Research Project
Computer Vision for Medical Imaging

Improving Brain MRI quality using Deep Neural Networks: Multi-Contrast Super Resolution

Author : Alexandre Bodinier
Promotion 2021
Computer Science (minor in AI)

Supervisors :
Dr Wenjia Bai (Imperial College London)
Dr Natalia Díaz Rodríguez (ENSTA Paris)

From 18/05/2020 to 07/08/2020

Non-confidential

Data Science Institute,
Imperial College London,
South Kensington Campus,
Exhibition Road,
London SW7 2AZ

Abstract

Super Resolution is a class of algorithms that increase the resolution of a given image. This post-processing technology is extremely useful in Medical Imaging because it overcomes hardware or practical limitations. For example, acquiring high resolution MR images is really expensive and time consuming. However, brain tumor diagnosis and treatment highly depends on Magnetic Resonance images quality and increasing its resolution is therefore of importance.

The purpose of this study is to propose a Super Resolution model fitted for Brain Magnetic Resonance Imaging (MRI) data. To this extent, Deep Convolutional Neural Networks will be investigated.
First, this work aims at establishing of a first Single-Contrast Super Resolution model adapted from state of the art architectures.
Then, the model is improved using Multi-Contrast Super Resolution.

The proposed models have achieved better performances on the conducted experiments in both qualitative and quantitative terms. It does not only exceeds Single-Contrast methods (+0.4dB on PSNR in average) but also presents good learning properties.

Keywords: Super Resolution - 2D Convolutional Neural Networks - Brain MRI

Acknowledgements

I would like to thank Mr. Bai for allowing me to join his team at Imperial College London and for his supervision all along the project. He introduced me to Research thinking and always gave me pertinent and rigorous advice. He involved me into the medical vision reading group where I had exciting discussions with the other researchers, who were always kind and welcoming.

I would also like to thank Chengliang and my sister Barbara for showing me how to use the cluster at Imperial College.

Finally, I would like to thank my family and friends for the moral support in periods of doubt.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

With 227,000 deaths worldwide in 2016 and a five-year survival rate that barely reaches 33%, **Brain Tumors** are amongst the most deadly cancers.[1]
Cancer is formed by a group of abnormal cells growing uncommonly fast and with potential to spread to parts of the body.
As brain tumors are hidden inside the brain cavity, they require medical imaging techniques to be diagnosed.

In practice, various neuro-imaging protocols are employed for (i) diagnosis, (ii) monitoring of the tumour's progression and treatment planning, and (iii) assessment of the treatment effects. **Magnetic resonance imaging (MRI)** is widely used in both clinical routine and research studies. It facilitates tumour analysis by allowing estimation of the extent, location and investigation of its sub-components [2].
MRI have the particularity to simultaneously capture multiple contrasts of a same scene (like multiple version of a same image). Some brain abnormalities such as inflammations or tumors are easier to be diagnosed in specific contrasts.

The accuracy of tumor diagnosis and monitoring highly depends on MR images quality. Unfortunately, capturing a high resolution MR image takes a lot of time and the trade-off between scanning time, costs and patient comfort often leads to low resolution acquisitions. It has been demonstrated that stronger magnetic fields lead to better Signal to Noise Ratio (SNR) performances, but this involves huge expenses for hospitals who already suffer from dangerous budget cuts. Also, it could present risks for the patient. [3]

Fortunately, loads of post-processing techniques are emerging to restore a decent image quality. Among them: denoising, demosaicing, contrast-correction or **Super Resolution**. Super Resolution is a class of algorithms aiming at increasing image resolutions.

1.2 Motivation

Super Resolution has received increasing attention from the research community in recent years.
In many fields, it has allowed to transcend physical or material limitations (the diffraction limit of optical imaging systems like microscopy or imaging sensors limits in photography) [4].
Briefly, Super Resolution is very useful for technologies where it is getting harder and harder to improve quality (because of the required industrial precision or of material costs for instance).

1.3 Project Scope

Super Resolution gathers a wide range of algorithms able to increase resolution (defined later in this report). I focus here on a category of algorithms that has been showed to outperform the others: **Deep Learning-based** algorithms and in particular Convolutional Neural Networks.
After demonstrating how Convolutional Neural Networks are capable of increasing image resolution we will identify the most relevant architectures for the Super Resolution task. Finally, we will try to take full advantage of MRI specificity, Multi-Contrast acquisition by proposing a Multi-Contrast Super Resolution model.

1.4 Contribution to the field

My personal contributions to the field will be:

- The proposal of a novel method for Brain MRI Super Resolution based on Multi-Contrast Super Resolution which outperforms existing techniques including Bicubic interpolation or Single-Contrast EDSR.
- An evaluation of the contribution of each MRI contrast for Super Resolution and the impacts of Multi-Contrast on the network training.

The idea is as follows: *To reconstruct a specific MRI contrast, why not use the information contained in other contrasts?*

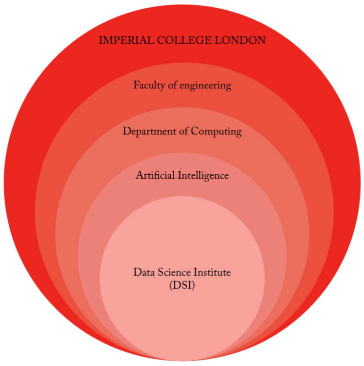
1.5 Structure of the thesis

Chapter one is the introduction of the thesis. It is followed by the Contextual Setting in chapter 2 which relates the organization of the project. Then, the preliminary, chapter 3, will present the fundamental topics of this project: **Magnetic Resonance Imaging** and **Super Resolution**. Then, building upon previous research and established models, we will construct a Super Resolution model using Convolutional Neural Networks in chapter 4. We will then evaluate its performances and analyze the results in chapter 5. Afterwards, a novel method based on Multi-Contrast Super Resolution will be proposed to further increase the performances in chapter 6. Finally, chapters 7 and 8 will conclude the research project and outline future directions.

Chapter 2

Contextual Setting

2.1 About the Data Science Institute (DSI)



Formed in April 2014, **DSI** is one of the six global institutes of Imperial College London: Fig 2.1.
Based on the observation that current issues are often too big to be tackled by individual academics or even departments, it promotes multi-disciplinary collaboration.
The DSI gathers a wide variety of researchers in Analytics, Biomedical Informatics, Image Informatics, and Visualization who work hand in hand.
Its main objectives are:

- Developing data management and analysis technologies and services that support research.
- Translating data science into innovation through collaborations with industrial partners and public institutions, and to support commercialisation.
- Promoting data science and its applications outside academia and to influence policy makers.

2.2 Team

Throughout my project, I was supervised by Dr Bai, lecturer jointly at *Data Science Institute* and *Department of Brain Sciences* at Imperial College.
Dr Bai introduced me to his team of 5 PhD students and post doctoral researchers (all working on computer vision).
Every Monday I attended the *Medical Vision Reading Group* where someone from the group presented a recent paper related to their own project. I actively participated to the reading group, notably by presenting two articles on Super Resolution. At the end of my traineeship, I also had the opportunity to give a presentation on my own research project to the team, who provided me useful feedback.

2.3 Computing resources

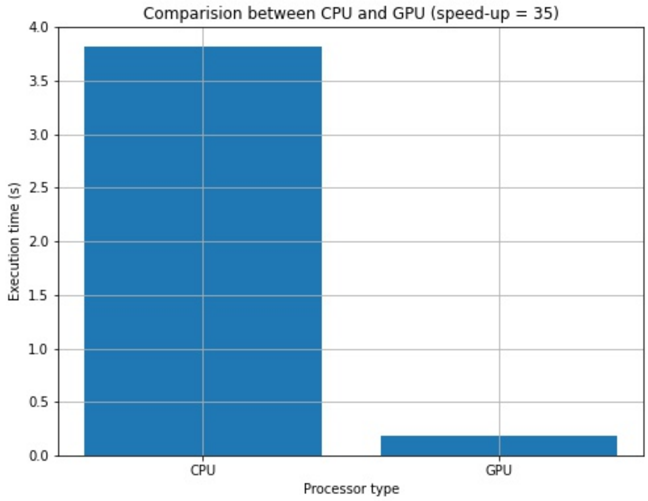


Figure 2.1: Execution time comparison: CPU (Central Processing Unit) vs GPU (on a basic convolution operation)

Training a Deep Neural Network involves heavy computational costs. Training time can thus be significantly reduced using adapted computing resources like powerful (and expensive!) GPUs (Graphics Processing Units) (Fig 2.1).
A solution is to use cloud services that provide access to high performance computing resources (HPC).

I used two cloud services:

- **Google Colab**: User friendly, it enables to run jobs directly (no queuing needed) on Google GPUs. But there are some limitations with a free membership (12h max per job, memory allocation, random GPU allocation (which leads to erratic execution times).
- Imperial College's cluster, the **HPC (High Performance Computing)**: HPC uses a PBS queuing system which, depending on the resources requested, can take several hours before the job starts running. It is then important to carefully check the code before. It is less user friendly than Colab (accessing HPC requires a VPN, all instructions are made using command line and shell scripts). But it has far less resource limitations.

Depending on my needs, I used either one or the other.
All codes are implemented with PyTorch.

2.4 Organization of the project

As instructed by Dr Bai, the project was really my own to define, lead, and organize. A simplified Gantt diagram is showed at Fig 2.2 Unfortunately, due to the Covid-19 pandemic, the research project was carried out remotely, which made interactions with the research team more complicated.
Each week was punctuated by three checkpoints:

- A weekly written report to share my progress and issues with Mr Bai.
- The Medical Vision Reading Group on Mondays.
- A bi-weekly online meeting. It was a privileged time to discuss further approaches.

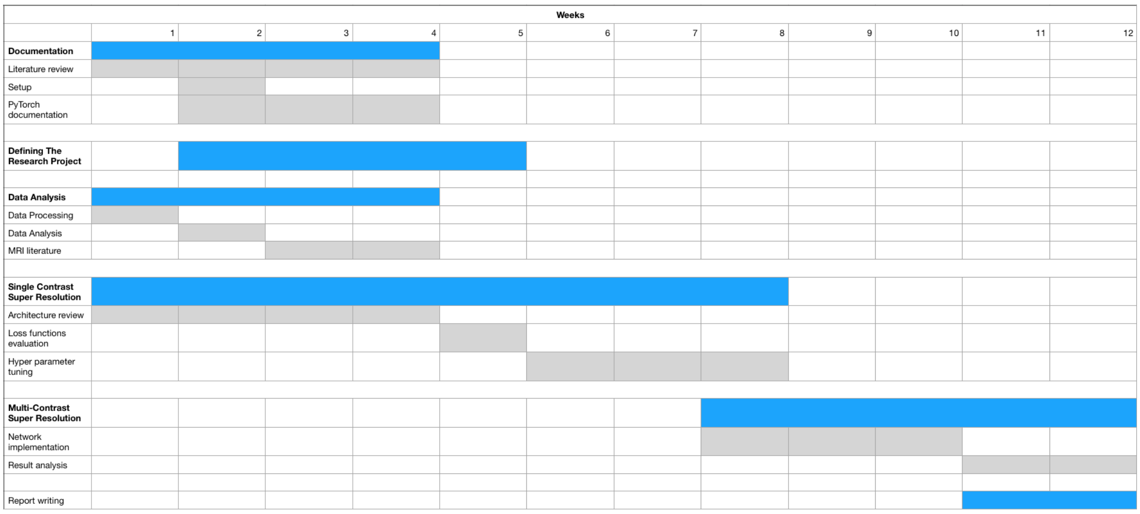


Figure 2.2: Schedule of the project

Chapter 3

Preliminaries

3.1 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a medical imaging technique widely used in neuroimaging. It allows for intracranial disease diagnosis (such as tumors or injuries).

