

**UNICAMP – UNIVERSIDADE ESTADUAL DE CAMPINAS**  
**FT – FACULDADE DE TECNOLOGIA**

**RELATÓRIO TÉCNICO: DESENVOLVIMENTO DE SOFTWARE  
PARA PROCESSAMENTO DE IMAGENS USANDO TÉCNICAS  
DE SUPER-RESOLUÇÃO**

**FÁBIO AUGUSTO ALVES DINIZ**

**LIMEIRA - SP**

**2022**

**FÁBIO AUGUSTO ALVES DINIZ**

**RELATÓRIO TÉCNICO: DESENVOLVIMENTO DE SOFTWARE  
PARA PROCESSAMENTO DE IMAGENS USANDO TÉCNICAS  
DE SUPER-RESOLUÇÃO**

Relatório Técnico realizado para o Trabalho de  
Conclusão de Curso apresentado ao Curso de  
Graduação em Sistemas de Informação da  
Universidade Estadual de Campinas.

Orientador: João Roberto Bertini Junior

**LIMEIRA – SP**

**2022**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Faculdade de Tecnologia  
Luiz Felipe Galeffi - CRB 8/10385

D615r      Diniz, Fábio Augusto Alves, 1997-  
Relatório técnico : desenvolvimento de software para processamento de  
imagens usando técnicas de super resolução / Fábio Augusto Alves Diniz. –  
Limeira, SP : [s.n.], 2022.

Orientador: João Roberto Bertini Junior.  
Trabalho de Conclusão de Curso (graduação) – Universidade Estadual de  
Campinas, Faculdade de Tecnologia.

1. Software - Desenvolvimento. 2. Python (Linguagem de programação de  
computador). 3. Resolução (Ótica). I. Bertini Junior, João Roberto, 1981-. II.  
Universidade Estadual de Campinas. Faculdade de Tecnologia. III. Título.

Informações adicionais, complementares

**Palavras-chave em inglês:**

Computer software - Development  
Python (Computer program language)  
Resolution (Optics)

**Titulação:** Bacharel

**Banca examinadora:**

João Roberto Bertini Junior [Orientador]  
Marcos Augusto Francisco Borges  
Plinio Roberto Souza Vilela

**Data de entrega do trabalho definitivo:** 29-07-2022

## **AGRADECIMENTOS**

Agradeço primeiramente a minha família por terem me dado apoios durante os momentos que mais precisei, a Deus por ter me dado saúde e força para superar todas as dificuldades. Obrigado a todos os professores que eu tive durante a minha graduação, o conhecimento obtido em cada disciplina cursada foi necessário para a conclusão desse projeto, e principalmente ao Prof. João Bertini pela paciência e pela oportunidade me dada para a realização desse trabalho. Meus agradecimentos aos amigos e a todos que fizeram parte da minha formação, seja indireta ou diretamente.

## RESUMO

Apesar da tecnologia estar muito avançada e as câmeras fotográficas tirarem fotos praticamente sem defeitos, em determinados ambientes e com uma luminosidade fraca essas fotos podem sair com alguns serrilhados, e se for necessária uma ampliação de alguma parte da foto pode ser que ela não saia nítida. Portanto, o projeto de desenvolvimento de um software que utilizando de técnicas de Super Resolução fará com que essas imagens fiquem melhores visualmente. Esse tipo de software pode ser usado para ajudar em investigações policiais como reconhecimento facial, na medicina entre outras áreas.

Este relatório tem como objetivo mostrar a minha experiência durante o processo de desenvolvimento de software, usando Python como linguagem, e um algoritmo de Super Resolução.

**Palavras chaves:** Super Resolução, Desenvolvimento de Software, Python, *Deep Learning*, Inteligência Artificial

## LISTA DE FIGURAS

|                                                               |    |
|---------------------------------------------------------------|----|
| Figura 1 - Diagrama de Casos de uso .....                     | 10 |
| Figura 2 - Comparação de imagem de entrada LR e saída HR..... | 12 |
| Figura 3 - Comparação de imagem de entrada LR e saída HR..... | 13 |

