



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS, SOCIAIS E APLICADAS
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

ELANNIO DE SOUSA SANTIAGO

**Recuperando a Beleza do Passado: Restauração do Cinema Clássico através da
Inteligência Artificial**

**PATOS
2024**

ELANNIO DE SOUSA SANTIAGO

**Recuperando a Beleza do Passado: Restauração do Cinema Clássico através da
Inteligência Artificial**

Trabalho de Conclusão de Curso
apresentado ao Curso de Graduação em
Ciência da Computação da Universidade
Estadual da Paraíba.

Área de concentração: Inteligência
Artificial.

Orientador(a): Profa. Dra. Jannayna Domingues Barros Filgueira

**PATOS
2024**

É expressamente proibida a comercialização deste documento, tanto em versão impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que, na reprodução, figure a identificação do autor, título, instituição e ano do trabalho.

S235r Santiago, Elannio de Sousa.
Recuperando a beleza do passado [manuscrito] :
restauração do cinema clássico através da inteligência artificial
/ Elannio de Sousa Santiago. - 2024.
49 f. : il. color.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Ciência da computação) - Universidade Estadual da Paraíba, Centro de Ciências Exatas e Sociais Aplicadas, 2024.

"Orientação : Prof. Dra. Jannayna Domingues Barros Filgueira, Coordenação do Curso de Computação - CCEA".

1. Cinema. 2. Denoising convolutional neural network (DnCNN). 3. Inteligência artificial. I. Título

21. ed. CDD 006.3

ELANNIO DE SOUSA SANTIAGO

RECUPERANDO A BELEZA DO PASSADO: RESTAURAÇÃO DO CINEMA
CLÁSSICO ATRAVÉS DA INTELIGÊNCIA ARTIFICIAL

Trabalho de Conclusão de Curso
apresentado à Coordenação do Curso
de Ciência da Computação da
Universidade Estadual da Paraíba,
como requisito parcial à obtenção do
título de Bacharel em Ciência da
Computação

Aprovada em: 18/11/2024.

Documento assinado eletronicamente por:

- **Suéllen Rodrigues Ramos da Silva** (**.528.604-**), em **25/11/2024 20:25:01** com chave **7e4c8becab8411ef997b06adb0a3afce**.
- **Jannayna Domingues Barros Filgueira** (**.837.144-**), em **25/11/2024 21:35:22** com chave **527f167eab8e11ef8e9b06adb0a3afce**.
- **Harllem Alves do Nascimento** (**.796.924-**), em **26/11/2024 17:14:14** com chave **020b9988ac3311ef89ed06adb0a3afce**.

Documento emitido pelo SUAP. Para comprovar sua autenticidade, faça a leitura do QrCode ao lado ou acesse https://suap.uepb.edu.br/comum/autenticar_documento/ e informe os dados a seguir.

Tipo de Documento: Termo de Aprovação de Projeto Final

Data da Emissão: 02/12/2024

Código de Autenticação: 26b125



AGRADECIMENTOS

À minha família, especialmente à minha mãe, que mesmo com as dificuldades sempre tentou fazer o melhor para mim e minhas irmãs.

Ao meu namorado, José Osvaldo, que sempre me incentivou a continuar e não desistir, sempre me ajudou em diversos problemas nesses cinco anos e mostrou como é ser uma pessoa tão especial. Obrigado por tudo.

Aos meus amigos, especialmente Dayane, Eduarda, Lucas, Danilo, Daniel, Emerson e outros amigos que conversam e me incentivam todo dia no nosso grupo do Whatsapp, principalmente Antônio e Robson. Todos me ajudaram a passar por essa bagunça que é a minha vida. Obrigado Dayane e Eduarda pelos dias que passamos e os dias que ainda vamos passar juntos, conversando e vendo filmes.

À minha orientadora, Profa. Dra. Jannayna Domingues Barros Filgueira e todas as professoras que passaram por mim e me ajudaram a conquistar toda a minha base de conhecimento. Obrigado a Rísia, Gorete, Luzia e Corrinha. Vocês são luz na vida de muitos alunos.

Este trabalho é dedicado a todos que passaram pela minha vida e marcaram de forma positiva.

“Feliz é o destino da inocente vestal!
Esquecendo-se do mundo e sendo por ele
esquecida, Brilho eterno de uma mente sem
lembranças! Toda prece é ouvida, toda graça
se alcança.”

(Brilho eterno de uma mente sem
lembranças, 2004).

RESUMO

Com o passar dos anos, a tecnologia vem atribuindo qualidade em diversas áreas do conhecimento. Em contrapartida, o cinema vem como uma forma de arte e um mercado que está sempre em evolução. Os meios computacionais são relativamente novos, então, o cinema não recebeu a ajuda necessária no passado. Deste modo, este trabalho busca explicar como os filmes foram se perdendo ao longo dos anos por falta de cuidado ou acidentes, muitos deles não existem mais ou estão sendo recuperados e melhorados de forma manual. Tendo em vista que a tecnologia pode ajudar outros campos a sobreviver, se preservar e evoluir, este trabalho tem como objetivo a discussão da Inteligência Artificial na restauração de filmes clássicos, buscando detectar técnicas de aprendizado profundo que possam ajudar um filme a se preservar ao longo dos anos e, principalmente, recuperar outros que foram muito prejudicados pelo tempo. Este estudo tem como foco principal a aplicação de uma Rede Neural, *Denoising Convolutional Neural Network* (DnCNN), que busca a remoção de ruídos (alteração de pixels em uma imagem) de forma eficaz. Utilizando um filme como base para testes e recuperação de informações danificadas da película. A aplicação envolveu a coleta de *frames* em um filme danificado e a utilização do algoritmo para remoção do ruído da imagem. Desse modo, as imagens foram submetidas para teste e foram atribuídos parâmetros de semelhança com o resultado final, assim, aplicando a remoção de ruído nas imagens.

Palavras-chave: cinema; denoising convolutional neural network (DnCNN); inteligência artificial.

ABSTRACT

Over the years, technology has been contributing to the enhancement of various fields of knowledge. On the other hand, cinema stands as an art form and a market that is constantly evolving. Computational methods are relatively new, so cinema did not receive the necessary support in the past. Thus, this work seeks to explain how films have been lost over the years due to negligence or accidents; many no longer exist or are being manually recovered and improved. Considering that technology can help other fields survive, preserve themselves, and evolve, this study aims to discuss the role of Artificial Intelligence in restoring classic films, focusing on deep learning techniques that can assist in preserving films over time and, most importantly, recovering those severely damaged by time. This research primarily centers on the application of a Neural Network, Denoising Convolutional Neural Network (DnCNN), designed to effectively remove noise (pixel alterations in an image). The study involved using a film as the basis for testing and recovering damaged frame information. Frames from a damaged film were collected, and the algorithm was applied to remove noise from the images. Subsequently, the images were tested, and similarity parameters were established for the final result, effectively applying noise removal to the frames.

Keywords: cinema; denoising convolutional neural network (DnCNN); artificial intelligence.

LISTA DE ILUSTRAÇÕES

Figura 1: Processo de tratamento e restauração de imagens.....	19
Figura 2: Representação das redes neurais.....	21
Figura 3: A arquitetura da rede DnCNN proposta.....	27
Figura 4: Resultado do treino de 50 épocas.....	35
Figura 5: Remoção de ruído de quatro imagens diferentes em formato jpg.....	37
Figura 6: Imagem 1 restaurada do filme Nosferatu (1922), antes e depois.....	38
Figura 7: Representação da Imagem 1 com ruído gaussiano e a imagem restaurada.	39
Figura 8: Imagem 2 restaurada do filme Nosferatu (1922), antes e depois.....	40
Figura 9: Representação da imagem 2 com ruído gaussiano e a imagem restaurada..	41
Figura 10: Imagem 3 restaurada do filme Nosferatu (1922), antes e depois.....	41
Figura 11: Representação da Imagem 3 com ruído gaussiano e a imagem restaurada	42
Figura 13: Representação da Imagem 4 com ruído gaussiano e a imagem restaurada.....	43

LISTA DE QUADROS

Quadro 1: Bibliotecas e API.....	19
----------------------------------	----

LISTA DE ABREVIATURAS E SIGLAS

API	<i>Application Programming Interface</i>
BMP	<i>Bitmap</i>
BN	<i>Batch Normalization</i>
CCD	<i>Charge-Coupled Device</i>
CNNs	<i>Convolutional Neural Networks</i>
Conv	<i>Convolutional Layer</i>
CPU	<i>Central Processing Unit</i>
CUDA	<i>Compute Unified Device Architecture</i>
cuDNN	<i>CUDA Deep Neural Network library</i>
DAE	<i>Denosing Autoencoder</i>
DB	<i>Decibéis</i>
DnCNN	<i>Denosing Convolutional Neural Network</i>
EDSR	<i>Enhanced Deep Super-Resolution</i>
ESRGAN	<i>Enhanced Super-Resolution Generative Adversarial Networks</i>
GANs	<i>Generative Adversarial Networks</i>
GIF	<i>Graphics Interchange Format</i>
GPU	<i>Graphics Processing Unit</i>
IA	<i>Inteligência Artificial</i>
JPEG	<i>Joint Photographic Experts Group</i>
ML	<i>Machine Learning</i>
NPCs	<i>Non-Player Characters</i>
NFPB	<i>National Federation of Professional Bullriders</i>
PDF	<i>Portable Document Format</i>
PDS	<i>Processamento Digital de Sinais</i>
PNG	<i>Portable Network Graphics</i>
PSNR	<i>Peak Signal-to-Noise Ratio</i>
ReLU	<i>Rectified Linear Unit</i>
RNNs	<i>Recurrent Neural Networks</i>
SISR	<i>Single Image Super-Resolution</i>
SR	<i>Super-Resolution</i>
SRCNN	<i>Super-Resolution Convolutional Neural Network</i>

