

UNIVERSIDADE FEDERAL DO PARANÁ

VIVIANE DA ROSA SOMMER

ESTUDO DE MODELOS DE SUPER RESOLUÇÃO APLICADOS A IMAGENS DE  
MICROSCOPIA DE CÉLULAS DA SÉRIE BRANCA

CURITIBA PR

2023

VIVIANE DA ROSA SOMMER

ESTUDO DE MODELOS DE SUPER RESOLUÇÃO APLICADOS A IMAGENS DE  
MICROSCOPIA DE CÉLULAS DA SÉRIE BRANCA

Trabalho apresentado como requisito parcial à conclusão do Curso de Bacharelado em Ciência da Computação, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: David Menotti Gomes.

CURITIBA PR

2023

**Universidade Federal do Paraná**  
**Setor de Ciências Exatas**  
**Curso de Ciência da Computação**

Ata de Apresentação de Trabalho de Graduação II

**Título do Trabalho:** ESTUDO DE MODELOS DE SUPER RESOLUÇÃO APLICADOS A  
IMAGENS DE MICROSCOPIA DE CÉLULAS DA SÉRIE BRANCA

**Autor(es):**

GRR 20182564 Nome: VIVIANE DA ROSA SOMMER

GRR \_\_\_\_\_ Nome: \_\_\_\_\_

GRR \_\_\_\_\_ Nome: \_\_\_\_\_

Apresentação: Data: 11 / 12 / 2023 Hora: 09:30 Local: <https://meet.google.com/dmw-qpwv-xrs>

Orientador: DAVID MENOTTI GOMES

Membro 1: RAYSON BARTOSKI LAROCA DOS SANTOS

Membro 2: VALFRIDE WALLACE DO NASCIMENTO

(nome)



Rayson Laroca



(assinatura)

<b>AVALIAÇÃO – Produto escrito</b>	<b>ORIENTADOR</b>	<b>MEMBRO 1</b>	<b>MEMBRO 2</b>	<b>MÉDIA</b>
Conteúdo (00-40)				32
Referência Bibliográfica (00-10)				10
Formato (00-05)				05
<b>AVALIAÇÃO – Apresentação Oral</b>				
Domínio do Assunto (00-15)				15
Desenvolvimento do Assunto (00-05)				05
Técnica de Apresentação (00-03)				01
Uso do Tempo (00-02)				02
<b>AVALIAÇÃO – Desenvolvimento</b>				
Nota do Orientador (00-20)		*****	*****	20
<b>NOTA FINAL</b>	*****	*****	*****	90

Pesos indicados são sugestões.

Conforme decisão do colegiado do curso de Ciência da Computação, a entrega dos documentos comprobatório de trabalho de graduação 2 deve respeitar os seguintes procedimentos: Orientador deve abrir um processo no Sistema Eletrônico de Informações (SEI – UFPR); Selecionar o tipo: Graduação: Trabalho Conclusão de Curso; informar os interessados: nome do aluno e o nome do orientador; anexar esta ata escaneada e a versão final do pdf da monografia do aluno.; Tramita o processo para CCOMP (Coordenação Ciência da Computação).

## **AGRADECIMENTOS**

Agradeço à minha mãe, cujo apoio tem sido constante desde a escolha deste curso, proporcionando-me anos de escuta e carinho dedicados a mim.

Agradeço ao meu pai por guiar-me nos estudos e despertar meu interesse pelas disciplinas exatas.

Agradeço à minha irmã por seu auxílio no ambiente universitário, por me orientar e por apoiar-me em todas as situações ocorridas ao longo desses anos.

Agradeço ao meu namorado, Henrique, que dedicou inúmeras noites estudando ao meu lado, ouvindo-me e oferecendo apoio em todos os momentos. Sua contribuição foi fundamental e essencial.

Agradeço aos meus cachorros, Teddy, Oliver, Bertholdo e Alphonse, que sempre gostaram de me acompanhar enquanto estudava e proporcionaram momentos de descontração com suas brincadeiras.

Agradeço a todos os amigos que fiz durante este curso, os quais não apenas me apoiaram em provas e trabalhos, mas também contribuíram para que eu desfrutasse plenamente da vida.

Agradeço ao meu orientador, que demonstrou grande paciência ao orientar-me durante todo o processo de escrita e pesquisa, respondendo às minhas dúvidas e oferecendo auxílio em todos os momentos necessários.

Assim, expresso meu agradecimento a todas as pessoas que tive o privilégio de conhecer ao longo deste curso, as quais desempenharam um papel fundamental na formação da pessoa que sou hoje e naquela que me tornarei no futuro.

## RESUMO

A super resolução consiste no processo de obtenção de imagens de alta resolução a partir de imagens de baixa resolução, com o objetivo de compensar degradações que ocorrem tanto nas imagens provenientes de dispositivos quanto nas imagens com degradações JPG. Nesse contexto, uma das áreas proeminentes é a da saúde, na qual modelos de super resolução estão sendo implementados e analisados para aprimorar imagens médicas, como de raio-x, tomografia computadorizada e microscopia. Portanto, o objetivo deste estudo é investigar modelos de super resolução para imagens de microscopia de células da série branca, realizando a inferência em modelos com pesos previamente treinados e treinando modelos para avaliar os resultados obtidos. Como base de imagens, foram utilizadas 1065 imagens fornecidas pela empresa *CellaVision*, e foram geradas cinco bases com tamanhos diferentes, utilizando interpolação bicúbica implementada pelas bibliotecas *OpenCV* e *Pillow* do *Python*. Os resultados indicam que tanto os modelos com pesos pré-treinados quanto o modelo treinado com *Fine Tuning* não apresentaram uma melhoria estatisticamente significativa nas imagens em comparação ao aumento da imagem apenas com a interpolação bicúbica. Este estudo abrange a fundamentação teórica da super resolução, redes *Generative Adversarial Network* e redes *Transformer*, uma breve revisão de trabalhos relacionados, as metodologias dos experimentos realizados e os resultados obtidos.

Palavras-chave: Super Resolução. Células da Série Branca. *Generative Adversarial Networks*.

## **ABSTRACT**

Super resolution involves the process of obtaining high-resolution images from low-resolution ones, aiming to compensate for degradations occurring in images from devices as well as those with JPG artifacts. In this context, a prominent application area is healthcare, where super resolution models are being implemented and analyzed to enhance medical images such as X-rays, computed tomography, and microscopy. Therefore, the objective of this study is to investigate super resolution models for white blood cell microscopy images, conducting inference on models with pre-trained weights and training models to evaluate the obtained results. A dataset of 1065 images provided by CellaVision was used, and five datasets of varying sizes were generated using bicubic interpolation implemented by the Python libraries OpenCV and Pillow. The results indicate that both models with pre-trained weights and the Fine Tuning-trained model did not show a statistically significant improvement in images compared to simply upscaling the image with bicubic interpolation. This study encompasses the theoretical foundation of super resolution, Generative Adversarial Network and Transformer networks, a brief review of related works, the methodologies of the conducted experiments, and the obtained results.

**Keywords:** Super-Resolution. White Blood Cells. Generative Adversarial Networks.

## LISTA DE FIGURAS

1.1	Exemplo de aplicação de um modelo de super resolução em imagens de células. . . . .	9
2.1	Exemplo de células da série branca (Hoffbrand et al., 2009). . . . .	12
2.2	Exemplo de modelo GAN (Cai et al., 2020). . . . .	13
2.3	Degradação realizada pela rede <i>RealESRGAN</i> . (Wang et al., 2021a). . . . .	14
2.4	Representação da arquitetura da rede <i>RealESRGAN</i> . (Wang et al., 2021a). . . . .	14
2.5	Modelo de arquitetura <i>Transformer</i> . (Vaswani et al., 2023) . . . . .	16
2.6	Arquitetura proposta pela <i>Swin2SR</i> .(Conde et al., 2022) . . . . .	17
4.1	Exemplo de mesma imagem em todos os datasets: original e degradados. . . . .	26
5.1	Resultados visuais obtidos nos modelos pré-treinados da <i>RealESRGAN</i> com fator de ampliação 4. . . . .	29
5.2	Resultados visuais obtidos nos modelos pré-treinados da <i>RealESRGAN</i> com fator de ampliação 2. . . . .	29
5.3	Resultados visuais obtidos nos modelos pré-treinados da <i>Swin2SR</i> com fator de ampliação 4 . . . . .	30
5.4	Resultados visuais obtidos nos modelos pré-treinados <i>Swin2SR</i> com fator de ampliação 2 . . . . .	30
5.5	Resultados visuais obtidos no modelo treinados <i>RealESRGAN</i> . . . . .	31

## LISTA DE TABELAS

3.1	Especificações das bases de dados dos artigos apresentados . . . . .	21
3.2	Especificações dos modelos de super resolução dos artigos apresentados . . . . .	23
4.1	Bases de imagens criadas a partir das imagens fornecidas pela <i>CellaVision</i> . . . . .	25
4.2	Modelos pré-treinados disponíveis da rede <i>RealESRGAN</i> . . . . .	26
4.3	Modelos pré-treinados disponíveis do modelo <i>RealESRGAN</i> . . . . .	27
5.1	Resultados obtidos como <i>baseline</i> . . . . .	28
5.2	Resultados obtidos com a rede <i>RealESRGAN</i> com modelos pré-treinados. . . . .	29
5.3	Resultados obtidos com a rede <i>Swin2SR</i> com modelos pré-treinados. . . . .	30
5.4	Resultados obtidos com a rede <i>Swin2SR</i> com com <i>Fine Tuning</i> . . . . .	31

## LISTA DE ACRÔNIMOS

SR	Super Resolução
AR	Alta Resolução
BR	Baixa Resolução
GANs	Generative Adversarial Network
ESRGAN	Enhanced Super-Resolution Generative Adversarial Networks
FSRCNN	Fast Super-Resolution Convolutional Neural Network
SISR	Super-Resolução de Imagem Única
FAWDN	Feedback Adaptive Weighted Dense Network
MSRCF	Multi-SR and Classification Framework
SOUP-GAN	Super-Resolution MRI Using Generative Adversarial Networks
RRDB	Residual-in-Residual Dense Blocks

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>9</b>
1.1	OBJETIVOS	10
1.2	ORGANIZAÇÃO DO DOCUMENTO	10
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>11</b>
2.1	CÉLULAS DA SÉRIE BRANCA	11
2.2	SUPER RESOLUÇÃO	12
2.3	ALGORITMOS	13
2.3.1	<i>Generative Adversarial Networks</i>	13
2.3.2	<i>RealESRGAN</i>	14
2.3.3	<i>Redes Transformers</i>	15
2.3.4	<i>Swin2SR</i>	16
2.4	MÉTRICAS DE AVALIAÇÃO	17
2.4.1	<i>Peak Signal-to-Noise Ratio</i>	17
2.4.2	<i>Structural Similarity Index</i>	18
2.4.3	<i>Mean Square Error</i>	18
<b>3</b>	<b>TRABALHOS RELACIONADOS</b>	<b>19</b>
3.1	TRABALHOS UTILIZANDO BASES DE IMAGENS DE MICROSCOPIA DE CÉLULAS DA SÉRIE BRANCA	19
3.2	TRABALHOS DE SUPER RESOLUÇÃO APLICADOS A PROBLEMAS NA ÁREA DA SAÚDE	21
3.3	SUPER RESOLUÇÃO APLICADA A IMAGENS DE CÉLULAS DA SÉRIE BRANCA	24
<b>4</b>	<b>MATERIAIS E MÉTODOS</b>	<b>25</b>
4.1	BASES DE IMAGENS	25
4.2	EXPERIMENTOS	26
<b>5</b>	<b>RESULTADOS</b>	<b>28</b>
5.1	<i>BASELINE</i>	28
5.2	MODELOS PRÉ-TREINADOS	28
5.3	MODELO COM FINE-TUNING	30
<b>6</b>	<b>CONCLUSÃO</b>	<b>32</b>
	<b>REFERÊNCIAS</b>	<b>33</b>

## 1 INTRODUÇÃO

Uma das definições de super resolução (SR), consiste no processo de se obter uma imagem de alta resolução (AR) a partir de uma imagem de baixa resolução (BR) (Nasrollahi e Moeslund, 2014). A imagem de SR tem como objetivo superar ou compensar as limitações e deficiências de um dispositivo, quando faz a aquisição de imagens e as condições desta estão mal definidas, com o intuito de produzir uma imagem de resolução mais alta com base em um conjunto de imagens que foram adquiridas anteriormente da mesma cena (Tian e Ma, 2011).