# SUMÁRIO

|   |                                              |    |
|---|----------------------------------------------|----|
| 1 | Introdução .....                             | 5  |
|   | Organização do trabalho .....                | 6  |
| 2 | Super-Resolução e <i>Deep Learning</i> ..... | 7  |
|   | Super-Resolução.....                         | 7  |
|   | Super-Resolução com Deep Learning .....      | 7  |
| 3 | Projeto .....                                | 9  |
|   | Análise e Levantamento de Requisitos .....   | 9  |
|   | Desenvolvimento do Sistema.....              | 9  |
| 4 | Testes e Resultados .....                    | 12 |
| 5 | Requisitos de sistema .....                  | 14 |
|   | Requisitos mínimos .....                     | 14 |
|   | Requisitos Recomendados .....                | 14 |
|   | Dependências.....                            | 14 |
| 6 | Tutorial .....                               | 15 |
| 7 | Trabalhos Futuros .....                      | 17 |
|   | Referências: .....                           | 18 |

## 1 Introdução

Super-resolução de imagem única (SISR), como um problema fundamental de visão de baixo nível, tem atraído cada vez mais atenção na comunidade de pesquisa e empresas de IA. O SISR visa recuperar uma imagem de alta resolução (HR) de uma única imagem de baixa resolução (LR). Desde o trabalho pioneiro do SRCNN proposto por Dong et al. [1, 2], abordagens de rede neural de convolução (CNN, do inglês *Convolutional Neural Network*) trouxeram um desenvolvimento próspero. A tarefa de estimar uma imagem de alta resolução a partir de sua contraparte de baixa resolução é conhecida como super-resolução (SR). SR recebeu atenção substancial da comunidade de pesquisa em visão computacional e tem uma ampla gama de aplicações [3, 4, 5].

*Deep Learning* ou Aprendizado profundo é uma área de pesquisa que atualmente está sendo bem ativa, uma grande quantidade de aplicações estão obtendo sucesso, reconhecimento de fala, como a Siri, Cortana e Alexa e reconhecimento facial e até mesmo em diagnósticos no setor de saúde e identificação de fake News. *Deep Learning (DL)* vem sendo utilizada em diversas áreas de estudos como: Reconhecimento de imagens, de áudio e para reconhecimento facial [6, 7]. Hoje várias corporações vêm utilizando dessa técnica em seus projetos, como a Apple e a Siri, Microsoft e a Cortana, Samsung com seus desbloqueios de tela por reconhecimento facial, entre outras.

Atualmente uma das maiores diversões das pessoas é tirar foto para postar nas redes sociais e atrair mais seguidores, porém dependendo da câmera fotográfica utilizada, a imagem que é entregue para o usuário pode não ter a qualidade que ele deseja e por isso ele pode querer editar essa foto para que fique melhor. Visando esse objetivo, uma técnica que aumenta e melhora detalhes em uma imagem, ao pegar uma imagem de baixa resolução e utilizar de técnicas de *Deep Learning* ela pode ser transformada em uma imagem de alta resolução.



## **Organização do trabalho**

Este relatório está organizado com o segundo capítulo falando sobre o que é Super-Resolução e *Deep Learning* e *alguns conceitos sobre eles*, o terceiro capítulo discorre sobre como o projeto foi realizado, a partir da análise e levantamento de requisitos e desenvolvimento do sistema, o quarto capítulo onde é falado sobre os testes, resultados obtidos e um vídeo demonstrativo do sistema, o quinto capítulo dedicado a um tutorial de como utilizar o software e o sétimo para os requisitos mínimos para o sistema funcionar corretamente.

## 2 Super-Resolução e *Deep Learning*

### Super-Resolução

Super-Resolução (SR) pode ser definida como a obtenção de uma imagem de maior resolução a partir de uma imagem de baixa resolução [8], como por exemplo ter a entrada de uma imagem de dimensões 125p x 120p e gerar uma imagem de saída com a resolução 500p x 480p. Um dos métodos para aumentar a resolução de imagens é utilizar a SR. O que torna a SR viável é justamente os deslocamentos em nível de sub-pixel.

O objetivo da super resolução é aumentar a resolução de uma imagem. Resolução é uma medida do conteúdo de frequência em uma imagem: imagens de alta resolução (HR) são limitadas por banda a uma faixa de frequência maior do que imagens de baixa resolução (LR). Na verdade, o hardware necessário processar imagens de RH é caro. A resolução das fotografias digitais é limitada pelo sistema de imagens. Em câmeras convencionais, por exemplo, a resolução depende da densidade do sensor CCD (Dispositivo de carga acoplada), que pode não ser suficiente para fornecê-la. Da mesma forma, os dispositivos infravermelhos e de raios-X têm suas próprias limitações. Super-Resolução é uma abordagem que tenta resolver esse problema com software em vez de hardware.

