



THE CONTRIBUTION OF SUPER-RESOLUTION TO TUMOR CLASSIFICATION

KAYRA ECESU ERTIZMAN

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF SCIENCE IN COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE

DEPARTMENT OF
COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

STUDENT NUMBER

2038356

COMMITTEE

prof. dr. Eric Postma
dr. Peter Hendrix

LOCATION

Tilburg University
School of Humanities and Digital Sciences
Department of Cognitive Science &
Artificial Intelligence
Tilburg, The Netherlands

DATE

June 24, 2022

PREFACE

I would like to begin by thanking my supervisor, prof. dr. Eric Postma, my dearly beloved family, friends, and dr. Peter Hendrix. They all helped and supported me a lot throughout my thesis journey.

I am really happy that I have worked on image super-resolution and on a very vital subject, which is cancer. Breast cancer is a crucial problem for women, and as a woman, I am very honored that I had a chance to have a hand in the improvement of this crucial illness' research processes.

Sometimes, I had tough times in my thesis journey. Nonetheless, I can say that with this thesis, I have improved myself and my skills significantly and learned a lot. I was able to 'deeply' improve my knowledge in the Deep Learning field and enjoy it. Because of my experiences in this thesis project, I was able to decide on my future career plans and to realize what my skills and passions are. I hope you will also enjoy reading this thesis as much as I enjoyed writing it.

Kayra Ecesu Ertizman

THE CONTRIBUTION OF
SUPER-RESOLUTION TO TUMOR
CLASSIFICATION

KAYRA ECESU ERTIZMAN

CONTENTS

1	Introduction	4
2	Related Work	6
2.1	Image Super-Resolution	6
2.2	Tumor Classification	9
3	Methods and Approach	10
3.1	Generative Adversarial Networks (GANs)	10
3.2	Super Resolution Generative Adversarial Networks (SR-GANs) Architecture	11
3.3	Enhanced Super Resolution Generative Adversarial Networks (ESRGANs)	12
3.4	Convolutional Neural Networks (CNNs)	14
4	Experiment Details	15
4.1	Dataset	15
4.2	Dataset Pre-processing	16
4.3	Experimental Procedure	17
4.3.1	ESRGAN Model	17
4.3.2	Convolutional Neural Network (CNN) Model for Classification	17
4.4	Software	18
5	Results	18
6	Discussion and Future Work	21
7	Conclusion	23

Abstract

Breast cancer is one of the leading cancer types seen among women (Luo et al., 2022). Breast cancers are found to be causing 22% of cancer-related mortality (Motlagh et al., 2018). Diagnosing cancer at an early stage is a crucial step toward preventing cancer to spread, thus, histopathology is essential to analyze tissues to see whether a patient has cancer or not. Machine learning methods in histopathology for cancer detection have been a common practice in recent years (Kourou, Exarchos, Exarchos, Karamouzis, & Fotiadis, 2015). However, the problem of low-quality microscopic tumor images is an obstacle to the correct classification of tumors in cancer tissues. For this reason, it is important to enhance the quality of cancer tissue images. Image super-resolution (SR) is a way to enhance the quality of images. SR has been shown to be effective in enhancing the resolution of histopathological images (Celik & Talu, 2020). The Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) is one of the most powerful SR models. In this thesis, the ESRGAN model is applied to breast-cancer tumor images. The goal is to determine whether ESRGAN-based enhancement improves tumor-classification performance. The research question addressed in this thesis reads: 'To what extent can ESRGANs improve the accuracy of the classification task of benign and malignant breast cancer tumors?' We answer the research question by applying an ESRGAN to low-resolution breast-cancer images. The breast-cancer tumor images are obtained from the dataset called the 'BreakHis dataset' Spanhol, Oliveira, Petitjean, and Heutte (2016a). The results of the experiments conducted in this thesis reveal that applying ESRGANs enhances CNN classification performance. This leads us to conclude that ESRGANs enhance the tumor classification accuracy.

1 INTRODUCTION

Cancer is defined as the uncontrollable growth and spread of a group of cells (Nain, 2015). There are different types of cancers such as lung, breast, brain, stomach and so more. Breast cancer is one of the most common cancer types (Motlagh et al., 2018). A group of cancer cells divides quickly, gathers together, and makes up a structure called a "tumor". Tumors can be divided into two main classes: malignant and benign (Patel, 2020). Malignant tumors have cells that grow uncontrollably and require special treatments to stop this growth. In addition, they invade other organs, leading to cancer cells spreading to many different parts of the body. Unlike malignant cells, benign cells do not spread but can still grow and exert pressure on other organs (Patel, 2020). It is, therefore, crucial to detect cancer cells and classify tumors into malignant and benign to stop their growth through specific treatments, at early stages. Histopathology is the

study and diagnosis of tissue diseases. Histopathologists (people analyzing cancer cells) focus on the diagnosis of cancer by investigating changes in biological cells and tissues (Musumeci, 2014). Nonetheless, using only human power in histopathology has the disadvantage of making high numbers of tumor misclassifications. This leads to administering wrong treatments to cancer patients. The successful outcomes of using Machine Learning (ML) methods in medical data analysis resulted in the use of these methods in cancer detection to minimize errors encountered by human pathologists in cancer diagnosis (Bardou, Zhang, & Ahmad, 2018; Shahidi, Mohd Daud, Abas, Ahmad, & Maarop, 2020). For example, ML methods are used in breast-cancer tumor classification tasks to correctly classify malignant and benign breast tumors with low error rates. In addition to error minimization, image processing methods are proved to be detecting cancer at earlier stages (Kalaiyarasi, Dhanasekar, Ram, & Vaishnavi, 2020).

While machine learning methods help in classifying cancer cells into malignant and benign, sometimes ML methods can fail in the correct classification. This is generally due to the quality of medical images. Classification performances can be degraded when the image looks distorted, blurry (Sert, Ozyurt, & Dogantekin, 2019). Improving image quality is therefore crucial. Image quality enhancement can either be performed using expensive medical imaging tools or by applying relatively cheap image super-resolution methods. Image super-resolution (SR) is a method for constructing high-resolution, fine-detailed images from low-resolution images that are blurred or distorted (Yue et al., 2016). The aim of super-resolution is to recover fine details from the blurry low-resolution input images. Super-resolution can be applied in many different areas, such as healthcare, astronomical imaging, satellite imagery, and so on (Malczewski & Stasiński, 2009).

Super-resolution methods have recently been proposed to be effective in optimizing the quality of histopathology images in cancer research and analysis (Celik & Talu, 2020; Shahidi, 2021). Image SR techniques are preferred over overpriced medical imaging tools because they are not expensive and are fast and easy to apply. Using SR on medical images results in better classification of tumors in deciding whether those tumors are benign or malignant. An example paper on the application of super-resolution on cancer images uses Generative Adversarial Networks (GANs). This method is named as Super-Resolution Generative Adversarial Networks (SRGANs) (Celik & Talu, 2020) which was proposed by Ledig et al. (2017). However, according to the study by X. Wang et al., while SRGANs are powerful in increasing the quality/resolution of images, they are found to be yielding hallucinated details. When zoomed in, it has been found that the details of the images are lost. Therefore, Enhanced Generative

Adversarial Networks (ESRGANs) were proposed to optimize the artifacts encountered by SRGANs and to produce images with details that are less likely to be hallucinated (X. Wang et al., 2018).