SSIM	<i>Structural Similarity Index</i>
TCNs	<i>Temporal Convolutional Networks</i>
VDSR	<i>Very Deep Super-Resolution</i>

SUMÁRIO

1 INTRODUÇÃO.....	12
1.1 Objetivos.....	14
1.1.1 Objetivo geral.....	14
1.1.2 Objetivos Específicos.....	14
1.1 Problemática.....	14
1.2 Justificativa.....	16
2 REVISÃO BIBLIOGRÁFICA.....	17
2.1 Processamento Digital de Sinais.....	17
2.2 Processamento Digital de Imagens.....	18
2.2.1 Imagens analógicas.....	18
2.2.2 Imagens digitais.....	19
2.2.3 Processo de tratamento e restauração de imagens.....	20
2.2.4 Extração de características/dados.....	22
2.3 Inteligência Artificial.....	23
2.3.1 Técnicas de Inteligência Artificial aplicadas à restauração de filmes.....	25
2.3.2 Algoritmos para restauração de vídeo.....	26
2.4 Denoising Convolutional Neural Network (DnCNN).....	27
2.4.1 Arquitetura do DnCNN.....	29
2.4.2 PSNR e SSIM.....	30
2.5 Trabalhos relacionados.....	31
3 METODOLOGIA.....	33
3.1 Treinamento do algoritmo.....	33
3.1.1 Python.....	34
3.1.2 TensorFlow e CUDA.....	36
3.1.3 Controladores de GPU.....	36
3.1.4 Configurações do treinamento.....	36
3.2 Teste do algoritmo treinado.....	38
4 RESULTADOS E DISCUSSÕES.....	40
5 CONSIDERAÇÕES FINAIS.....	46
REFERÊNCIAS.....	48

1 INTRODUÇÃO

Para Ferro (1992), o cinema é um modo de expressão histórico, em que as obras cinematográficas refletem a sociedade da qual emanam, suas mentalidades, suas crises e evoluções. Como parte de um momento histórico do século XX, conseguiu atrair milhões de pessoas que estavam à procura de diversos tipos de experiências. Cada indivíduo possui um momento marcante relacionado a filmes, podendo permanecer ou não nessa cultura. Diante desse fato, aqueles que permanecem guardam lembranças especiais sobre cada filme assistido.

Em consequência do cenário cinematográfico e tecnológico do século XXI, os filmes clássicos perderam o apelo que anteriormente possuíam diante do público. Com a alta tecnologia do século XXI, o mundo vive momentos gradativos e rápidos de evolução: a televisão e o *streaming* de vídeo cresceram em popularidade por terem quadros e *frames* maiores. O cinema teve que evoluir também para que os filmes não caiam no esquecimento.

Para entender o contexto devastador da história do cinema, em um informe feito pela *National Federation of Professional Bullriders* (NFPB), é afirmado que dos 11 mil filmes de ficção do século XX, apenas 14% deles sobreviveram até os dias atuais. Já um estudo nomeado "A sobrevivência dos filmes mudos norte-americanos: 1912-1929" diz que cerca de 11% dos filmes só foram preservados em outros países ou em diferentes formatos.

De acordo com Martin Scorsese, em 2013, diretor de cinema famoso pelos filmes *Taxi Driver* (1976) e *Bons Companheiros* (1990), nós devemos nos preocupar com o que ainda nos resta. Aqueles filmes que foram perdidos ao longo dos anos devem ser usados para pensar em como vamos definir o futuro do que restou, assim ajudando a entender que o passado importa.

Scorsese também se preocupa com a perda dos filmes clássicos e a falta de preocupação das pessoas em relação a saber suas origens e que possivelmente não entendam que o passado é importante para a compreensão do mundo atual. Essa preocupação não está, somente, ligada à conservação dos acervos cinematográficos do século XX, mas em criar uma prática saudável. Para que assim, com o passar dos anos, tenham uma preocupação maior com as relíquias que restam do passado.

A determinação dos cineastas em preservar o passado é algo que deve ser elogiado. Tendo em vista que muitos filmes não se adaptaram às novas tecnologias e necessitam de mais cuidados para não serem perdidos, a ideia de uma restauração e preservação do cinema é um avanço significativo no âmbito tecnológico.

Deve-se preservar a grande quantidade de filmes antigos de forma mais inteligente, melhorando imagem, ruído e áudio. Nesse contexto, é importante fazer uso de um dos meios mais atuais da tecnologia, a Inteligência Artificial (IA), tendo em vista que os métodos de restauração utilizam a IA.

A Inteligência Artificial (IA), no pensamento de Stuart Russell e Peter Norvig (2013), está dividida em 4 partes: o “Pensando como um humano”, que devem ser máquinas com mente, no sentido total e literal; em seguida, o “Pensando racionalmente”, em que elas devem raciocinar e agir; por consequência, em seguida temos o “Agindo como seres humanos”; e, por último, o “Agindo racionalmente”, em que o computador deve seguir os passos do ser humano em aprender e executar racionalmente tudo.

Os meios a serem destacados com a evolução do conhecimento da IA e a entrada dela no restauro de filmes são as principais fontes de pesquisa para uma restauro bem feito e coeso: como os métodos de super-resolução de imagem, redução de ruído, restauração de cores, reconstrução de imagens faltantes, restauro de áudio, estabilização de vídeo e reconstrução de quadros perdidos.

Levando em consideração o parágrafo anterior, é impossível deixar de destacar a discussão sobre essa prática de restauro nos anos de 1980, de onde surgiram uma enorme quantidade de comentários sobre a questão ética na prática de duplicação dos filmes, principalmente em relação ao cinema mudo. Ela não foi bem vista nesta época, como foi destacado em um dos primeiros artigos sobre o tema, *Silent Films for Contemporary Audiences*, de John. B. Kuiper (1982).

Kuiper (1982), em seu trabalho, destaca várias questões importantes na exibição de filmes mudos da George Eastman House, um museu internacional de fotografia e cinema. Em seu artigo, ele deixa explícita a preocupação com a imprecisão histórica que o restauro pode trazer. Criar mais uma cópia de um filme clássico pode assim distrair a narrativa pretendida do filme original, tais como a redução de uma emoção entregue a versão original do filme, principalmente do cinema mudo.

A relação entre Inteligência Artificial (IA) e cinema permanece em uma linha tênue, adicionando problemas em relação ao seu uso dentro dos filmes. A utilização da IA no mundo vem crescendo cada vez mais e o seu uso pode ser necessário para melhoria dos filmes clássicos, então, durante o projeto, a melhoria de qualidade de imagem e ruído será mais aprofundada.

A discussão entre o restauro e o não restauro dos filmes antigos nunca vai deixar de existir. Porém, a proposta deste trabalho é destacar os benefícios da Inteligência Artificial nas aplicações de restauração de imagens. A utilização da IA de forma respeitosa com o valor histórico e cultural do cinema no mundo é o primeiro passo para a restauração de filmes.

Em destaque a essa problemática, deve se ter um preparo para conciliar, estudar e avaliar novas técnicas de restauro que vão surgindo com o tempo. Este trabalho busca compreender uma dessas técnicas, a redução de ruído DnCNN *Denoising Convolutional Neural Network*. A aplicação dessa técnica serve de início para um estudo mais aprofundado sobre restauração de filmes clássicos. Para isso, é necessário coletar os dados, treinar e testar.

1.1 Objetivos

1.1.1 Objetivo geral

Realizar uma aplicação para redução de ruídos em cenas de filmes clássicos utilizando Redes Convolucionais DnCNN (Denoising Convolutional Neural Network).

1.1.2 Objetivos Específicos

- Realizar pesquisa bibliográfica sobre o uso da Inteligência Artificial no restauro de filmes;
- Entender, discutir e aplicar os conceitos de Redes Neurais Convolucionais (CNNs);
- Coletar e selecionar quadros de filmes a que possam ser aplicado a técnica DnCNN com Aprendizado de Máquina;

1.1 Problemática

É importante focar na relevância histórica e cultural do cinema para as pessoas e o avanço significativo da IA. Devemos entender que o uso da IA é

necessário para a melhoria de filmes que marcaram uma geração, como *Psicose* (1960), de Alfred Hitchcock, ...*E o vento levou* (1939), de Victor Fleming e *A lista de schindler* (1993), de Steven Spielberg.

Esses filmes foram importantes para a construção do cinema e ainda influenciam as mídias nos dias atuais, sendo um meio de aprendizado para muitas pessoas, deixando assim um legado para a indústria que não pode ser perdido. As técnicas de restauro têm, em sua base, uma duplicação de um filme, melhorando a versão original para que ela seja vista por muitas pessoas no futuro.

O trabalho de Mayer (2001) promove uma aprendizagem multimídia, que é útil para começar uma teoria que se baseia em pesquisas sobre como as pessoas aprendem com as imagens. Com essa aprendizagem, é importante destacar que os filmes podem ajudar as pessoas a compreender o mundo em que vivemos.