As tarefas de Super-Resolução de imagem geralmente pegam uma imagem de baixa resolução como entrada e geram uma de alta resolução com detalhes nítidos.

Super-Resolução também se refere à tarefa de restaurar imagens de alta resolução de uma ou mais observações de baixa resolução da mesma cena. Ele pode ser classificado em Super-Resolução de imagem única (SISR) e Super-Resolução de imagem múltipla (MISR).

### Super-Resolução com *Deep Learning*

Métodos SISR baseados em aprendizagem, também conhecidos como métodos baseados em exemplos, são trazidos em foco por causa de sua computação rápida e excelente desempenho. Esses métodos geralmente

utilizam algoritmos de aprendizado de máquina para analisar relações estatísticas entre os LR e sua contraparte de RH correspondente de exemplos de treinamento substancial. Inspirados na teoria de recuperação de sinal esparso, os pesquisadores aplicaram métodos de codificação esparsos a problemas de SISR. Enquanto isso, muitas pesquisas combinaram os méritos dos métodos baseados na reconstrução com os baseados na aprendizagem para reduzir ainda mais os artefatos trazido por exemplos de treinamento externo.

Redes neurais convolucionais profundas (DCNNs) demonstraram desempenho excelente em SISR. Dong [2] apresentou a CNN à Super-resolução de imagem e demonstrou que o aprendizado profundo pode alcançar uma imagem de qualidade superior do que outros métodos baseados em aprendizagem. Eles projetaram uma rede neural totalmente convolucional simples que diretamente aprende um mapeamento ponta a ponta entre imagens de baixa e alta resolução. Além disso, eles apontaram que as três camadas convolucionais podem ser abstraídas em extração e representação de patch, mapeamento não linear e reconstrução, respectivamente. Vários outros modelos são apresentados para melhorar o desempenho de métodos de aprendizagem profunda baseados em CNNs.

### **3 Projeto**

#### **Análise e Levantamento de Requisitos**

O primeiro passo do projeto foi levantar e analisar os principais requisitos para o desenvolvimento do software. Foi perguntado para um grupo de pessoas que gostam de tirar fotos e que gostam de divulgar uma imagem na melhor qualidade possível e foi dito que era necessário apenas um software que selecionasse uma imagem, melhorasse a qualidade e resolução dela e que salvasse no local selecionado.

#### **Desenvolvimento do Sistema**

Para o processo de implementação e desenvolvimento a partir dos requisitos passados pelos usuários, foi necessário que houvesse, inicialmente, um planejamento e para isso foi criado um diagrama de casos de uso. Com esse planejamento foi possível passar para a parte de desenvolvimento do *software*.

Esses diagramas são úteis para facilitar na hora da programação, e para que os usuários saibam quais recursos foram implantados no sistema.