It has two majors assets: it is considered to be **clearer** (it produces better images) but also **safer** than related techniques because it does not involve X-rays or ionizing radiation, which distinguishes it from *CT* (computerized tomography) and *PET* (Positron emission tomography) scans [5].

Later in this project, Multi-Contrast Super Resolution will be investigated, so it is important to understand what those multiple contrasts are. This will be explained in the following section: 3.1.1.

3.1.1 How does MRI work ?

> 60% of the human body is made up of water molecules, which consist of hydrogen and oxygen atoms. Each Hydrogen atom has a **proton** in its center, which possess a spin making it act like a small **magnet very sensitive to magnetic fields.**
>
> MRI scanners apply a strong magnetic field which makes the protons in your body line up in the same direction. One can think of how a magnet can pull the needle of a compass.
>
> Then, **Radio Frequency pulses** are applied to specific areas to disturb protons' equilibrium. When the radio waves are turned off, the protons realign. While realigning, they send out radio signals, which are picked up by receivers.
>
> These signals provide information about the **exact location** of the protons in the body. They also help to **distinguish between the various types of tissue** in the body, since the protons in different tissue types realign at different speeds and produce distinct signals.
>
> In the same way that millions of pixels on a computer screen can create complex pictures, the signals from the millions of protons in the body are combined to create a detailed image of the inside of the body.

(National Institute of Biomedical Imaging and Bioengineering: *Magnetic resonance imaging* [6]). The basic steps of a MRI scan are illustrated in 3.1.

3.1.2 Capturing multiple contrasts with MRI

At this point, we know how MRI is able to re-create an image of the inside of the body, but not how it is possible to obtain different contrasts (different versions) of a same scene (see Fig 3.2.

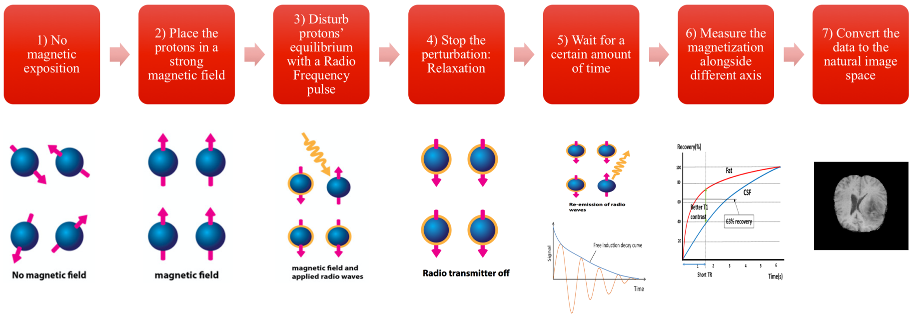


Figure 3.1: Basic steps of a MRI scan

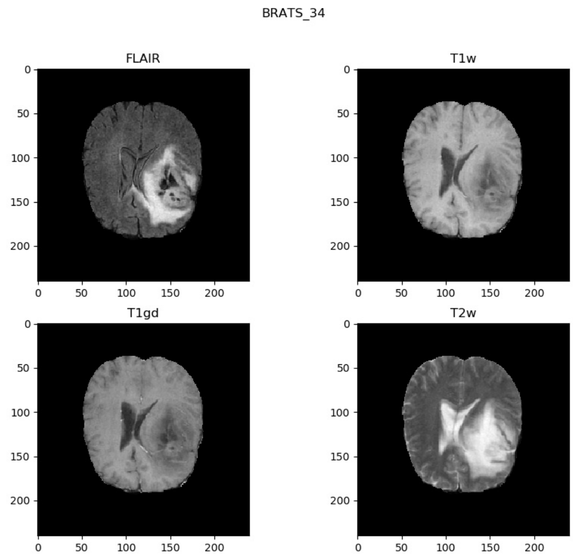


Figure 3.2: Common Brain MRI contrasts

Radio Frequency Pulse (RF pulses) have been evoked, and this is how it becomes possible to capture different MRI contrasts. RF pulses are set with two parameters : **Time Echo (TE)** and **Time Repetition (TR)**. And different combination will lead to different contrasts as shown in 3.3:

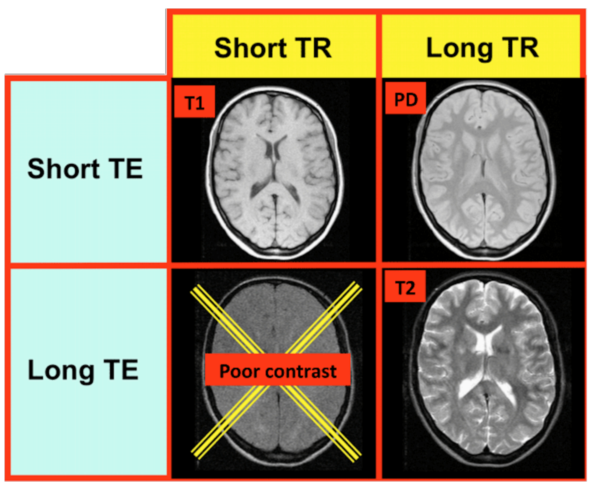
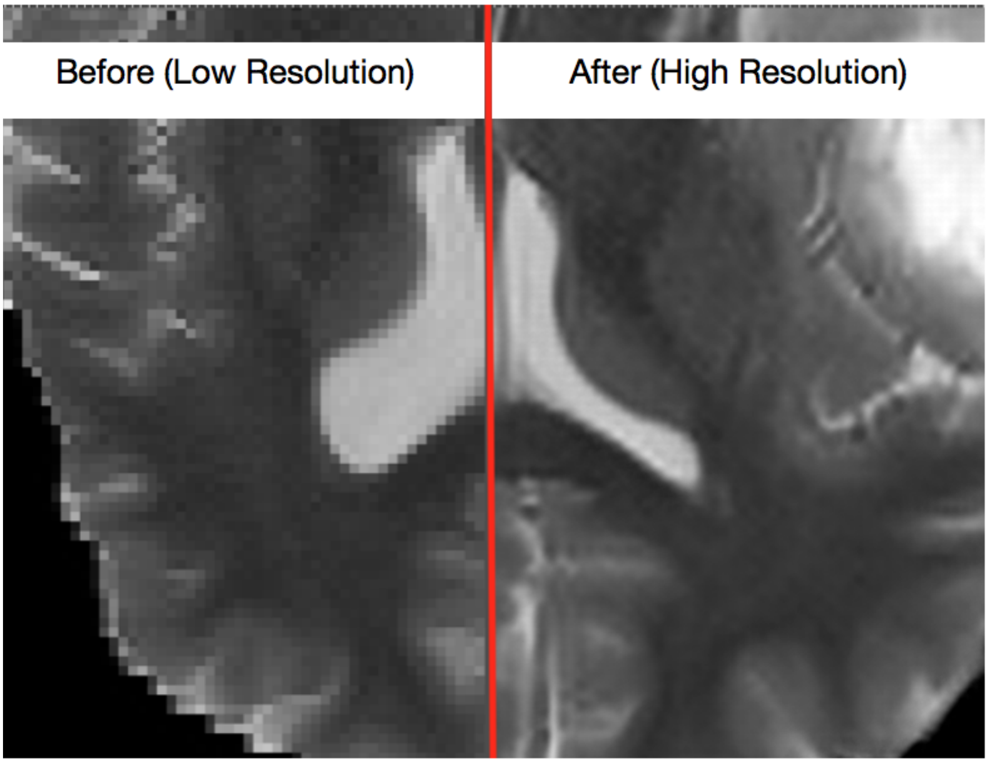


Figure 3.3: TR/TE combination and MRI contrasts (source: MRIQuestions.com)

3.2 Super Resolution

3.2.1 Overview



In detective movies, there are often scenes where detectives use enhancement programs to zoom in a suspect's face on surveillance videos. This is no longer a science fiction fantasy but an active research field nowadays and it is called **Super Resolution**.

Super Resolution refers to all techniques that increase the resolution (number of pixels per inch, measured in DPI) of a given image (Fig 3.2.1). It aims at finding a mapping from a low resolution image space LR $\mathbb{R}^{m \times n}$ to a high resolution $\mathbb{R}^{M \times N}$ image space HR (Fig 3.2.1), where $M = m \times s$ and $N = n \times s$, s being the up-scaling factor.

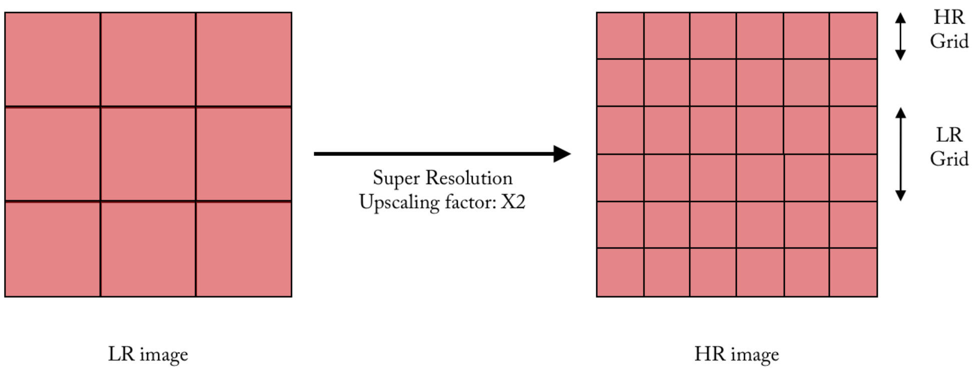


Figure 3.4: Super Resolution with an up-sampling factor of s=2. The number of missing pixels to predict increases in s^2.

3.2.2 Methods

Super Resolution algorithms are mainly divided into three categories [7]:

- **Interpolation-based**: like Spline, Bilinear or Bicubic interpolation.
- **Reconstruction-based**: like Maximum Likelihood (ML) in a frequentist framework or Maximum A Posteriori (MAP) in a Beyesian context.
- **Example-based** using Deep Neural Networks that approximate a mapping function from the LR space to the HR space.

Benefiting from the strong capacity of extracting high-level abstractions that bridge the LR and HR space, Convolutional Neural Networks (CNNs) have been showed to outperform competing algorithms in terms of both image quality and computation time. More over they don't require expert domain-knowledge which is convenient in such cutting-edges fields.

3.2.3 Problem formulation

The process of obtaining a low-resolution image from its high-resolution counterpart can be modeled as in Fig 3.5:

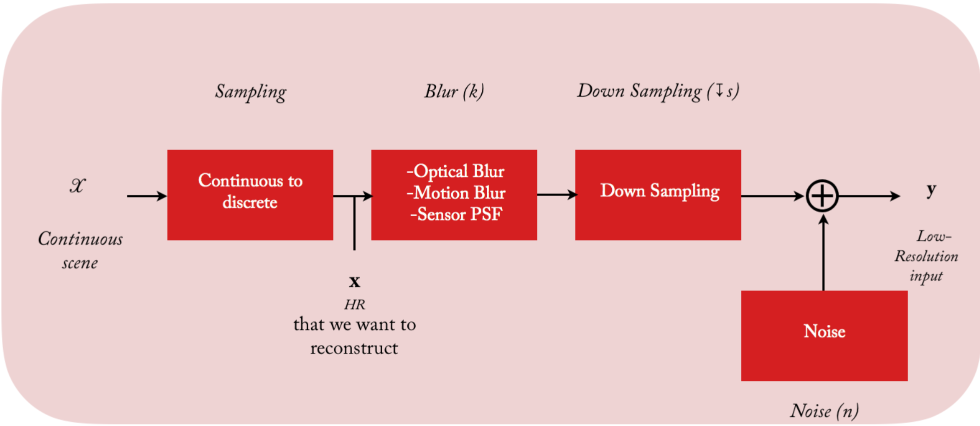


Figure 3.5: Observation model

$$y = (x \otimes k) \downarrow_s + n \tag{3.1}$$

where y is the low resolution image, $x \otimes k$ is the convolution between the blurry kernel k and the unknown HR image x, $\downarrow s$ is the down-sampling operator with a down-scale factor of s and n is an independent noise term [7].

