function. A popular optimization method is gradient descent. It guides the model in the direction of reducing errors to reach an optima. The development of back-propagation (Rumelhart, Hinton, and Williams, 1986) has been instrumental in successful implementation of optimization algorithms for neural networks.

1.4.1 Convolutional Neural Networks

The neural network depicted in 1.2 is an example of densely connected network, where all the neighboring nodes are connected with one another. As the size of data increases (say large image data), and the network becomes more complex, the number of parameters increases exponentially. To address this and also to be more suitable for image data CNNs were formulated. CNNs are extensively used in computer vision tasks like image classification, object detection, image segmentation (Voulodimos et al., 2018). The three main building blocks of a CNN are Convolution, Activation and Pooling. Each of these layers are discussed below:

1.4.2 Convolution

Images are digitally stored in the form of 2D or 3D matrices depending on the format. A convolution kernel (also known as filter) is a matrix that operates on these images and transforms them based on the kernel values. These kernel values are also known as weights in the neural network terminology. Typically, the size of the kernel is much smaller than that of the image. Many sets of these kernels form the convolution layer of the CNN. The movement of the kernel over the image can be made either by a single pixel or multiple pixels. This step size is called stride (s). The resulting output of a convolution between filter and image is called a feature map. Consider a kernel h and input image f with m rows and n columns. Convolution between h and f results in a feature map g:

$$g[m,n] = (h * f)[m,n] = \sum_i \sum_j h[i,j] f[m-i, n-j] \tag{1.18}$$

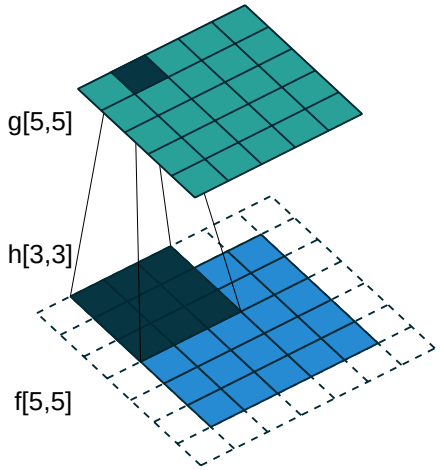

FIGURE 1.3: Convolution of an input image of dimensions 5×5 with a filter of dimensions 3×3.

Given in Fig 1.3 is a representation of the convolution operation. Zero padding is used to manipulate the dimensions of the feature maps. In the above Figure above it is indicated with dotted lines. The function of padding here is to maintain same dimensions in the input image f and the feature map g. A CNN learns features from the input through many convolutional layers. The earlier layers learn general features like edges, contrast, the deep layers learn more abstract and finer details.

1.4.3 Activation Layer

The activation layer that follows the convolution layer in a CNN is most commonly the ReLU activation function, depicted in Fig 1.4.

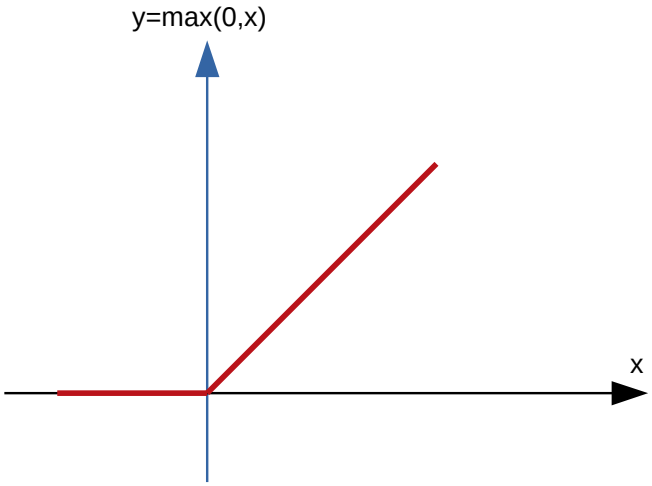

FIGURE 1.4: The ReLU function

Many of the task based on images are non-linear in nature. Whether a computer vision task like identifying objects in an image or a medical imaging task involving tumor detection, the relationships are far from being linear. The ReLU function increases this required non-linearity in the CNN.

1.4.4 Pooling Layer

The third building block of a CNN is the pooling layer. Pooling operation is mainly used to reduce the dimensions of a tensor which enables faster computation. Max pooling is the most commonly used pooling operation. A max pooling operator of a particular size returns the maximum value of a selected region in the feature map. Similar to a filter it is implemented with a specific stride. A max pooling filter with $s = 2$ is depicted in Fig 1.5.