Entendendo melhor os problemas envolvendo a perda de filmes no Brasil, em 13 de fevereiro de 1996, o *Jornal do Brasil* fez uma matéria intitulada *O diário da tragédia* relatou enchente de grandes proporções, provocada pela tempestade torrencial que caiu sobre a cidade do Rio de Janeiro, e um dos bairros mais afetados, Jacarepaguá, onde existiam os estúdios da Cinédia. A água acabou entrando nos depósitos de filmes e atingindo o acervo, assim fazendo perder mil rolos de película.

Neste contexto, vale destacar que os originais se perderam em momentos da história e o dever que temos de preservar o que ainda existe. A duplicação ou restauro dos arquivos existentes originais de vários filmes, é necessária para a preservação desses ao longo dos anos e a constante distribuição para as novas gerações.

Para Paolo Cherchi Usai (2000), a restauração audiovisual é o conjunto de procedimentos técnicos que compensam a perda e degradação da imagem em movimento, trazendo assim de volta ao estado mais próximo possível do arquivo original. Portanto, a IA vem como uma ideia a acrescentar nos cuidados com o restauro feitos anteriormente à mão, sem tanto auxílio da tecnologia que temos hoje.

No momento em que há avanços tecnológicos, a IA surge como alternativa de identificar e resolver problemas, atribuindo métodos de restauro em vários trabalhos, como melhorando a resolução de imagens, adição de *pixels* faltantes ou iluminação em imagens muito escuras, problemas como a melhoria de imagem que antes era feita manualmente, agora é mecanizado.

1.2 Justificativa

Esta é uma proposta de conscientização do público sobre a preservação da cultura cinematográfica com o uso da Inteligência Artificial. É sobre salvar o trabalho produzido por mentes brilhantes que influenciaram não apenas o cinema, mas o mundo fora da tela, respeitando sempre a versão original, sem mudar muito o que foi feito.

O uso das técnicas de Inteligência Artificial no cuidado e tratamento dos filmes clássicos, adicionando o cinema mudo, deve ser implementado de forma que transpareça a ideia central do projeto. Também serão apresentados os problemas da falta de cuidado com as obras cinematográficas e como a IA pode ajudar a manter o cinema preservado e adaptado às novas tecnologias. Será necessário o levantamento de métodos com ênfase na IA e técnicas especializadas para desenvolvimento do projeto.

A ideia de preservar o patrimônio cinematográfico também está ligada ao peso emocional que os filmes colocam na vida das pessoas. Aqueles que acompanharam o momento de ouro do cinema ainda querem reviver aquele mesmo sentimento da época.

Com isso, este estudo tem por finalidade a melhoria de um filme importante para a cultura brasileira, para que haja entendimento da importância do restauro e a conservação de filmes que poderiam ou não serem perdidos com o tempo. A intenção se baseia na preservação e progressão temporal de um filme que poderia ser esquecido e nunca mais assistido no futuro.

2 REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta os conceitos teóricos fundamentais para a compreensão dos métodos e técnicas aplicados neste trabalho.

2.1 Processamento Digital de Sinais

Em termos de eficiência, a IA vem contribuindo significativamente para todas as áreas em que ela está contida. Na melhoria de ruído em vídeos, vale destacar o modelo descrito pela equipe de Kai Zhang (2017), que é capaz de melhorar e remover ruído gaussiano (aleatório) com nível de ruído desconhecido (ou seja, remoção de ruído gaussiano cego).

Entende-se como parte importante da revisão deste trabalho, o uso das redes neurais, que são modelos da computação inspirados no funcionamento do cérebro humano. É uma coleção de unidades que se conectam, neurônios artificiais. Eles trabalham de forma organizada para resolver problemas complexos de aprendizado de máquina.

Segundo Goodfellow, Bengio e Courville (2016, p. 29):

Embora os tipos de redes neurais usadas para aprendizado de máquina às vezes tenham sido usadas para entender a função cerebral (Hinton e Shallice, 1991), elas geralmente não são projetadas para serem modelos realistas da função biológica. A perspectiva neural sobre aprendizado profundo é motivada por duas ideias principais. Uma ideia é que o cérebro fornece uma prova, por exemplo, de que o comportamento inteligente é possível, e um caminho conceitualmente direto para construir inteligência é engenharia reversa nos princípios computacionais por trás do cérebro e duplicar sua funcionalidade.

É basicamente o entendimento do cérebro humano e os princípios da inteligência humana para a aprendizagem de máquina. Com intuito de copiar atitudes humanas para as ações de máquina, fazendo assim com que elas aprendam com as ações humanas.

Para Steven W. Smith (1997), o Processamento Digital de Sinais (PDS) se difere das outras áreas da Ciência da Computação, pois o tipo único de dados que ele utiliza são os sinais, que, na maioria dos casos, servem como dados sensoriais do mundo real, como as vibrações sísmicas, imagens visuais, ondas sonoras, etc.

Para contextualizar Steven (1997), no geral, o PDS envolve matemática, algoritmos e técnicas para manipular os sinais e serem convertidos em formato digital; incluindo diversos tipos de objetivos, como fazer o aprimoramento de

imagens visuais, reconhecimento e geração de fala, como também a compressão de dados para armazenar e transmitir.

O primeiro passo na restauração de um filme é entender as subáreas que envolvem todo o contexto de restauro. O Processamento Digital de Sinais é uma área complexa que envolve o processamento de áudio, processamento de imagem e ecolocalização (um eco que fornece informações sobre distância e tamanho).

De acordo com Steven W. Smith (1999, p.15), no *The Scientist and Engineer's Guide to Digital Signal Processing*:

O Processamento Digital de Sinais se distingue de outras áreas da ciência da computação pelo tipo único de dados que utiliza: sinais. Na maioria dos casos, esses sinais são dados sensoriais do mundo real, como vibrações sísmicas, imagens visuais, ondas sonoras, etc. O PDS envolve as matemáticas, os algoritmos e as técnicas usadas para manipular esses sinais após serem convertidos em formato digital. Isso inclui uma variedade de objetivos, como aprimoramento de imagens visuais, reconhecimento e geração de fala, compressão de dados para armazenamento e transmissão, entre outros.

O objetivo de entender sobre processamento digital de sinais, neste trabalho, é para compreender como ele age na restauração de uma imagem, contribuindo na remoção de ruídos, correção de cores, restauração de quadros danificados, estabilização de imagem, e melhorando detalhes.

2.2 Processamento Digital de Imagens

2.2.1 Imagens analógicas

No passado, o cinema se baseava muito em tecnologia analógica (películas analógicas). Em questão de fotografia nos filmes, as câmeras capturavam imagens em rolos de películas analógicas que, diferentemente das imagens digitais, as analógicas não necessitavam de meios tecnológicos mais avançados, eram mais físicas, já que os rolos eram processados em laboratórios utilizando processos químicos.

Segundo Mulvey (2006, p.18):

A especificidade do cinema, a relação entre sua base material e sua poética, se dissolve enquanto outras relações, intertextuais e intermédias, começam a desaparecer e surgir. Além disso, o digital, como uma informação abstrata sistema, rompeu com as imagens analógicas, finalmente varrendo a relação com a realidade, que tinha, em geral, dominado a tradição fotográfica.

As imagens analógicas eram usadas em câmeras fotográficas ou de vídeo antigas, a partir da luz. A câmera se baseava na luz e armazenava em tubos de captura, que são dispositivos que convertem a luz refletida em cenas com sinais elétricos, logo em seguida, representam a imagem capturada, podendo variar em diversos tipos de tubo.

Existem também os dispositivos de carga acoplada (CCD), que são sensores usados em câmeras modernas, tanto analógicas como digitais, para capturar imagens, convertendo também luz em cargas elétricas, que, logo após, são convertidas em sinais digitais, como nas câmeras digitais ou em sinais analógicos.

2.2.2 Imagens digitais

Com o passar dos anos, as imagens digitais tomaram conta do cinema com alguns benefícios, como, por exemplo, a qualidade dos filmes. A alta resolução e detalhes, a consistência na edição dos filmes depois de serem gravados, custos foram altamente reduzidos mesmo com equipamentos caros. Ao passar dos anos, essa tecnologia se tornou mais barata.

Uma imagem pode ser definida como uma função bidimensional, $f(x,y)$, onde x e y são coordenadas espaciais e a amplitude de f em qualquer par de coordenadas é chamada de intensidade ou nível de cinza da imagem nesse ponto. Quando x , y e os valores de intensidade de f são quantidades finitas e discretas, a imagem é digital. O processamento digital de imagens refere-se ao tratamento de imagens digitais por meio de um computador. Uma imagem digital é composta por um número finito de elementos, chamados pixels. (Gonzalez; Woods, 2018, p.18).

Os sensores das câmeras são compostos por milhões de fotosensores individuais (pixels) que são sensíveis à luz, convertendo a luz em sinais elétricos. O pixel é a menor unidade de uma imagem digital, cada um contém uma cor específica e quando estão juntos formam uma imagem completa.

Para identificar cada pixel é atribuído um valor numérico em cada amostra de imagem e armazenado em um formato digital, esses formatos podem variar em *Bitmap* (BMP), *JPG*, *PNG*, *GIF*. Eles podem ser manipulados facilmente e editados, armazenados e reproduzidos em vários dispositivos modernos, diferente dos analógicos que necessitam de uma complexidade na sua estrutura física para reprodução.