It can also be written as:

$$\begin{aligned} y &= DB_k X + n \\ &= H_k X + n \end{aligned} \tag{3.2}$$

Example-based algorithms seek to recover an approximation $\hat{x}$ of the high resolution x as:

$$\hat{x} = \underset{x}{\mathrm{argmin}}(\|Hx - y\| + \lambda R(x)) \tag{3.3}$$

Where $\lambda R(x)$ is a regularization term [8].

Super resolution is therefore an inverse problem. Unfortunately the mapping from LR to HR is extremely complex and intractable. The problem is also ill-posed because for a specific LR image, plenty of HR images are plausible.

In practice, it is then challenging to enlarge an image without compromising its perceptual quality, sharpness and textures and staying close to the truth, especially at high up-sampling factors.

3.3 Super Resolution Networks

Convolutional Neural Networks (CNN) is a class of Deep Learning algorithms (Fig 3.6) that have generated amazing results in computer vision tasks. They are based on convolution operations that:

- Preserve the spatial relationship between pixels by learning image features using small squares of input data.
- Significantly reduce the number of parameters of the networks.

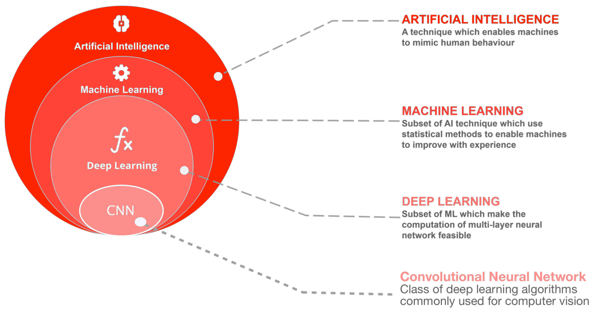


Figure 3.6: Convolutional Neural Networks among the AI family

3.4 Relative Work

3.4.1 SRCNN

The 2014 paper: *Image Super Resolution using Convolutional Neural Networks* (SRCNN) was the first attempt to use CNN for Super Resolution [9]. Authors proposed a three-layer architecture, presented in Fig 3.7: the first layer acts as a patch extractor with a large convolution filter of size (9×9), the second layer acts as a feature extractor from the low resolution patches and the third layer acts as a reconstruction layer with a filter size of (5x5).

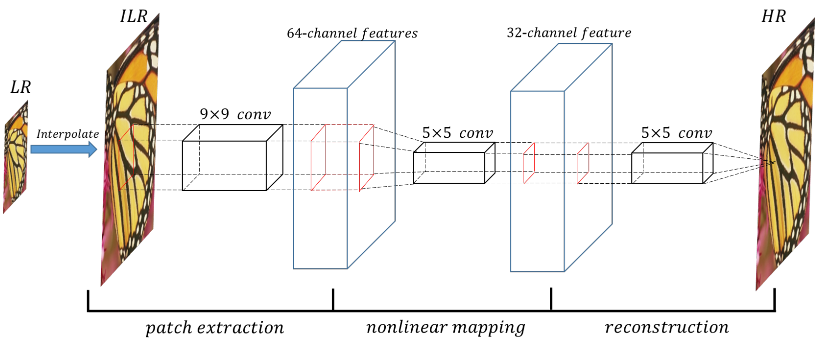


Figure 3.7: *Image Super-Resolution Using Deep Convolutional Networks, C. Dong et al., 2014*

SRCNN has an *Early-Upsampling Design* because it takes an interpolated LR image as input, contrary to *Late-Upsampling Designs* like ESPCN [10]. **ESPCN** does not take an interpolated version of LR as input, it increases the resolution at the end of the process using **Sup-Pixel Convolution** 9.
Based on this successful attempt at using SRCNN, deeper networks were built:

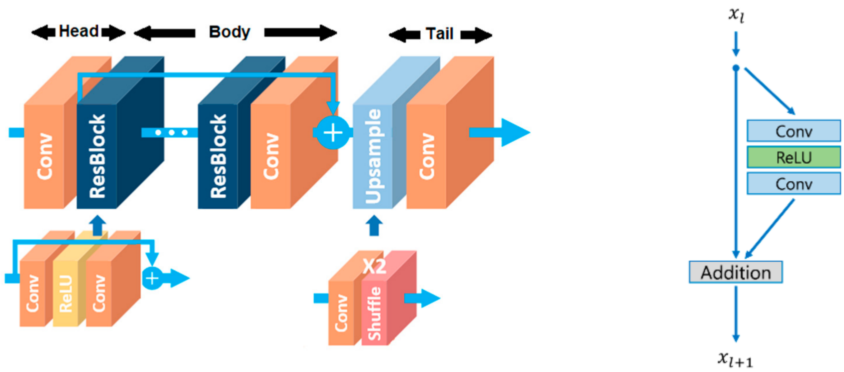
3.4.2 VDSR

VDSR (Accurate Image Super-Resolution Using Very Deep Convolutional Networks), inspired by *VGG-net*, is a SRCNN variant with more layers (20 layers) [11]. In this paper, the authors did not

only extended the number of layers, they proposed two effective strategies to avoid slow convergence in deep networks. Firstly, instead of directly learning to generate a HR image, they learned **a residual mapping** that generates the difference between the HR and LR image. This trick, known as residual learning, allows the network to focus on high-frequency information. Secondly, they implemented **gradient clipping** which enables very high learning rates hence speeding up the training process.
Their results supported the argument that deeper networks provide better performances.

3.4.3 Skip connection

In contrast to linear networks, residual networks use skip connections 9 (blue arrows in the following illustration) to avoid the **vanishing gradients** issue involved in very deep networks. This advance has made it possible to train deeper and deeper networks.
EDSR (Enhanced Deep Residual Networks for Single Image Super Resolution) [11] who won the NTIRE (New Trends in Image Restoration and Enhancement) 2017 Super Resolution challenge is one of the most popular architectures and is presented in Fig 3.8:



(**a**) Enhanced deep super resolution (EDSR) architecture (single-scale) (**b**) Residual block

Figure 3.8: EDSR uses skip connections to train more than 60 convolution filters, (*Enhanced Deep Residual Networks for Single Image Super-Resolution , B. Lim et al., 2017)*

A common problem is that all of these networks aim at minimizing the mean-squared error (MSE) between the generated super-resolved (SR) images and the corresponding ground-truth HR images. But, minimizing MSE leads to a blurring effect and hampers textures too. In fact, in the case of MSE, the theoretical solution is the weighted pixel-wise average of the set of realistic images that downscale properly to the LR input [12].

3.4.4 Perceptual Loss

SRGAN (Super Resolution Generative Adversarial Network) [13] proposed to use Generative Adversarial Networks for Super Resolution to overcome the blurring issues of MSE.
The image is first super resolved by a Generator (which is often a VGG network) and then, a Discriminator is charged to tell whether the image is a real HR image or an image produced by the Generator as shown in Fig 3.9.

SRGAN introduces a discriminative loss which is not pixel-wise and represents how realistic the reconstruction is.
By combining both MSE and discriminative loss (respectively called content and perceptual loss), we get a new loss function best suited for Super Resolution.
GANs have enabled the generation of photo-realistic Super Resolution, a highly challenging task, particularly at high up-sampling factors. However, training a GAN can be hard in practice. Convergence is not guaranteed and depends on many factors including the nature of the data and network model being used.

To summarize, today's Super Resolution architectures can be grouped as shown in Fig 3.10 [8]:

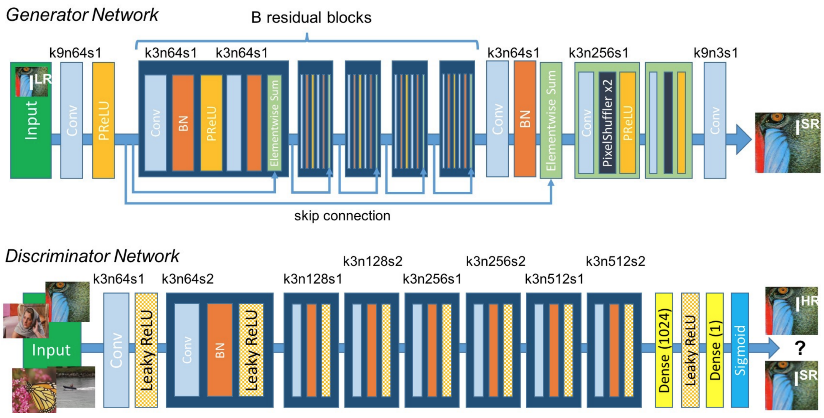


Figure 4: Architecture of Generator and Discriminator Network with corresponding kernel size (k), number of feature maps (n) and stride (s) indicated for each convolutional layer.

Figure 3.9: *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, C. Ledig et al., 2017*

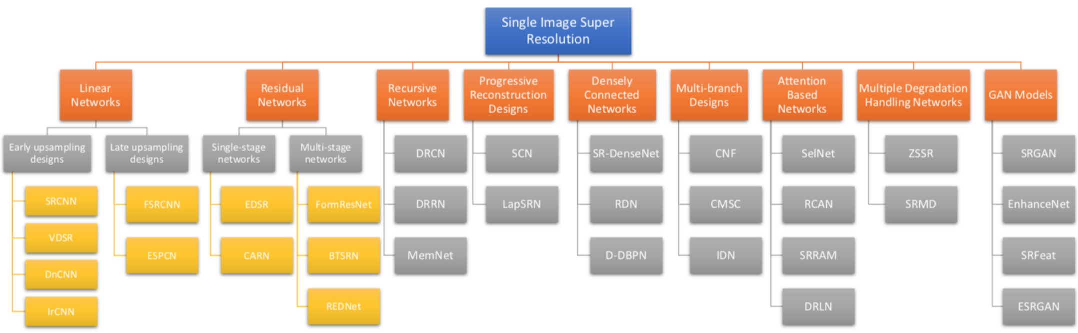


Figure 3.10: Taxonomy of the existing SR techniques (*A Deep Journey into Super-resolution: A Survey, S. Anwar et al., 2020*)

Chapter 4

Building a Super Resolution Network

4.1 Dataset

The data used in this study comes from the MICCAI BraTS challenge [14]:
The BraTS (*Brain Tumor segmentation*) challenge focuses on the evaluation of state-of-the-art methods for the segmentation of brain tumors in MRI scans.

4.1.1 Data description

The BraTS dataset provides 750 scans i.e $750 \times 155(slices) \times 4(contrasts) = 465.000$ images in total.

- File format: NIFTI file (nii.gz)
- Data type: *float64*
- Image Shape: 240(Width) $\times$ 240(Height) $\times$ 155(Number of Slices) $\times$ 4(Multi-contrast)
- Contrasts available: T1w, T1wgd, T2w and FLAIR
- Resolution: $1mm^3$
- Anatomical plane: Axial

Fig 4.1 is an example of what the data look like. We can see the 4 contrasts captured by MRI: FLAIR 9, T1W 9, T1GD 9 and T2W 9.

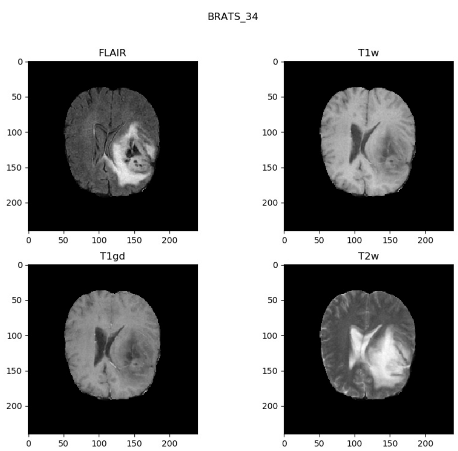


Figure 4.1: Multi-contrast brain MRI (patient $n°34$ and slice $n°85$)

4.1.2 Pre-processing

Data were acquired with different clinical protocols and various scanners from 19 different institutions. They are distributed after the following pre-processing: co-registration to the same anatomical template, interpolation to the same resolution ($1mm^3$) and skull-stripped.
All the images have been normalized between 0 and 1 to smooth over the differences between scanners.