De acordo com Tian e Ma (2011), o rápido desenvolvimento e implementação de processamento de imagem para comunicações visuais e compreensão de cena geram uma forte demanda por fornecer ao espectador imagens de alta resolução não apenas para fornecer uma melhor visualização, mas também para extrair detalhes adicionais de informações.

Considerando o alto custo e as limitações do aprimoramento de resolução por meio de técnicas “de hardware“, especialmente para dispositivos de imagem em grande escala, os métodos de processamento de sinal de SR se tornaram uma forma potencial de obter imagens de alta resolução (Yue et al., 2016).

Segundo Nasrollahi e Moeslund (2014), esse processo cada vez é mais utilizado em inúmeras áreas, como imagens de satélites (Wang et al., 2022), imagens de faces humanas (Gao et al., 2023) e imagens de texto (Honda et al., 2022).

Contudo, um dos campos em que atualmente é amplamente usado e estudado é na área da saúde. Alguns exemplos de imagens utilizadas são as de raio-X, tomografia computadorizada, ressonância magnética e microscopia (El-Shafai et al., 2021a; Ma et al., 2021; Qiu et al., 2021).

No caso da microscopia, as técnicas de super resolução de imagem podem ser utilizadas para aumentar a resolução de imagens adquiridas em menor magnificação, revelando assim estruturas finas que só poderiam ser observadas utilizando uma lente de maior ampliação, conforme ilustrado na Figura 1.1. No entanto, a aplicação em imagens médicas requer a preservação de características diagnosticamente relevantes e a não introdução de quaisquer artefatos que possam causar confusão no diagnóstico (Tom et al., 2019).

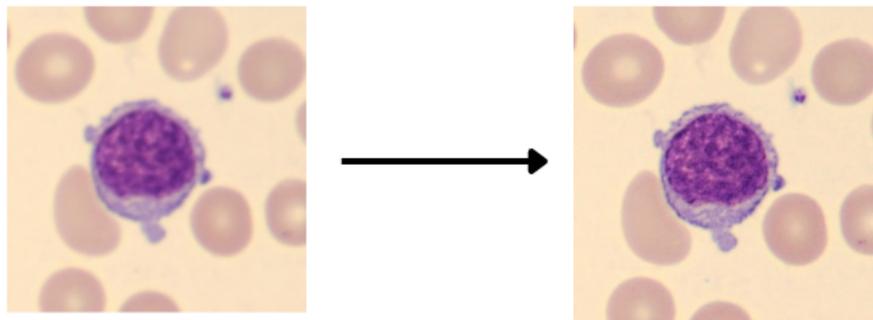


Figura 1.1: Exemplo de aplicação de um modelo de super resolução em imagens de células.

Por tais razões, esta pesquisa tem como objetivo comparar os modelos de super resolução publicados nos últimos anos, aplicando os mesmos na área da saúde, especialmente em imagens de microscopia de células da série branca, com o propósito de avaliar a capacidade desses modelos em aprimorar as imagens desta área de forma satisfatória e sem comprometer características e propriedades importantes das células.

## 1.1 OBJETIVOS

O objetivo deste trabalho consiste em realizar um estudo comparativo de métodos do estado-da-arte em super resolução, que foram avaliados com imagens de microscopia, mais especificamente de células da série branca. Dessa forma, busca-se avaliar qual desses modelos melhor se adequa ao problema de transformar imagens de baixa resolução em alta resolução, sem deformá-las ou criar artefatos na imagem.

Este estudo será conduzido utilizando imagens de alta resolução degradadas como fonte de imagens de baixa resolução. A degradação empregada nas imagens foi a interpolação bicúbica, e os modelos foram avaliados utilizando as métricas *Peak Signal-to-Noise Ratio*, *Structural Similarity Index* e *Mean Square Error*.

Dessa maneira, empresas, laboratórios e pesquisadores que possuem acesso a ferramentas de microscopia de custo mais baixo e que apresentam limitações, poderão avaliar e, possivelmente, utilizar um modelo de super resolução para aprimorar as imagens obtidas para análise clínica posterior. Para atingir este objetivo geral, foram definidos os seguintes objetivos específicos:

- Selecionar modelos de super resolução atuais;
- Inferir os modelos escolhidos com um conjunto de dados de células da série branca;
- Avaliar os resultados obtidos

## 1.2 ORGANIZAÇÃO DO DOCUMENTO

O presente trabalho é estruturado em seis capítulos distintos. O Capítulo 1 tem por objetivo proporcionar uma introdução ao processo de super resolução, destacando sua aplicação em diversas áreas do conhecimento, com ênfase na área da saúde. Além disso, são apresentados o problema a ser abordado, a motivação que conduziu à realização do estudo, assim como os objetivos gerais e específicos, estabelecendo o arcabouço conceitual necessário para a compreensão do estudo em questão.

O Capítulo 2 explora os fundamentos das células da série branca, os conceitos de super resolução, a definição de *Generative Adversarial Networks* e redes *Transformers*, além da explicação dos modelos *RealESRGAN* e *Swin2SR*, abrangendo também as métricas utilizadas.

No Capítulo 3, realiza-se uma análise dos trabalhos correlatos disponíveis na literatura, promovendo uma comparação com o trabalho atual e problemas similares.

O Capítulo 4 aborda as especificações da implementação do estudo proposto, descrevendo como as bases de imagens foram obtidas e geradas, assim como os modelos pré-treinados utilizados.

O Capítulo 5 apresenta os resultados obtidos a partir dos experimentos realizados com as imagens de células da série branca.

O Capítulo 6 encerra o presente estudo com uma análise do que foi efetivamente realizado, ao mesmo tempo em que explora potenciais direções para trabalhos futuros nesta área específica.

## 2 FUNDAMENTAÇÃO TEÓRICA

O objetivo deste capítulo é estabelecer os fundamentos teóricos e conceituais que sustentam a pesquisa acerca do estudo de modelos de super resolução aplicados a imagens de microscopia de células da série branca. Para tal propósito, serão abordados conceitos relacionados às células da série branca, bem como a fundamentação teórica da super resolução. Além disso, serão apresentadas as redes *Generative Adversarial Network* e redes *Transformers* empregados para a obtenção de super resolução, e discutidas as métricas utilizadas na avaliação dos resultados obtidos.

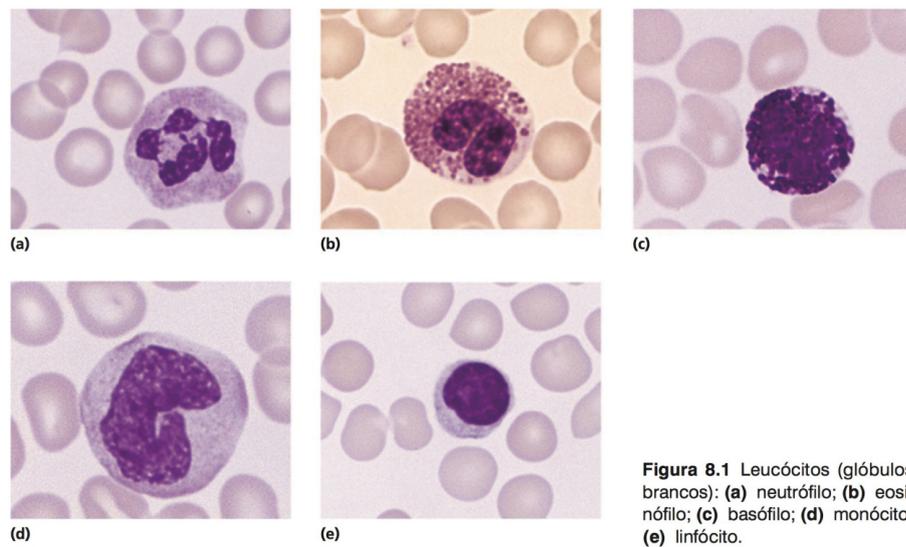
### 2.1 CÉLULAS DA SÉRIE BRANCA

Os leucócitos, também conhecidos como células da série branca, ou ainda glóbulos brancos, são componentes sanguíneos cujo propósito é proteger o organismo humano contra doenças e infecções. Essas células são distribuídas por todo o corpo por meio da corrente sanguínea, sendo uma das suas origens a medula óssea, onde são formadas e desenvolvidas. A série branca possui cinco principais tipos de célula, sendo elas: neutrófilos, eosinófilos, monócitos, basófilos e linfócitos (VIVAS, 2014; Hoffbrand et al., 2009).

De acordo com Hoffbrand et al. (2009) e VIVAS (2014), os neutrófilos constituem aproximadamente 50% a 70% das células circulantes e têm a capacidade de atravessar as paredes dos vasos sanguíneos para adentrar nos tecidos, onde desempenham a função de proteger o corpo através da fagocitose de bactérias e substâncias estranhas ao organismo. Os eosinófilos assemelham-se aos neutrófilos e possuem um papel especial nas resposta alérgicas e na defesa contra parasitos.

Os monócitos, de maior dimensão em relação aos demais leucócitos, possuem apenas um núcleo e têm a capacidade de se diferenciar em células maiores que englobam bactérias, apresentando funções específicas em distintos tecidos, como pele, intestino, fígado, entre outros. Os basófilos, presentes em menor quantidade, têm a propriedade de liberar histamina, que contribui para respostas alérgicas ao dilatar e tornar mais permeáveis os vasos sanguíneos, além de liberar heparina, que previne a coagulação do sangue. Os linfócitos, o segundo tipo mais abundante de leucócitos, desempenham um papel crucial na defesa do organismo contra infecções ao produzirem anticorpos, sendo originados no timo e na medula óssea (VIVAS, 2014; Hoffbrand et al., 2009).

A Figura 2.1 ilustra os cinco tipos de células da série branca.



**Figura 8.1** Leucócitos (glóbulos brancos): (a) neutrófilo; (b) eosinófilo; (c) basófilo; (d) monócito; (e) linfócito.

Figura 2.1: Exemplo de células da série branca (Hoffbrand et al., 2009).

## 2.2 SUPER RESOLUÇÃO

Segundo Tian e Ma (2011), a super resolução é uma técnica de processamento de imagens cujo propósito consiste em reconstituir uma imagem de alta resolução a partir de uma ou mais imagens de baixa resolução da mesma cena. Tal abordagem se revela fundamental para transpor as limitações do processo de aquisição de imagens. A super resolução é um problema clássico que ainda é considerado um problema em aberto e desafiador na visão computacional, principalmente devido à existência de múltiplas soluções para uma mesma imagem de baixa resolução e à crescente complexidade da solução à medida que o fator de ampliação aumenta (Anwar et al., 2020).

Diversas abordagens têm sido desenvolvidas para a aplicação da super resolução, abrangendo métodos que se fundamentam em interpolação e reconstrução de imagem, que são considerados mais clássicos. Entretanto, o atual estado-da-arte encontra-se concentrado em técnicas baseadas em *Deep Learning*. O *Deep Learning* é um ramo do *Machine Learning* que tem como objetivo aprender automaticamente a relação entre a entrada e a saída diretamente dos dados. Além de sua aplicação na super resolução, os algoritmos de *Deep Learning* têm demonstrado resultados promissores em diversos outros subcampos da inteligência artificial, tais como classificação e detecção de objetos, processamento de linguagem natural, processamento de imagens e processamento de sinais de áudio (Anwar et al., 2020).

Em uma imagem de baixa resolução, o processo de degradação normalmente é desconhecido e pode ser afetado por diversos fatores, tais como artefatos de compressão, degradações anisotrópicas, ruído do sensor e ruído granular. Atualmente, há uma variedade de bases de dados disponíveis para a super resolução de imagens, as quais diferem significativamente em quantidade de imagens, qualidade e resolução, dentre outros aspectos. Algumas dessas bases fornecem pares de imagens de baixa resolução e alta resolução, enquanto outras oferecem apenas imagens de alta resolução, sendo que as imagens de baixa resolução são obtidas aplicando-se um algoritmo de degradação, como por exemplo a interpolação bicúbica (Wang et al., 2021b).

## 2.3 ALGORITMOS

Conforme mencionado na seção anterior, existem duas abordagens para o problema de super resolução. Entretanto, este trabalho concentrará seu enfoque nos modelos de *Deep Learning* propostos. Nas subseções a seguir, serão apresentados alguns tipos de algoritmos e modelos que se baseiam em *Deep Learning* e que servirão como fundamentação dos próximos capítulos.