So far, there have been SR techniques applied to histopathology images in cancer research, such as Super-Resolution Convolutional Neural Networks, SRGANs (Celik & Talu, 2020; Umehara, Ota, & Ishida, 2017). Nonetheless, previous work did not investigate the effect of using ESRGAN on breast-cancer tumor images. Thus, in this study, the ESRGAN model proposed by X. Wang et al. (2018) will be applied to breast tumor images. Then, a classification task will be performed to classify tumors into malignant and benign using Convolutional Neural Networks (CNNs). This study aims to investigate whether ESRGANs can help us to classify more accurately between malignant and benign breast tumors compared to low-resolution, non-ESRGAN applied images.

RQ: To what extent can ESRGANs improve the accuracy of classification task of benign and malignant breast tumors?

The sub-question is as follows:

RQ1: How does the ESRGAN-enhanced tumor-classification performance compare to the classification performance obtained on the low-resolution images and the original high-resolution images?

2 RELATED WORK

2.1 Image Super-Resolution

Until now, in the image super-resolution (SR) field, different models have been proposed e. g., interpolation-based, learning-based, and reconstruction-based (Ooi & Ibrahim, 2021). However, with the advancements in Machine Learning (ML) and Deep Learning, learning-based models have gained much more attention compared to other previously proposed traditional models.

Learning-based models are based on a branch of Machine Learning called Deep Learning (DL). Deep Learning models aim to learn several different representations of data with the use of multiple deep layers. Hence, DL models are known as Deep Networks. Deep Networks are comprised of multiple layers that are stacked upon each other (LeCun, Bengio, & Hinton, 2015). In the image super-resolution field, deep networks make statistical comparisons between original, low-resolution images and SR model applied high-resolution images (Yang, Zhang, Tian, Wang, & Xue, 2018) using deep layers. Learning-based SR models are fast to apply

and they can produce high-frequency images with sharp details. The most popular, pioneering work in the image SR field by utilizing deep networks belongs to [Dong, Loy, He, and Tang \(2015\)](#). The model used in [Dong et al.](#)'s work is a learning-based SR model named as Super-Resolution Convolutional Neural Networks (SRCNNs). SRCNNs exploit the network architecture of Convolutional Neural Networks (CNNs). The advantage of using SRCNN is that this model was designed intentionally to have a basic architecture. Its architecture is composed only of three parts: Patch extraction and representation from bicubic interpolation upscaled images; Non-linear mapping; Reconstruction. Nevertheless, SRCNN has some drawbacks. One of the most important drawbacks of this model is the application of bicubic interpolation upscaling. These images have smoothed details. This leads to making wrong estimations about the actual image ([Yang et al., 2018](#)).

To overcome the drawbacks of SRCNNs, recently, Super-Resolution Generative Adversarial Networks (SRGANs) and Enhanced Super Resolution Generative Adversarial Networks (ESRGANs) were proposed. SRGANs and ESRGANs are Generative Adversarial Networks (GANs), constructed specifically for the SR task. The details of GANs, SRGANs and the main deep learning model of this paper, which is ESRGANs, will be discussed in the "Methods and Approach" section more specifically. However, it is worth reviewing these models briefly.

Super-Resolution Generative Adversarial Networks (SRGANs) are Generative Adversarial Networks created specifically for the image super-resolution research ([Ledig et al., 2017](#)). As stated by the authors of the SRGAN paper, previously designed SR models, such as SRCNNs, lack in creating fine-detailed, perceptually-salient images. In contrast to SRCNNs, SRGANs have two different parts: Generator and Discriminator. Generator takes an input of a low-resolution image. This low-resolution image is obtained by convolving a Gaussian kernel over the high-resolution (HR) image. The generator aims to create an output SR image from the LR image. The discriminator then takes the generated SR image and the original HR image and aims to differentiate between the SR image and the HR image ([Ledig et al., 2017](#)). This model's performance was tested using an evaluation metric called "Mean Opinion Score" (MOS). MOS is a metric used to evaluate "perceptually" by rating the image quality from one-to-five. The authors have asked 26 participants to rate SR images and other images (for instance: bicubic interpolation applied images, SRCNN applied images). It has been found that SRGAN outperforms the original HR image, bicubic interpolation applied image and SRCNN applied images with a higher MOS rating score. This means that the SRGAN model is good at generating perceptually satisfactory high-quality images ([Ledig et al., 2017](#)).

SRGANs are widely applied in healthcare, especially in cancer research (Song, Chowdhury, Yang, & Dutta, 2020; Sood, Topiwala, Choutagunta, Sood, & Rusu, 2018). One example study focuses on the application of the SRGAN model to the histopathological breast-cancer images. This study is conducted by Celik and Talu (2020) and the authors applied SRGANs for the tasks of image quality change when resizing the image with different magnifications and image noise cleaning. In the noise removal task, they have compared the effect of SRGAN to the other traditional noise removal methods. In the second task, they have compared bicubic interpolation and SRGAN on images with different magnifications, 2x; 4x; 8x, and 16x. For both noise removal and image magnification tasks, authors have used evaluation metrics to evaluate the performance of SRGAN. It has been found that, for images with higher magnification rates and also for the noise removal, SRGAN is able to outperform classical approaches by increasing the resolution and cleaning the noise (Celik & Talu, 2020).

Although, Super-Resolution Generative Adversarial Networks, SRGANs, significantly improved the image quality, according to X. Wang et al. (2018) there is still a big gap between the super-resolved output images of the SRGAN and the original high-resolution images in terms of image quality. When compared with the original images, SRGAN applied images have been demonstrated to have hallucinated details when images are zoomed in. To overcome the quality problems of SRGANs, Enhanced Super-Resolution Generative Adversarial Networks (ESRGANs) were proposed. ESRGANs are also an example of Generative Adversarial Networks, like SRGANs. They are built upon SRGANs to reduce the problem of hallucinated details and to produce higher-quality SR images. The network structure is practically the same as SRGANs. However, to yield fine-detailed images, there have been some minor modifications. The problems of SRGANs and the improvements for those problems with the use of ESRGANs will further be discussed in the Methods Section 4.3.1: ESRGAN Model. In addition to similar network structure, the application areas of ESRGANs are also similar to that of SRGANs, such as healthcare, satellite imagery.

One example application of ESRGAN, which concerns the healthcare domain like this study, is for 3D brain-MRI. Even though there are high-resolution MRI scanning technologies, image SR techniques can also be used in MRI scanings. The high-resolution MRI technologies have been proved to be useful in accurate brain scan analysis. However, these technologies tend to be slow (J. Wang, Chen, Wu, Shi, & Gee, 2019). The reason for applying ESRGAN is to see whether the brain tissue analysis can be made faster. Authors of the ESRGAN for brain MRI paper proposed to add a new generator, namely "Memory efficient Residual Dense

Generator" (MRDG), to the existing ESRGAN architecture. This ESRGAN model's performance was compared to other previously proposed popular SR models such as SRCNN, SRResNet, and also to the original HR image. It has been found that ESRGAN with Memory efficient Residual Dense Generator outperforms all the other images by yielding images with better resolutions.

Another application of ESRGAN concerns the area of earth sciences. Landslide is a crucial natural disaster and it's monitoring is important. To carefully monitor and evaluate these landslides, Interferometric Synthetic Aperture Rader (InSAR) technology is used (Wu & Ma, 2020). After capturing the slope and terrain information of the land by InSAR technologies, information is shown visually in digital elevation models (DEMs). However, DEM models can lack high resolution in the areas with steep terrains. This can lead to difficulties in the correct monitoring of the possible landslides. Therefore, Wu and Ma, investigated the effect of applying ESRGAN on DEMs to address the limited resolution problem of DEMs. It has been found that DEMs with ESRGANs, outperformed the original DEMs in correctly monitoring the slope deformations and possible landslides.