4.2 The Neural Network

4.2.1 Establishing a training set

To train a super resolution network in a supervised manner, we need to feed it with (LR, HR) pairs, LR being the input and HR the target.
As only HR images were obtained from the BraTS dataset, the following model was used to create LR versions (Fig 4.2) :

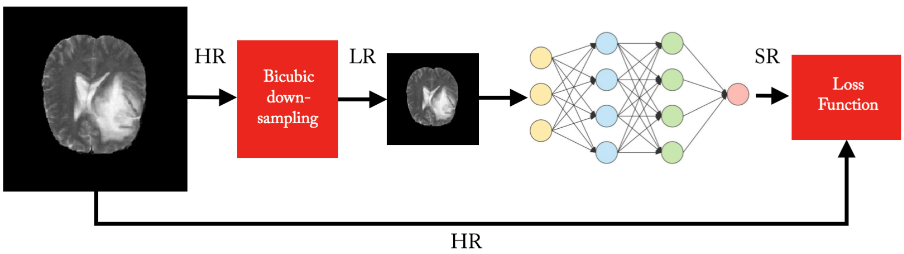


Figure 4.2: Image processing pipeline

We make a copy of HR that is down-sampled to get a LR version. To improve model's robustness, it is also possible to degrade LR image with noise, blur, etc.

4.2.2 Network Architecture

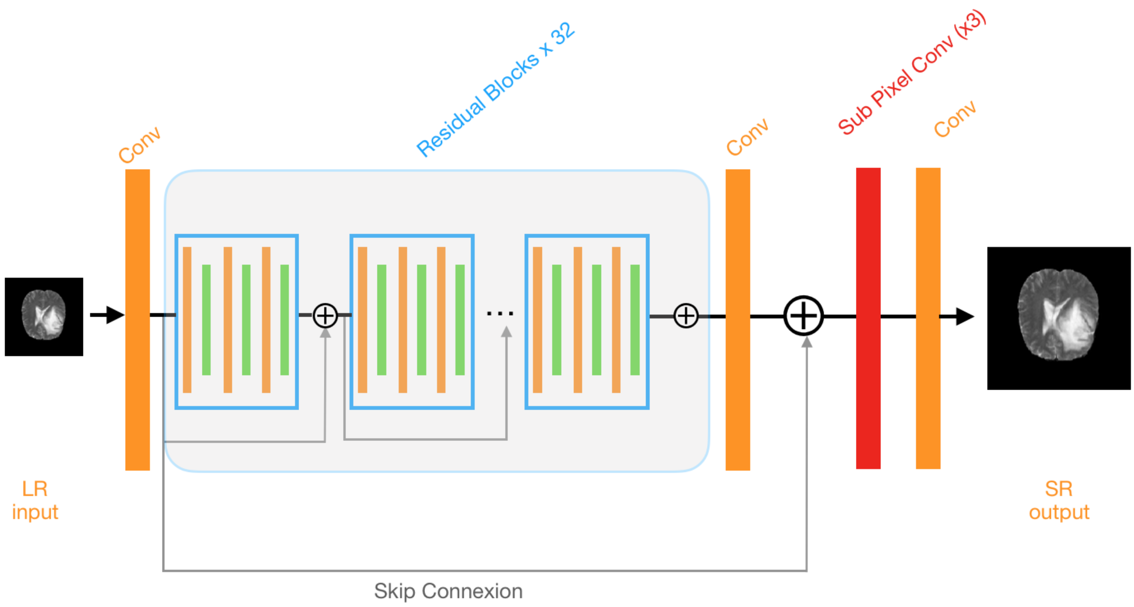
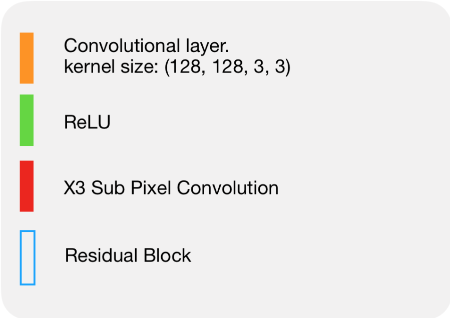


Figure 4.3: Network architecture adapted from EDSR



After a comparative study on existing super resolution networks, I retained EDSR [11] architecture for its superior performances.
An ablation study revealed that removing Batch Norm layers could increase its performances [11].
For memory considerations, I had to reduce original feature maps size from 256 to 128, but I kept the 32 residual blocks depth.
The architecture of the network is presented in Fig 4.3

4.2.3 Loss Function

Among the set of possible loss functions [15] the most commonly used ones are: $\mathcal{L}1_{Loss}$, $\mathcal{L}2_{Loss}$ and SSIM [16].

- L2-Loss (MSE).

$$\mathcal{L}2_{Loss} = \sum_{n=1}^{N} (\hat{y}_n - y_n)^2 \tag{4.1}$$

- L1-Loss (MAE).

$$\mathcal{L}1_{Loss} = \sum_{n=1}^{N} |\hat{y}_n - y_n| \tag{4.2}$$

- Structural Similarity Loss ($SSIM$).

$$SSIM_{Loss}(\hat{y}, y) = 1 - \frac{(2\mu_{\hat{y}}\mu_y + C_1) + (2\sigma_{\hat{y}y} + C_2)}{(\mu_{\hat{y}}^2 + \mu_y^2 + C_1)(\sigma_{\hat{y}}^2 + \sigma_y^2 + C_2)} \tag{4.3}$$

After comparing all of them, I retained the $\mathcal{L}2_{Loss}$ for its good trade-off between performances and computational costs. Minimizing the $\mathcal{L}2_{Loss}$ directly optimizes Peak Signal to Noise Ratio (PSNR) which is convenient since PSNR is the most widespread evaluation metric for Super Resolution.

4.3 Gradient Descent Algorithm

Mini-batch Gradient Descent:
I trained the network using Mini-batch Gradient Descent. Mini-batch is a good trade-off between the robustness of Stochastic Gradient Descent and the stability of Batch Gradient Descent [17]. It also allows to save a lot of memory by loading only a subset of the training set for each batch.

However, it requires to set an additional hyper-parameter: the mini-batch-size. Among the configurations tested: 8, 16, 32, 64 and 128, a batch size of **32** images gave the best performances according to PSNR.

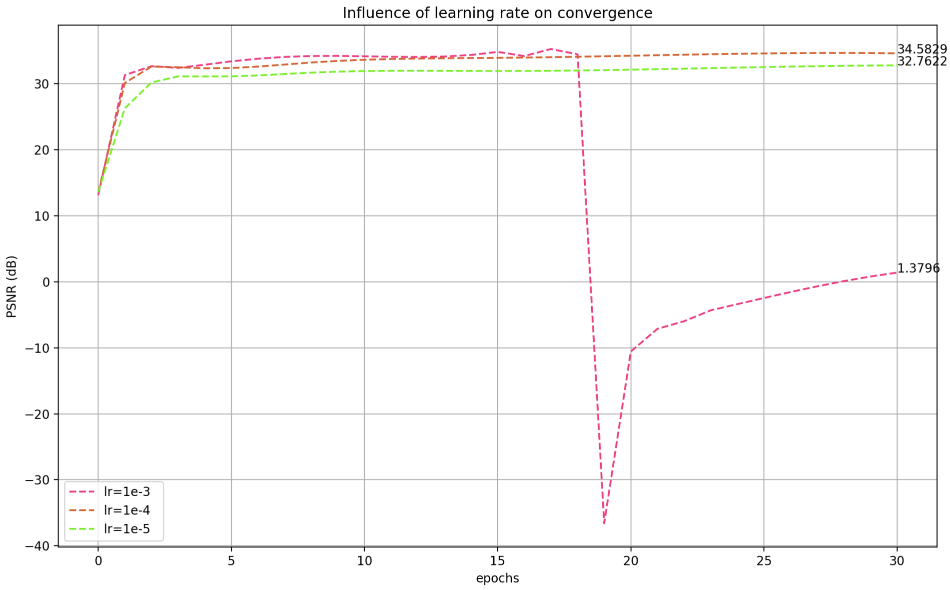


Figure 4.4: Influence of the learning rate on the network stability

Optimizer:
The **AdamW optimizer** [18] (a variant of Adam optimizer that implements L2-regularization) has been showed to outperform competing optimizers.
I set the learning rate to 10^{-4} (10^{3} being to unstable and 10^{-5} too slow as shown in Fig 4.3).

Once the architecture is selected and the hyper-parameter tuning done, it is time to evaluate the model performances. This is the subject of the following chapter.

Chapter 5

Single Contrast Super Resolution

The purpose of this chapter is to evaluate our first model, which can be qualified as a Single-Contrast model: as shown in Fig 4.3, a single LR image is taken as input to predict its HR counterpart.

5.1 Evaluation metrics

To evaluate the performances, I chose the PSNR [19] and SSIM (Structural Similarity Index Measure) [16] which are the two widespread metrics for Super Resolution:

- **PSNR**:
$$PSNR = 10 \times log(\frac{1}{MSE}) \tag{5.1}$$
between 0 and $+\infty$.

- **SSIM**:
$$SSIM = \frac{(2\mu_{\hat{y}}\mu_y + C_1) + (2\sigma_{\hat{y}y} + C_2)}{(\mu_{\hat{y}}^2 + \mu_y^2 + C_1)(\sigma_{\hat{y}}^2 + \sigma_y^2 + C_2)} \tag{5.2}$$
between 0 and 1.

PSNR is pixel-wise while SSIM is patch-wise and a better indicator of texture reconstruction quality [16].

5.2 Implementation and Evaluation

The dataset provides images in 4 distinct contrasts (T1W, T2W, T1GD and FLAIR). For the sake of completeness, I trained 4 models (one for each contrast). Nevertheless, to avoid repetition, I will mostly focus on T1W and T2W Super Resolution in this chapter. T1GD and FLAIR Super Resolution results can be found in annex.

5.2.1 T2W Super Resolution

Visual Results

Fig 5.1 reveals that our first model Single-Contrast Super Resolution, or SCSR, achieves a good reconstruction even at x3 up-sampling. However, high frequency information like edges are harder to recover. In fact, networks with good PSNR are explicitly penalized for attempting to hallucinate details they are unsure about, thus optimizing PSNR results in blurring and lacking of details.

This experiment revealed the presence of unintended stripes in SCSR's output, as shown in Fig 5.2. This is detrimental for texture quality. Yet, texture quality is of highest importance for other image processing tasks like Automatic brain tumor segmentation.