Contextualizando, as imagens digitais são as usadas nos dias de hoje por serem de fácil manipulação. As tecnologias atuais ajudam na compreensão dos erros que podem ocorrer com imagens distorcidas ou com falta de pixels. A partir da conversão de imagens analógicas para digitais, podemos melhorar o que foi danificado no passado.

2.2.3 Processo de tratamento e restauração de imagens

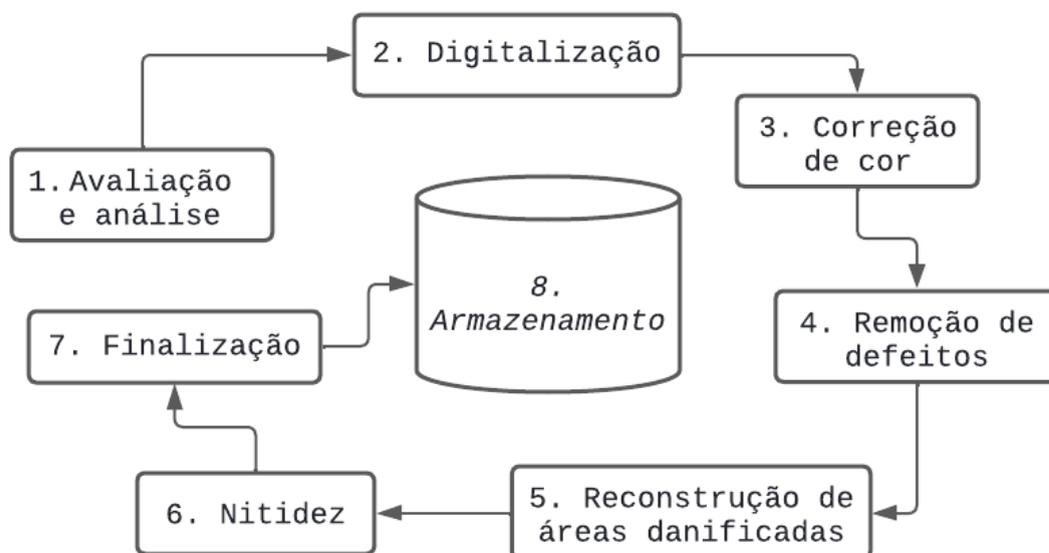
Para Rafael C. Gonzalez e Richard E. Woods (2018), as principais fontes de ruído em imagens digitais surgem durante a aquisição ou transmissão de imagens. Durante essa aquisição da imagem, os fatores mais comuns que podem afetá-la são o nível de luz disponível, a temperatura do sensor das imagens e a quantidade de elementos sensores.

O desempenho dos sensores de imagem é afetado por uma variedade de fatores ambientais durante a aquisição da imagem e pela qualidade dos próprios elementos sensores. Por exemplo, ao adquirir imagens com uma câmera CCD, os níveis de luz e a temperatura do sensor são fatores importantes que afetam a quantidade de ruído na imagem resultante. As imagens são corrompidas durante a transmissão principalmente por interferências no canal de transmissão. Por exemplo, uma imagem transmitida usando uma rede sem fio pode ser corrompida por relâmpagos ou outras perturbações atmosféricas. (Gonzalez; Woods, 2018, p. 318).

O processo de tratamento segue várias regras para chegar em uma restauração bem feita, mas, para entender melhor, é necessário compreender toda a complexidade do momento em que a foto foi tirada. O ambiente e a iluminação são fatores essenciais para uma imagem sem ruído ou sem impacto significativo no filme. A estrutura da câmera é um dos fatores que não deve ser esquecido.

O processo de tratamento e restauração de imagens se baseia na avaliação e análise da condição que a imagem está, partindo para a digitalização, correção de cor, remoção de defeitos, reconstrução de áreas danificadas, nitidez e detalhe, finalização e armazenamento. A Figura 1 ilustra o fluxograma do processo de tratamento e restauração de uma imagem.

Figura 1: Processo de tratamento e restauração de imagens



Fonte: Elaborado pelo autor

Como ilustrado acima, esses passos devem ser seguidos para um processamento inteligente no tratamento e restauro das imagens:

- **Avaliação e análise:** É o passo da identificação dos danos e imperfeições da imagem. Deste modo você consegue visualizar quais são as falhas;
- **Digitalização:** Aqui é o momento de criar uma versão digital da imagem original, que pode ser trabalhada depois em algum *software* de edição;
- **Correção de cor:** nesta etapa, ocorre a melhoria das cores, em que deve se aproximar da versão original;
- **Remoção de defeitos:** Aqui começa a remoção de arranhões, manchas, poeira e rachaduras;
- **Reconstrução da áreas danificadas:** Neste momento, são usadas as técnicas de clonagem, preenchimento sensível ao conteúdo para restaurar imagens com partes faltantes ou muito danificadas;
- **Nitidez:** A adição de filtros de nitidez são necessários para contornar e detalhar;
- **Finalização:** Após todos esses processos, têm que verificar se tudo está como foi planejado;
- **Armazenamento:** E, por último, no armazenamento, é guardada a imagem restaurada em diversos formatos de alta qualidade que não perdem os dados.

2.2.4 Extração de características/dados

O início da extração é um passo importante em várias aplicações do processamento de imagens. Nesse momento, a análise de elementos específicos dentro da imagem é feita, assim, conseguindo compreender quais são os elementos mais visíveis. Após a extração, é possível utilizar técnicas como reconhecimento facial, que usam algoritmos de redes neurais convolucionais (CNNs) para processamento de dados estruturados, como imagens.

Ademais, as técnicas adicionais se baseiam também na identificação de formas e texturas, com análise de contornos utilizando redes neurais profundas (que aprendem dados complexos). A partir disso, a análise de padrões de movimento é feita utilizando redes neurais recorrentes (RNNs) para dados sequenciais em imagens.

A extração de características começa com a digitalização do filme para converter as imagens analógicas para digitais utilizando *scanners* de alta resolução. Logo em seguida, são transformadas em formato digital. Depois é feita a limpeza e correção de defeitos, passando o filme para um processo de correção de alguns defeitos. Em seguida, restaurando a cor e a nitidez.

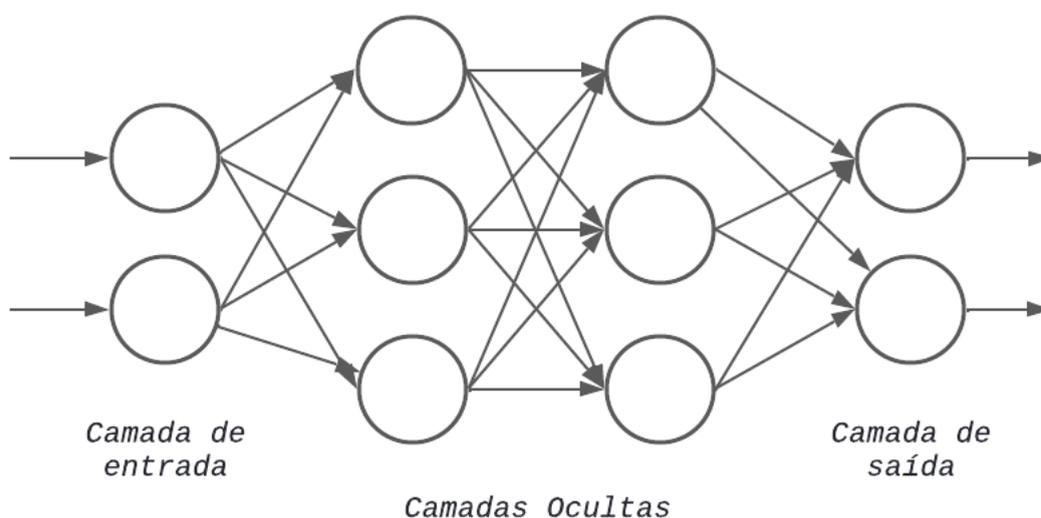
Para o entendimento das redes neurais nesses processos deve-se entender como elas se comportam e o que são. Para melhor compreensão, as redes neurais são modelos computacionais parecidos com o cérebro humano, que foram projetadas para reconhecer padrões e realizar tarefas complexas de aprendizado de máquina. Vale ressaltar que o termo pode ser bastante reproduzido de diferentes formas por autores distintos.

Uma rede neural é um conjunto interconectado de elementos de processamento simples, unidades ou nós, cuja funcionalidade é vagamente baseada no neurônio animal. A capacidade de processamento da rede é armazenada nas forças de conexão entre as unidades, ou pesos, obtidos por um processo de adaptação ou aprendizado a partir de um conjunto de padrões de treinamento. (Gurney, 1997, p.13).

Para Gurney (1997), os equivalentes artificiais dos neurônios biológicos são os nós ou unidades, como definido na Figura 2. As sinapses vão ser modeladas por um único número ou peso, em que cada entrada seja multiplicada por um peso antes de ser enviada para a sua equivalente. Assim, eles são somados por adição aritmética para fornecer um nó de ativação. A ativação é então comparada com um

limite; se a ativação excede o limite, a unidade produz uma saída de alto valor, caso contrário, produzirá zero.

Figura 2: Representação das redes neurais



Fonte: Elaborado pelo autor

- **Camada de entrada (*Input Layer*):** aqui são recebidos os dados brutos que vão ser processados pela rede neural. Cada neurônio, nesta camada, é uma característica;
- **Camadas Ocultas (*Hidden Layers*):** neste momento, executam-se cálculos intermediários para extrair características complexas dos dados da camada de entrada;
- **Camada de saída:** É a saída final da rede neural, que pode ser uma classe, uma probabilidade ou um valor contínuo, variando para cada problema.