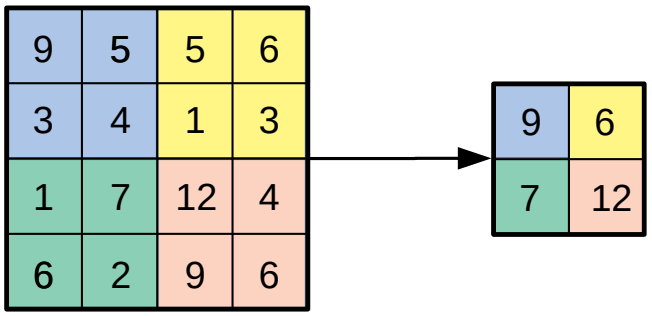

FIGURE 1.5: Max pooling with 2×2 filter and stride 1

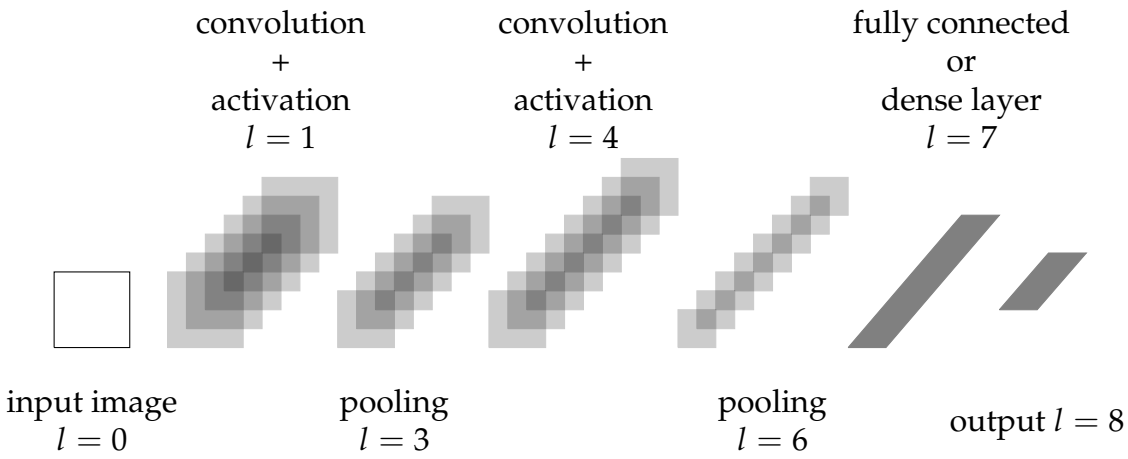

FIGURE 1.6: Architecture of a typical CNN. This representation was first proposed by LeCun and Bengio, 1995.

A CNN with 2 convolutional layers, 2 activation layers and 2 pooling layers is represented in Fig 1.6. Layer number is given by l. The first and the last layer are the input and output respectively. Usually the last set of layers in a CNN used for classification, regression tasks is a fully-connected layer which is similar to the neural network represented in Fig 1.2.

1.4.5 Fully Convolutional Networks

Image to image translation tasks require the CNN to map from image in one domain to an image in another related domain. This requires the design of the CNN to be quite different from the one depicted in Fig 1.6. The shape of the output in an image to image translation task like reconstruction is an image. In a fully convolutional network

Appendix A

Frequently Asked Questions

A.1 How do I change the colors of links?

The color of links can be changed to your liking using:

`\hypersetup{urlcolor=red}`, or

`\hypersetup{citecolor=green}`, or

`\hypersetup{allcolor=blue}`.

If you want to completely hide the links, you can use:

`\hypersetup{allcolors=.}`, or even better:

`\hypersetup{hidelinks}`.

If you want to have obvious links in the PDF but not the printed text, use:

`\hypersetup{colorlinks=false}`.

Bibliography

[Amy+19] A Amyar et al. "3-D RPET-NET: development of a 3-D PET imaging convolutional neural network for radiomics analysis and outcome prediction". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.2 (2019), pp. 225–231.

[Che+17] Hu Chen et al. "Low-dose CT denoising with convolutional neural network". In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE. 2017, pp. 143–146.

[Cui+18] Sunan Cui et al. "Artificial Neural Network With Composite Architectures for Prediction of Local Control in Radiotherapy". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.2 (2018), pp. 242–249.