To conclude, ESRGANs have been verified to be a powerful SR algorithm in many different areas, such as satellite imagery, earth sciences and healthcare (Malczewski & Stasiński, 2009; Song et al., 2020; Wu & Ma, 2020). In this study, the main focus will be on the application of this model to breast-cancer tumor images.

2.2 Tumor Classification

Cancer is considered one of the main causes of high illness-related mortality rates for long periods of time (Akash Kumar, Mishra, & Arora, 2021). Since there are different kinds of cancer occurring in different parts of the human body, e. g., breasts, lungs and brain, detecting the cancer type and applying specific treatments accordingly is crucial. Other than where the cancer is, identifying whether the cancer is benign or malignant is also critical since malignant cells spread very quickly. If this spread is not prevented, cancer cells lead to the death of patients.

For the classification of cells into benign and malignant different classification models have been used throughout the literature. The study of Spanhol, Oliveira, Petitjean, and Heutte focused on using traditional machine learning algorithms such as Random Forests and Support Vector Machines (SVMs). By building upon those models, a recent paper was published by the same authors. In the 2019 paper of Spanhol et al., a deep learning approach was preferred and Convolutional Neural Networks (CNNs) have been used. One of the main reasons CNNs are used for

tumor classification is because they are fast to perform and these network models yield good accuracy results (Akash Kumar et al., 2021; Hasan, Fatemi, Khan, Kaur, & Zaguia, 2021). They are proven to be outperforming traditional machine learning models in the classification tasks. In the paper of Spanhol et al., CNNs provided a 6% higher accuracy than the models in the previous study.

3 METHODS AND APPROACH

In this part, the main model of this research, Enhanced Super Resolution Generative Adversarial Networks (ESRGANs) and Convolutional Neural Networks (CNNs), will be discussed. The network architectures of these models will be the main focus of this section. To understand ESRGANs clearly, firstly Generative Adversarial Networks (GANs) and then the Super-Resolution Generative Adversarial Networks (SRGANs) will be reviewed since ESRGANs are built upon GANs and SRGANs.

3.1 *Generative Adversarial Networks (GANs)*

Generative Adversarial Networks (GANs) are an example of deep neural networks, consisting of deep layers. GANs have been proposed in 2014 by Goodfellow et al.. The aim of GANs is to create indistinguishable duplicates of the input images. GANs are comprised of two parts to achieve this aim. These two parts are known as the Generator and the Discriminator and have the following properties:

1. Generator: In image processing tasks, generator tries to duplicate the input training images that is as realistic as the original image.
2. Discriminator: The aim of the discriminator is to distinguish between the real images and the images created by the discriminator (Goodfellow et al., 2014).

Generator, however, does not have direct interaction with the real images. It tries to optimize the generating process by getting feedback from the discriminator (Creswell et al., 2018). If the discriminator can distinguish well between the generated (fake) image and the real image, the generator tries to optimize its processes. If the discriminator cannot distinguish between the fake and real images, it can be understood that the generator did a good job in "fooling" the discriminator and creating duplicates of the input images.

3.2 Super Resolution Generative Adversarial Networks (SRGANs) Architecture

Super-Resolution Generative Adversarial Networks, usually known as SRGANs, are proposed by [Ledig et al. \(2017\)](#). SRGANs utilize the network structure of GANs. The aim of this network model is to create higher-quality versions, duplicates of the low-resolution input images. The generation and discrimination process occur in a similar way as in the regular GANs. According to the creators of the SRGAN model, this model is the first model able to capture photo-realistic features and details of the input images ([Ledig et al., 2017](#)).

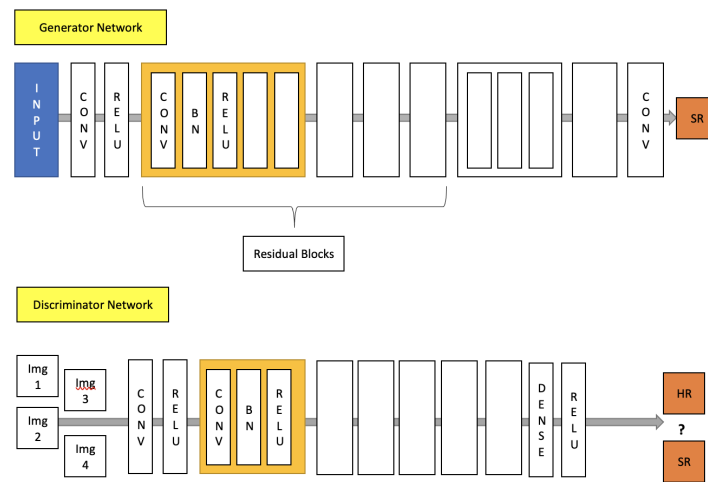


Figure 1: Generator and Discriminator structures of the SRGAN model.

Figure 1 shows the Generator and Discriminator architectures of the SRGAN. Generator, as mentioned earlier in Section 3.2: Generative Adversarial Networks, are responsible for creating duplicates of the low-resolution input data. This low-resolution input (LR) downsampling an HR image using the Gaussian Kernel ([Ledig et al., 2017](#)). The generator creates an SR image that is indistinguishable and as realistic as possible as the original LR inputs. While making the output SR image indistinguishable and realistic, the main aim is to produce it with as many fine details as possible, with better perceptual quality.

As it can be seen, the Generator takes the Low-Resolution image as input and after a couple of stacked layers, outputs a Super-Resolved image. The discriminator then takes multiple input images (the original/real HR image, generated SR image, and other possible real images) to distinguish between the original High-Resolution image and the SRGAN's Super-Resolved output image. The discriminator then tries to distinguish between the SR image and the original image.

As mentioned earlier by [X. Wang et al.](#), SRGAN's network architecture have some drawbacks. In the next subsection, the drawbacks of the network structure of SRGANs will be discussed through the explanation of the network architecture of the optimized version of SRGAN, which is ESRGAN.

3.3 *Enhanced Super Resolution Generative Adversarial Networks (ESRGANs)*

Enhanced Super Resolution Generative Adversarial Networks are built upon SRGANs to overcome the drawbacks encountered when applying SRGAN. The network architecture of the ESRGAN model is build in light of the existing SRGAN architecture ([X. Wang et al., 2018](#)). ESRGANs are "enhanced" versions of the SRGANs that have been created to overcome the image quality problem of SRGANs.

Output SR images produced by the SRGAN model are proved to have hallucinated details rather than fine details ([X. Wang et al., 2018](#)). When zoomed in, the details of the original images still seem to be blurred. [X. Wang et al.](#), proposed that these hallucinated details are due to the following parts of the SRGAN architecture:

1. presence of Batch Normalization layer in the Residual Blocks in the generator structure,
2. type of the discriminator,
3. use of VGG features of the perceptual loss before activation layers.