### 2.3.1 *Generative Adversarial Networks*

Segundo Goodfellow et al. (2014), as *Generative Adversarial Networks* (GANs) são modelos complexos compostos por dois componentes fundamentais: um gerador e um discriminador. O gerador, alimentado por um vetor de ruído aleatório, tem como propósito criar amostras que mimetizem a distribuição dos dados de treinamento. Por outro lado, o discriminador é incumbido de avaliar se uma amostra é genuína, proveniente de um conjunto de dados, ou se é artificial, gerada pelo mencionado gerador.

Estes dois modelos são treinados simultaneamente em um processo adversarial. O gerador busca constantemente aprimorar suas capacidades, gerando amostras cada vez mais realistas com o intuito de iludir o discriminador. Em contrapartida, o discriminador trabalha incessantemente para distinguir de maneira mais precisa entre amostras reais e geradas (Goodfellow et al., 2014).

Ao longo do processo de treinamento, tanto o gerador quanto o discriminador passam por atualizações alternadas. Inicialmente, o discriminador é treinado utilizando conjuntos de amostras reais e falsas, calculando seu erro. Subsequentemente, o gerador é treinado para gerar amostras que minimizem o erro previamente computado do discriminador. Este ciclo é repetido iterativamente até que o gerador alcance o patamar em que suas amostras se tornam indistinguíveis das amostras reais pelo discriminador. Ao término do treinamento, o gerador adquire a capacidade de gerar novas amostras que se assemelham significativamente aos dados de treinamento (Goodfellow et al., 2014).

A Figura 2.2 ilustra o funcionamento básico de um modelo GAN.

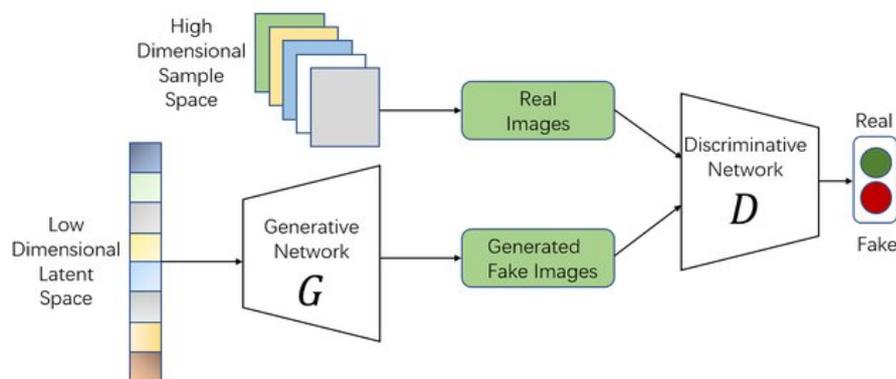


Figura 2.2: Exemplo de modelo GAN (Cai et al., 2020).

A temática da super resolução, empregando *Generative Adversarial Networks*, tem sido objeto de intensa pesquisa nos últimos anos. Em decorrência desse interesse crescente, observa-se a existência de uma variedade significativa de modelos e suas respectivas variações, desenvolvidos com propósitos distintos (Wang et al., 2021a; Ledig et al., 2017; Denton et al., 2015; Zhu et al., 2019; Song et al., 2022).

Portanto, escolhemos o modelo *Real-ESRGAN* como objeto de estudo de super resolução, considerando seu vasto uso em vários problemas recentes (Agarwal et al., 2022; Septian et al., 2022; Lv et al., 2023).

### 2.3.2 *RealESRGAN*

A rede *RealESRGAN* é uma tecnologia de restauração de imagens que se destaca em relação a tentativas anteriores de super resolução, permitindo a restauração de imagens degradadas do mundo real. As degradações complexas reais geralmente surgem de combinações complicadas de diferentes processos, como o sistema de imagem das câmeras, a edição de imagens e a transmissão pela internet. Dessa maneira, a rede adotou um processo de degradação de segunda ordem para alcançar um bom equilíbrio entre simplicidade e eficácia (Wang et al., 2021a).

A rede foi treinada com dados puramente sintéticos e é capaz de restaurar a maioria das imagens do mundo real, apresentando um desempenho visual superior em comparação com trabalhos anteriores, o que a torna mais prática em aplicações do mundo real. Para simular uma degradação mais realista, a imagem é submetida a algoritmos de *Blur*, *Resize*, *Noise* e compressão JPEG, conforme ilustrado na Figura 2.3 (Wang et al., 2021a).

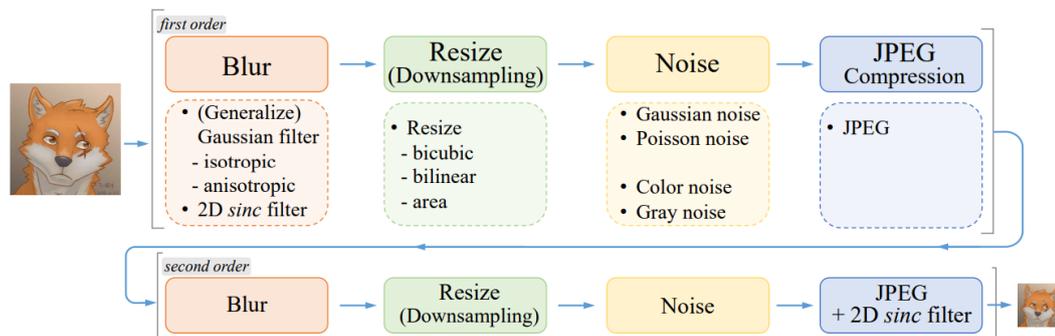


Figura 2.3: Degradação realizada pela rede *RealESRGAN*. (Wang et al., 2021a)

Para o gerador, foi utilizada a rede profunda Residual-in-Residual Dense Blocks (RRDB). Já para o discriminador, foi usado uma rede U-net (Schönfeld et al., 2021) com normalização espectral, pois gera valores de autenticidade para cada pixel e pode fornecer um *feedback* detalhado por *pixel* ao gerador. A Figura 2.4 mostra uma representação da arquitetura utilizada (Wang et al., 2021a).

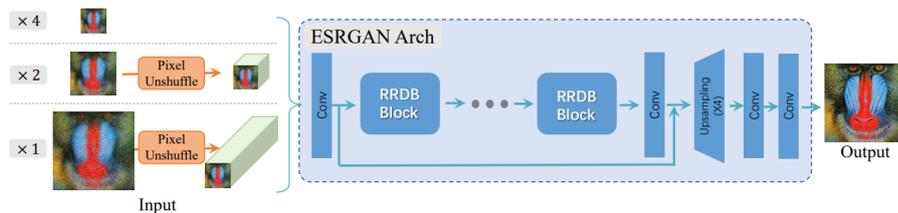


Figura 2.4: Representação da arquitetura da rede *RealESRGAN*. (Wang et al., 2021a)

O processo de treinamento do modelo foi dividido em duas etapas. Primeiramente, foi treinado um modelo orientado para PSNR com a perda L1. O modelo obtido é denominado *Real-ESRNet*. Em seguida, foi utilizado o modelo treinado orientado para PSNR como uma inicialização do gerador e treinado o *Real-ESRGAN* com uma combinação de perda L1, perda perceptual (Johnson et al., 2016) e perda GAN (Goodfellow et al., 2014; Wang et al., 2021a).

### 2.3.3 Redes *Transformers*

Segundo Vaswani et al. (2023), as redes *Transformers* representam uma arquitetura de rede neural desenvolvida primariamente para tarefas de processamento de linguagem natural. Entretanto, sua versatilidade se estende para diversas outras áreas, incluindo visão computacional e análise de sequências temporais (Parvaiz et al., 2023; Bi et al., 2021; Xu et al., 2021; Liu et al., 2021).

A estrutura essencial de uma rede *Transformer* é composta por um codificador e um decodificador, cada um consistindo em múltiplas camadas de auto-atenção e camadas totalmente conectadas. Cada camada de atenção proporciona ao modelo a capacidade de focalizar-se em segmentos específicos da entrada, atribuindo pesos distintos a cada componente. A pontuação de atenção, calculada para cada parte da entrada, orienta a ponderação da importância relativa de cada segmento na representação final. Este método permite ao modelo concentrar-se nas partes mais relevantes da entrada, descartando aquelas irrelevantes ou redundantes para a tarefa em questão (Vaswani et al., 2023).

O codificador recebe uma sequência de entrada, gerando uma representação contínua dessa sequência. Cada palavra na sequência é representada por um vetor de palavras, submetido a múltiplas camadas de auto-atenção e camadas totalmente conectadas. O decodificador, por sua vez, utiliza a representação contínua da sequência de entrada para produzir uma sequência de saída. A geração da sequência de saída ocorre iterativamente, com o modelo utilizando os dados gerados anteriormente como entrada para a produção da próxima sequência (Vaswani et al., 2023).

Durante a geração da sequência de saída, o decodificador emprega a auto-atenção para capturar dependências de longo alcance entre os dados na sequência gerada. O treinamento do *Transformer* é realizado por meio de um processo supervisionado, alimentando a rede com pares de sequências de entrada e saída e ajustando seus parâmetros para minimizar a discrepância entre a sequência de saída gerada pelo modelo e a sequência de saída correta (Vaswani et al., 2023).

A Figura 2.5 ilustra a arquitetura de um modelo *Transformer*.

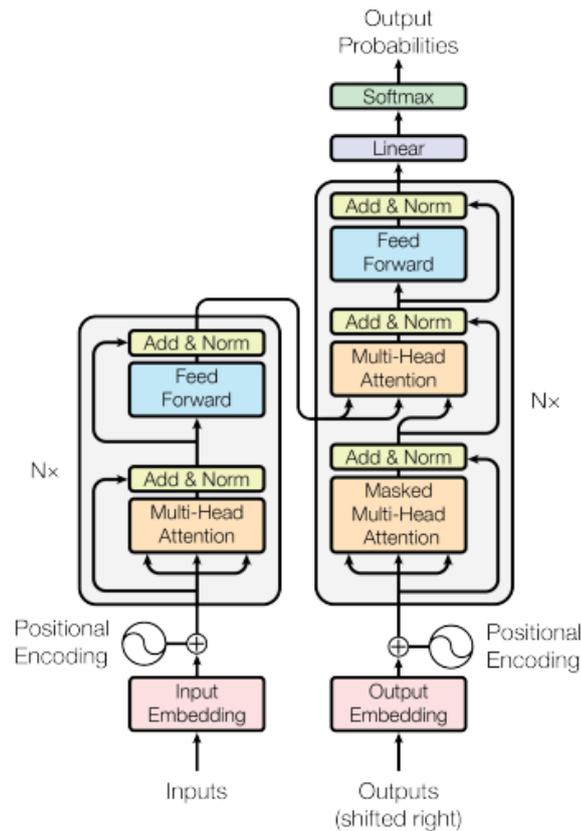


Figura 2.5: Modelo de arquitetura *Transformer*. (Vaswani et al., 2023)

### 2.3.4 *Swin2SR*

O modelo *Swin2SR* é fundamentado no *Swin V2 Transformer* (Liu et al., 2022) para super resolução e restauração de imagens comprimidas. Este representa uma possível evolução em relação ao *SwinIR* (Liang et al., 2021), proporcionando treinamento e convergência mais eficientes, além de maior capacidade e resolução. O *Swin2SR* alcançou resultados de última geração na remoção de artefatos de compressão JPEG, na super resolução de imagens clássica e na super resolução de imagens comprimidas. Adicionalmente, obteve resultados competitivos no *AIM 2022 Challenge on Super-Resolution of Compressed Image and Video* (Yang et al., 2022), figurando entre os cinco melhores classificados (Conde et al., 2022).

Para o método *Swin2SR*, foram propostas algumas modificações em relação ao *SwinIR* a fim de aprimorar as capacidades do modelo para super resolução, especialmente quando aplicado a entradas comprimidas. O *Residual Transformer Block* original foi atualizado, incorporando as camadas de atenção do *SwinV2 Transformer*, visando aumentar a capacidade e a resolução (Conde et al., 2022).