2.3 Inteligência Artificial

O campo da inteligência artificial vem avançando como uma ciência desde o século XX. Na computação, a IA ajuda no desenvolvimento de sistemas e máquinas que realizam tarefas que não necessitam da inteligência humana, já que a última é replicada pela máquina.

Em seu famoso ensaio inicial, Turing (1950) considera a ideia de realmente programar suas máquinas inteligentes à mão. Ele estima quanto trabalho isso poderia exigir e conclui que “algum método mais eficiente parece desejável”. O método que ele propõe é construir máquinas com aprendizagem e depois ensiná-las. Em muitas áreas de IA, esse é agora o método preferido para se criar sistemas do estado da arte. O aprendizado tem outra vantagem, como observamos antes: ele permite ao agente operar em ambientes inicialmente desconhecidos e se tornar mais competente do que seu conhecimento inicial sozinho poderia permitir. (Russell; Norvig, 2013, p. 18).

O estudo de Turing serviu para o humano e uma máquina se comunicarem através de um terminal, de modo que um terceiro não sabe quem está falando, homem ou máquina. Nesse momento, é possível entender como o procedimento funciona e que características foram transferidas para a máquina com o tempo.

Hoje, a IA vem tendo um avanço significativo desde 1950, considerando todos os equipamentos eletrônicos, os processamentos computacionais, os veículos autônomos, todos os sistemas de recomendação, avanços na saúde e medicina, indústria, jogos e educação.

Por exemplo:

- **Equipamentos eletrônicos:** os celulares podem utilizar IA para reconhecimento facial e alguns equipamentos vêm com assistentes virtuais integradas, como *echo dot* da *Amazon*;
- **Processamentos computacionais:** plataformas como a *AWS* da *Amazon* e o *Google Cloud* oferecem serviços de IA, como reconhecimento de imagem;
- **Veículos Autônomos:** Alguns carros da *Tesla Autopilot*.
- **Sistemas de recomendação:** serviços de *streaming* usam da IA em recomendações por conta individual para entender como cada usuário se comporta e assiste;
- **Avanços na saúde e medicina:** a IA é usada nessa área em imagens médicas para detectar condições como fraturas, diagnósticos médicos e outros;
- **Indústria:** robôs industriais foram feitos para ajudar as pessoas a produzir melhor e com eficiência, como robôs da *KUKA* e *ABB*, que são usados para montagem;
- **Jogos:** nos jogos, a IA serve para criar NPCs, que são personagens fictícios para acrescentar na história, geração de mundo e outros;

- **Educação:** na educação, alguns modelos são usados para personalizar o conteúdo de aprendizado para cada estudante.

2.3.1 Técnicas de Inteligência Artificial aplicadas à restauração de filmes

Na literatura, é possível encontrar diversas técnicas de inteligência artificial aplicadas à restauração de filmes, dentre elas destaca-se: *Denoising Autoencoder* (DAE), *Inpainting*, *Super-Resolution* (SR), *Frame Interpolation*. Essas técnicas serão descritas a seguir.

Segundo Goodfellow, Bengio e Courville (2016, p. 510):

O autoencoder de remoção de ruído (DAE) é um autoencoder que recebe um ponto de dados corrompido como entrada e é treinado para prever o ponto de dados original e não corrompido como saída.

O *Denoising Autoencoder* (DAE) é usado como base para entendimento de outras técnicas, já que pode ser uma parte crucial da restauração e melhoria de filmes. O significado das DAEs se baseia nas variantes dos *autoencoders*, que são projetados para aprender uma grande quantidade de dados (Vincent *et. al*, 2010).

Em consequência ao avanço de técnicas para melhoria de imagem utilizando IA, surgiu a técnica de *Inpainting*. Para Bertalmio *et al*, (2000), essa técnica está diretamente ligada a modificar uma imagem de forma indetectável, e é tão antiga quanto a própria arte. As aplicações de pintura interna são muitas, iniciando na restauração de pinturas danificadas e fotografias para a remoção ou substituição de objetos selecionados.

O *Inpainting* consegue preencher partes faltantes ou danificadas de uma imagem, em que a restauração seja a mais próxima possível do natural, conseguindo assim não mudar drasticamente toda a fonte do arquivo e permanecer fiel à base. Caso a imagem tenha manchas ou algo danificado, o *inpainting* pode preencher essas áreas faltantes.

Em contrapartida, divergindo os algoritmos citados até o momento, as técnicas para melhoria de imagem avançam no momento em que surge a ideia de aumentar a resolução delas. Essa prática vem avançando no meio acadêmico em diversas ideias de autores que tentam melhorar cada vez mais uma imagem.

Para Lim *et al*, (2017), em muitos algoritmos de super-resolução baseados em aprendizado profundo, a imagem de entrada é ampliada via interpolação bicúbica (método que calcula os valores de novos pixels em uma imagem) antes de

ser alimentada na rede, em vez de simplesmente usar a imagem interpolada como entrada. Assim, pode-se também treinar módulos de ampliação no final da rede. Após isso, é possível reduzir muito os cálculos sem perder a capacidade do modelo, pois o tamanho das características diminui.

No entanto, existem desvantagens nessa abordagem. Alguns problemas não podem ser resolvidos com múltiplas escalas em um único *framework* como no *Very Deep Super-Resolution* (VDSR), que é uma técnica de super-resolução de imagem com utilização de redes neurais profundas para aumentar a resolução.

O *frame interpolation* é outra técnica usada para criar quadros intermediários em uma base sequencial de vídeo. É um processo importante em muitas aplicações, como a retemporização de vídeo.

A interpolação de quadros é outra aplicação amplamente utilizada da estimativa de movimento, frequentemente implementada no mesmo circuito que o hardware de desentrelaçamento necessário para ajustar um vídeo recebido à taxa de atualização atual do monitor. Assim como no desentrelaçamento, informações dos novos quadros intermediários precisam ser interpoladas a partir dos quadros anteriores e subsequentes. (Szeliski, 2022, p.418-419).

Ele funciona de uma forma em que estima o fluxo óptico, calculando os quadros e prevendo os intermediários que simulam uma transição suave entre os quadros originais de um vídeo.

2.3.2 Algoritmos para restauração de vídeo

Existem alguns algoritmos que serão abordados para mais aprofundamento durante este trabalho, tais como *Deep Convolutional Neural Networks* (CNNs), que são redes neurais convolucionais profundas, especializadas no processamento de dados com uma grade estruturada como imagens, usando convoluções (operações matemáticas que extraem características de dados) para pegar características importantes.

O *Generative Adversarial Networks* (GANs), são duas redes neurais, uma que gera, criadora de dados falsos, e a outra discriminadora, que tenta distinguir esses dados dos reais, até que, com o seu treinamento, ele possua dados muito realistas.

Contribuindo para mais algoritmos que serão usados, há o *Enhanced Super-Resolution Generative Adversarial Networks* (ESRGAN), que é uma variante

do GANs, que melhora a resolução das imagens usando uma rede que gera imagens de alta resolução, aumentando a nitidez e os detalhes.

O *Single Image Super-Resolution* (SISR) serve para aumentar a resolução de uma imagem única de baixa resolução para alta resolução. Para entender melhor suas funcionalidades, ele utiliza técnicas de aprendizado profundo para adicionar detalhes finos que não existiam anteriormente nas imagens.

Como adição, temos o *Temporal Convolutional Networks* (TCNs), que são redes convolucionais projetadas para lidar com dados sequenciais, como vídeos, preenchendo lugares sem informações nos vídeos. E o *Super-Resolution Convolutional Neural Network* (SRCNN), também sendo um tipo de rede neural que realiza super-resolução de imagens e reconstruindo detalhes faltantes.

2.4 Denoising Convolutional Neural Network (DnCNN)

Nesta seção, é abordada a técnica de DnCNN e a importância na remoção de ruídos em imagens ruidosas. Na visão computacional, a restauração de cenas é um passo importante para ter um filme completamente restaurado. Esse modelo foi escolhido por ser uma forma eficaz de compreender a importância da Inteligência Artificial e Redes Neurais Convolucionais Profundas (CNNs) no restauro de um filme. Para o modelo descrito, é importante entender que ele considera o ruído nas imagens como algo discriminativo, em uma imagem ruidosa, é separado o ruído da imagem. Antes desse modelo, seguia-se uma prática comum de modelar explicitamente o prior de imagem.

Devido ao grande espaço de conteúdo natural da imagem, informações prévias sobre as estruturas da imagem são cruciais para regularizar o espaço de solução e produzir uma boa estimativa da imagem latente. A modelagem e o aprendizado prévios da imagem são, então, questões-chave na pesquisa de RI. (Zhang et al., p.1).

Para Zhang *et al.* o prior da imagem é um método tradicional que funciona como um conjunto de suposições prévias sobre as propriedades de uma imagem que guia na reconstrução dela. Já no DnCNN, o modelo aprende a remover ruídos diretamente com dados de treinamento. Assim, o prior é aprendido e não imposto no modelo.