## Sistema de Super-resolução

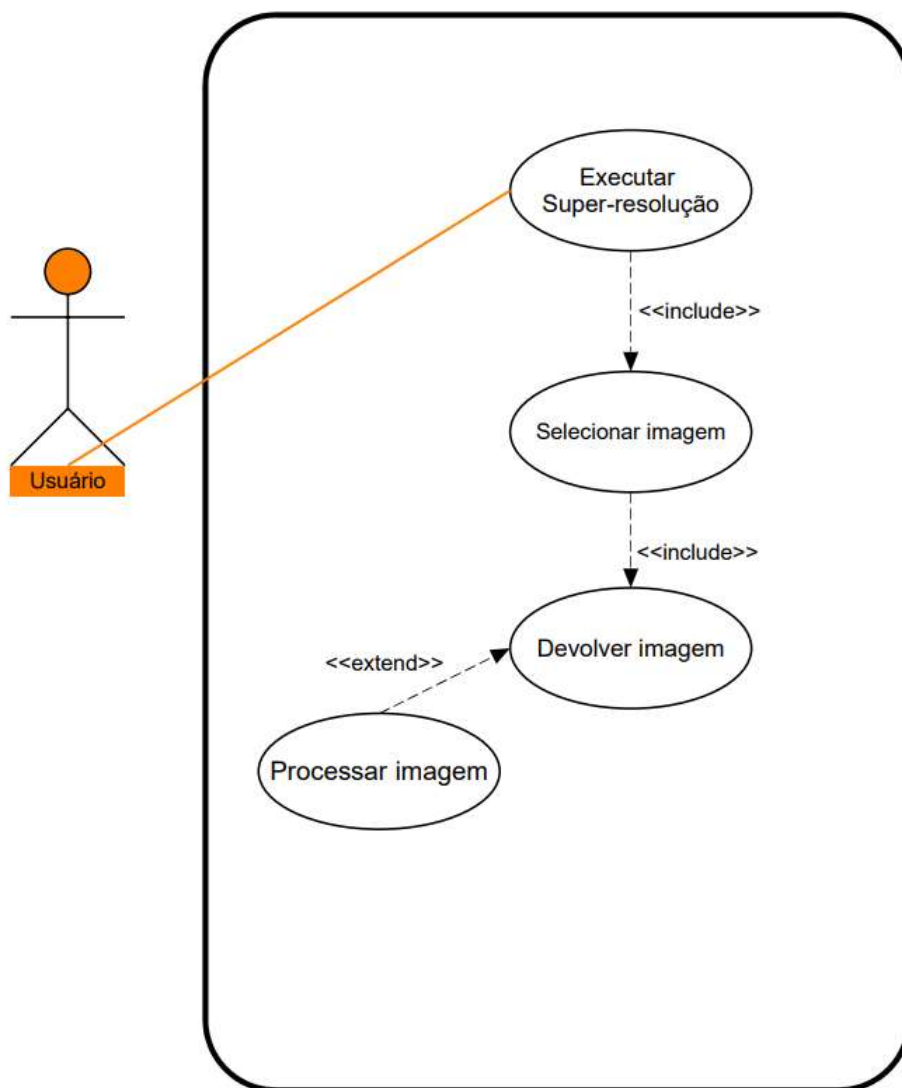


Figura 1 - Diagrama de Casos de uso

A maior dificuldade do desenvolvimento do sistema foi encontrar um algoritmo de Super resolução que resultasse em uma saída satisfatória para o usuário, e que esse tipo de algoritmo normalmente necessita de um computador com especificações não tão acessíveis ao público, fazendo com que dependendo do tamanho da foto, a máquina não tenha memória gráfica o suficiente para processar a imagem.

Após algumas pesquisas foi encontrado o algoritmo ESRGAN (*Enhanced Super-Resolution Generative Adversarial Networks*) [9] e nele foi trabalhado para que os requisitos do sistema fossem atribuídos.

Seguindo SRGAN [10], todos os experimentos são realizados com um fator de escala de 4x entre imagens LR e HR. Obtemos imagens LR reduzindo a taxa de amostragem da imagem HR usando a função kernel bicúbica do MATLAB. O tamanho do mini-batch é definido como 16. O tamanho espacial do patch HR recortado é  $128 \times 128$ .

Observamos que treinar uma rede mais profunda se beneficia de um tamanho de patch maior, uma vez que um campo receptivo ampliado ajuda a capturar mais informações semânticas. No entanto, isso custa mais tempo de treinamento e consome mais recursos de computação. Este fenômeno também é observado em métodos orientados PSNR.

O processo de treinamento é dividido em duas etapas. Primeiro, treinamos um modelo orientado PSNR. A taxa de aprendizado é inicializada como  $2 \times 10^{-4}$  e diminuída por um fator de 2 a cada  $2 \times 10^5$  de atualizações de mini-batch. Em seguida, empregamos o modelo orientado por PSNR treinado como uma inicialização para o gerador. O gerador é treinado usando a função de perda na equação  $\lambda = 5 \times 10^{-3}$  e  $\eta = 1 \times 10^{-2}$ .