O modelo apresenta um ramo de aumento clássico que emprega interpolação bicúbica, capaz por si só, de recuperar informações estruturais básicas. Por esta razão, a saída do modelo é adicionada à imagem aumentada básica, buscando aprimorá-la. Além disso, foram exploradas diferentes funções de perda para conferir maior robustez ao modelo frente aos artefatos de compressão JPEG, possibilitando a recuperação de detalhes de alta frequência em imagens comprimidas de baixa resolução, e, conseqüentemente, alcançando um desempenho superior. Na figura 2.6, está representada a arquitetura proposta pela *Swin2SR* (Conde et al., 2022).

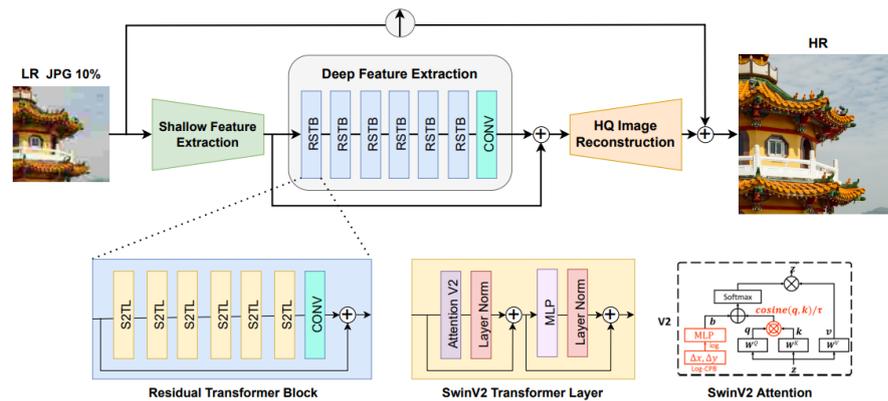


Figura 2.6: Arquitetura proposta pela *Swin2SR*.(Conde et al., 2022)

A *Swin2SR* apresenta os seguintes elementos, semelhantes ao *SwinIR*: módulos de extração de características rasas, extração de características profundas e reconstrução de imagem de alta qualidade. O módulo de extração de características rasas utiliza uma camada de convolução para extrair características, que são transmitidas diretamente ao módulo de reconstrução para preservar informações de baixa frequência. O módulo de extração de características profundas é composto principalmente por blocos *Residual Transformer Block*, cada um dos quais utiliza várias camadas do *Transformer SwinV2* para atenção local e interação entre janelas. Finalmente, tanto as características rasas quanto as profundas são fundidas no módulo de reconstrução para a reconstrução de imagens de alta qualidade. Para ampliar a escala da imagem, foi utilizada uma operação padrão de redistribuição de *pixels*.

Na avaliação do modelo, foram conduzidas três tarefas: remoção de artefatos de compressão JPEG, super-resolução de imagem clássica e super-resolução de imagem comprimida.

## 2.4 MÉTRICAS DE AVALIAÇÃO

Com o propósito de avaliar os modelos em estudo, empregamos três métricas amplamente reconhecidas no campo de super resolução, o *Peak Signal-to-Noise Ratio* (PSNR), o *Structural Similarity Index* (SSIM) e o *Mean Square Error* (MSE). Assim, para cada uma das imagens sujeitas à análise, as referidas métricas foram computadas e, a fim de efetuar a avaliação dos modelos, procedeu-se ao cálculo da média, mediana e desvio padrão de cada uma delas.

### 2.4.1 *Peak Signal-to-Noise Ratio*

O *Peak Signal-to-Noise Ratio* (PSNR) representa a relação entre a potência máxima possível de um sinal e a potência do ruído que afeta a fidelidade de sua representação. Em geral, o PSNR é expresso em decibéis e serve como uma estimativa aproximada da percepção humana em relação à qualidade da reconstrução. Embora, frequentemente, um valor mais elevado de PSNR sugira uma qualidade superior na reconstrução, especialmente no contexto de compressão de imagens, há situações em que é necessário que o PSNR seja menor para obter resultados adequados, como em aplicações de detecção de bordas (Poobathy e Chezian, 2014).

Uma das principais vantagens do PSNR é a sua simplicidade, pois ele pode ser calculado diretamente a partir do *Mean Square Error* (MSE) entre os *pixels* da imagem original e da imagem predita. No entanto, é importante notar que o PSNR é particularmente sensível a distorções que causam desalinhamento de *pixels*, como deslocamento espacial, rotação ou redimensionamento (Korhonen e You, 2012).

O PSNR é calculado com base no MSE pela Equação 2.1, onde R é a variação máxima nos dados da imagem de entrada. (Poobathy e Chezian, 2014)

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{R^2}{\text{MSE}} \right) \quad (2.1)$$

#### 2.4.2 Structural Similarity Index

O *Structural Similarity Index* (SSIM) é constituído por três termos e estima o impacto visual das alterações na luminosidade da imagem, nas mudanças de contraste e em eventuais erros remanescentes, coletivamente identificados como mudanças estruturais (Dosselmann e Yang, 2011). Para as imagens originais e codificadas x e y, respectivamente, o índice SSIM é definido conforme explicitado na Equação 2.2, onde  $\alpha > 0$ ,  $\beta > 0$  e  $\gamma > 0$  controlam a importância relativa de cada um dos três termos do índice, portanto, sendo uma métrica que varia de -1 a 1 (Dosselmann e Yang, 2011).

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (2.2)$$

Na maioria dos casos, contudo, é atribuída uma pontuação no intervalo de [0, 1], onde valores próximos de 0 representam níveis mais baixos de qualidade de imagem, enquanto valores mais próximos de 1 indicam níveis mais elevados de qualidade visual. Além disso, é importante observar que existe um máximo único, o que significa que  $\text{SSIM}(x, y) = 1$  se e somente se  $x = y$  (Dosselmann e Yang, 2011).

#### 2.4.3 Mean Square Error

O *Mean Square Error* (MSE) é calculado como a média dos quadrados das diferenças entre os valores preditos e os valores reais. Ao elevar as diferenças ao quadrado, o MSE atribui um peso maior a erros maiores, o que pode ser apropriado em muitos contextos (James et al., 2013).

O MSE é calculado da seguinte maneira, onde n é o número total de observações,  $Y_i$  é o valor real da i-ésima observação e  $\hat{Y}_i$  é o valor predito da i-ésima observação (James et al., 2013), conforme a Equação 2.3.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.3)$$

No entanto, o MSE tem algumas desvantagens, como a falta de captura de variações correlacionadas das quantidades medidas e a necessidade de uma comparação adicional de magnitude contra as próprias medições para qualquer interpretação significativa. Além disso, o MSE é criticado por sua falta de limites e convexidade da função (Pandit e Schuller, 2019).

### 3 TRABALHOS RELACIONADOS

Neste capítulo, realizamos uma revisão da literatura existente que aborda previamente a problemática apresentada neste trabalho. Inicialmente, apresentamos estudos que fazem uso de bases de imagens de microscopia de células da série branca (Seção 3.1). Em seguida, discutimos estudos relacionados à área da saúde que aplicam a técnica de super resolução (Seção 3.2). Por fim, apresentamos um artigo que aplica a técnica de super resolução em imagens de células da série branca (Seção 3.3).

#### 3.1 TRABALHOS UTILIZANDO BASES DE IMAGENS DE MICROSCOPIA DE CÉLULAS DA SÉRIE BRANCA

Existem diversos estudos recentes que empregam imagens de células da série branca com o objetivo de classificação e segmentação. Assim, analisamos trabalhos que utilizam métodos de *Machine Learning* e *Deep Learning* para aprofundar o conhecimento sobre as bases de imagens empregadas nessa área e alcançar um desempenho superior na solução desses problemas.

A abordagem de Prinyakupt e Pluempitiwiriwawej (2015) propôs um método para segmentar e classificar células da série branca utilizando classificadores lineares e *Naïve Bayes*. Para avaliar a eficácia do método, foram utilizadas duas bases de imagens, a primeira fornecida pela *Rangsit University* e a segunda denominada *CellaVision*. Os resultados demonstraram que o processo de segmentação em ambas as bases de imagens foi rápido, robusto, eficiente e coerente. Além disso, a classificação das células normais da série branca em cinco tipos apresentou alta sensibilidade. No entanto, ao analisar a conclusão do artigo, nota-se que a precisão da classe *Eosinophil* é de 0.750 e 0.444, no classificador linear e modelo *Naïve Bayes*, respectivamente, o que indica que o modelo tem dificuldade em fornecer resultados corretos ou precisos para esta classe.

O trabalho de Liu et al. (2017) concentra-se na segmentação de células da série branca, levando em consideração a complexidade desse problema devido às características intrínsecas das células e a fatores externos, como iluminação e diferentes vistas microscópicas. O algoritmo proposto foi também avaliado usando duas bases de dados, o *CellaVision* e *Jiashan*. Os resultados obtidos demonstram consistentemente que o algoritmo proposto supera outras abordagens de localização e segmentação apresentados no artigo.

O trabalho de Andrade et al. (2019) também se concentra na segmentação de células da série branca, mas com o objetivo de fazer um levantamento dos métodos desenvolvidos nos últimos anos, visando a identificação de pacientes com leucemia. Foram utilizadas cinco bases de imagens para analisar o desempenho dos métodos: *ALL-IDB 2*, *BloodSeg*, *Leukocytes*, *JTSC Database* e *CellaVision*. Como conclusão, é citado que embora os algoritmos avaliados tenham apresentado uma eficácia promissora, ainda existem desafios a serem superados nessa área devido a grande diferença de *background* nas imagens entre as bases de dados.

Vogado et al. (2016) seguem na mesma linha, avaliando as técnicas de segmentação para leucócitos. Foram utilizadas três bases públicas para efetuar a avaliação das metodologias, a *ALL-IDB 2*, *BloodSeg* e *Leukocytes*. Constatou-se que os métodos como um todo apresentaram resultados promissores, entretanto necessitam de maior robustez para uma melhor performance em diferentes bases.

O artigo de Banik et al. (2020) aborda o problema da segmentação e classificação de células da série branca, propondo um novo modelo de rede neural convolucional. Os

pesquisadores realizaram experimentos utilizando as bases de dados *BCCD*, *ALL-IDB2*, *JTSC* e *CellaVision*, nos quais aplicaram o método proposto para a segmentação das células brancas, sendo que a base de imagens *BCCD* também foi utilizada para a etapa de classificação. O modelo desenvolvido pelos autores é notável por sua baixa complexidade computacional e por apresentar um desempenho superior em relação aos outros modelos mencionados no estudo.

Shahin et al. (2019) adotou uma abordagem baseada em redes neurais convolucionais para desenvolver um sistema de identificação de células da série branca em imagens de esfregaços sanguíneos. Três bases de dados públicos, não identificados, foram utilizados, cada um contendo cinco classes: neutrófilo, eosinófilo, linfócito, monócito e basófilo. O estudo concluiu apresentando uma taxa de acurácia de 92,9% ao combinar as bases de dados, demonstrando que o modelo possui um desempenho sólido, mesmo quando aplicado a bases de dados distintas.

O trabalho realizado por Yao et al. (2021) concentra-se na classificação de células brancas do sangue utilizando redes neurais convolucionais deformáveis otimizadas com pesos. Os pesquisadores utilizaram uma base de imagens privada fornecido por uma empresa parceira, bem como a base de imagens pública *BCCD*. Dentre os nove modelos avaliados, o modelo *TWO-DCNN* apresentou um dos melhores desempenhos com uma precisão de 95,7%, enquanto o modelo *Inception-V3* obteve um dos piores desempenhos, com uma precisão de 22,1%, ambos para bases de imagens de baixa resolução.

Dong et al. (2022) seguem na mesma linha de pesquisa ao propor um modelo que combina recursos de aprendizado profundo com recursos artificiais. Os pesquisadores compilaram uma base de imagens própria, contendo cinco tipos de células brancas do sangue: monócitos, linfócitos, basófilos, eosinófilos e neutrófilos. A metodologia proposta demonstrou baixa complexidade e alta precisão, o que a torna uma valiosa referência para outros problemas de detecção de imagens médicas, levando em conta que o modelo foi treinado com apenas 500 imagens.

Anita e Yadav (2021) propõem um algoritmo de detecção automática de células brancas do sangue utilizando uma abordagem de detecção de elipses. Para os experimentos, foram utilizadas amostras de sangue obtidas do *ASH Image Bank*. Além disso, um teste de robustez foi realizado, o qual justificou que a técnica é capaz de lidar com diferentes tipos de ruídos presentes nas imagens. Os resultados experimentais demonstram, portanto, a eficiência do esquema proposto em termos de precisão de detecção e estabilidade.