Vale ressaltar que o DnCNN funciona para diversos tipos de imagem, desde sistemas de segurança até um patrimônio afetivo, como as fotos em álbuns de família. São várias as vantagens na restauração de uma imagem. Para muitos, uma

imagem serve como apego emocional e, para outros, como patrimônio histórico. Cada indivíduo compreende uma imagem de formas diferentes. As imagens conseguem transportar uma pessoa para diversos lugares, conseguindo armazenar diferentes emoções. Elas podem não só documentar realidades históricas, mas também podem criar conexões emocionais entre a pessoa e o evento que a imagem representa, segundo Markham (2019).

O sentido de coexistência entre passado e presente está explícita no trabalho de Batchen (2004), onde ele exemplifica a ligação que as pessoas possuem com o próprio passado, as fotografias são ditas como artefato emocional. O sentimento de luto é um dos focos em seu projeto, a ideia das fotografias produzirem sentimento de saudade e luto. A imagem segue como um dos meios principais de retratar o luto e eternizar a lembrança.

No artigo intitulado *Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising* proposto pela equipe de Kai Zhang (2017), é resumido em três diferentes contribuições: propor uma CNN profunda treinável de ponta a ponta para remoção de ruídos gaussianos; descobrir que a aprendizagem residual e a normalização em lote pode beneficiar muito a aprendizagem de CNN, e usar o DnCNN para lidar com tarefas de remoção de ruídos em imagem de forma fácil. Para eles, o modelo não segue apenas para remoção de ruídos, mas também para a outras tarefas de processamento de imagem, como super-resolução de imagem única (SISR) e desbloqueio de JPEG. Sendo possível, pois, o modelo consegue lidar com diferentes tipos de degradação de imagem, fazendo com que se torne uma forma versátil para diversas tarefas de remoção de ruídos.

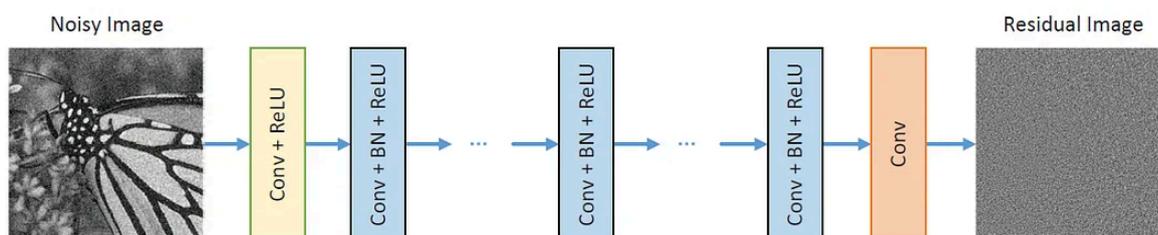
Os autores do DnCNN, desenvolveram uma CNN para ser treinada de forma *end-to-end* (realizado de forma direta) para remoção de ruídos gaussianos de imagens. Diferente de outros métodos de redes neurais, o algoritmo utiliza uma aprendizagem residual que prevê a diferença entre uma imagem ruidosa e uma imagem limpa.

No estudo de Zhang *et al*, o ruído gaussiano é chamado de ruído de amplificadores, conhecido como ruído estatístico que possui uma função de densidade de probabilidade (PDF), compatível com a distribuição gaussiana. Para melhor compreensão, o ruído gaussiano serve para amplificar o ruído em imagens para melhor aprendizagem.

2.4.1 Arquitetura do DnCNN

No trabalho da equipe de Kai Zhang (2017), é definida a arquitetura do DnCNN como mostrado na Figura 3.

Figura 3: A arquitetura da rede DnCNN proposta.



Fonte: Traduzido de Kai Zhang *et al.*, (2017).

Desse modo, deve-se entender da seguinte forma: após a imagem de entrada, existem três tipos de camadas na figura, (i) Conv+ReLU para a primeira camada, (ii) Conv+BN+ReLU, e então a normalização do lote [21] é adicionada entre a convolução e a ReLU. E a última camada (iii) conv, com filtros de tamanho 3 x 3 x 64 são usados para reconstruir a saída. O ReLU (*Rectified Linear Unit*) é uma função de ativação que não adiciona linearidade na rede, é altamente utilizado na aprendizagem de padrões mais complexos, ele irá anular valores negativos, assim mantendo apenas valores positivos ($f(x) = \max(0,x)$). Já o BN (*Batch Normalization*) é uma famosa técnica que normaliza a saída da camada convolucional que melhora, estabiliza e acelera o treinamento. É feita antes de passar por ReLU. Nessa camada se usa nas camadas 2 a (D-1) em que D é a profundidade total da rede, com 64 filtros de 3x3.

O tamanho dos filtros convolucionais é definido como 3x3 e as camadas de *pooling* serão removidas, fazendo assim o campo receptivo de DnCNN com profundidade de d sendo (2 d+1) (2 d+1).

A partir do contexto exposto pela equipe de Kai Zhang (2017), exemplificando melhor as camadas:

- (i) Conv+ReLU: é usado na primeira camada, com 64 filtros de tamanho 3 x 3 x c geram 64 mapas de características. c = 1 para imagem cinza e c = 3 para imagem colorida.

- (ii) Conv+BN+ReLU: são para as camadas 2 a (D-1), são usados 64 filtros de tamanho $3 \times 3 \times 64$, e a normalização em lote será adicionada entre a convolução e o ReLU.
- (iii) Conv: e, na última camada, os filtros x de tamanho $3 \times 3 \times 64$ são usados para reconstruir a saída.

Após a ativação do ReLU, o DnCNN consegue separar a estrutura da imagem de observação ruidosa por meio de camadas ocultas, e, em seguida, é treinado ponta a ponta.

2.4.2 PSNR e SSIM

Para entender os passos de restauro usando o DnCNN, há a necessidade de conhecer os métodos de *Peak Signal-to-Noise Ratio* (PSNR) e *Structural Similarity Index Measure* (SSIM), onde tem um método de comparação entre a imagem base e a imagem após a remoção de ruído.

Segundo Adrian Dziembowski, Dawid, Jakub e Adam Grzelka (2021, p.3):

O maior grupo de métodos de avaliação de qualidade de imagem disponíveis é baseado no cálculo do PSNR. Esses métodos têm uma grande vantagem sobre outras métricas, pois são amplamente utilizados em uma ampla gama de pesquisas relacionadas à imagem (por exemplo, compressão), tornando-os muito intuitivos para os pesquisadores. Alguns métodos baseados em PSNR foram aprimorados para serem mais úteis em vídeos imersivos, como WS-PSNR [13], CPP-PSNR [14] e OV-PSNR [15]. Esses métodos levam em conta a possibilidade de que as imagens testadas sejam omnidirecionais, permitindo uma melhor comparação direta de vídeos ERP em codificação de mídia imersiva. No entanto, tal comparação não é muito prática, pois o vídeo final gerado para o espectador após a decodificação é um vídeo regular em perspectiva.

Quanto maior o valor de PSNR, melhor é a qualidade da imagem restaurada, significando que ela está mais próxima da imagem original. O valor PSNR é medido em decibéis (dB).

Para Adrian Dziembowski, Dawid, Jakub e Adam Grzelka (2021), outro grupo de métrica importante é o SSIM, onde foca na extração de informações estruturais das imagens comparadas, seguindo a hipótese de que o sistema visual humano é altamente adaptado a mudanças neste domínio. Esses métodos proporcionam um desempenho relativamente bom em comparação ao PSNR.

O SSIM avalia a qualidade da imagem com base em fatores como luminância, contraste e estrutura, tentando se alinhar melhor com a percepção humana do que o

PSNR. O valor do SSIM varia entre -1 e 1, onde 1 indica que as duas imagens são idênticas em termos de estrutura. Valores mais próximos de 1 indicam maior similaridade.

Segundo Adrian Dziembowski, Dawid, Jakub, Adam Grzelka (2021, p.3):

A utilidade desses métodos é limitada na avaliação de vídeos sintetizados. Em tais vídeos, pequenos deslocamentos podem ser introduzidos pelo processo de síntese. Embora tais pequenas distorções não sejam significativas para o espectador do vídeo, elas ainda podem influenciar fortemente a pontuação do SSIM (assim como outros métodos baseados na comparação de estrutura).

O PSNR e o SSIM são muito importantes na comparação de métricas do modelo após ser testado com as imagens de entrada. A necessidade de comparação é entender o tratamento feito em uma imagem e se o modelo está sendo eficiente.

2.5 Trabalhos relacionados

Este trabalho tem como base, pesquisas a partir de revisões bibliográficas de artigos, livros, monografias, dissertações, *sites*, jornais e algumas plataformas digitais confiáveis, assim, ajudando a entender todo o contexto sociocultural e tecnológico relacionado à restauração de filmes nos dias de hoje.

A pesquisa está baseada primeiro na parte ética da melhoria de um filme, onde explica como respeitar sem modificar tanto o conteúdo original de um projeto estruturado da forma única dele. Em seguida, avaliar como essa prática pode afetar a sociedade. Logo após, é necessária a busca de um filme que precise de restauração e coletar as informações necessárias para buscar melhorias futuras para a película. O filme escolhido será com base no critério de relevância histórica, estado de conservação e disponibilidade.

As redes neurais artificiais entram em foco neste trabalho, ajudando a compreender como o *denoising* (remoção de ruído) funciona em imagens danificadas. Nesse sentido, o artigo e o código fonte disponível no [github](#) serão os pontos de estudo e aplicação principais para treinar e testar o modelo.