$$\lambda = 5 \times 10^{-3} \text{ e } \eta = 1 \times 10^{-2}$$

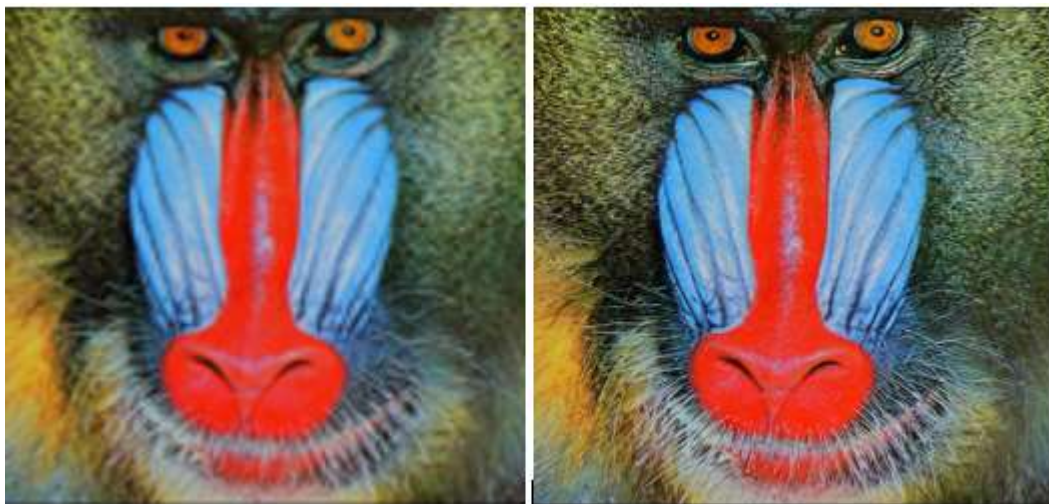
A taxa de aprendizado é definida como  $1 \times 10^{-4}$  e reduzida à metade nas iterações [50k, 100k, 200k, 300k]. O pré-treinamento com perda em pixels ajuda os métodos baseados em GAN a obter resultados visualmente mais agradáveis.

Para treinamento, usamos principalmente o conjunto de dados DIV2K, que é um conjunto de dados de alta qualidade (resolução 2K) para tarefas de restauração de imagens. Além do conjunto de treinamento do DIV2K que contém 800 imagens, também buscamos outros conjuntos de dados com texturas ricas e diversas para nosso treinamento. Para este fim, usamos ainda o conjunto de dados Flickr2K que consiste em 2650 imagens 2K de alta resolução coletadas no site do Flickr e o conjunto de dados OutdoorSceneTraining (OST) para enriquecer nosso conjunto de treinamento.

## 4 Testes e Resultados

Foi utilizado um modelo treinado ESRGAN 4x para que não seja necessário fazer um treinamento do algoritmo toda vez que for realizar uma nova requisição.

As imagens usadas para os testes são do conjunto de dados DIV2K [11], e para os resultados foram usadas imagens sem direitos autorais encontradas pela internet.



*Figura 2 - Comparação de imagem de entrada LR e saída HR*

Acima é possível ver na esquerda, a imagem de entrada de um babuíno com resolução 125x120 e na direita a saída com resolução 500x480.



*Figura 3 - Comparação de imagem de entrada LR e saída HR*

Nestas imagens acima é possível ver uma imagem de um cachorro sentado na areia da praia, na esquerda a entrada com resolução 251x280 e a saída na direita de resolução 1004x1120



## **5 Requisitos de sistema**

### **Requisitos mínimos**

- Placa de vídeo: Nvidia GTX 1660 Super – 6GB VRAM
- Processador: Intel i5 9600KF
- Memória RAM: 8GB
- Windows 10 e Windows 11

Para o software funcionar corretamente, é necessária uma placa gráfica de no mínimo 6 GB de memória de vídeo, mas para alguns tipos de imagens é necessário pelo menos 8 GB de VRAM.

### **Requisitos Recomendados**

- Placa de vídeo: Nvidia RTX 3060 – 12GB VRAM
- Processador: Intel i7 9700K
- Memória RAM: 16GB ou mais
- Windows 10 e Windows 11

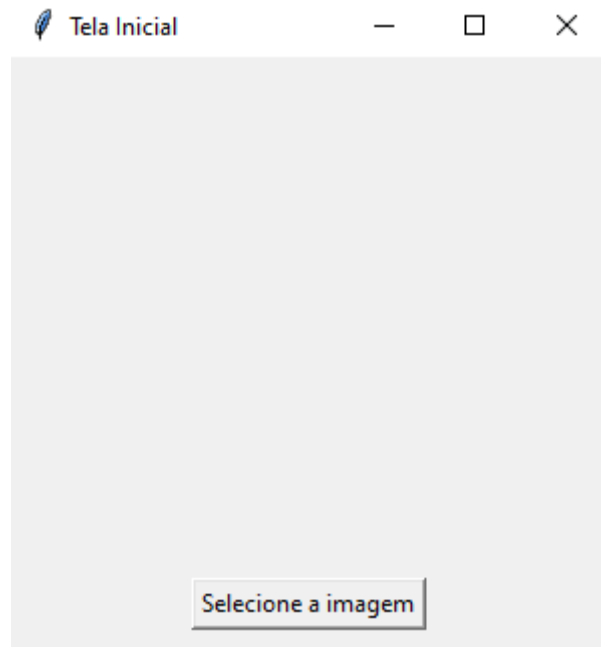
Por conta de o algoritmo de Super-Resolução utilizar a plataforma de computação paralela desenvolvida pela Nvidia, é recomendado que o computador que seja executado o programa contenha uma placa da Nvidia que tenha essa plataforma.