A Tabela 3.1 sumariza os trabalhos estudados nesta seção. É importante destacar que muitos dos estudos mencionados combinam diversas bases de imagens, embora todas sejam compostas por células da série branca. No geral, os modelos obtiveram acurácias satisfatórias. Porém, em imagens de baixa qualidade, percebe-se uma redução na precisão da segmentação e classificação, resultando em um aumento nos casos de falsos positivos e falsos negativos. Portanto, é evidente que imagens de alta qualidade são essenciais para obter uma classificação, segmentação e uso médico precisos.

Tabela 3.1: Especificações das bases de dados dos artigos apresentados

Artigo	Bases de dados	Número de Imagens	Tamanho da imagem (Pixels)
(Prinyakupt e Pluempitiwiriyaewej, 2015)	<i>Rangsit University, CellaVision</i>	555, 477	960 × 1.280, 360 × 363
(Liu et al., 2017)	<i>CellaVision, Jiashan</i>	60, 800	360 x 360, 500 x 500
(Andrade et al., 2019)	<i>ALL-IDB 2, BloodSeg, Leukocytes, JTSC Database, CellaVision</i>	260, 367, 107, 300, 100	2592 × 1944, 640 × 480, None, 120 × 120, 300 × 300
(Vogado et al., 2016)	<i>ALL-IDB 2, BloodSeg, Leukocytes</i>	260, 368, 107	2592 x 1944, 640 x 480, 198 x 189
(Banik et al., 2020)	<i>BCCD, ALL-IDB 2, JTSC, CellaVision</i>	367, 12444, 260, 300, 100	640 x 480 320 x 240 2592×1944 120×120 300 x 300
(Shahin et al., 2019)	<i>Dataset_ALL</i>	125, 254, 2172	2592 × 1944, 2592 × 1944, 720 × 576 70 × 70
(Yao et al., 2021)	<i>Private, BCCD</i>	11865, 375	224 x 224, 640 x 480
(Dong et al., 2022)	<i>Private</i>	500	299 × 299
(Anita e Yadav, 2021)	<i>ASH Image Bank</i>	60	None
(Yamin et al., 2023)	<i>Blood Cell Images-Kaggle</i>	12500	640 x 480

### 3.2 TRABALHOS DE SUPER RESOLUÇÃO APLICADOS A PROBLEMAS NA ÁREA DA SAÚDE

Com o crescente aumento do uso de algoritmos de super resolução nos últimos anos, destaca-se sua ampla aplicação na área da saúde. Nesse contexto, foram examinados estudos que empregaram tais algoritmos com o intuito de aprofundar o entendimento dos modelos no estado da arte e dos resultados alcançados em diversas problemáticas.

O trabalho realizado por Avci et al. (2023) apresenta um estudo sobre imagens de sedimentos de urina, aplicando a técnica de super resolução *Faster R-CNN*. Essa abordagem tem como objetivo melhorar a precisão na detecção e classificação de objetos em imagens de baixa resolução. No método proposto, foi utilizado o conceito de super resolução de imagem única baseada em auto-similaridade no pré-processamento, visando aprimorar a qualidade de resolução das imagens. O modelo final alcançou uma acurácia de 98,6%. No entanto, é importante ressaltar que ainda há espaço para melhorias, uma vez que a base de imagens utilizada no estudo apresenta desequilíbrio entre as classes, o que pode impactar os resultados.

Deng et al. (2020) propõem o uso do modelo *Enhanced Deep Super-Resolution Network* como algoritmo de super resolução para reconstruir imagens de lâminas inteiras em patologia. O algoritmo é um modelo de *Deep Learning* que utiliza redes neurais convolucionais profundas, amplamente empregado em tarefas de super resolução de imagens. O estudo incluiu uma avaliação com patologistas, exibindo imagens geradas por super resolução e imagens de alta resolução.

Observou-se que os patologistas não foram capazes de distinguir facilmente a diferença entre esses dois tipos de imagens. No entanto, o artigo apresenta algumas limitações, pois as imagens geradas foram restritas a um *zoom* duplo, o que não seria condizente com uma avaliação realista em algumas situações.

O estudo realizado por Sun e Ng (2022) emprega o modelo avançado de rede neural artificial chamado *Enhanced Super-Resolution Generative Adversarial Networks* (ESRGAN) para aprimorar a qualidade das imagens de angiografia por tomografia computadorizada coronariana. Os resultados demonstraram que as imagens processadas pelo ESRGAN apresentaram uma melhoria na especificidade em todas as amostras das artérias coronárias quando comparadas às imagens originais. Além disso, observou-se uma redução nas taxas de falso positivo nas imagens processadas pelo ESRGAN em comparação com as imagens originais. Entretanto, ainda são necessárias melhorias adicionais no modelo para reduzir ainda mais essas taxas.

A abordagem de Qiu et al. (2021) tem como objetivo aprimorar a qualidade da reconstrução de imagens médicas, utilizando o algoritmo de rede neural convolucional de super-resolução rápida (FSRCNN) em conjunto com múltiplas redes residuais aprimoradas. O foco do artigo está em imagens médicas, incluindo ressonância magnética e tomografia computadorizada de áreas como cabeça, joelho e cérebro. Os resultados mostraram que o método proposto superou outras abordagens existentes, gerando imagens reconstruídas com melhores efeitos visuais. No entanto, sugere-se a exploração de estratégias de otimização mais avançadas e a expansão da largura da estrutura da rede para obter ainda melhores resultados.

El-Shafai et al. (2021b) utiliza a técnica de super resolução de imagem única (SISR) em conjunto com modelos de redes neurais convolucionais para realizar a detecção automatizada de *COVID-19*. Os dados utilizados no estudo consistem em imagens de tomografia computadorizada e radiografias obtidas a partir de quatro bases de dados diferentes. Os resultados obtidos demonstraram que o modelo possui um bom grau de generalização, com uma acurácia de 98.53% nas imagens de radiografia e 99.05% nas imagens de tomografia computadorizada. Esses resultados indicam a eficácia do modelo proposto na detecção automatizada de *COVID-19* em diferentes tipos de imagens médicas.

O estudo realizado por Ma et al. (2021) propõe o uso de uma rede generativa adversarial média relativística para aprimorar a qualidade de imagens de ressonância magnética e tomografia computadorizada. Esse método utiliza a probabilidade de que as amostras reais sejam consideradas mais autênticas do que as amostras geradas para avaliar o progresso do modelo na melhoria das imagens. Os resultados do estudo demonstram que o modelo consegue equilibrar efetivamente a complexidade, com aproximadamente 4989 mil parâmetros ajustáveis, em comparação com outros modelos que possuem um número maior de parâmetros, conforme descrito no artigo.

O artigo de Chen et al. (2020) aborda o método de super resolução de imagens médicas chamado *Feedback Adaptive Weighted Dense Network* (FAWDN). O FAWDN utiliza um mecanismo de *feedback* baseado em uma rede neural recorrente para corrigir erros gerados pela rede em etapas temporais anteriores. No estudo, foram utilizadas imagens de ressonância magnética de diferentes partes do corpo humano, incluindo pulmão, abdômen, joelho, cabeça e coluna vertebral. Observou-se que, em comparação com outros modelos mencionados no artigo, o FAWDN apresentou uma velocidade de processamento mais lenta ao lidar com uma única imagem.

Ahmad et al. (2022) propõe uma rede generativa adversarial que utiliza um gerador de múltiplos caminhos e progressivo para aprimorar a resolução de imagens médicas. A arquitetura do gerador é baseada em uma rede *ResNet34* e é treinada em conjunto com um discriminador para gerar imagens de alta qualidade a partir de imagens de baixa resolução. No estudo, foram utilizadas imagens retinianas, imagens de segmentação de lesões de pele, imagens de ressonância

magnética de tumores cerebrais e imagens de ultrassom cardíaco. O artigo utiliza a métrica *Peak Signal-to-Noise Ratio* para analisar os resultados do modelo, demonstrando resultados superiores em comparação com outros dois métodos. No entanto, ainda há espaço para melhorias em comparação com modelos mais recentes.

O trabalho realizado por El-Shafai et al. (2022) apresenta o *Multi-SR and Classification Framework* (MSRCF), um modelo desenvolvido para melhorar a eficiência de armazenamento de imagens médicas e apoiar o diagnóstico eficiente. O MSRCF utiliza técnicas de aprendizado profundo e redes adversariais generativas para superar as limitações das abordagens tradicionais. No estudo, foram utilizadas imagens histopatológicas de câncer de mama e imagens de ultrassom de câncer de mama. Os resultados obtidos demonstraram uma precisão de classificação de 99,74%, evidenciando a eficiência do modelo proposto para o problema abordado.

A abordagem proposta por Zhang et al. (2022) apresenta um novo método chamado *Super-Resolution MRI Using Generative Adversarial Networks* (SOUP-GAN) para aprimorar a resolução de imagens de ressonância magnética. O objetivo é gerar fatias mais finas com maior resolução no plano 'Z', por meio do uso de desfocagem. O método demonstrou potencial em aplicações clínicas, como a redução do tempo de aquisição e a obtenção de resolução adicional em exames de ressonância magnética em estudos retrospectivos, além de ser aplicável em tarefas de pesquisa.

O objetivo deste estudo é explorar e analisar os modelos atualmente utilizados para atividades de super resolução aplicadas na área da saúde. A Tabela 3.2 sintetiza os artigos estudados nesta seção.

Assim, destaca-se que a área da super resolução é dinâmica e continua a evoluir, apresentando potencial de aplicação em várias modalidades de imagens médicas. Fica evidente que a adoção da super resolução é fundamental para a melhoria da qualidade das imagens médicas, oferecendo um suporte substancial aos profissionais da saúde para a realização de diagnósticos mais precisos.

Tabela 3.2: Especificações dos modelos de super resolução dos artigos apresentados

<b>Artigo</b>	<b>Tipo de Imagem Médica</b>	<b>Modelo de super resolução</b>
Avci et al. (2023)	Imagens de sedimentos de urina	<i>Single Image Super-resolution with Self-similarity</i>
Deng et al. (2020)	Imagens de lâminas histopatológicas inteiras	<i>Enhanced Deep Super-Resolution Network</i>
Sun e Ng (2022)	Imagens de angiografia por tomografia computadorizada coronariana	ESRGAN
Qiu et al. (2021)	Imagens de ressonância magnética e tomografia computadorizada	FSRCNN
El-Shafai et al. (2021b)	Imagens de tomografia computadorizada e radiografias	SISR
Ma et al. (2021)	Imagens de ressonância magnética e tomografia computadorizada	Rede Generativa Adversarial Média Relativística
Chen et al. (2020)	Imagens de ressonância magnética	FAWDN
Ahmad et al. (2022)	Imagens retinianas, imagens de segmentação de lesões de pele, imagens de ressonância magnética e imagens de ultrassom	Rede Generativa Adversarial com ResNet34WDN
El-Shafai et al. (2022)	Imagens histopatológicas e imagens de ultrassom	MSRCF
Zhang et al. (2022)	Imagens de ressonância magnética	SOUP-GAN

### 3.3 SUPER RESOLUÇÃO APLICADA A IMAGENS DE CÉLULAS DA SÉRIE BRANCA

Dado o caráter recente da aplicação da super resolução na área da saúde, foi identificado poucos artigos relacionados ao problema abordado neste trabalho. Isso indica que a problemática ainda possui um vasto espaço para ser explorado e investigado em maior profundidade.

O artigo de Tom et al. (2019) propõe o uso de uma rede convolucional profunda aprimorada para super resolução (SRNet) com o objetivo de reconstruir de forma realista, regiões de interesse de células brancas do sangue em esfregaços de sangue periférico ou células epiteliais em histopatologia de tecidos de biópsia cancerígena.

O trabalho utiliza os parâmetros de *Peak Signal-to-Noise Ratio* e *Structural Similarity Index* para avaliar os resultados do modelo. Observou-se que o modelo obteve um *Structural Similarity Index* melhor nas bases de imagens utilizadas em comparação com outros modelos, porém não apresentou resultados superiores nas demais métricas avaliadas.