Para Wazlawick (2009), áreas que são mais amadurecidas que as anteriores exigem que qualquer abordagem seja comparada quantitativamente com outras literaturas. O trabalho da equipe de Kai Zhang (2017) aborda uma estratégia de remoção de ruído comparativa com diversas outras que fazem o mesmo, mas de maneiras diferentes.

Primeiramente, o necessário para a execução deste projeto será o artigo *Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising* feito pela equipe de Kai Zhang (2017), nesse trabalho está disponível a técnica proposta por eles de DnCNN, em que a remoção de ruído é uma das partes mais importantes na restauração de um filme, adicionando a inteligência artificial como foco do trabalho.

Esse trabalho visa o entendimento da construção completa de um filme restaurado, buscando entender como o DnCNN é um algoritmo eficaz para a remoção de ruídos, neste sentido o foco não está em comparar modelos de denoising, mas, sim, estudar os conceitos de restauro e a aplicação de um modelo específico. A estrutura do trabalho está no conhecimento de uma técnica que pode ajudar na restauração de um filme, mas se deve entender que existem outras técnicas que podem ajudar. A base estrutural de um projeto que visa procurar várias técnicas diferentes é bastante complexa, pois um filme deve ser distribuído em diferentes tipos de algoritmo, não apenas na remoção de ruído.

Tendo em vista que a restauração de um filme abrange diversas técnicas, no DnCNN, é visto um meio de compreender o funcionamento da remoção de ruído em um frame de um filme. Os meios para o restauro de um filme usando IA demanda muito tempo e diversas técnicas diferentes, para entendimento de uma delas, nesta seção irei abordar o DnCNN, um modelo CNN para remoção de ruídos.

O trabalho a partir de agora segue um estudo de algoritmo, primeiro é explicado o que é e como o DnCNN funciona. Logo após, o modelo é treinado e testado com alguns dados de entrada, é explicado as dificuldades durante a execução do treinamento e no funcionamento da *Unidade de Processamento Gráfico* (GPU) da máquina e como é feito para resolver os diversos problemas que surgem.

3 METODOLOGIA

Nessa seção, será necessário a coleta de dados, em relação a qual cena de filme utilizar, logo após, será detalhado o treinamento do algoritmo definido anteriormente, o DnCNN. Nesse caso, o artigo utilizado foi proposto pela equipe de Kai Zhang (2017) para aplicação. Após o treinamento será definido os parâmetros de teste e utilizadas as imagens do filme escolhido para, logo em seguida, definir a qualidade da imagem restaurada.

3.1 Treinamento do algoritmo

A utilização do DnCNN é significativa na restauração de imagens e vídeos. Até o momento, não foi especialmente configurado para restauração de um filme completo. Neste trabalho será explicado porque a inteligência artificial pode entrar em uma área delicada da história.

Nesse caso, o filme escolhido tem que estar danificado. Pensando nisso, foi decidida a utilização de uma versão que não foi modificada do filme *Nosferatu* (1922). A restauração completa desse filme estaria na utilização dos quadros disponíveis por segundo. Como exemplo de restauro, foram escolhidos cinco *frames* do filme que seria trabalhado.

Após a coleta de quadros, é a fase de testes e verificar se o algoritmo está funcionando corretamente. Para isso, as imagens vão servir de parâmetro com o antes e o depois dela, e, com isso, será possível visualizar alguma diferença.

Para treinamento foi necessário usar o algoritmo disponível no github da equipe de Kai Zhang (2017). Dentro do código está disponível o *Training Codes*, essa pasta contém o DnCNN Keras que será o modelo utilizado para treino e teste.

Para treinar o algoritmo, é necessário uma máquina com GPU. Foi importante encontrar uma máquina com GPU para reduzir o tempo que era feito com a Unidade Central de Processamento (CPU), o algoritmo demanda horas para treino. Principalmente com a CPU. Antes das configurações feitas no código, o treinamento estava sendo feito durante dias.

Foi escolhido o *TensorFlow*, que é uma biblioteca de código aberto, para aprendizado de máquina aplicável a uma grande variedade de tarefas. O *TensorFlow* é uma biblioteca que funciona bem, mas é necessário que tudo que for utilizado com ele seja compatível com a versão. Foi necessário adaptar o código para

compatibilidade com a versão *TensorFlow* 2.8.0, já que algumas funções como “tf.Session” e “tf.compat.v1” eram específicas para essa versão.

Para entendimento das mudanças, o código foi reestruturado para corrigir erros relacionados às APIs mais antigas, para que funcionasse corretamente. Foi utilizado um gerador de dados personalizado, que adiciona ruído às imagens de treinamento, com um parâmetro *sigma* (nível do ruído) controlado, que é utilizado para preparar os dados em tempo real durante o treinamento, evitando assim uma sobrecarga de memória.

Foi ajustada também a taxa de aprendizado e foi feito o tratamento de imagens com *OpenCV* que foi utilizada como biblioteca compatível com Python como também o *Pillow* (biblioteca de manipulação de imagens em Python), servindo como processamento de imagem usadas no treinamento, com base na mesma linguagem de programação.

Após muitas dificuldades na implementação de um modelo que utilizaria a GPU, atualizações no código original para versões atuais e o tempo para treinamento, por fim, o objetivo foi finalizado, e, nas próximas sessões, é explicado tudo que necessita para o código funcionar da maneira correta para a máquina específica.

Durante o uso da GPU, foi verificado com o gerenciador de tarefas do *Windows* e observado que a GPU estava sendo usada, mas não com sua força completa. O uso estava com uma taxa de porcentagem baixa, mas funcionando. Foram mudados os parâmetros do tamanho do lote (*Batch Size*), mas não modificou muito em relação ao original. Os tamanhos de lote foram variando entre 32 e 128, tentando encontrar um equilíbrio que permitisse um melhor uso da GPU sem causar problemas de memória ou sobrecarga do processamento.

3.1.1 Python

Python é uma linguagem de programação amplamente usada em aplicações *Web*, desenvolvimento de *software*, Ciência de Dados e *Machine Learning* (ML). Ele é uma das linguagens mais utilizadas atualmente por ser simples e muito útil. A versão que foi utilizada foi a 3.7.9, pois servia de compatibilidade com o *TensorFlow* 2.8.0.

Para evitar problemas de compatibilidade, essa versão do *Python* era a mais recomendada. Foram utilizadas outras versões, mas tiveram bastantes problemas na execução do código.

Foi criado também um ambiente virtual específico (*tf_env*) para isolar todas as dependências do projeto.

Algumas bibliotecas foram instaladas para que o código funcionasse corretamente, assim como a API do *TensorFlow*, com suas funções específicas:

Quadro 1: Bibliotecas e API

Bibliotecas e API	Descrição
<i>TensorFlow</i>	<i>Framework</i> principal para a construção e treinamento do modelo de aprendizado de máquina.
<i>NumPy</i>	Utilizado para operações matemáticas e manipulação de <i>arrays</i> numéricos, fundamental para o processamento de dados de entrada.
<i>Pillow</i> (PIL)	Biblioteca usada para abrir, manipular e salvar diferentes formatos de imagem. Nesse caso, foi utilizada para carregar imagens de treino no formato .jpg.
<i>OpenCV</i> (cv2)	É uma biblioteca usada para processar imagens e vídeos, muito comum em tarefas de visão computacional. Nesse caso, foi necessário carregar e processar imagens usadas no treinamento.
<i>TensorFlow Keras</i>	API utilizada para construir, compilar e treinar seu modelo DnCNN. O <i>Keras</i> oferece uma interface mais simples para construir redes neurais usando o <i>TensorFlow</i> como <i>backend</i> .
<i>Adam Optimizer</i>	Foi utilizada essa API para construir, compilar e treinar seu modelo DnCNN. O <i>Keras</i> oferece uma interface mais simples para construir redes neurais usando o <i>TensorFlow</i> como <i>backend</i> .
<i>Glob</i>	Para trabalhar com arquivos e buscar imagens no diretório de dados.
<i>Argparse</i>	Para passar parâmetros de configuração via linha de comando, como o caminho dos dados de treino, número de épocas, tamanho do lote, entre outros.

Fonte: Elaborado pelo autor

3.1.2 *TensorFlow* e *CUDA*

Foi utilizada a versão 2.8.0 do *TensorFlow* para utilização da Arquitetura Unificada de Dispositivos de Cálculo (*CUDA*), com versão 11.2 e Biblioteca de Redes Neurais Profundas *CUDA* (*cuDNN*), com a versão 8.1.

O *CUDA* da *NVIDIA* (2006), é uma plataforma de computação paralela e modelo de programação que ajuda os desenvolvedores a acelerarem suas aplicações aproveitando a estrutura da *GPU*.

3.1.3 Controladores de *GPU*

A *GPU* utilizada foi a *NVIDIA GeForce RTX 3060*. E, para garantir compatibilidade com o *TensorFlow* e o *CUDA*, é necessário atualizar todos os *drivers*. O *TensorFlow* precisa se comunicar diretamente com o *hardware* da *GPU*, e versões desatualizadas dos *drivers* causam falhas e baixo desempenho.

3.1.4 Configurações do treinamento

Primeiro, foi necessária a criação de uma máquina virtual (*tf_env*) para conter todas as dependências importantes para a execução do treino. Após a criação da máquina virtual, é necessário ajustar o código de treinamento e teste. O local será treinado a partir do arquivo do *main_train.py*, que está localizado dentro da pasta *dncnn_keras* dentro de outra pasta chamada *Training Codes*.