To further improve the SRGAN model, [X. Wang et al. \(2018\)](#) suggested to make modifications to the network architecture of the SRGAN model. These modifications are:

1. Rather than using Residual Blocks, using "Residual in Residual Blocks" and removing the Batch Normalization layers.
 - The reason for changing residual blocks with Residual in Residual Blocks are to increase the capacity of the network. More blocks mean the networks have more layers and it is able to perform more.
 - By removing Batch Normalization (BN) layers, performance of the network is increased and computational complexity was reduced. In addition to these, it was proven that the artifacts like eliminating fine details of images are reduced after the removal of BN layers.
2. type of the discriminator,

- The discriminator used in SRGAN is named the "standard" discriminator. In the ESRGAN model, the standard discriminator is changed with the Relativistic Discriminator.
 - The previous discriminator was predicting whether the image is fake or real, it was a binary task. Now, with the Relativistic Discriminator, the aim is to see whether the real image is more realistic than the fake image. With this change, it is seen that the model is able to reveal realistic details of images more compared to the SRGAN model.
3. use of VGG features of the perceptual loss after activation layers.
- Perceptual loss is specific to the SR problem. The perceptual loss is based on the idea of human observation, not the standard evaluation metrics. Rather than minimizing error in the pixel space, it is minimized in image representations. The pixel values of images are not compared one by one but their perceptual representations are compared.
 - It has been found that, using perceptual loss features before activation but not after, yields images with better brightness, more detailed texture and edges that are sharper.

With these proposed changes, ESRGANs were created. It has been found that, ESRGANs were able to outperform SRGANs in the single image super-resolution task by yielding images with better qualities. Figure 2 shows the network architecture of the ESRGAN model. The authors employ the SRResNet structure, which is also the network structure of SRGAN (Ledig et al., 2017; X. Wang et al., 2018). However, the authors of the ESRGAN paper made changes to the Residual Block structure depicted in Figure 3.

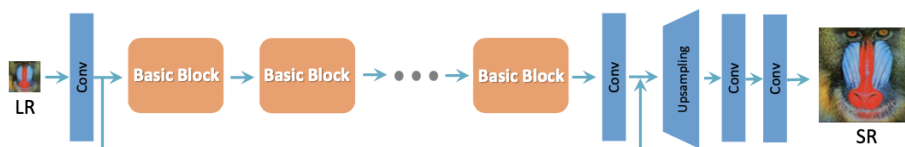


Figure 2: The network structure of the ESRGAN model. SRResNet structure is employed. Basic blocks are also known as the Residual in Residual Blocks. This image is from the original ESRGAN paper (X. Wang et al., 2018)

In Figure 3, on the right, the Residual Block structure of the SRGAN model is shown. On the left, the modified Residual Block structure named as Residual-in-Residual Block of the ESRGAN model is shown. Residual-in-Residual Blocks contain more layers compared to Residual-Blocks and

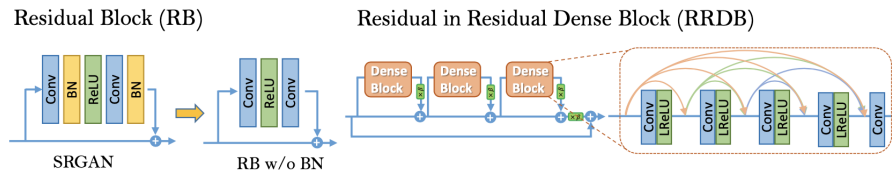


Figure 3: While the left part shows the Residual Block structure of SRGAN, the right part shows the modified Residual Blocks of the ESRGAN model, named as Residual in Residual Blocks. This image was used from the original ESRGAN paper (X. Wang et al., 2018)

there are no Batch Normalization (BN) layers. According to X. Wang et al., BN layers are one of the main reasons for the SRGAN model yielding hallucinated details and the model having a high computational complexity. In addition to these, it is believed that, adding more layers to the Residual-in-Residual Blocks boost the performance of the ESRGAN model resulting in higher quality output images.

3.4 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are one of the powerful deep learning networks with multiple (deep) layers. The name "convolutional" comes from the fact that these networks apply a mathematical operation, convolution, to input values.

CNNs are composed of multiple stacked layers and Figure 4, shows the architecture with the layers mentioned below:

- **Input layer:** In this layer, the values are fed into the CNNs. In the case of this research, these are pixel values of images fed as arrays.
- **Convolutional layer:** This is the layer where the convolution operation is applied to pixel values.
 - Convolution operation: Filter, also known as the 'kernel', is applied to input values via dot product.
 - The aim is to collect relevant features and reduce the number of features that will be passed to the next layer, which is the pooling layer. Key features are captured in convolution.
- **Pooling layer:** the layer where downsampling is applied to the values that are obtained from the convolutional layer. The aim here is to reduce the dimensionality so that the computational complexity is also reduced.

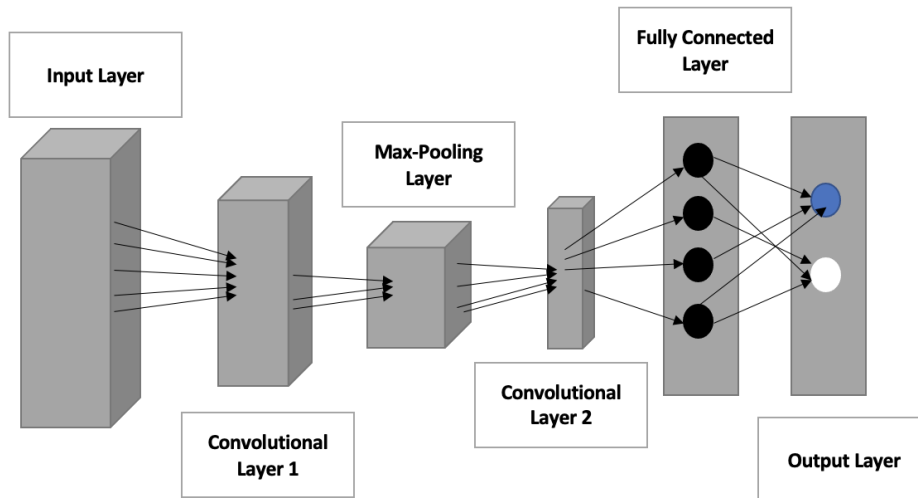


Figure 4: Structure of a Convolutional Neural Network (CNN). Input image pixels are processed by the "Input Layer" and those pixel values are then passed onto the "Convolutional Layer". In the Convolutional Layer, kernel is passed through the pixel values and the output is fed into the "Pooling Layer". After values are downsampled in the Pooling layer, Fully Connected gets the values and makes the classification task and the output is obtained.

- **Fully-connected layer:** This layer takes the output of the pooling layer as an input, flattens them, and produces scores per each class in the classification task (Ertizman, 2021).

CNNs are powerful in many different tasks such as image processing, image classification and speech recognition (Spanhol et al., 2016a). In the scope of this paper, CNNs will be used for image classification.

4 EXPERIMENT DETAILS

4.1 Dataset

The dataset used in this research is the 'Breast Cancer Histopathological Database (BreakHis)'. This dataset contains 9019 histopathological images collected from 82 breast-cancer patients (Spanhol et al., 2016b). The data is obtained from the platform called Kaggle, where many deep and machine learning related datasets can be found ¹.

The dataset contains 2480 benign, and 5429 malignant breast cancer tumor images with a size/resolution of 700X460. The BreakHis dataset

¹ <https://www.kaggle.com/datasets/ambarish/breakhis>

is useful for the purpose of this study because this dataset is specifically created for the classification of malignant and benign tumors. Therefore, images are labeled as malignant and benign beforehand. Four different magnifying factors are present in the dataset: 40X, 100X, 200X, and 400X. The tumors are also separated into different histological subtypes, however, the subtypes are outside the scope of this study.