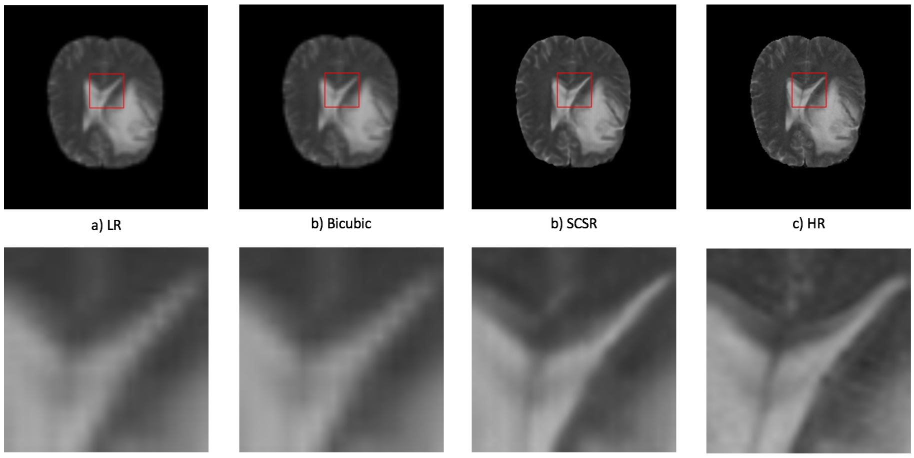


Figure 5.1: T2W super resolution example (zoom on edges)

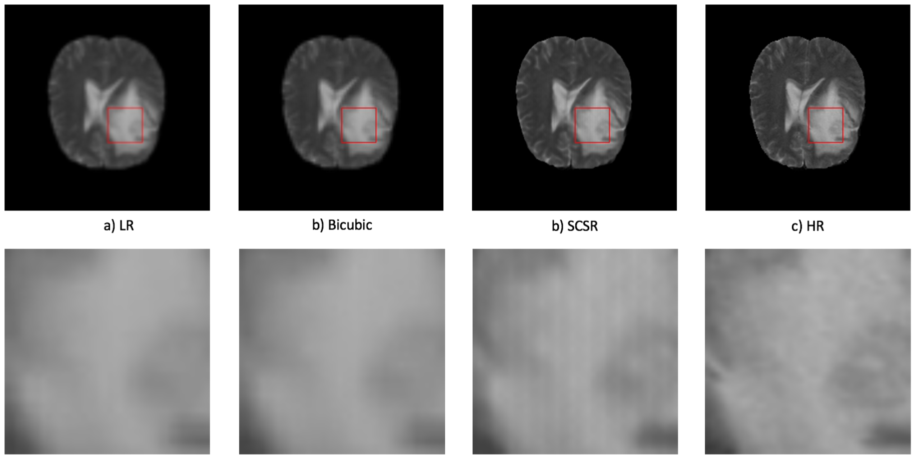


Figure 5.2: T2W super resolution example (zoom on tumor)

Model	PSNR	SSIM
Bicubic	31.761	0.931
SCSR	**35.878**	**0.983**

Figure 5.3: PSNR/SSIM evaluation

PSNR/SSIM evaluation

SCSR achieves attractive performances in terms of PSNR and SSIM, close to state-of-the-art performances for real images Super Resolution.

5.2.2 T1W Super Resolution

Visual Results

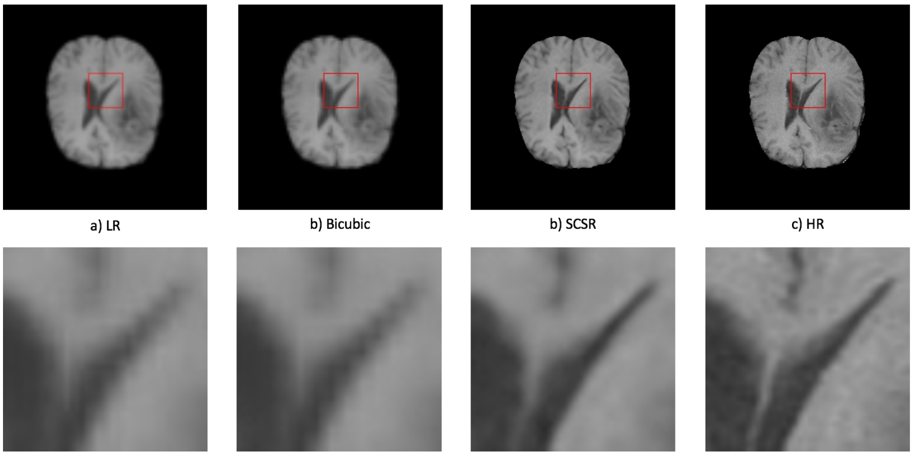


Figure 5.4: T1W super resolution example (zoom on edges)

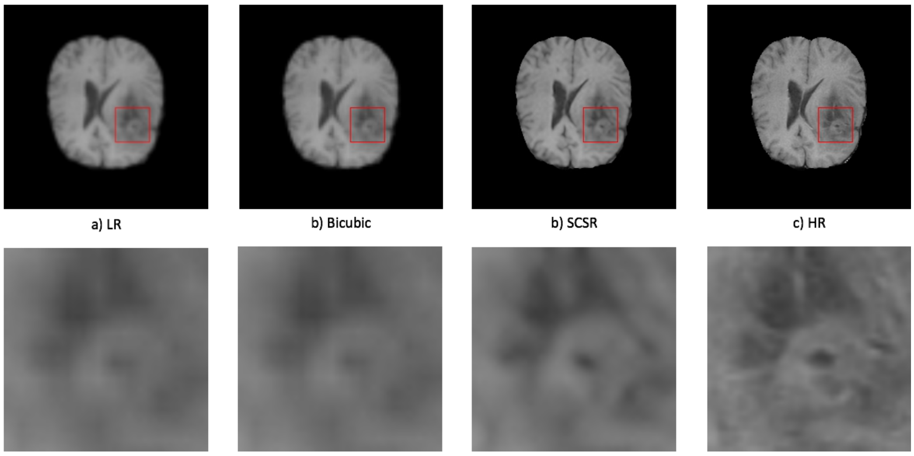


Figure 5.5: T1W super resolution examples (zoom on tumor)

The same observations can be done with T1W reconstruction except for the fact that textures are better recovered with T1W, as it can be observed in Fig 5.5.

PSNR/SSIM evaluation

Model	PSNR	SSIM
Bicubic	32.257	0.939
SCSR	**37.896**	**0.986**

Figure 5.6: PSNR/SSIM evaluation

PSNR and SSIM performances Fig 5.6 suggest that T1W is "easier" to reconstruct than T2W.

5.2.3 T1GD and FLAIR Super Resolution

To avoid repetition, T1GD and FLAIR results are shown in annex.

5.3 Discussion

The goal of those experiments was to evaluate our baseline model and see whether an EDSR-like architecture could provide good performances for MRI Super Resolution.

The experiments showed promising results in both qualitative (visual quality) and quantitative (PSNR and SSIM) terms.

In the meantime, they revealed two weaknesses: a lack of sharpness and bad performances on textures.

Those observations led to the elaboration of a new model: **Multi-Contrast Super Resolution** to overcomes these two issues.

Chapter 6

Enhanced MRI Super Resolution: Multi-Contrast Super Resolution

6.1 Introduction

6.1.1 Motivation

In fact, super resolution methods can be divided in two categories: *Single-Contrast Super Resolution* (SCSR) and *Multi-Contrast Super Resolution* (MCSR).

For automatic brain tumor segmentation, Multi-Contrast-based models are the norm because they have achieved far superior performances on Tumor Semantic Segmentation [20]. This can be explained by the fact that the complementarity between MRI contrasts is essential to distinguish tissue types.
But surprisingly, Multi-Contrast-based models are far less frequently explored for Super Resolution.

Previously, we have seen that our first model suffers from several weaknesses which may be tackled by mixing the information of all the available contrasts.

For example, T1W super resolution provides good texture quality which is a weakness of T2W super resolution. So, could T1W be used to guide T2W texture reconstruction?

6.1.2 Presentation of the novel method

There are many ways to incorporate the information of additional contrasts in the model ([21], [22]), the easier being to concatenate the images and directly feed the network with T1W, T2W, T1GD and FLAIR images as shown in Fig 6.1:

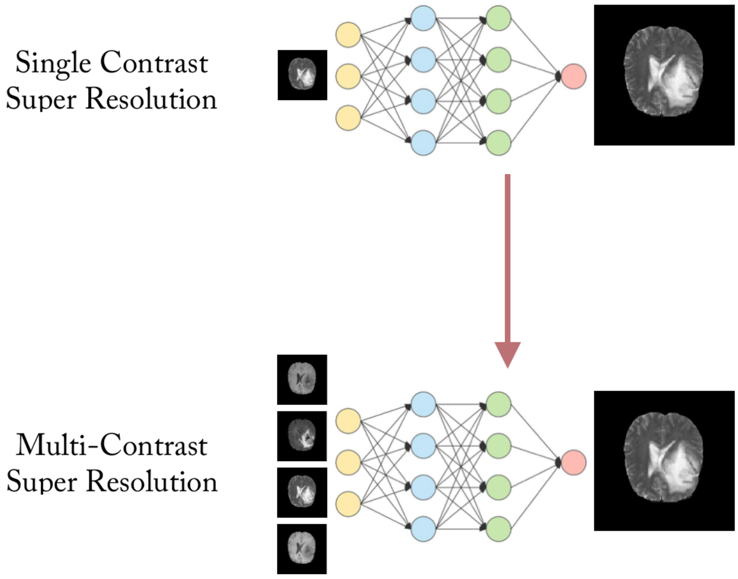


Figure 6.1: From Single-Contrast to Multi-Contrast Super Resolution

The network will learn on its own how to use the information contained on additional contrasts to super resolve the targeted contrast.
We can expect contrasts to help each other because although they share a similar structure (close enough to help each other reconstruction), the information is not completely redundant [23] (Fig 6.3) and each tissue type has its specific visual characteristics depending on the MRI contrast (Fig 6.2):

Some tissue types are indistinguishable with only one contrast.

Tissue	T1w	T2w	FLAIR
CSF	Dark	Bright	Dark
White Matter	Light	Dark Gray	Dark Gray
Cortex	Gray	Light Gray	Light Gray
Fat	Bright	Light	Light
Inflammation	Dark	Bright	Bright

Tissue appearances depending on MRI Sequence (T1W, T2W and FLAIR).

Figure 6.2: Each tissue type has its specific appearance based on a specific contrast

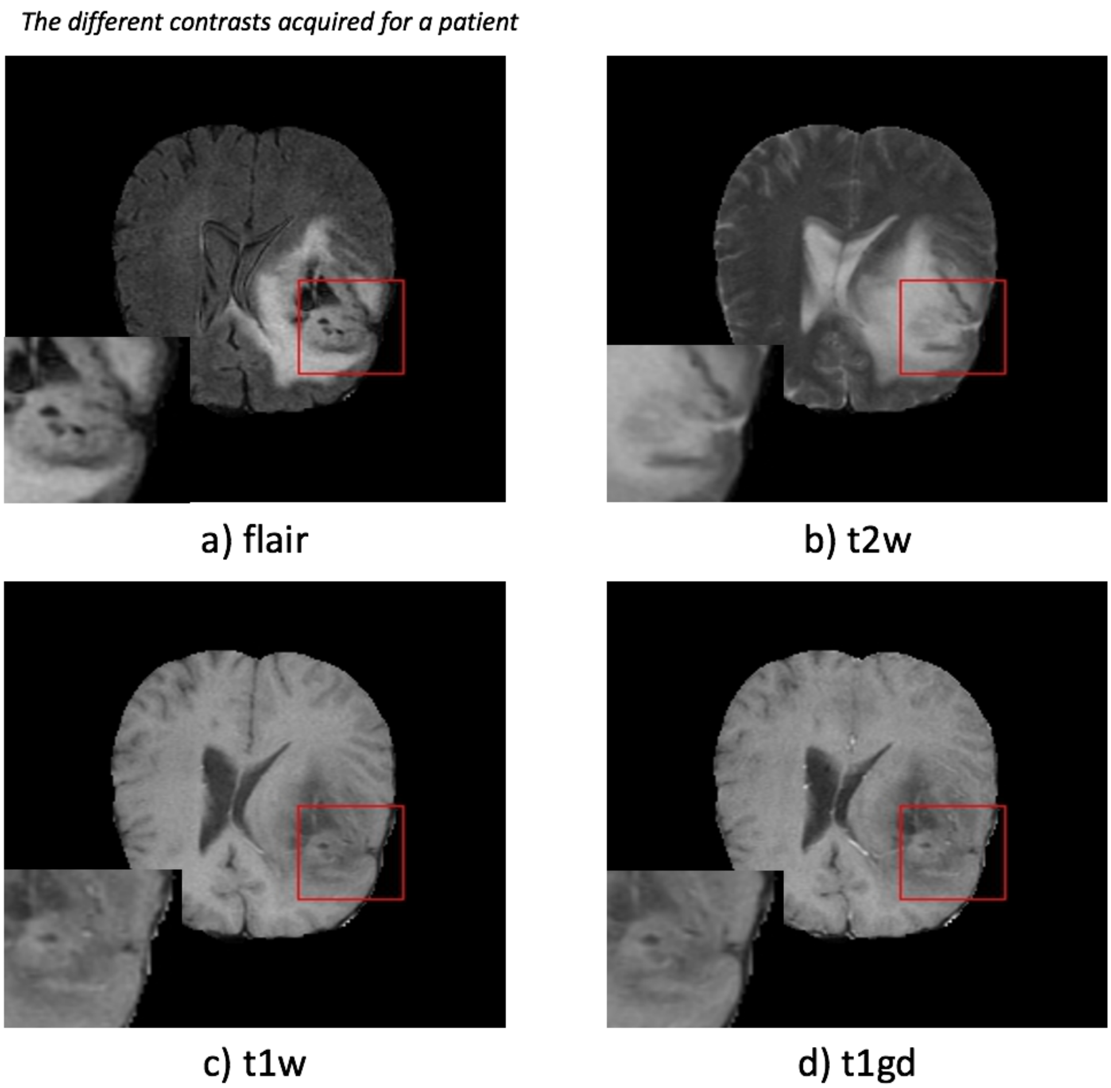


Figure 6.3: Some details that are invisible is some contrasts can be observed in others

6.2 Presentation of the experiments

To verify this hypothesis, I have trained a model for each possible input combination (8 combinations in total).
Again, for sake of completeness, an experiment will be conducted for each contrast super resolution: T2W, T1W, T1GD and FLAIR (respectively EXP A, B, C and D).
There are then $4 \times 8 = 32$ models to train, as shown in Fig 6.4.