[Dol+18] Jose Dolz et al. "HyperDense-Net: a hyper-densely connected CNN for multi-modal image segmentation". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 38.5 (2018), pp. 1116–1126.

[DP95] A. R. De Pierro. "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography". In: *IEEE Transactions on Medical Imaging* 14.1 (1995), pp. 132–137.

[EF02] I. A. Elbakri and J. A. Fessler. "Statistical image reconstruction for polyenergetic X-ray computed tomography". In: *IEEE Transactions on Medical Imaging* 21.2 (2002), pp. 89–99.

[FSF00] J. A. Fessler, M. Sonka, and J. M. Fitzpatrick. "Statistical image reconstruction methods for transmission tomography". In: *Handbook of medical imaging* 2 (2000), pp. 1–70.

[Guo+16] Yanming Guo et al. "Deep learning for visual understanding: A review". In: *Neurocomputing* 187 (2016), pp. 27–48.

[Guo+19] Zhe Guo et al. "Deep learning-based image segmentation on multimodal medical imaging". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.2 (2019), pp. 162–169.

[Hat+18] Mathieu Hatt et al. “The first MICCAI challenge on PET tumor segmentation”. In: *Medical image analysis* 44 (2018), pp. 177–195.

[Iso+17] Phillip Isola et al. “Image-to-image translation with conditional adversarial networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134.

[Kad+18] Venkata S Kadimesetty et al. “Convolutional neural network-based robust denoising of low-dose computed tomography perfusion maps”. In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.2 (2018), pp. 137–152.

[KRF14] Donghwan Kim, Sathish Ramani, and Jeffrey A Fessler. “Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction”. In: *IEEE Transactions on Medical Imaging* 34.1 (2014), pp. 167–178.

[LB+95] Yann LeCun, Yoshua Bengio, et al. “Convolutional networks for images, speech, and time series”. In: *The handbook of brain theory and neural networks* 3361.10 (1995), p. 1995.

[Led+17] Christian Ledig et al. “Photo-realistic single image super-resolution using a generative adversarial network”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 4681–4690.

[Leu+21] Johannes Leuschner et al. “Quantitative Comparison of Deep Learning-Based Image Reconstruction Methods for Low-Dose and Sparse-Angle CT Applications”. In: *Journal of Imaging* 7.3 (2021), p. 44.

[Li+20] Meng Li et al. “SACNN: Self-Attention Convolutional Neural Network for Low-Dose CT Denoising with Self-supervised Perceptual Loss Network”. In: *IEEE Transactions on Medical Imaging* (2020).

[Lim+17] Bee Lim et al. “Enhanced deep residual networks for single image super-resolution”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017, pp. 136–144.

[Lit+17] Geert Litjens et al. “A survey on deep learning in medical image analysis”. In: *Medical image analysis* 42 (2017), pp. 60–88.

[Liu+13] Y. Liu et al. "Total variation-Stokes strategy for sparse-view X-ray CT image reconstruction". In: *IEEE Transactions on Medical Imaging* 33.3 (2013), pp. 749–763.

[Nuy+98] John Nuyts et al. "Iterative reconstruction for helical CT: a simulation study". In: *Physics in Medicine & Biology* 43.4 (1998), p. 729.

[Rea+20] Andrew J Reader et al. "Deep learning for PET image reconstruction". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 5.1 (2020), pp. 1–25.

[RFB15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241.

[RHW86] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. "Learning representations by back-propagating errors". In: *nature* 323.6088 (1986), pp. 533–536.

[SD19] Ashish Sinha and Jose Dolz. "Multi-scale guided attention for medical image segmentation". In: *arXiv preprint arXiv:1906.02849* (2019).

[SV82] L. A. Shepp and Y. Vardi. "Maximum Likelihood Reconstruction for Emission Tomography". In: *IEEE Transactions on Medical Imaging* 1.2 (1982), pp. 113–122.

[Vou+18] Athanasios Voulodimos et al. "Deep learning for computer vision: A brief review". In: *Computational intelligence and neuroscience* 2018 (2018).

[Yan+18] Qingsong Yang et al. "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 37.6 (2018), pp. 1348–1357.

[ZD20] Hai-Miao Zhang and Bin Dong. "A review on deep learning in medical image reconstruction". In: *Journal of the Operations Research Society of China* (2020), pp. 1–30.

[Zhu+17] Jun-Yan Zhu et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 2223–2232.