4.2 Dataset Pre-processing

In this study, the tumor classification performances of three different kinds of images are compared:

1. Original images
2. Low-Resolution images
3. ESRGAN-applied images

To obtain the images mentioned above, the following data pre-processing steps are followed:

- **Selecting an appropriate magnification factor:** As mentioned in the previous subsection (Subsection 4.1), the BreakHis dataset contains different magnification factors. The previous studies using this BreakHis dataset used all of the present magnification factors. However, to limit the scope of the study to especially on the effect of ESRGANs and not focus on its effect on different magnification factors, a single magnification factor was chosen. Out of all other magnifications, there were more images in 200x magnification. Therefore, in order to test the ESRGAN model efficiently with enough images at hand, this magnification factor was chosen.
- **Image Downscaling:** Since the ESRGAN model requires low-resolution images as inputs, images are resized from 700x460 to 128x128. The reason behind choosing the size of 128x128 is the fact that the original the ESRGAN paper also downscales high-resolution images to this size. In order to apply the ESRGAN model correctly, all of the images are transformed to this size (Ledig et al., 2017).
- **Image Upscaling:** Low-resolution images with the size of 128x128 are upscaled to 512x512.
- **Resizing original images:** Changing the size from 700x460 to 512x512.

The third step and the last step is performed to make a reasonable comparison of LR images and original images to ESRGAN images. ESRGAN model outputs images of size 512x512. To compare CNN accuracies

correctly, all of the images need to be of the same size. Therefore, the sizes of both LR images and original images were changed to 512x512 as ESRGAN images.

4.3 *Experimental Procedure*

This study is composed of three parts:

1. Dataset Pre-Preprocessing: Data is pre-processed as mentioned in the Subsection 4.2 to obtain LR images and also original images with an appropriate size for the CNN classification after the ESRGAN application.
2. ESRGAN Model Application to Low-Resolution (Downscaled) Images
3. CNN Classification

4.3.1 *ESRGAN Model*

In this step of the study, using the GitHub code of the original ESRGAN paper written by [X. Wang et al.](#), the image super-resolution model ESRGAN was applied on low-resolution (downscaled) images.

The step-by-step guide about how to clone the repository was provided in the GitHub repository. Using the instructions of the repository, the ESRGAN model was applied on low-resolution images using MacOS Terminal.

4.3.2 *Convolutional Neural Network (CNN) Model for Classification*

After the application of the ESRGAN, CNN is used for binary the classification task of malignant and benign tumors.

Originally, images are named with their respective cancer subtypes (ductal carcinoma, lobular carcinoma...). However, the focus of this paper was not on the subtypes but on whether the images are benign or malignant. Therefore, images were renamed with the respective cancer types (benign/malignant) and a number for better readability (e.g. 'benign_1').

After renaming, before feeding the images into the CNN, train-test-validation datasets were created. A function was written to separate images into training-test-validation sets. This function randomly chose the images to avoid overfitting. After the train-test-validation split, there were 783 images in the training set, 615 images in the test, and the validation sets containing both benign and malignant images.

During the experiment, image data were fed into the CNN using the "image_dataset_from_directory" method belonging to the Keras Pre-processing Library.

A GitHub repository ² was created for the CNN code and detailed code can be found there.

The best performance for the CNN Classification was obtained with the following hyperparameters for both SR and LR images:

- Batch size: 32
- Number of epochs: 30
- Optimizer choice: RMSProp
- Loss function: Binary Cross Entropy (since we have a binary classification task)

4.4 Software

All of the programming tasks were performed using Python 3. The ESRGAN model was applied using MacOS terminal when cloning a GitHub repository as discussed in the "ESRGAN" section of the subsection "Experimental Procedure".

CNN Classification was implemented using Google Colaboratory (also known as Google Colab). The reason for using Google Colab is due to CNN algorithm requires high-speed GPU support, and Google Colab can handle this with its built-in GPU service.

Deep learning models, both ESRGAN and CNN, were applied using the Keras Library. Keras Library is a high-level Application Programming Interface (API) of Tensorflow 2.0 (John Joseph, Nonsiri, & Monsakul, 2021).

In addition, the following Python libraries were used throughout the experimental procedure:

- Numpy (Harris et al., 2020)
- OpenCV, also known as cv2 (Bradski, 2000)

5 RESULTS

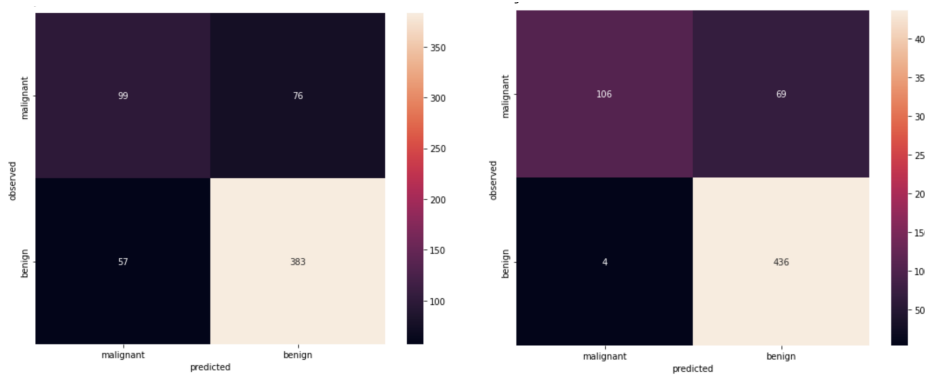
In this section, the tumor-classification results without and with ESRGAN enhancement are reported. The detailed CNN model with the number of layers involved can be found in Appendix A. Table 1 lists the results. The results are as expected: the ESRGAN enhanced images elevate the

² <https://github.com/keces/Tilburg-University-BSc.-CSAI-Thesis>

accuracy by 8%, from 0.7837 to 0.8634. By comparing the enhanced image performance to the optimal performance obtained by the original high-resolution images (last row), it is clear that the ESRGAN was effective in enhancing the image quality, therefore the tumor-classification performance (0.8634 vs 0.7967).

Image Type (512 x 512 pixels)	Accuracy
Low-resolution images without enhancement	0.7837
Super-Resolved images with the ESRGAN enhancement	0.8634
High-resolution images without enhancement	0.7967

Table 1: Classification Accuracy for low-resolution images with and without ESRGAN enhancement. The performance on the high-resolution images (without enhancement) is included for reference.



(a) The Confusion Matrix of the classification of Low-Resolution images (b) The Confusion Matrix of the classification of ESRGAN enhanced images

Figure 5: The Confusion Matrices for the CNN performance on LR images and ESRGAN enhanced images.

Figure 5 shows the Confusion Matrix with the predicted and observed values of CNN classification performance per image class (benign and malignant) of the LR and ESRGAN-enhanced images. For example, while correctly classified benign LR tumor images are 383 on the Matrix (a), the number increases to 436 on the Matrix (b) when ESRGAN is applied. It can also be seen that, 57 of the benign tumors were predicted as malignant when the images are LR images. However, when the image quality was enhanced, incorrect classification lowered down to four.

Tumor Class	Precision	Recall	F1-Score
Malignant	0.63	0.57	0.60
Benign	0.83	0.87	0.85

Table 2: Precision, Recall and F1-Scores for the low-resolution images, separately for malignant and benign classes.