To avoid repetition, only T2W and FLAIR Super Resolution will be presented and T1W and T1GD results can be found in annex.

EXP. A		
	INPUT	OUTPUT
EDSR	T2W	
A.1	T2W + T1W	
A.2	T2W + T1GD	
A.3	T2W + FLAIR	
A.4	T2W + T1W + T1GD	T2W
A.5	T2W + T1W + FLAIR	
A.6	T2W + T1GD + FLAIR	
A.7	T2W + T1W + T1GD + FLAIR	

EXP. B		
	INPUT	OUTPUT
EDSR	T1W	
B.1	T1W + T2W	
B.2	T1W + T1GD	
B.3	T1W + FLAIR	
B.4	T1W + T2W + T1GD	T1W
B.5	T1W + T2W + FLAIR	
B.6	T1W + T1GD + FLAIR	
B.7	T1W + T2W + T1GD + FLAIR	

EXP. C		
	INPUT	OUTPUT
EDSR	T1GD	
C.1	T1GD + T2W	
C.2	T1GD + T1W	
C.3	T1GD + FLAIR	
C.4	T1GD + T2W + T1W	T1GD
C.5	T1GD + T2W + FLAIR	
C.6	T1GD+ T1W + FLAIR	
C.7	T1GD+ T2W + T1W + FLAIR	

EXP. D		
	INPUT	OUTPUT
EDSR	FLAIR	
D.1	FLAIR + T2W	
D.2	FLAIR+ T1GD	
D.3	FLAIR+ T1W	
D.4	FLAIR + T2W + T1GD	FLAIR
D.5	FLAIR+ T2W + T1GD	
D.6	FLAIR + T1W + T1GD	
D.7	FLAIR+ T2W + T1W+ T1GD	

Figure 6.4: The 32 models to evaluate

6.3 Experimental results

6.3.1 Visual Results

Impacts on textures

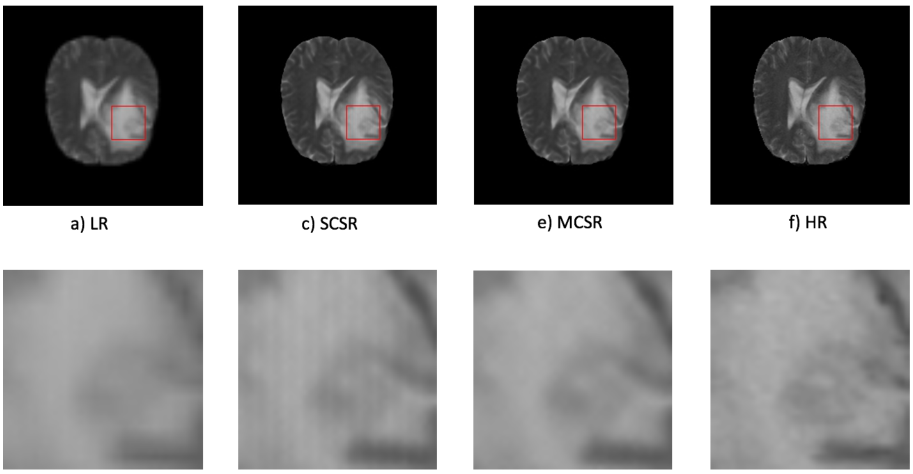


Figure 6.5: T2W super resolution, zoom on tumor

Fig 6.5 shows that contrary to SCSR, our new model reconstructs high quality textures (there are no longer stripes in the tumor region).

Impacts on edges

We can see in Fig 6.6 and Fig 6.7 that the novel method produces sharper edges and finer details.

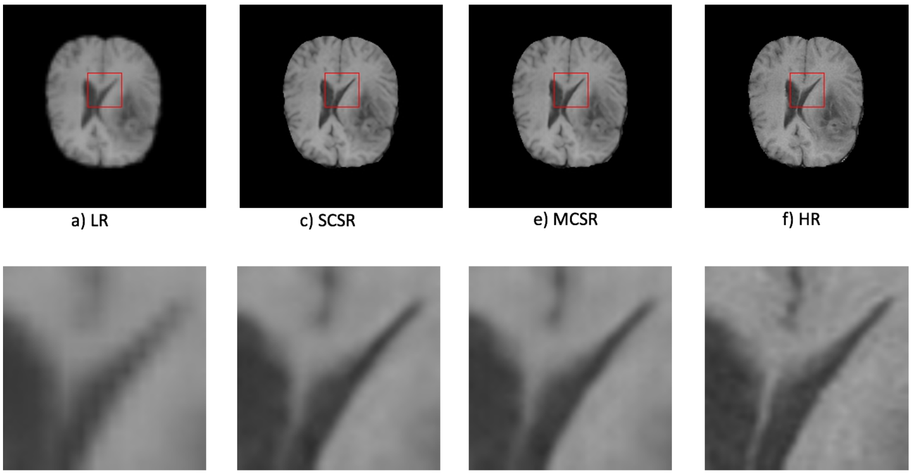


Figure 6.6: T1W super resolution, zoom on the edges

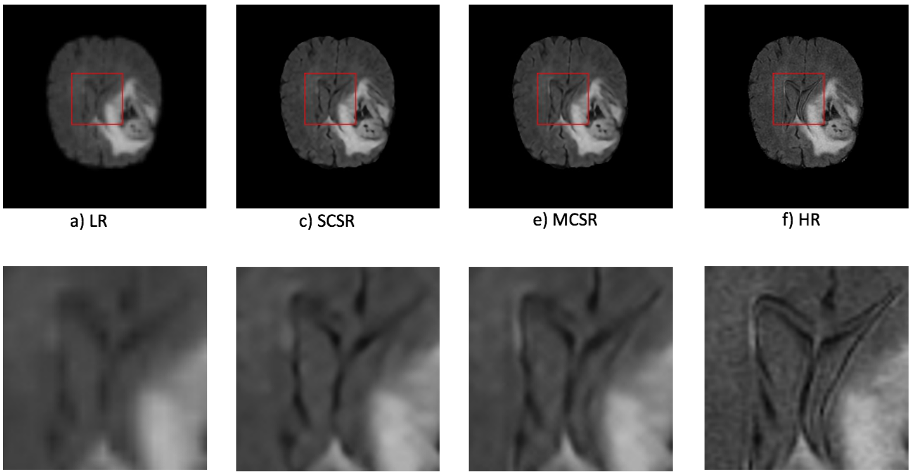


Figure 6.7: FLAIR super resolution, zoom on the edges

6.3.2 Evaluation metrics

The evaluation metrics are consistent with the qualitative evaluation:
It is important to note that every Multi-Contrast combination has outperformed SCSR.
Another interesting fact is that MCSR is more stable (see Fig 6.10:
MCSR's stability enables to train networks with a higher learning rate, therefore it learns faster and is more unlikely to get stuck in local minima.
But counter-intuitively, more input contrasts is not associated with better performances. One possible explanation is over-fitting.

The quantity of examples contained in the dataset is fixed. However, models with many input contrasts have to learn more parameters than the other with the same amount of examples, which is prone to over-fit. This can be observed in 6.11:
The gap between T2W+T1W+T1GD+FLAIR and T2W+T1W+FLAIR narrows over time and finally, even if T2W+T1W+T1GD+FLAIR shows better performances on the training set, T2W+T1W+FLAIR ends up exceeding its performances on the test set, showing a better generalizability.

It legitimizes doing a comparative study in order to find the best trade-off between the amount of information and generalizability.

(Note: for each experiment, the best model is shown in bold in the evaluation metrics tables Fig 6.8 and Fig 6.9. They are called MCSR in the visual results figures Fig 6.5, Fig 6.6, and Fig 6.7).

Model	PSNR	SSIM
Bicubic	31.761	0.931
SCSR	35.878	0.983
T2W+T1W	36.202	0.986
T2W+T1GD	36.198	0.984
T2W+FLAIR	36.138	0.984
T2W+T1W+T1GD	36.417	0.985
T2W+T1W+FLAIR	36.450	0.986
T2W+T1GD+FLAIR	36.330	0.985
T2W+T1W+T1GD+FLAIR	36.384	0.982

Figure 6.8: T2W Super Resolution evaluation

Model	PSNR	SSIM
Bicubic	32.257	0.939
SCSR	37.896	0.986
T1W+T2W	38.312	0.986
T1W+T1GD	38.264	0.984
T1W+FLAIR	38.069	0.989
T1W+T2W+T1GD	38.253	0.984
T1W+T2W+FLAIR	38.305	0.986
T1W+T1GD+FLAIR	38.242	0.986
T1W+T2W+T1GD+FLAIR	38.223	0.988