### **Dependências**

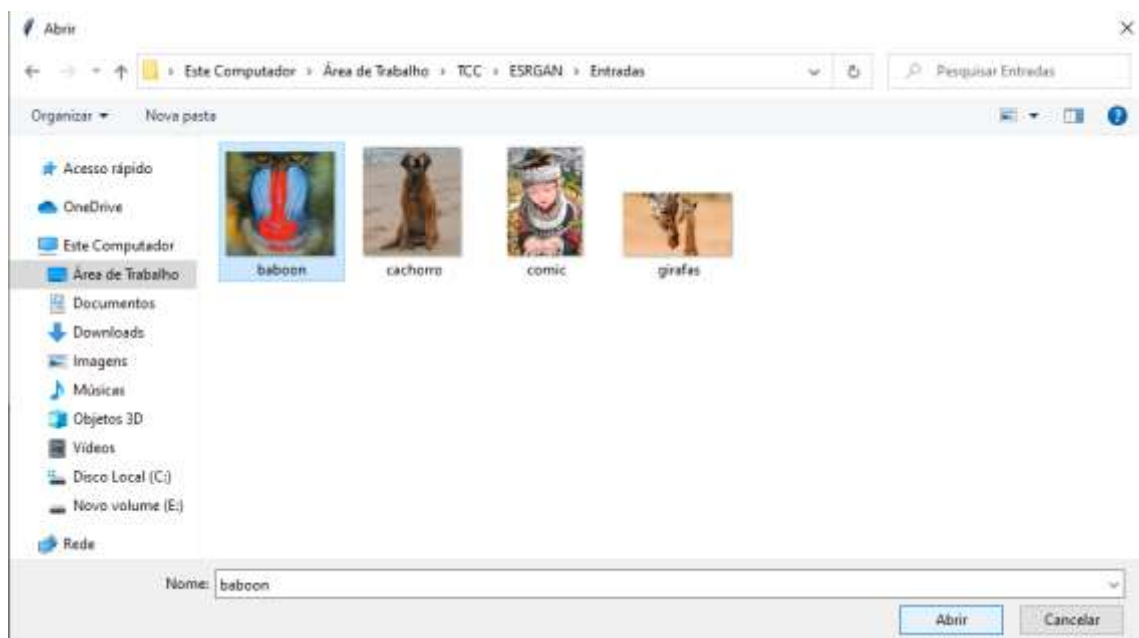
- Python >= 3.7 (Recomendado utilizar Anaconda ou Miniconda)
- PyTorch >= 1.0+ (CUDA 7.5+, se instalado com CUDA)
- Pacotes Python: pip install numpy opencv-python.

## 6 Tutorial

1 - Ao abrir o aplicativo abrirá a tela principal dele onde terá um botão para escolher qual imagem que será aplicada a Super Resolução.