## 4 MATERIAIS E MÉTODOS

Neste capítulo, apresentamos os procedimentos que nortearam a condução do estudo e dos experimentos. Na seção inicial, fornecemos detalhes adicionais sobre a base de imagens e o processo de obtenção. Em seguida, elucidamos os pormenores relativos à execução dos experimentos, abrangendo as ferramentas e modelos pré-treinados utilizados.

### 4.1 BASES DE IMAGENS

Para a condução do estudo presente neste trabalho, após a análise dos conjuntos de dados mencionados no capítulo anterior, optou-se por estabelecer contato, por *email*, com a empresa *CellaVision*, com o intuito de obter uma base de imagens referente a células da série branca. Após a comunicação com a referida empresa, obteve-se a base composta por 1065 imagens, cada uma delas retratando uma única célula ou artefato.

Essas imagens apresentam dimensões diversas, o que motivou a realização da média das medidas de altura e largura, seguida pelo redimensionamento de todas elas para o formato de  $360 \times 360$  *pixels*. Isso se deve ao fato de que a largura média das imagens foi de 359.71 e a altura média foi de 363.46. Vale ressaltar que todas as imagens foram adquiridas por meio do software utilizado pela *CellaVision*, destinado à realização da microscopia hematológica digital.

Para o treinamento do modelo, foi criado um conjunto de dados composto pelas imagens originais redimensionadas para  $360 \times 360$  *pixels* cada. Estas imagens foram divididas, sendo que o conjunto de treinamento conta com 873 imagens, enquanto 192 imagens foram reservadas para a validação dos resultados. Uma vez que o modelo escolhido incorpora uma ferramenta própria para a geração de imagens de baixa resolução, nenhuma biblioteca em *Python* foi utilizada pela autora.

Para a realização da inferência dos modelos já treinados, foram criadas quatro bases de imagens. Para a geração das imagens de baixa resolução, optou-se por utilizar duas bibliotecas em *Python*: *OpenCV* e *Pillow*. Cada uma dessas bibliotecas disponibiliza uma implementação da interpolação bicúbica que resulta em imagens ligeiramente diferentes.

A Tabela 4.1 apresenta os conjuntos de dados gerados para os experimentos. É relevante ressaltar que todos os conjuntos de dados possuem, ao final, 1065 imagens.

Tabela 4.1: Bases de imagens criadas a partir das imagens fornecidas pela *CellaVision*.

Nome da base de imagens	Biblioteca utilizada para degradação	Tamanho das imagens
<i>OpenCV-90</i>	<i>OpenCV</i>	$90 \times 90$ <i>pixels</i>
<i>OpenCV-180</i>	<i>OpenCV</i>	$180 \times 180$ <i>pixels</i>
<i>Pillow-90</i>	<i>Pillow</i>	$90 \times 90$ <i>pixels</i>
<i>Pillow-180</i>	<i>Pillow</i>	$180 \times 180$ <i>pixels</i>
<i>Original-360</i>	<i>None</i>	$360 \times 360$ <i>pixels</i>

A Figura 4.1 ilustra a mesma imagem presente na base original, assim como nas quatro bases degradadas.

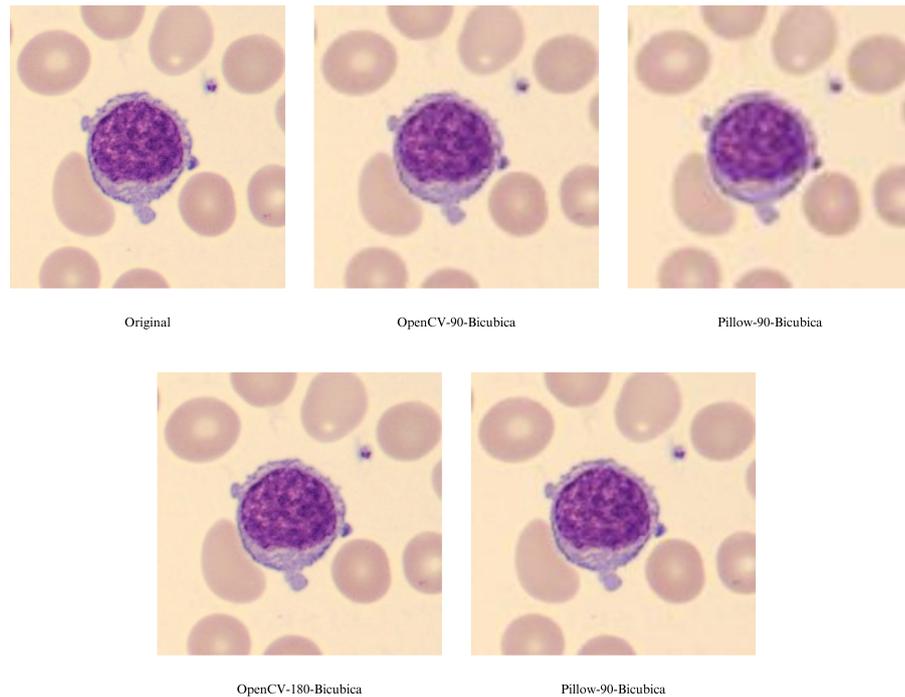


Figura 4.1: Exemplo de mesma imagem em todos os datasets: original e degradados.

## 4.2 EXPERIMENTOS

Para os modelos pré-treinados, os experimentos foram conduzidos no ambiente do Google Colab, utilizando uma GPU T4. No caso dos experimentos relacionados à rede *RealESRGAN*, as bibliotecas indicadas no arquivo *requirements.txt*, fornecido pelo autor, foram empregadas. Além disso, para os experimentos envolvendo essa série de modelos, foram selecionados três modelos pré-treinados apresentados na Tabela 4.2. O fator de ampliação representa em quantas vezes o modelo pode ampliar a imagem que recebe como entrada.

Tabela 4.2: Modelos pré-treinados disponíveis da rede *RealESRGAN*.

Modelo pré-treinado	Fator de Ampliação
RealESRGAN_x2plus	2
RealESRNet_x4plus	4
RealESRGAN_x4plus	4

Portanto, para os modelos que possuem fator de ampliação quatro, utilizaram-se as bases contendo imagens de dimensões  $90 \times 90$  *pixels*. No caso do modelo que ampliava em duas vezes, empregou-se as bases com imagens de  $180 \times 180$  *pixels*, a fim de que as imagens resultantes atingissem as dimensões de  $360 \times 360$  *pixels*, viabilizando assim uma comparação com as imagens originais.

Para os experimentos relacionados a inferência da rede *Swin2SR*, as bibliotecas e suas versões também foram utilizadas as fornecidas pelos autores, no arquivo *requirements.txt*. Com isso, foram escolhidos cinco modelos pré-treinados, apresentados na Tabela 4.3.

Da mesma maneira, para os modelos que ampliavam a resolução da imagem em quatro vezes, utilizaram-se as bases contendo imagens de dimensões  $90 \times 90$  *pixels*. No caso do modelo que ampliava em duas vezes a imagem, empregou-se as bases com imagens de  $180 \times 180$  *pixels*.

Tabela 4.3: Modelos pré-treinados disponíveis do modelo *RealESRGAN*.

Modelo pré-treinado	Fator de Ampliação
ClassicalSR_X2_64	2
Lightweight_X2_64	2
ClassicalSR_X4_64	4
CompressedSR_X4_64	4
RealworldSR_X4_64_BSRGAN_PSNR	4

Para o treinamento do modelo *RealESRGAN*, foi utilizado um *notebook* com placa de vídeo *GeForce RTX 4060* de 8GB, processador *Intel Core i7-13700H*, 64GB de memória RAM e SSD M.2 NVME de 1TB, geração 4.

Foi optado por realizar o processo de *Fine Tuning* do modelo utilizando o conjunto de dados *Original-360*. O *Fine Tuning* tem sido empregado com sucesso na transferência entre tarefas semelhantes (Howard e Ruder, 2018) e constitui uma técnica de treinamento que envolve a reutilização de arquiteturas pré-definidas e pré-treinadas de redes neurais convolucionais (Torres, 2023). A validação do modelo treinado foi feita no ambiente do Google Colab, utilizando uma GPU T4.

O treinamento teve duração aproximada de um dia e oito horas, compreendendo 400.000 iterações. Foram empregadas as configurações fornecidas pelos próprios autores do artigo do *RealESRGAN*. Utilizou-se o modelo pré-treinado *RealESRNet\_x4plus* como base para o treinamento, mencionado anteriormente.

## 5 RESULTADOS

O presente estudo tem como objetivo investigar modelos de super resolução aplicados a imagens de microscopia de células da série branca. Nesse contexto, foram avaliados os modelos *RealESRGAN* e *Swin2SR*, visando analisar o desempenho na aprimoração das imagens. Almeja-se que os resultados obtidos superem aqueles provenientes de uma ampliação por meio da interpolação bicúbica.

Os experimentos estão organizados da seguinte maneira: inicialmente, foi configurado o ambiente Google Colab para os modelos pré-treinados *RealESRGAN* a fim de realizar a inferência nas bases de imagem. No segundo experimento, utiliza-se a rede *Swin2SR* com os modelos pré-treinados para a análise das métricas e do desempenho no problema proposto. O terceiro experimento envolve o processo de *Fine Tuning* do modelo *RealESRGAN*, utilizando a base de dados original dividida em conjuntos de treinamento e validação.

Importa destacar que os modelos foram testados tanto com a degradação proporcionada pela biblioteca *OpenCV* quanto pela *Pillow* do *Python*, visto que as imagens resultantes são ligeiramente diferentes.

### 5.1 BASELINE

Para estabelecer uma *baseline*, optou-se por realizar uma ampliação das imagens de cada um dos conjuntos de dados criados, utilizando a função fornecida pela biblioteca *OpenCV* com interpolação bicúbica, para se obter imagens de tamanho  $360 \times 360$  pixels. As mesmas métricas obtidas na avaliação dos modelos foram extraídas para permitir uma comparação. A Tabela 5.1 apresenta os resultados obtidos como *baseline*.

Tabela 5.1: Resultados obtidos como *baseline*

Dataset	PSNR			SSIM			MSE		
	Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão
Pillow-90	29,76	29,66	1,29	0,89	0,89	0,02	103,36	102,97	20,61
OpenCV-90	31,37	30,98	1,57	0,93	0,93	0,02	73,25	71,71	16,52
Pillow-180	34,34	34,36	1,20	0,96	0,96	0,01	57,37	55,87	13,37
OpenCV-180	37,38	37,42	1,09	0,98	0,98	0,00	32,91	32,04	7,62

Assim, com base nos valores obtidos, é possível concluir que uma ampliação de duas vezes resulta em um valor de SSIM na faixa de 0,96 a 0,98, indicando uma ampliação próxima à imagem original. Contudo, ao realizar uma ampliação de quatro vezes, nota-se uma perda de nitidez, resultando em uma diminuição na métrica SSIM para o intervalo de 0,89 a 0,93. Cabe ressaltar que, na métrica PSNR, foram registrados valores superiores na ampliação de duas vezes, sendo a biblioteca *OpenCV* aquela que obteve o melhor desempenho.

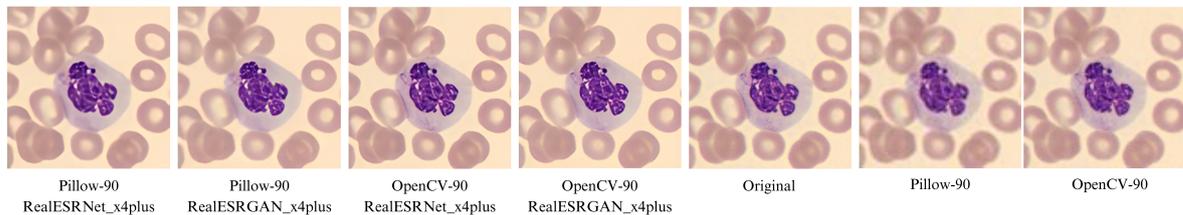
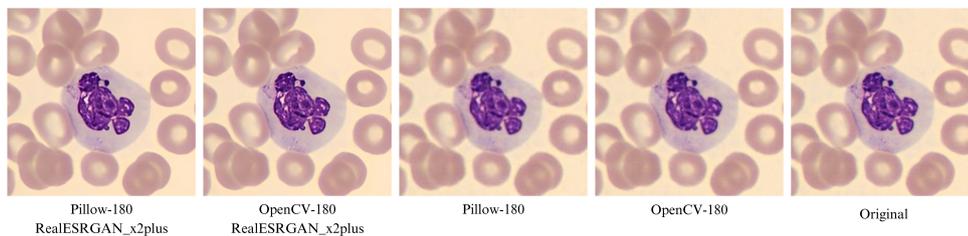
### 5.2 MODELOS PRÉ-TREINADOS

O primeiro experimento consistiu na inferência de todos os modelos da rede *RealESRGAN* utilizando as quatro bases de imagens criadas. Os resultados deste experimento estão apresentados na Tabela 5.2. As quatro primeiras linhas da tabela mostram os resultados obtidos no *baseline*, afim de auxiliar na comparação dos resultados.