Figure 6.9: T1W Super Resolution evaluation

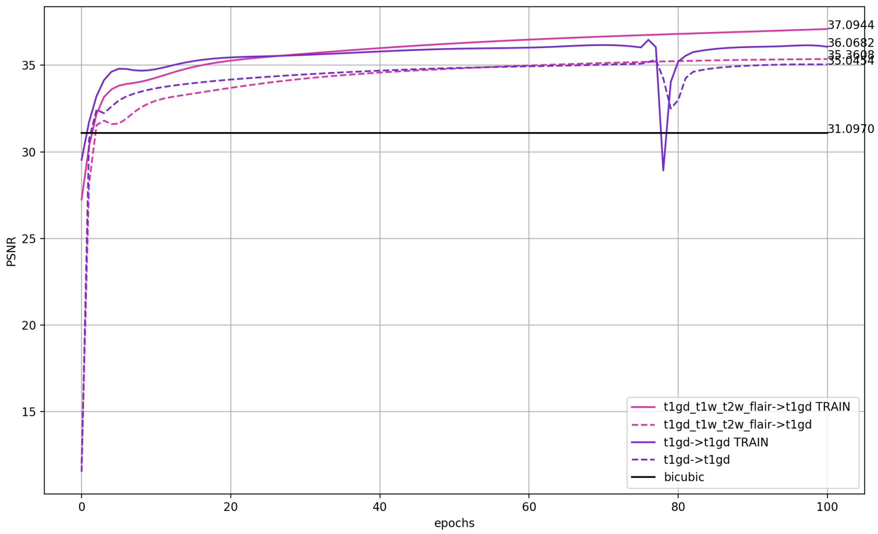


Figure 6.10: SCSR instability is fixed with Multi-Contrast

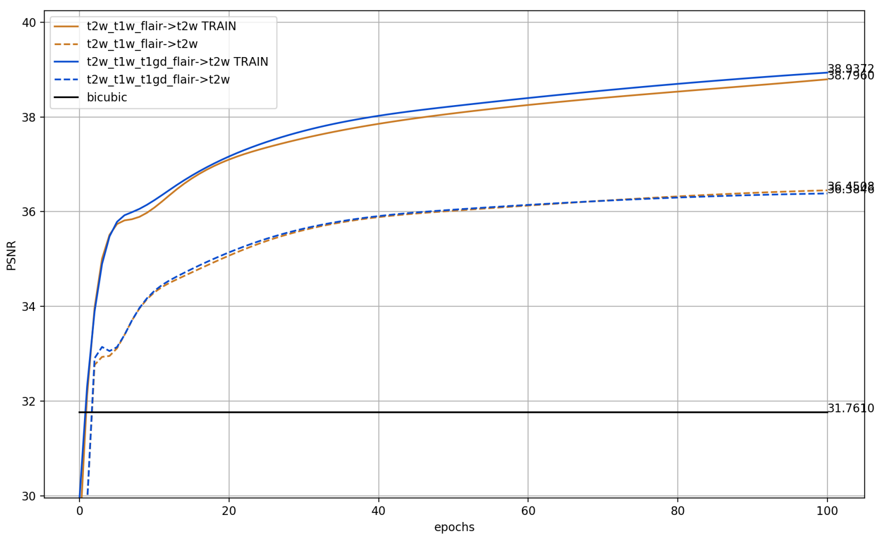


Figure 6.11: Models with many input contrasts are prone to over-fit

Chapter 7

Discussion and future directions

7.1 Discussion

7.1.1 Loss function

Those models have been trained to minimize the MSE Loss. Yet pixel-wise MSE Loss is not optimal for Super Resolution since it has blurring effects and fails to recover finer details. The image quality could be improved by adding a perceptual loss (a measure of how realistic the image is), consequently to train a Generative Adversarial Network [13].

7.1.2 Training set acquisition

As explained in 4.2.1, the strategy to obtain LR-HR pairs is questionable. We are learning the mapping between a HR image and its Bicubic-degraded LR counterpart. However this is just a simplification of reality since the real world degradation is much more complex (it depends on the device, motion blur, ...). This study is then conducted in a simple case.

To avoid over-fitting, it is preferable to use a large dataset, which can be problematic if the region of the body studied is not well documented.

7.1.3 Real World Data

It may be noted that Multi-Contrast methods assume the required contrasts to be available, however this is not always achievable in a clinical setting (limitations of the equipment).

7.2 Future directions

7.2.1 Generability of the method

Applicability on other architectures

During this research project, all experiments were conducted with a deep residual network whose architecture is close to that of EDSR. But it would be interesting to see if the same observations can be made with other architectures, in particular with GANs.
Similarly, a future work would be to conduct experiments with other up-sampling factors like x2, x4, x6.

Adaptability and robustness

Adaptability and robustness are vital for an implementation in clinical setting, yet those two aspects have not been covered during the project.
To improve the model's robustness, a solution is to implement Data Augmentation [24].
When it comes to adaptability, it could be assessed using additional datasets.

7.2.2 Influence of MCSR on Brain Image Segmentation

Manual Tumor segmentation of brain tumor extent from brain 3D volumes is vital for tumor detection and treatment but also extremely time consuming and subject to high variability inter and intra-rater variability. In this context, Automatic Brain Tumor Segmentation constitutes a valuable tool.

The results of this study have proven that Multi-Contrast helps super resolution and the next step would be to study the influence of MCSR on Brain Image Segmentation. If it turns out that MCSR also improve segmentation performances, Multi-Contrast would be a simple way to enhance not only Super Resolution but also Image Segmentation.

Chapter 8

Conclusion

To summarize, we have seen that Magnetic Resonance Imaging is a valuable technology for brain tumor diagnosis. However, obtaining high quality images remains costly: post-processing methods, like Super Resolution, are currently under development to improve MRI quality.

The following steps were completed during the research project:

- Investigation of the state-of-the-art Super Resolution algorithms available in the literature among which Convolutional Neural Networks stood out for their performances in terms of PSNR.
- A training set have been built: HR images were provided by BraTS dataset and LR images have been processed.
- A first model, denoted by **SCSR**, has been implemented in Python and trained to super resolve MR Images, yielding to promising results.
- After a more thorough study of MRI technology and inspired by Automatic Brain Tumor Segmentation, I decided to extend that baseline model with Multi-Contrast Super Resolution: **MCSR**.
- Comparative study to identify the best model, based on real HR images and simulated LR versions.
- Interpretation and discussion of the experimental results in order to establish a foundation for further study in the future.

The proposed model, MCSR has outperformed one of the state-of-the-art model, EDSR, in terms of both PSNR and SSIM (up to +0.4dB on PSNR). MCSR achieves simultaneously a better texture recovering and finer details which are two challenging aspects in super resolution.

Furthermore, MCSR appears to be an interesting line of research for potential improvement of Image Segmentation.

Chapter 9

Glossary

- **Super Resolution** [7]: Class of algorithms that increase the resolution of images.
- **MRI** [5]: Magnetic Resonance Imaging is a medical imaging technology that enables to create a detailed image of the inside of the body.
- **T2W**: One of the MRI contrasts (produced using long TR and long TE times).
- **T1W**: One of the MRI contrasts (produced using short TR and short TE times).
- **T1GD**: T1 Gadolinium enhanced. Because of its paramagnetic properties, gadolinium is useful to enhance T1W acquisition. Before scanning, the patient is injected small concentrations and these agents increase T1 signals. Pathological tissues (tumors, areas of inflammation / infection) will demonstrate accumulation of contrast and therefore appear as brighter than surrounding tissue.
- **FLAIR**: One of the MRI contrasts (produced using very long TR and very long TE times).
- **Convolutional Neural Networks**: Class of deep learning algorithm extremely useful on computer vision tasks.
- **BraTS Dataset** [14]: The MICCAI Brain Tumor Segmentation Challenge Dataset.
- **Single Contrast Super Resolution**:
- **Multi Contrast Super Resolution**:
- **Residual Blocks** [11]: Residual neural networks use skip connections, or shortcuts to jump over some layers. One motivation for skipping over layers is to avoid the problem of vanishing gradients.
- **Sup-Pixel Convolution** [10]: Sub-pixel convolution (or pixel shuffle) is an operation which rearranges the elements of $\mathrm{H} \times \mathrm{W} \times \mathrm{C} \cdot \mathrm{r}^2$ tensor to form $\mathrm{rH} \times \mathrm{rW} \times \mathrm{C}$ tensor.
- **PSNR** [19]: Peak Signal To Noise Ratio (PSNR) is used to measure the quality of image reconstruction.
- **SSIM** [16]: Structural Similarity Index Measure (SSIM) is a method that predicts the perceived image quality.