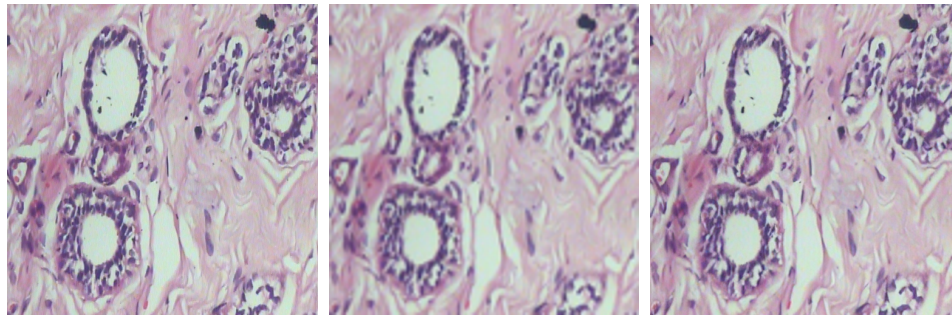
Tumor Class	Precision	Recall	F1-Score
Malignant	0.94	0.55	0.70
Benign	0.85	0.99	0.91

Table 3: Precision, Recall and F1-Scores for the ESRGAN applied images, separately for malignant and benign classes.

Above tables, namely Table 2 and Table 3, show the detailed analysis of results of the CNN performances on low-resolution and ESRGAN applied images. As it can be seen, the scores for benign and malignant images differ from each other substantially. This is mostly due to the imbalance between the number of benign and malignant tumor images. This is discussed again in the Section 6: Discussion and Future Work.

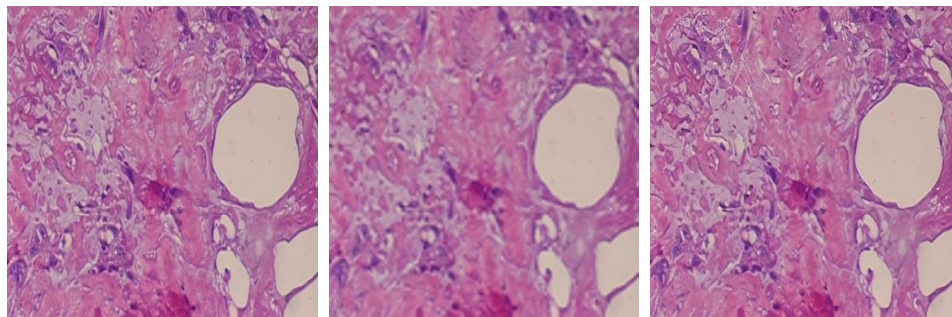
The training-validation accuracy plots and training-validation loss plots can be found in Appendix B and C, respectively. However, it is worth explaining these plots briefly. As it can be seen, for both low-resolution and ESRGAN-applied images, the training accuracy increases by time, since the model gets adapted to the images. Since the model's accuracy increases for the images, the loss for both training and validation sets decreases over time.

Figures 6 and 7 are examples of cancer images produced in this study. Figure 6 consists of benign images of three different image types, namely, original images, low-resolution image and the ESRGAN applied image. Figure 7 is the same as 6, but in this case an example of a malignant cancer image is shown. As it can be seen, the first and last images have a better quality compared to the image in the middle, which is the low-resolution image. The third image is the ESRGAN applied, super-resolved image.



(a) Original image (b) Low-Resolution image (c) ESRGAN applied image

Figure 6: Three different image types of the same benign cancer image. All have a size of 512×512 . First image is an original image, only resized to 512×512 . Second image is the low-resolution benign cancer image, downsampled to 128×128 and then upsampled to 512×512 . The third image is the ESRGAN applied benign cancer image. As it can be seen, the second image does not have a good quality and looks blurred in contrast to other high-resolution images.



(a) Original image (b) Low-Resolution image (c) ESRGAN applied image

Figure 7: Three different image types of the same malignant cancer image. All have a size of 512×512 . The first image is an example of an original malignant cancer image resized to 512×512 . Second image is the low-resolution malignant cancer image, downsampled to 128×128 and then upsampled to 512×512 . The third image is the ESRGAN applied malignant cancer image. As it can be seen, the second image does not have a good quality with a blurry look when compared to other high-resolution images.

6 DISCUSSION AND FUTURE WORK

The goal of this study was to show the effect of applying an image super resolution technique called Enhanced Super-Resolution Generative Adversarial Networks (ESRGANs) on breast cancer tumor images.

As discussed before, histopathological cancer images are analyzed by histopathologists. The use of human power in histopathology can sometimes be disadvantageous since there can be mistakes in the correct

analysis and detection of cancerous cells. In addition to this, there exists another problem in histopathology, which is the quality of histopathological images. Images that are analyzed in cancer research can sometimes have a bad quality due to imaging procedures and the systems used. Therefore, low-resolution histopathological cancer tumor images can lead to the incorrect classification of benign and cancer tumors. While there exists medical imaging tools to solve the issue of low-resolution histopathology images, they tend to be expensive and time-consuming. Recently proposed models that use deep-learning networks, known as image super-resolution (SR) models, have been demonstrated to be effective in cancer image research for increasing image quality (Celik & Talu, 2020). In addition, it has also been shown that, SR models are effective in correctly classifying benign and malignant tumors (Shahidi, 2021). SRGAN and WA-SRGAN were the previously applied SR models on breast tumor images. A recently proposed, powerful SR model, ESRGAN were not tested on breast tumor images. Therefore, in this study, ESRGAN model proposed by X. Wang et al., were applied to the BreakHis breast cancer tumor dataset.

Results that are presented in Section 5: Results, have supported the hypothesis of this study. A classification task was performed using Convolutional Neural Networks (CNNs) to classify between malignant and benign breast cancer tumors. In this classification task, low-resolution images scored 0.7837 accuracy. Where the ESRGAN-applied breast-tumor images scored 0.8634 accuracy. SR images outperformed LR images by 7.97%, which is a big difference in terms of CNN classification accuracy. In addition to that, SR images' performances also outperformed the original HR images, in terms of accuracy.

The precision, recall and F1-score per tumor classes show that there is a performance difference between classes. The reason behind this difference is due to having an imbalance between the number of benign and malignant tumor images. As a future work, data augmentation methods can be applied on images to obtain more images for the less frequent image class. After the increase, the CNN classification can be performed and new scores can be obtained.

In this study, from the BreakHis dataset, in order to limit the scope of this study to focusing only on the effect of the SR model ESRGAN and not to compare its effect on different magnification factors, a single magnification factor, 200x, was preferred. However, as a future work, the effect of ESRGAN on different magnification factors can further be investigated.

As mentioned in the introduction, before this study, the ESRGAN model was not tested on tumor classification tasks. However, there exists other SR models that agree with the hypothesis of this study. One example

study conducted by [Shahidi](#). In the study of [Shahidi](#), it has been found that SR models such as SRGAN outperform LR images in tumor classification tasks. Authors have found that, while LR images performed 95.82% in image classification, for example, the SRGAN performed 99.11% accuracy.

ESRGANs were proposed in 2018. Until then, new image super-resolution models have been proposed. Some of them were actually build in light of the SRGAN model just like ESRGAN, such as Real-ESRGAN and RankSRGAN ([X. Wang, Xie, Dong, & Shan, 2021](#); [Zhang, Liu, Dong, & Qiao, 2021](#)). Real-ESRGAN is the new version of the ESRGAN model, proposed by the authors of the ESRGAN paper. The difference of this model to the previously proposed ESRGAN model is that, this model focuses on the "real degradations" in the images. With real degradations, authors wanted to emphasize the unknown and complex resolution disruptors. They occur naturally because of the imaging systems of cameras, transmission of images through the internet, and image edits ([X. Wang et al., 2021](#)). As stated by the authors, the performance of the Real-ESRGAN model is tested by taking an input of naturally low-resolution images, not bicubically downsampled images. The original papers of SRGANs and ESRGANs focused on the "perceptual" evaluation of these models. They used perceptual metrics for evaluation. Rank-SRGANs, however, were intended to evaluate the model with a "ranker". The Ranker works in the following way: it has been trained to learn the behavior perceptual evaluation metrics. Then, there is the loss function just like in the SRGAN and ESRGAN model, namely the rank-content loss. This rank-content loss function tries to optimize the perceptual quality of the images. This model was evaluated against ESRGAN and SRGAN. It was found that this newly proposed Real-SRGAN model outperforms both models. To conclude, as a future work, Real-ESRGANs and Rank-SRGANs can be tested on histopathological breast-tumor images. These model's CNN accuracies can then be compared to the results of this study.