Tabela 5.2: Resultados obtidos com a rede *RealESRGAN* com modelos pré-treinados.

Dataset	Modelo	PSNR			SSIM			MSE		
		Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão
Pillow-90	<i>Baseline</i>	29,76	29,66	1,29	0,89	0,89	0,02	103,36	102,97	20,61
OpenCV-90	<i>Baseline</i>	31,37	30,98	1,57	0,93	0,93	0,02	73,25	71,71	16,52
Pillow-180	<i>Baseline</i>	34,34	34,36	1,20	0,96	0,96	0,01	57,37	55,87	13,37
OpenCV-180	<i>Baseline</i>	37,38	37,42	1,09	0,98	0,98	0,00	32,91	32,04	7,62
Pillow-90	RealESRNet_x4plus	31,19	31,09	0,84	0,92	0,92	0,02	73,97	75,18	11,66
Pillow-90	RealESRGAN_x4plus	28,91	28,98	1,02	0,89	0,89	0,03	95,31	97,29	16,43
OpenCV-90	RealESRNet_x4plus	31,44	31,38	1,05	0,92	0,92	0,02	68,13	68,71	12,38
OpenCV-90	RealESRGAN_x4plus	29,10	29,20	1,11	0,90	0,90	0,02	92,86	94,35	16,51
Pillow-180	RealESRGAN_x2plus	30,26	30,26	1,49	0,92	0,92	0,02	71,27	70,36	13,70
OpenCV-180	RealESRGAN_x2plus	30,59	30,58	1,53	0,93	0,93	0,02	66,70	65,62	13,56

As Figuras 5.1 e 5.2 mostram visualmente os resultados obtidos com os modelos pré-treinados da *RealESRGAN*.

Figura 5.1: Resultados visuais obtidos nos modelos pré-treinados da *RealESRGAN* com fator de ampliação 4.Figura 5.2: Resultados visuais obtidos nos modelos pré-treinados da *RealESRGAN* com fator de ampliação 2.

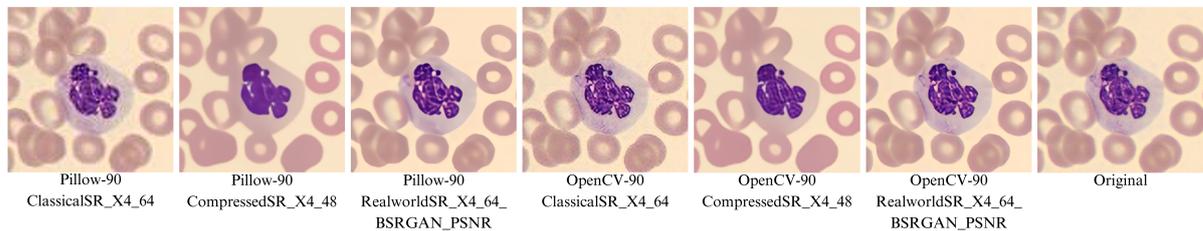
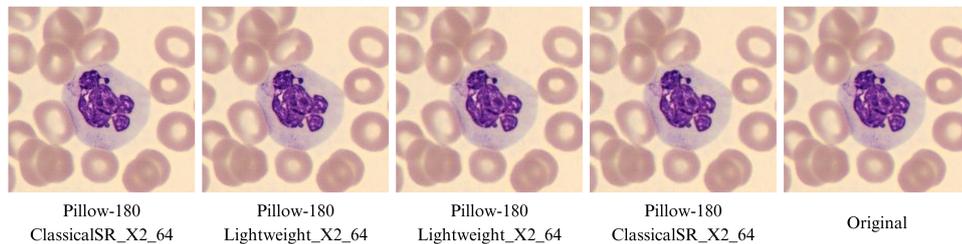
Dessa forma, é observável que o modelo *RealESRGAN* alcançou um resultado próximo à interpolação bicúbica nas ampliações de fator quatro, com o SSIM variando de 0,89 a 0,92. Entretanto, nas ampliações de fator dois, o modelo obteve resultados inferiores à interpolação bicúbica, em todas as três métricas. É relevante ressaltar que, neste experimento, tanto a biblioteca *Pillow* quanto a biblioteca *OpenCV* obtiveram resultados próximos, embora não idênticos.

No segundo experimento, foram conduzidas as inferências do modelo *Swin2SR* utilizando os pesos apresentados na seção anterior, e as bases de imagens geradas a partir do conjunto de dados fornecido pela *CellaVision*. A Tabela 5.3 exibe os resultados alcançados. As quatro primeiras linhas da tabela mostram os resultados obtidos no *baseline*, afim de auxiliar na comparação dos resultados.

As Figuras 5.3 e 5.4 mostram visualmente os resultados obtidos com os modelos pré-treinados da *RealESRGAN*.

Tabela 5.3: Resultados obtidos com a rede *Swin2SR* com modelos pré-treinados.

Dataset	Modelo	PSNR			SSIM			MSE		
		Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão
Pillow-90	<i>Baseline</i>	29,76	29,66	1,29	0,89	0,89	0,02	103,36	102,97	20,61
OpenCV-90	<i>Baseline</i>	31,37	30,98	1,57	0,93	0,93	0,02	73,25	71,71	16,52
Pillow-180	<i>Baseline</i>	34,34	34,36	1,20	0,96	0,96	0,01	57,37	55,87	13,37
OpenCV-180	<i>Baseline</i>	37,38	37,42	1,09	0,98	0,98	0,00	32,91	32,04	7,62
OpenCV-90	ClassicalSR_X4_64	29,54	29,45	1,70	0,89	0,89	0,03	86,07	85,52	19,30
OpenCV-90	CompressedSR_X4_64	29,01	28,97	1,07	0,89	0,89	0,03	112,05	113,91	20,92
OpenCV-90	RealworldSR_X4_64_BSRGAN_PSNR	31,77	31,74	1,04	0,93	0,93	0,02	61,17	61,51	11,62
Pillow-90	ClassicalSR_X4_64	30,09	30,01	1,22	0,89	0,89	0,02	104,99	105,61	20,88
Pillow-90	CompressedSR_X4_64	29,65	29,59	1,05	0,89	0,90	0,03	99,37	101,22	18,01
Pillow-90	RealworldSR_X4_64_BSRGAN_PSNR	31,23	31,20	1,04	0,92	0,92	0,02	70,50	71,34	13,60
OpenCV-180	ClassicalSR_X2_64	35,68	35,72	1,25	0,96	0,96	0,01	41,61	40,37	10,12
OpenCV-180	Lightweight_X2_64	35,69	35,74	1,28	0,96	0,96	0,01	41,56	40,14	10,28
Pillow-180	ClassicalSR_X2_64	33,91	33,90	1,24	0,95	0,95	0,01	60,11	58,46	14,04
Pillow-180	Lightweight_X2_64	33,93	33,94	1,24	0,95	0,95	0,01	60,02	58,42	14,04

Figura 5.3: Resultados visuais obtidos nos modelos pré-treinados da *Swin2SR* com fator de ampliação 4Figura 5.4: Resultados visuais obtidos nos modelos pré-treinados *Swin2SR* com fator de ampliação 2

Com base nesse resultado, é evidente que o modelo também apresentou métricas próximas à interpolação bicúbica ao aumentar a imagem em duas vezes, com SSIM variando entre 0,89 e 0,93. Ao analisar os resultados da ampliação da imagem em quatro vezes, o modelo *Swin2SR* obteve SSIM e PSNR superiores aos alcançados pela rede *RealESRGAN*, mas ainda inferiores em relação à interpolação bicúbica.

Neste experimento, observou-se que a biblioteca *Pillow* apresentou um desempenho de PSNR e SSIM inferior à biblioteca *OpenCV*, embora a diferença não seja estatisticamente significativa.

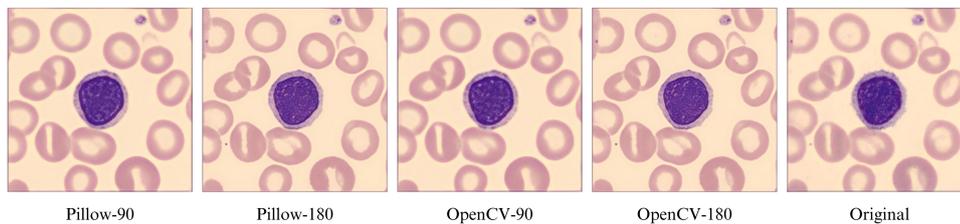
### 5.3 MODELO COM FINE-TUNING

Para avaliar o modelo *RealESRGAN* com *Fine Tuning*, foi realizada a inferência das 192 imagens separadas para validação. A Tabela 5.4 mostra os resultados obtidos, com cada uma das bases de imagens criadas. As quatro primeiras linhas da tabela mostram os resultados obtidos no *baseline*, afim de auxiliar na comparação dos resultados.

Tabela 5.4: Resultados obtidos com a rede *Swin2SR* com com *Fine Tuning*.

Dataset	PSNR			SSIM			MSE		
	Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão	Média	Mediana	Desvio Padrão
Pillow-90	29,76	29,66	1,29	0,89	0,89	0,02	103,36	102,97	20,61
OpenCV-90	31,37	30,98	1,57	0,93	0,93	0,02	73,25	71,71	16,52
Pillow-180	34,34	34,36	1,20	0,96	0,96	0,01	57,37	55,87	13,37
OpenCV-180	37,38	37,42	1,09	0,98	0,98	0,00	32,91	32,04	7,62
Pillow-90	30,06	30,26	0,94	0,88	0,89	0,02	93,15	91,75	11,04
OpenCV-90	30,14	30,32	0,93	0,89	0,89	0,01	92,41	91,23	10,59
Pillow-180	28,78	28,93	0,66	0,89	0,89	0,01	99,31	98,11	11,41
OpenCV-180	28,86	29,00	0,65	0,89	0,89	0,01	98,24	97,02	11,26

A Figura 5.5 mostram visualmente os resultados obtidos com o modelo treinado da *RealESRGAN*

Figura 5.5: Resultados visuais obtidos no modelo treinados *RealESRGAN*

Portanto, é notável que o modelo treinado apresentou um desempenho inferior em comparação com a *baseline*, obtendo um SSIM entre 0,88 e 0,89 nas quatro bases de imagens geradas. Pode-se inferir que as imagens exibem uma aparência mais borrada, com perda das finas estruturas das células.

## 6 CONCLUSÃO

O objetivo deste trabalho é realizar um estudo sobre os modelos de Super Resolução aplicados a imagens de microscopia de células da série branca, um campo ainda muito vasto e com grande potencial para estudos. Para construir a base de dados, estabeleceu-se contato com a empresa *CellaVision* e obteve-se imagens de alta resolução. A partir disso, foram geradas quatro bases de imagens degradadas utilizando interpolação bicúbica, por meio das bibliotecas *OpenCV* e *Pillow* no *Python*, gerando imagens com redução de tamanho em quatro e duas vezes.

Como *baseline*, foi realizado um aumento das imagens utilizando apenas a função *resize* da biblioteca *OpenCV*. No primeiro experimento, conduziu-se a análise da rede *RealESRGAN* para a inferência dessas bases de imagem. No segundo experimento, utilizou-se a rede *Swin2SR* para a análise das métricas e do desempenho do modelo no problema proposto. O terceiro experimento consistiu em fazer o *Fine Tuning* do modelo *RealESRGAN*, com a base de dados original dividida em treinamento e validação.

Os resultados obtidos indicam que os modelos não foram estatisticamente superiores à interpolação bicúbica isolada, sugerindo a conclusão de que o treinamento completo das redes poderiam resultar em desempenhos superiores.

Além disso, pretende-se, no futuro, criar uma base de dados das células da série branca que represente de maneira mais fiel uma ampliação de um microscópio, com o objetivo de obter resultados mais realistas e condizentes com o problema em questão.