Appendix A

Single-Contrast Super Resolution additional results

A.1 T1GD Super Resolution

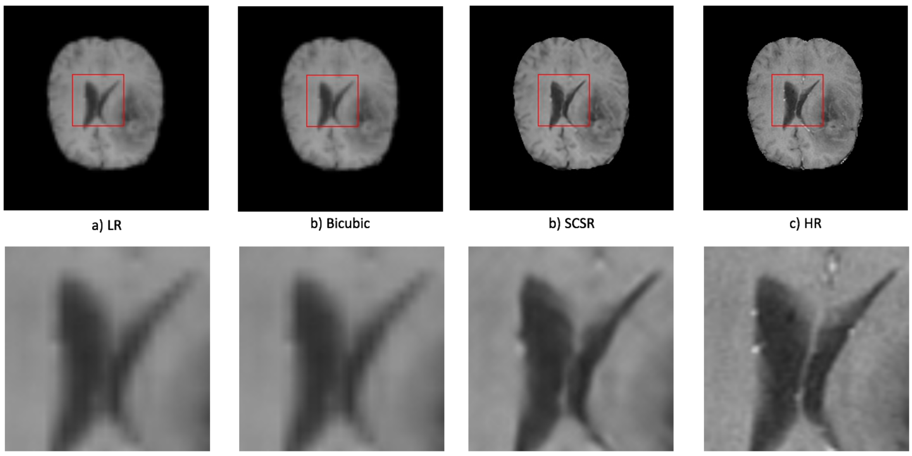


Figure A.1: T1GD super resolution, zoom on edges

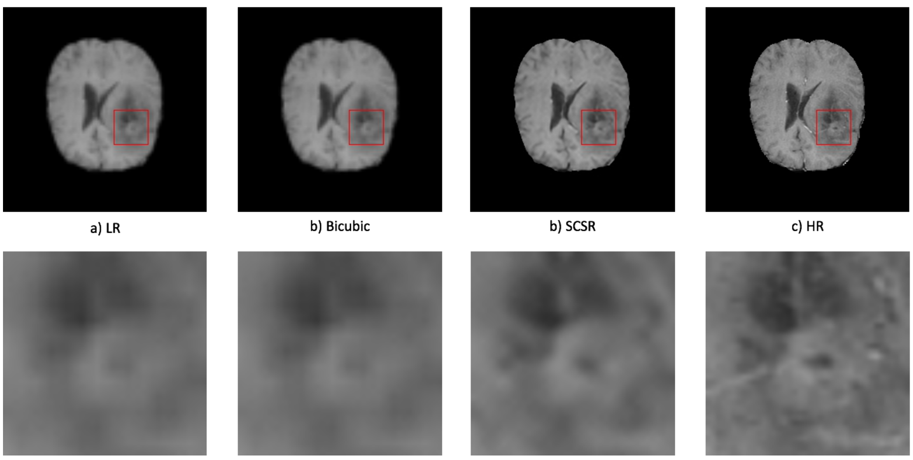


Figure A.2: T1GD super resolution, zoom on tumor

Model	PSNR	SSIM
Bicubic	31.097	0.929
SCSR	35.045	0.978

Figure A.3: T1GD Super Resolution evaluation

A.2 T1GD Super Resolution

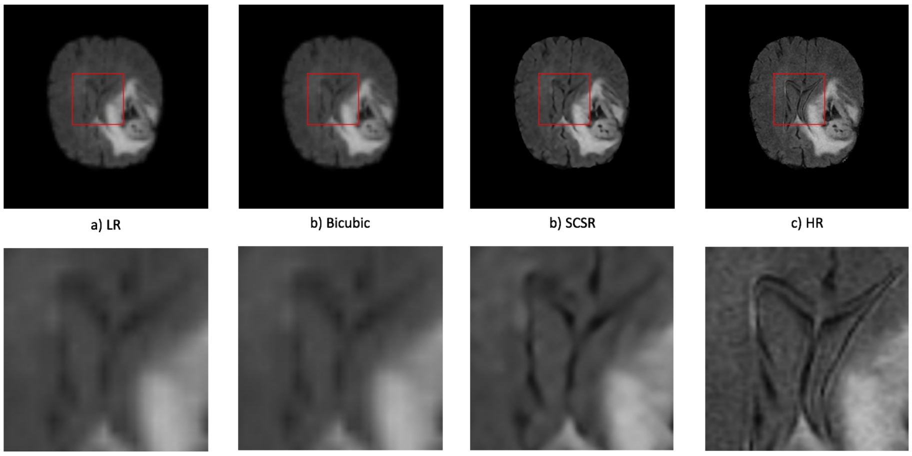


Figure A.4: FLAIR super resolution, zoom on edges

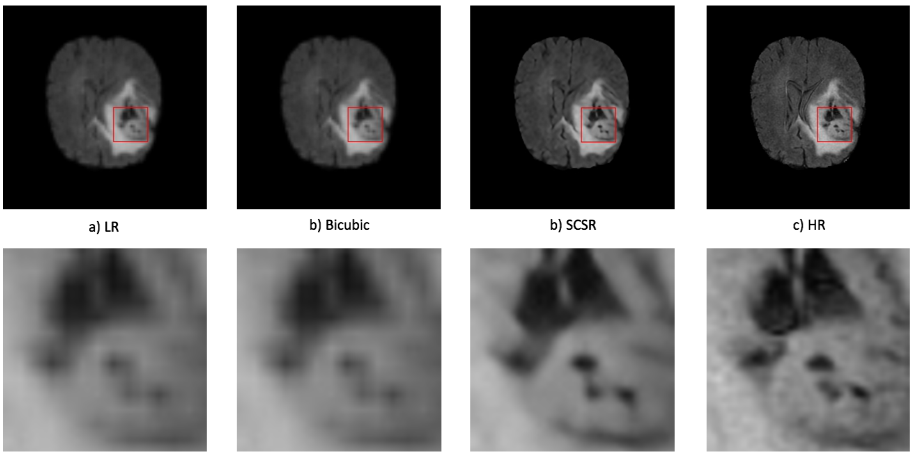


Figure A.5: FLAIR super resolution, zoom on tumor

Model	PSNR	SSIM
Bicubic	31.131	0.918
SCSR	36.357	0.965

Figure A.6: FLAIR Super Resolution evaluation

Appendix B

Multi-Contrast Super Resolution additional results

B.1 T1W Super Resolution

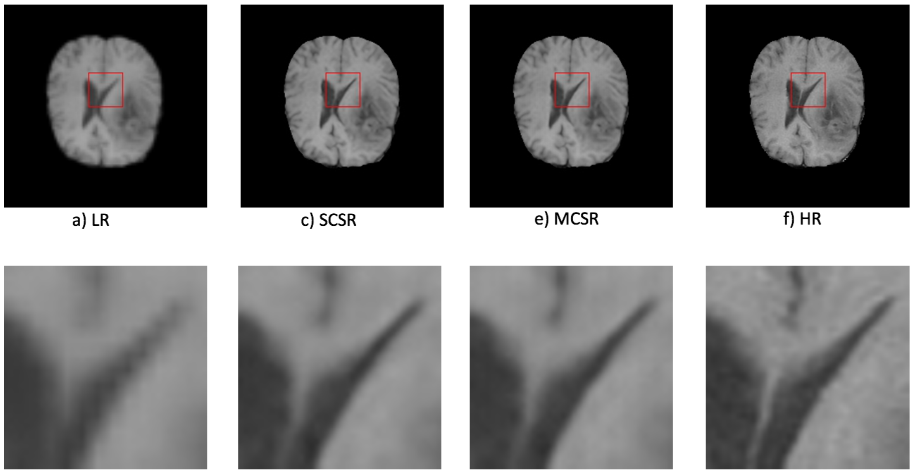


Figure B.1: T1W super resolution, zoom on edges

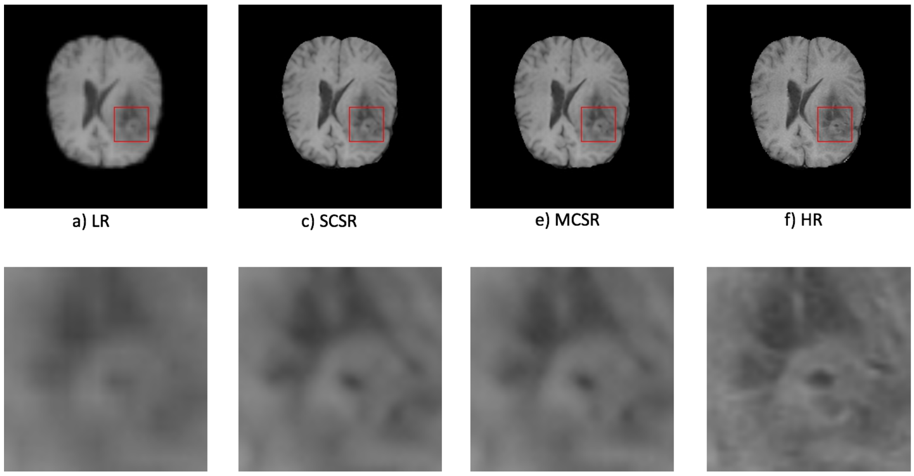


Figure B.2: T1W super resolution, zoom on tumor

Model	PSNR	SSIM
Bicubic	32.257	0.939
SCSR	37.896	0.986
T1W+T2W	38.312	0.986
T1W+T1GD	38.264	0.984
T1W+FLAIR	38.069	0.989
T1W+T2W+T1GD	38.253	0.984
T1W+T2W+FLAIR	38.305	0.986
T1W+T1GD+FLAIR	38.242	0.986
T1W+T2W+T1GD+FLAIR	38.223	0.988

Figure B.3: T1W Super Resolution evaluation

B.2 T1GD Super Resolution

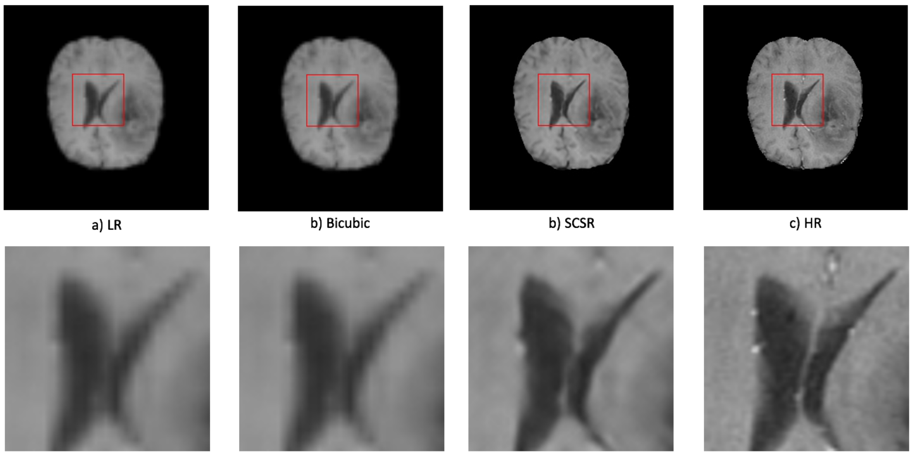


Figure B.4: T1GD super resolution, zoom on edges

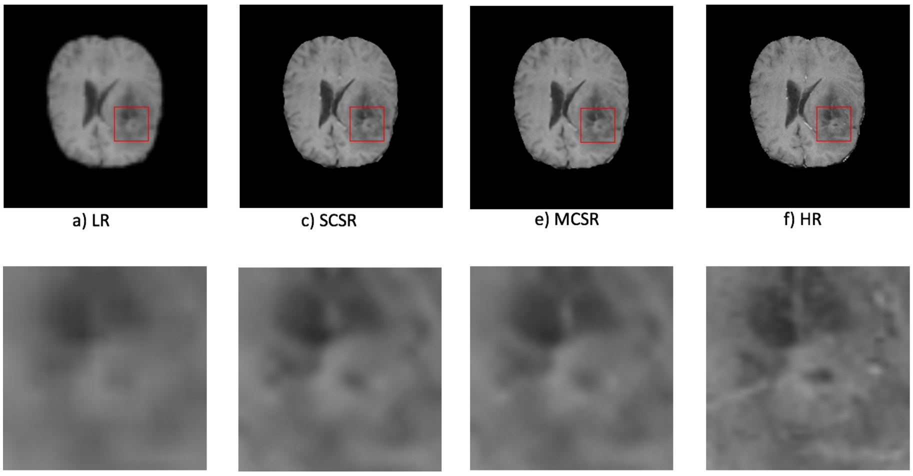


Figure B.5: T1GD super resolution, zoom on tumor

Model	PSNR	SSIM
Bicubic	31.097	0.929
SCSR	35.045	0.978
T1GD+T1W	35.352	0.973
T1GD+T2W	35.416	0.973
T1GD+FLAIR	35.240	0.977
T1GD+T1W+FLAIR	35.309	0.977
T1GD+T1W+T2W	35.464	0.978
T1GD+T2W+FLAIR	35.323	0.978
T1GD+T2W+T1W+FLAIR	35.360	0.979

Figure B.6: T1GD Super Resolution evaluation

Bibliography

[1] Lancet Neurol. Global, regional, and national burden of brain and other cns cancer, 1990–2016: a systematic analysis for the global burden of disease study 2016. http://dx.doi.org/10.1016, 2016.

[2] L.M. DeAngelis. Brain tumors. *New England Journal Of Medicine*, vol. 344(no. 2):p. 114–123, 2001.

[3] Anna Nowogrodzki. The world's strongest MRI machines are pushing human imaging to new limits. *Nature*, vol. 322(no. 563):p. 24–26, 2018.

[4] Wikipedia. Super resolution. https://en.wikipedia.org/wiki/Super-resolution_imaging.

[5] Wikipedia. Magnetic resonance imaging. https://en.wikipedia.org/wiki/Magnetic_resonance_imaging.

[6] National Institute of Biomedical Imaging and Bioengineering. Magnetic resonance imaging. https://www.nibib.nih.gov/science-education/science-topics/magnetic-resonance-imaging-mri.

[7] Wenming Yang, Xuechen Zhang, Yapeng Tian, Wei Wang, and Jing-Hao Xue. Deep learning for single image super-resolution: A brief review. *CoRR*, abs/1808.03344, 2018.

[8] Saeed Anwar, Salman Khan, and Nick Barnes. A deep journey into super-resolution: A survey. *CoRR*, abs/1904.07523, 2019.

[9] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *CoRR*, 2015.

[10] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. *CoRR*, abs/1609.05158, 2016.

[11] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. *CoRR*, abs/1707.02921, 2017.

[12] Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi, and Cynthia Rudin. Pulse: Self-supervised photo upsampling via latent space exploration of generative models, 2020.

[13] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew P. Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. *CoRR*, abs/1609.04802, 2016.

[14] University of Pennsylvania Perelman School Of Medicine. Multimodal brain tumor segmentation challenge 2018. https://www.med.upenn.edu/sbia/brats2018/data.html, 2018.

[15] H. Zhao, O. Gallo, I. Frosio, and J. Kautz. Loss functions for image restoration with neural networks. *IEEE Transactions on Computational Imaging*, 3(1):47–57, 2017.

[16] H. R. Sheikh Z. Wang, A. C. Bovik and E. P. Simoncelli. "image quality assessment: From error visibility to structural similarity". *IEEE Transactions on Image Processing*, 13.

[17] Genevieve B Orr Klaus-Robert Müller Yann A LeCun, Léon Bottou. Neural networks: Tricks of the trade. pages 9–48, 1998.

[18] Ilya Loshchilov and Frank Hutter. Fixing weight decay regularization in adam. *CoRR*, abs/1711.05101, 2017.

[19] Xikipedia. Peak-signal-to-noise ratio. https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio.

[20] Bauer S et al. Menze BH, Jakab A. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans Med Imaging.*, 2015.

[21] Zeng Kun, Hong Zheng, Congbo Cai, Yu Yang, Kaihua Zhang, and Zhong Chen. Simultaneous single- and multi-contrast super-resolution for brain mri images based on a convolutional neural network. *Computers in Biology and Medicine*, 99, 06 2018.

[22] Hélène Meunier Nathalie Bednarek Ronan Fablet et al. Chi-Hieu Pham, Carlos Tor-Díez. Multiscale brain mri super-resolution using deep 3d convolutional networks. *Elsevier*, 2019.

[23] Chi-Hieu Pham. Deep learning for medical image super resolution and segmentation. *Image Processing [eess.IV]*, 2018.

[24] Jaejun Yoo, Namhyuk Ahn, and Kyung-Ah Sohn. Rethinking data augmentation for image super-resolution: A comprehensive analysis and a new strategy. 04 2020.