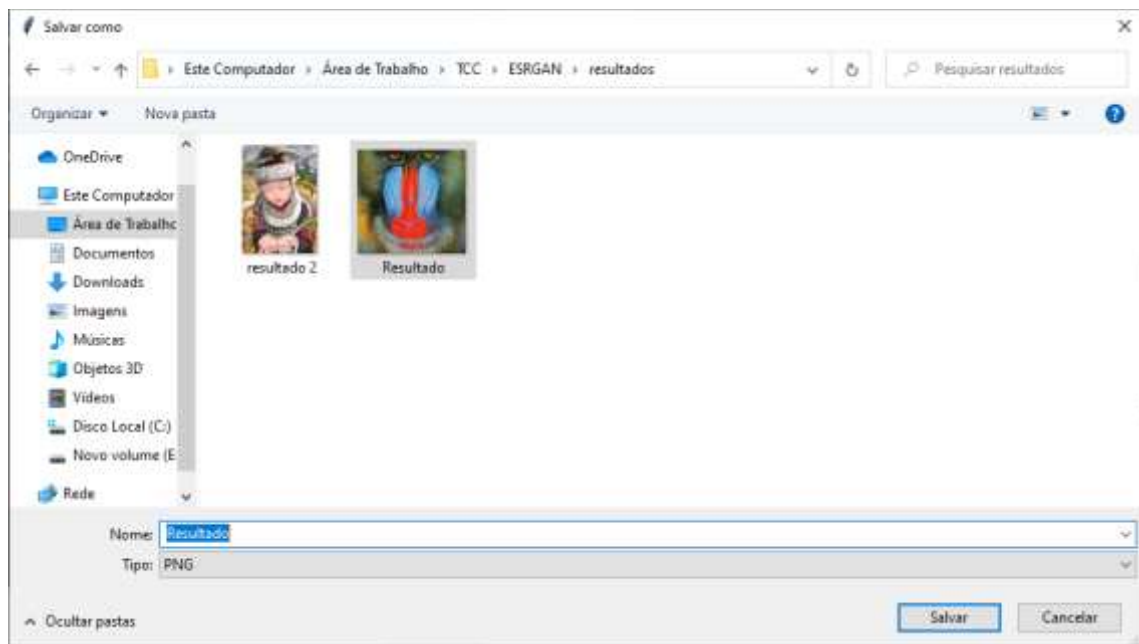


2 – Na tela que aparecer, vá até a pasta em que a imagem está salva e selecione-a.



3 – Após selecionar, o processo de Super Resolução será iniciado.

4 – Quando terminar irá aparecer uma outra tela para que seja escolhida a pasta ao qual deseja salvar a nova imagem, após selecionar é só escolher um nome para o arquivo e salvar.



5 – A imagem estará pronta na pasta em que foi escolhida no passo 4.

Link do projeto no GitHub:

[https://github.com/faabiodiniz/TCC\\_Super-Resolucao](https://github.com/faabiodiniz/TCC_Super-Resolucao)

## **7 Trabalhos Futuros**

Muitas linhas podem ser seguidas, seria interessante fazer mais experimentos, com aumento da resolução da imagem por diferentes fatores e uso de um número reduzido de imagens, para investigar as limitações dos algoritmos implementados. O efeito da utilização de mais níveis de decomposição não foi estudado neste trabalho e poderia ser vantajoso em alguns casos. Uma análise da complexidade computacional dos métodos com diferentes parâmetros, tais como número de imagens LR e fator de aumento de resolução, poderia ser realizada [12].

## **Referências:**

- [1] K. Simonyan and A. Zisserman. *Very deep convolutional networks for large-scale image recognition. In International Conference on Learning Representations (ICLR)*, 2015.
- [2] DONG, C. et al. *Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 38, n. 2, p. 295–307, Feb 2016. ISSN 0162-8828.
- [3] Q. Yang, R. Yang, J. Davis, and D. Nister. *Spatial-depth super resolution for range images. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.
- [4] W. Zou and P. C. Yuen. *Very Low-Resolution Face Recognition in Parallel Environment. IEEE Transactions on Image Processing*, 21:327–340, 2012
- [5] K. Nasrollahi and T. B. Moeslund. *Super-resolution: A comprehensive survey. In Machine Vision and Applications*, volume 25, pages 1423–1468. 2014
- [6] Copeland, Michel – What's the Difference Between Artificial Intelligence, Machine Learning and Deep Learning, Disponível em <<https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>> acesso: 20/05/2022.
- [7] Pacheco, César Augusto Rodrigues, Pereira, Natasha Sophie. *Deep Learning Conceitos e Utilização nas Diversas Áreas do Conhecimento*, 2018
- [8] Park, S. C., Park, K., Kang, M.G. M. *Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine*. V. 20, n. 3, p. 21-26, 2003.
- [9] Xintao Wang, Ke Yu, Kelvin C.K. Chan, Chao Dong and Chen Change Loy. *BasicSR: Open-Source Image and Video Restoration Toolbox*. <https://github.com/xinntao/BasicSR>, 2018.
- [10] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: *Photo-realistic single image superresolution using a generative adversarial network*. In: CVPR. (2017)

- [11] Agustsson, E and Timofte, R. NTIRE 2017 *Challenge on Single Image Super-Resolution: Dataset and Study*
- [12] J. B. Huang, A. Singh, and N. Ahuja. Single image super-resolution from transformed self-exemplars. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5197–5206, 2015. 2, 6, 8