## REFERÊNCIAS

- Agarwal, V., Lohani, M. C., Singh Bist, A., Rahardja, L., Hardini, M. e Mustika, G. (2022). Deep cnn–real esrgan: An innovative framework for lung disease prediction. Em *2022 IEEE Creative Communication and Innovative Technology (ICCIT)*, páginas 1–6.
- Ahmad, W., Ali, H., Shah, Z. e Azmat, S. (2022). A new generative adversarial network for medical images super resolution. *Scientific Reports*, 12(1):9533.
- Andrade, A. R., Vogado, L. H., de M.S. Veras, R., Silva, R. R., Araujo, F. H. e Medeiros, F. N. (2019). Recent computational methods for white blood cell nuclei segmentation: A comparative study. *Computer Methods and Programs in Biomedicine*, 173:1–14.
- Anita e Yadav, A. (2021). An intelligent model for the detection of white blood cells using artificial intelligence. *Computer Methods and Programs in Biomedicine*, 199:105893.
- Anwar, S., Khan, S. e Barnes, N. (2020). A deep journey into super-resolution: A survey. *ACM Comput. Surv.*, 53(3).
- Avcı, D., Sert, E., Dogantekin, E., Yildirim, O., Tadeusiewicz, R. e Plawiak, P. (2023). A new super resolution faster r-cnn model based detection and classification of urine sediments. *Biocybernetics and Biomedical Engineering*, 43(1):58–68.
- Banik, P. P., Saha, R. e Kim, K.-D. (2020). An automatic nucleus segmentation and cnn model based classification method of white blood cell. *Expert Systems with Applications*, 149:113211.
- Bi, J., Zhu, Z. e Meng, Q. (2021). Transformer in computer vision. Em *2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, páginas 178–188.
- Cai, L., Chen, Y., Cai, N., Cheng, W. e Wang, H. (2020). Utilizing amari-alpha divergence to stabilize the training of generative adversarial networks. *Entropy*, 22:410.
- Chen, L., Yang, X., Jeon, G., Anisetti, M. e Liu, K. (2020). A trusted medical image super-resolution method based on feedback adaptive weighted dense network. *Artificial Intelligence in Medicine*, 106:101857.
- Conde, M. V., Choi, U.-J., Burchi, M. e Timofte, R. (2022). Swin2sr: Swinv2 transformer for compressed image super-resolution and restoration.
- Deng, Y., Feng, M., Jiang, Y., Zhou, Y., Qin, H., Xiang, F., Wang, Y., Bu, H. e Bao, J. (2020). Development of pathological reconstructed high-resolution images using artificial intelligence based on whole slide image. *MedComm*, 1(3):410–417.
- Denton, E. L., Chintala, S., szlam, a. e Fergus, R. (2015). Deep generative image models using a laplacian pyramid of adversarial networks. Em Cortes, C., Lawrence, N., Lee, D., Sugiyama, M. e Garnett, R., editores, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc.

- Dong, N., Feng, Q., Zhai, M., Chang, J. e Mai, X. (2022). A novel feature fusion based deep learning framework for white blood cell classification. *Journal of Ambient Intelligence and Humanized Computing*.
- Dosselmann, R. e Yang, X. D. (2011). A comprehensive assessment of the structural similarity index. *Signal, Image and Video Processing*, 5(1):81–91.
- El-Shafai, W., Magdy, A., El-Rabaie, E.-S., Soliman, N., Algarni, A. e Abd El-Samie, F. (2021a). Automated covid-19 detection based on single-image super-resolution and cnn models. *Computers, Materials and Continua*, 70:1141–1157.
- El-Shafai, W., Magdy, A., El-Rabaie, E.-S., Soliman, N., Algarni, A. e Abd El-Samie, F. (2021b). Automated covid-19 detection based on single-image super-resolution and cnn models. *Computers, Materials and Continua*, 70:1141–1157.
- El-Shafai, W., Mohamed, E., Zeghid, M., Magdy, A. e Aly, M. (2022). Hybrid single image super-resolution algorithm for medical images. *Computers, Materials and Continua*, 72:4879–4896.
- Gao, G., Xu, Z., Li, J., Yang, J., Zeng, T. e Qi, G.-J. (2023). Ctcnet: A cnn-transformer cooperation network for face image super-resolution. *IEEE Transactions on Image Processing*, 32:1978–1991.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. e Bengio, Y. (2014). Generative adversarial networks.
- Hoffbrand, A., Moss, P. e Pettit, J. (2009). *Fundamentos em hematologia - 5.Ed.* Artmed Editora.
- Honda, K., Fujita, H. e Kurematsu, M. (2022). Improvement of text image super-resolution benefiting multi-task learning. Em Fujita, H., Fournier-Viger, P., Ali, M. e Wang, Y., editores, *Advances and Trends in Artificial Intelligence. Theory and Practices in Artificial Intelligence*, páginas 275–286, Cham. Springer International Publishing.
- Howard, J. e Ruder, S. (2018). Universal language model fine-tuning for text classification.
- James, G., Witten, D., Hastie, T. e Tibshirani, R. ([2013]). *An introduction to statistical learning : with applications in R*. New York : Springer, [2013] ©2013. Includes bibliographical references and index.
- Johnson, J., Alahi, A. e Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution.
- Korhonen, J. e You, J. (2012). Peak signal-to-noise ratio revisited: Is simple beautiful? Em *2012 Fourth International Workshop on Quality of Multimedia Experience*, páginas 37–38.
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. e Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network.
- Liang, J., Cao, J., Sun, G., Zhang, K., Gool, L. V. e Timofte, R. (2021). Swinir: Image restoration using swin transformer.
- Liu, Q., Li, J. e Lu, Z. (2021). St-tran: Spatial-temporal transformer for cellular traffic prediction. *IEEE Communications Letters*, 25(10):3325–3329.

- Liu, Y., Cao, F., Zhao, J. e Chu, J. (2017). Segmentation of white blood cells image using adaptive location and iteration. *IEEE Journal of Biomedical and Health Informatics*, 21(6):1644–1655.
- Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., Ning, J., Cao, Y., Zhang, Z., Dong, L., Wei, F. e Guo, B. (2022). Swin transformer v2: Scaling up capacity and resolution.
- Lv, H., Zhao, A., Wang, Y., Li, H., Li, Y. e Zou, G. (2023). Research on missing bolt sub-detection based on real-esrgan super-resolution reconstruction. Em *2023 4th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, páginas 520–523.
- Ma, Y., Liu, K., Xiong, H., Fang, P., Li, X., Chen, Y., Yan, Z., Zhou, Z. e Liu, C. (2021). Medical image super-resolution using a relativistic average generative adversarial network. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 992:165053.
- Nasrollahi, K. e Moeslund, T. B. (2014). Super-resolution: a comprehensive survey. *Machine Vision and Applications*, 25(6):1423–1468.
- Pandit, V. e Schuller, B. W. (2019). On many-to-many mapping between concordance correlation coefficient and mean square error. *CoRR*, abs/1902.05180.
- Parvaiz, A., Khalid, M. A., Zafar, R., Ameer, H., Ali, M. e Fraz, M. M. (2023). Vision transformers in medical computer vision—a contemplative retrospection. *Engineering Applications of Artificial Intelligence*, 122:106126.
- Poobathy, D. e Chezian, R. (2014). Edge detection operators: Peak signal to noise ratio based comparison. *International Journal of Image, Graphics and Signal Processing*, 6:55–61.
- Prinyakupt, J. e Pluempitiwiriyaewej, C. (2015). Segmentation of white blood cells and comparison of cell morphology by linear and naïve bayes classifiers. *BioMedical Engineering OnLine*, 14(1):63.
- Qiu, D., Zheng, L., Zhu, J. e Huang, D. (2021). Multiple improved residual networks for medical image super-resolution. *Future Generation Computer Systems*, 116:200–208.
- Schönfeld, E., Schiele, B. e Khoreva, A. (2021). A u-net based discriminator for generative adversarial networks.
- Septian, G., Wahiddin, D., Novita, H. Y., Handayani, H. H., Juwita, A. R. e Nur Masruriyah, A. F. (2022). The implementation of real-esrgan as an anticipation to reduce cer value in plate number extraction results employing easyocr. Em *2022 Seventh International Conference on Informatics and Computing (ICIC)*, páginas 1–5.
- Shahin, A., Guo, Y., Amin, K. e Sharawi, A. A. (2019). White blood cells identification system based on convolutional deep neural learning networks. *Computer Methods and Programs in Biomedicine*, 168:69–80.
- Song, K.-U., Shim, D., wook Kim, K., young Lee, J. e Kim, Y. (2022). Fs-ncsr: Increasing diversity of the super-resolution space via frequency separation and noise-conditioned normalizing flow.
- Sun, Z. e Ng, C. K. C. (2022). Artificial intelligence (enhanced super-resolution generative adversarial network) for calcium deblooming in coronary computed tomography angiography: A feasibility study. *Diagnostics*, 12(4).

- Tian, J. e Ma, K.-K. (2011). A survey on super-resolution imaging. *Signal, Image and Video Processing*, 5(3):329–342.
- Tom, F., Sharma, H., Mundhra, D., Dastidar, T. R. e Sheet, D. (2019). Learning a deep convolution network with turing test adversaries for microscopy image super resolution. Em *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, páginas 1391–1394.
- Torres, M. B. (2023). O que É o fine-tuning (ajuste fino) e como ele funciona nas redes neurais?
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. e Polosukhin, I. (2023). Attention is all you need.
- VIVAS, W. L. P. (2014). *Manual prático de hematologia*. Universidade Tirandentes.
- Vogado, L. H., Veras, R. e Lins, J. (2016). Avaliação de técnicas de segmentação para leucócitos em imagens de sangue. *Revista de Sistemas e Computação*, 6:65–73.
- Wang, P., Bayram, B. e Sertel, E. (2022). A comprehensive review on deep learning based remote sensing image super-resolution methods. *Earth-Science Reviews*, 232:104110.
- Wang, X., Xie, L., Dong, C. e Shan, Y. (2021a). Real-esrgan: Training real-world blind super-resolution with pure synthetic data.
- Wang, Z., Chen, J. e Hoi, S. C. H. (2021b). Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10):3365–3387.
- Xu, M., Dai, W., Liu, C., Gao, X., Lin, W., Qi, G.-J. e Xiong, H. (2021). Spatial-temporal transformer networks for traffic flow forecasting.
- Yamin, M., Basahel, A. M., Abusurrah, M., Basahel, S. M., Mohanty, S. N. e Lydia, E. L. (2023). Automated artificial intelligence empowered white blood cells classification model. *Computers, Materials & Continua*, 75(1):409–425.
- Yang, R., Timofte, R., Li, X., Zhang, Q., Zhang, L., Liu, F., He, D., li, F., Zheng, H., Yuan, W., Ostyakov, P., Vyal, D., Zhussip, M., Zou, X., Yan, Y., Li, L., Tang, J., Chen, M., Zhao, S., Zhu, Y., Qin, X., Li, C., Leng, C., Cheng, J., Rota, C., Buzzelli, M., Bianco, S., Schettini, R., Zhang, D., Huang, F., Liu, S., Wang, X., Jin, Z., Li, B., Li, X., Li, M., Liu, D., Zou, W., Dong, P., Ye, T., Zhang, Y., Tan, M., Niu, X., Ayazoglu, M., Conde, M., Choi, U.-J., Jia, Z., Xu, T., Zhang, Y., Ye, M., Luo, D., Pan, X. e Peng, L. (2022). Aim 2022 challenge on super-resolution of compressed image and video: Dataset, methods and results.
- Yao, X., Sun, K., Bu, X., Zhao, C. e Jin, Y. (2021). Classification of white blood cells using weighted optimized deformable convolutional neural networks. *Artificial Cells, Nanomedicine, and Biotechnology*, 49(1):147–155. PMID: 33533656.
- Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H. e Zhang, L. (2016). Image super-resolution: The techniques, applications, and future. *Signal Processing*, 128:389–408.
- Zhang, K., Hu, H., Philbrick, K., Conte, G. M., Sobek, J. D., Rouzrokh, P. e Erickson, B. J. (2022). Soup-gan: Super-resolution mri using generative adversarial networks. *Tomography*, 8(2):905–919.
- Zhu, M., Pan, P., Chen, W. e Yang, Y. (2019). Dm-gan: Dynamic memory generative adversarial networks for text-to-image synthesis.