7 CONCLUSION

This study aimed to investigate the effect of deep-learning based image super-resolution technique on breast-cancer tumor images. This technique was named as Enhanced Super-Resolution Generative Adversarial Networks, shortly ESRGANs, by the authors [X. Wang et al. \(2018\)](#).

Histopathology deals with the analysis of cancer tissues. The use of human power has been a common practice for long periods of time in histopathology and people analyzing these cancer tissues are called as histopathologists. However, it has been suggested that, histopathology is very error-prone due to using human power in this field ([Bardou et al.,](#)

2018). Therefore, the improvements in technology gave rise to applying Machine Learning (ML) algorithms in histopathology. An example is using Deep Learning Classification techniques in the classification of malignant and benign tumors (Shahidi et al., 2020).

Nonetheless, there exists another problem in histopathological tumor images which is the main reason why this paper has emerged. The problem is that the histopathological images can sometimes have a low-resolution and bad quality. While imaging systems in histopathology have been improved with the advancements in technology, these systems tend to be expensive to use. In addition to that these imaging systems can be time-consuming. Deep-learning based image super-resolution techniques have been proved to be useful in healthcare to minimize costs and time (Vahadane, Kumar, & Sethi, 2016). While there exist many different SR algorithms in the literature, this study focused on the powerful ESRGAN algorithm which is an example of a Generative Adversarial Network created specifically for the image super-resolution task (X. Wang et al., 2018). With the aforementioned problems and a possible solution for the task at hand, the following research question has been addressed: "To what extent can ESRGANs improve the accuracy of classification task of benign and malignant breast tumors?".

To conclude, it has been found that, ESRGANs are powerful SR algorithms that can be applied on breast-tumor images to enhance the quality so that the benign and malignant images can be classified correctly. This has been shown by using Convolutional Neural Networks and by ESRGANs outperforming Low-Resolution images in accuracy scores.

REFERENCES

- Akash Kumar, V., Mishra, V., & Arora, M. (2021). Convolutional neural networks for malignant and benign cell classification using dermatoscopic images. In *2021 third international conference on intelligent communication technologies and virtual mobile networks (icicv)* (p. 1040-1044). doi: 10.1109/ICICV50876.2021.9388605
- Bardou, D., Zhang, K., & Ahmad, S. M. (2018). Classification of breast cancer based on histology images using convolutional neural networks. *IEEE Access*, 6, 24680-24693. doi: 10.1109/ACCESS.2018.2831280
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.
- Celik, G., & Talu, M. F. (2020). Resizing and cleaning of histopathological images using generative adversarial networks. *Physica A: Statistical Mechanics and its Applications*, 554, 122652.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1), 53-65. doi: 10.1109/MSP.2017.2765202
- Dong, C., Loy, C. C., He, K., & Tang, X. (2015). *Image super-resolution using deep convolutional networks*.
- Ertizman, B. (2021). *In search of pain: A spatiotemporal model for the pain intensity assessment* (Unpublished master's thesis). Tilburg University.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... Bengio, Y. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, & K. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 27). Curran Associates, Inc. Retrieved from <https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... Oliphant, T. E. (2020, September). Array programming with NumPy. *Nature*, 585(7825), 357-362. Retrieved from <https://doi.org/10.1038/s41586-020-2649-2> doi: 10.1038/s41586-020-2649-2
- Hasan, M. R., Fatemi, M. I., Khan, M. M., Kaur, M., & Zaguia, A. (2021). Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks. *Journal of Healthcare Engineering*, vol. 2021. doi: 10.1155/2021/5895156
- John Joseph, F. J., Nonsiri, S., & Monsakul, A. (2021, 07). Keras and tensorflow: A hands-on experience. In (p. 85-111). doi: 10.1007/978-3-030-66519-7_4
- Kalaiyarasi, M., Dhanasekar, R., Ram, S. S., & Vaishnavi, P. (2020, nov).

- Classification of benign or malignant tumor using machine learning. *IOP Conference Series: Materials Science and Engineering*, 995(1), 012028. Retrieved from <https://doi.org/10.1088/1757-899x/995/1/012028> doi: 10.1088/1757-899x/995/1/012028
- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8-17. Retrieved from <https://www.sciencedirect.com/science/article/pii/S2001037014000464> doi: <https://doi.org/10.1016/j.csbj.2014.11.005>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015, 05). Deep learning. *Nature*, 521, 436-44. doi: 10.1038/nature14539
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., ... Shi, W. (2017). *Photo-realistic single image super-resolution using a generative adversarial network*.
- Luo, C., Li, N., Lu, B., Cai, J., Lu, M., Zhang, Y., ... Dai, M. (2022, 01). Global and regional trends in incidence and mortality of female breast cancer and associated factors at national level in 2000 to 2019. *Chinese Medical Journal*, 135, 42-51. doi: 10.1097/CM9.0000000000001814
- Malczewski, K., & Stasiński, R. (2009). Super resolution for multimedia, image, and video processing applications. In M. Grgic, K. Delac, & M. Ghanbari (Eds.), *Recent advances in multimedia signal processing and communications* (pp. 171-172). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from https://doi.org/10.1007/978-3-642-02900-4_8 doi: 10.1007/978-3-642-02900-4_8
- Motlagh, M. H., Jannesari, M., Aboulkheyr, H., Khosravi, P., Elemento, O., Totonchi, M., & Hajirasouliha, I. (2018). Breast cancer histopathological image classification: A deep learning approach. *bioRxiv*. doi: 10.1101/242818
- Musumeci, G. (2014). Past, present and future: overview on histology and histopathology. *Journal of Histology and Histopathology*, 1-5.
- Nain, S. (2015, 09). Cancer: An overview. *Academic Journal of Cancer Research*, 1-9.
- Ooi, Y. K., & Ibrahim, H. (2021). Deep learning algorithms for single image super-resolution: A systematic review. *Electronics*, 10(7). Retrieved from <https://www.mdpi.com/2079-9292/10/7/867> doi: 10.3390/electronics10070867
- Patel, A. (2020, 09). Benign vs Malignant Tumors. *JAMA Oncology*, 6(9), 1488-1488. Retrieved from <https://doi.org/10.1001/jamaoncol.2020.2592> doi: 10.1001/jamaoncol.2020.2592
- Sert, E., Ozyurt, F., & Dogantekin, A. (2019). A new approach for brain tumor diagnosis system: Single image super resolu-

- tion based maximum fuzzy entropy segmentation and convolutional neural network. *Medical Hypotheses*, 133, 109413. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0306987719310102> doi: <https://doi.org/10.1016/j.mehy.2019.109413>
- Shahidi, F. (2021). Breast cancer histopathology image super-resolution using wide-attention gan with improved wasserstein gradient penalty and perceptual loss. *IEEE Access*, 9(3), 32795–32809.
- Shahidi, F., Mohd Daud, S., Abas, H., Ahmad, N. A., & Maarop, N. (2020). Breast cancer classification using deep learning approaches and histopathology image: A comparison study. *IEEE Access*, 8, 187531–187552. doi: 10.1109/ACCESS.2020.3029881
- Song, T.-A., Chowdhury, S. R., Yang, F., & Dutta, J. (2020). Pet image super-resolution using generative adversarial networks. *Neural Networks*, 125, 83–91. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0893608020300393> doi: <https://doi.org/10.1016/j.neunet.2020.01.029>
- Sood, R., Topiwala, B., Choutagunta, K., Sood, R., & Rusu, M. (2018). An application of generative adversarial networks for super resolution medical imaging. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (p. 326–331). doi: 10.1109/ICMLA.2018.00055
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016a). Breast cancer histopathological image classification using convolutional neural networks. In *2016 International Joint Conference on Neural Networks (IJCNN)* (p. 2560–2567). doi: 10.1109/IJCNN.2016.7727519
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016b). A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7), 1455–1462. doi: 10.1109/TBME.2015.2496264
- Umehara, K., Ota, J., & Ishida, T. (2017, 11). Super-resolution imaging of mammograms based on the super-resolution convolutional neural network. *Open Journal of Medical Imaging*, 07, 180–195. doi: 10.4236/ojmi.2017.74018
- Vahadane, A., Kumar, N., & Sethi, A. (2016). Learning based super-resolution of histological images. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* (p. 816–819). doi: 10.1109/ISBI.2016.7493391
- Wang, J., Chen, Y., Wu, Y., Shi, J., & Gee, J. (2019). *Enhanced generative adversarial network for 3d brain mri super-resolution*. arXiv. Retrieved from <https://arxiv.org/abs/1907.04835> doi: 10.48550/ARXIV.1907.04835

- Wang, X., Xie, L., Dong, C., & Shan, Y. (2021). Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *IEEE/CVF international conference on computer vision workshops, ICCVW 2021, montreal, bc, canada, october 11-17, 2021* (pp. 1905–1914). IEEE. Retrieved from <https://doi.org/10.1109/ICCVW54120.2021.00217> doi: 10.1109/ICCVW54120.2021.00217
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., ... Tang, X. (2018). *Esrgan: Enhanced super-resolution generative adversarial networks*.
- Wu, Z., & Ma, P. (2020). Esrgan-based dem super-resolution for enhanced slope deformation monitoring in lantau island of hong kong. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B3-2020*, 351–356. Retrieved from <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLIII-B3-2020/351/2020/> doi: 10.5194/isprs-archives-XLIII-B3-2020-351-2020
- Yang, W., Zhang, X., Tian, Y., Wang, W., & Xue, J. (2018). Deep learning for single image super-resolution: A brief review. *CoRR, abs/1808.03344*. Retrieved from <http://arxiv.org/abs/1808.03344>
- Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., & Zhang, L. (2016). Image super-resolution: The techniques, applications, and future. *Signal Processing*, 128, 389–408. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0165168416300536> doi: <https://doi.org/10.1016/j.sigpro.2016.05.002>
- Zhang, W., Liu, Y., Dong, C., & Qiao, Y. (2021). Ranksrgan: Super resolution generative adversarial networks with learning to rank. *CoRR, abs/2107.09427*. Retrieved from <https://arxiv.org/abs/2107.09427>

APPENDIX A: CONVOLUTIONAL NEURAL NETWORK (CNN) LAYERS

```

Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
conv2d (Conv2D)              (None, 510, 510, 32)       896

max_pooling2d (MaxPooling2D) (None, 255, 255, 32)       0

batch_normalization (BatchN (None, 255, 255, 32)       128
ormalization)

conv2d_1 (Conv2D)            (None, 253, 253, 64)       18496

max_pooling2d_1 (MaxPooling (None, 126, 126, 64)       0
2D)

batch_normalization_1 (Batc (None, 126, 126, 64)       256
hNormalization)

conv2d_2 (Conv2D)            (None, 124, 124, 128)      73856

max_pooling2d_2 (MaxPooling (None, 62, 62, 128)        0
2D)

batch_normalization_2 (Batc (None, 62, 62, 128)        512
hNormalization)

conv2d_3 (Conv2D)            (None, 60, 60, 128)        147584

max_pooling2d_3 (MaxPooling (None, 30, 30, 128)        0
2D)

batch_normalization_3 (Batc (None, 30, 30, 128)        512
hNormalization)

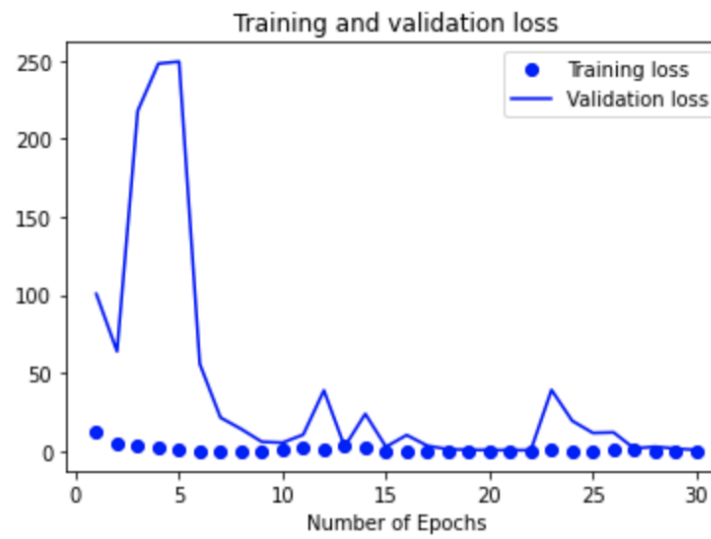
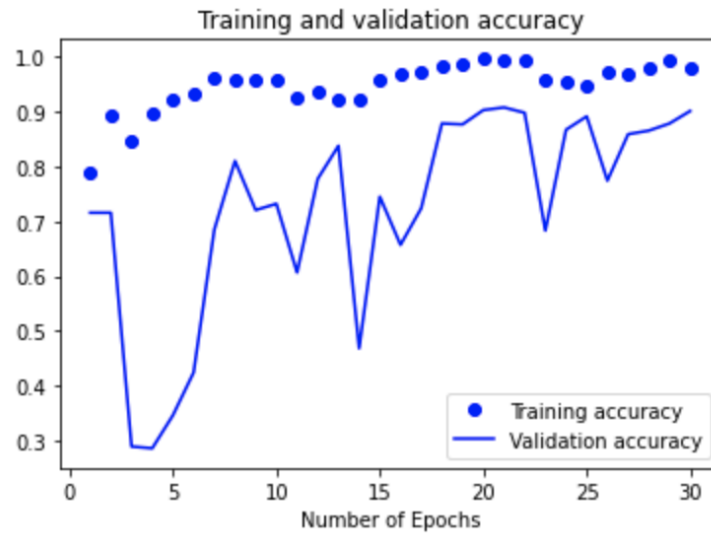
flatten (Flatten)            (None, 115200)              0

dense (Dense)                 (None, 512)                 58982912

dense_1 (Dense)               (None, 1)                   513
-----
Total params: 59,225,665
Trainable params: 59,224,961
Non-trainable params: 704

```

APPENDIX B: TRAINING-VALIDATION ACCURACIES AND LOSSES OF THE LOW RESOLUTION IMAGES



APPENDIX C: TRAINING-VALIDATION ACCURACIES AND LOSSES OF ESRGAN APPLIED SUPER RESOLVED IMAGES