Será necessária a importação de bibliotecas para treinar o modelo, variando entre processadores de imagem e manipuladores de formatos específicos, como *jpg*. É utilizado também o *TensorFlow* como *API* para construir, compilar e treinar o modelo.

Em questão de parâmetros, temos, primeiramente, o tipo do modelo, que, nesse caso, é o *DnCNN* para remoção de ruídos. O tamanho do lote que funciona como controlador do número de imagens que o modelo processa a cada passo de treinamento, assim, diminuindo o tempo de treino, assim, tentando não sobrecarregar o algoritmo. É necessário também definir o caminho dos dados de treinamento. Nesse caso foi utilizado o padrão *Train400*, contendo um total de 400 imagens em alta resolução para treinar, sendo usado no algoritmo original.

O σ ¹ irá definir o nível de ruído na imagem. Quanto maior for o σ , mais forte será o ruído. Assim, a tarefa do *denoising* será mais difícil. Nesse processo, o modelo remove diferentes tipos de ruído. Nesse caso, pode variar o σ entre 15, 25 e 50. O nível definido no algoritmo é o padrão, 25.

No *Epoch*², será definido o número de ciclos completos que o modelo deve treinar. Cada época representa um ciclo completo em que o modelo analisa todas as imagens do conjunto de treinamento. Com o aumento das épocas, o modelo se torna mais preciso. O modelo foi concluído com 50 épocas. Logo após, temos o controlador da taxa de aprendizado inicial para o otimizado Adam, que é um otimizador de gradiente descendente bastante utilizado no treinamento de redes neurais profundas. A taxa de aprendizado determina o quanto os pesos da rede são ajustados em cada iteração. Um valor muito alto pode fazer com que o modelo não convirja, e um valor muito baixo pode fazer com que o treinamento seja muito lento. E, por último, é determinada a frequência em que o modelo será salvo durante o treinamento. Nesse caso, o modelo será salvo a cada 1 época.

Figura 4: Resultado do treino de 50 épocas

```

=====
Total params: 558,016
Trainable params: 556,096
Non-trainable params: 1,920

-----
2024-10-14 16:52:14: Modelo treinado carregado a partir de models/DnCNN_sigma25\model_050.hdf5
Forma inicial de data: (400, 596, 40, 40)
Forma após reshape: (238400, 40, 40, 1)
^^-Dados de treinamento concluídos-^^
Forma inicial de data: (400, 596, 40, 40)
Forma após reshape: (238400, 40, 40, 1)
^^-Dados de treinamento concluídos-^^

```

Fonte: Elaborado pelo autor

Na Figura 6, conseguimos ver quando o modelo está totalmente treinado com as 50 épocas definidas. Essas 50 épocas foram o padrão treinado anteriormente pela equipe de Kai Zhang (2017). É possível identificar também os parâmetros

¹ É frequentemente usado para quantificar a intensidade do ruído gaussiano que se deseja remover da imagem. Um valor maior de σ indica uma maior intensidade de ruído.

² É a representação de uma iteração completa, em que o modelo ajusta seus parâmetros com base nos dados de entrada e nas saídas desejadas. Nesse sentido, o DnCNN aprende a diferenciar a imagem original e a imagem ruidosa, melhorando sua capacidade de restaurar imagens.

treinados e os que não foram possíveis treinar, a partir do conjunto de dados utilizados.

3.2 Teste do algoritmo treinado

Nesta seção, irei destacar como é possível testar o modelo treinado com o algoritmo DnCNN. Para isso foram definidas algumas bibliotecas e parâmetros necessários. O teste do modelo se baseia no treinamento. Nesse caso ele utilizará as 50 épocas treinadas como parâmetro para compreensão profunda do algoritmo.

Foram necessárias algumas bibliotecas, como as que servem para visualização de dados, controlador e medidor de tempo. É necessária a utilização de algumas bibliotecas em *Python*, como o *Numpy*, para cálculos numéricos, módulos para calcular erros quadráticos e algumas métricas para avaliar a qualidade de imagens como o *compare_psnr* e *compare_ssim*.

Temos que destacar também a utilização do *Skimage* para ler e salvar as imagens, sendo uma biblioteca *Python* para processamento de imagens. A partir das importações, é necessário definir os parâmetros para testar com qualidade as imagens.

É importante destacar que os parâmetros para teste se baseiam em técnicas básicas nos testes de redução de ruído principalmente com o algoritmo DnCNN:

- O *set_dir* irá especificar o diretório onde o conjunto de dados de teste está localizado.
- O *set_names* irá definir os nomes dos conjuntos de dados de teste que serão usados.
- O *sigma* define o nível de ruído gaussiano a ser adicionado nas imagens para teste. Geralmente se usam três tipos de ruído para teste, o 25 foi definido por padrão.
- O *model_dir* irá definir o diretório que o modelo treinado será armazenado.
- O *model_name* especifica o nome do arquivo do modelo treinado. Nesse caso como já passou no treino, o modelo está com 50 épocas treinadas.
- O *result_dir* é o diretório onde os resultados das imagens restauradas serão salvas, caso a opção de salvar esteja ativada.
- E o *save_result* especifica se as imagens serão restauradas ou não.

Os testes seguem, depois do treinamento. Foi adicionado um total de 50 épocas de treinamento, como mostrado na Figura 6, após aproximadamente 24h com as configurações atuais do sistema, foi concluído com as condições atuais da NVIDIA GeForce RTX 3060. O modelo foi carregado e adicionado na parte do teste.

Após a ativação do algoritmo é possível ver o carregamento dos modelos treinados, e as imagens que foram restauradas. Os modelos treinados estão relacionados às épocas decididas anteriormente, nesse caso 50. Portanto, após o carregamento dos modelos, o algoritmo irá remover o ruído das imagens ruidosas, e adicionar o tempo em que elas foram restauradas junto com o tamanho, como mostrado na Figura 10.

Figura 5: Remoção de ruído de quatro imagens diferentes em formato jpg

```
2024-11-14 23:38:16.932767: I tensorflow/stream_executor/cuda/cuda_dnn.cc:368] Loaded cuDNN version 8101
Set12 : 1.jpg : 7.2387 second
PSNR: 34.8765, SSIM: 0.8644
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
Processando imagem: 2.jpg, tamanho: (288, 368, 3)
Set12 : 2.jpg : 0.1010 second
PSNR: 34.3566, SSIM: 0.8212
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
Processando imagem: 3.jpg, tamanho: (288, 368, 3)
Set12 : 3.jpg : 0.1070 second
PSNR: 36.3328, SSIM: 0.9007
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
Processando imagem: 4.jpg, tamanho: (288, 368, 3)
Set12 : 4.jpg : 0.1040 second
PSNR: 35.0708, SSIM: 0.9039
```

Fonte: Elaborado pelo autor

Ainda na Figura 10, podemos ver como o modelo definiu as imagens adicionadas na etapa de restauração. Nesse momento, é verificada a imagem para teste e substituída por uma nova. As imagens utilizadas estão no formato jpg e o tempo que elas foram restauradas está ao lado em segundos.

Após o teste, o modelo irá armazenar as imagens em um local específico definido pelo usuário. Desse modo podemos passar para a próxima etapa, em que definimos a eficiência do modelo, os erros e a comparação das imagens antigas às imagens restauradas. Assim, podendo melhorar o modelo para cada caso específico.

4 RESULTADOS E DISCUSSÕES

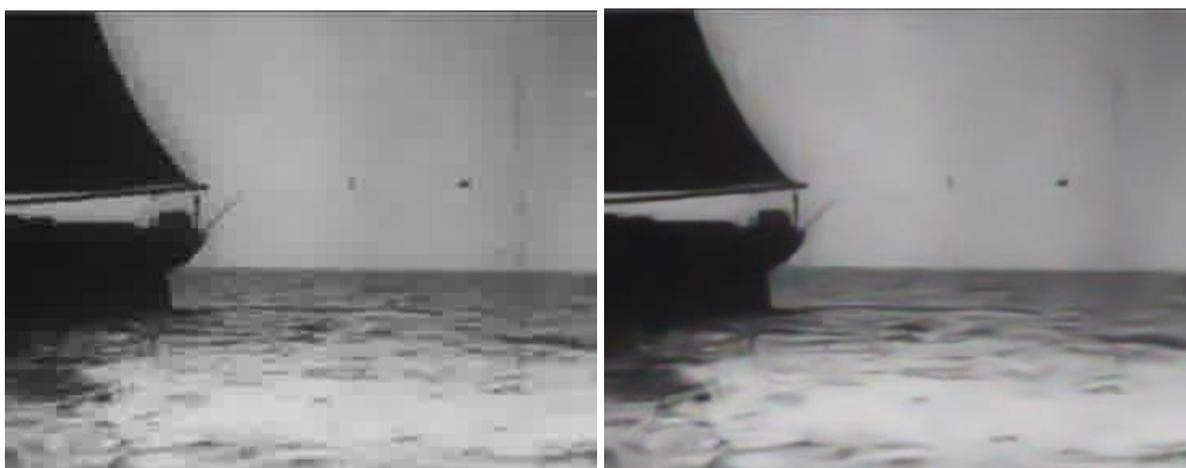
Nesta seção, será discutido e visualizado os resultados obtidos a partir do algoritmo com as configurações atuais e o nível de treinamento definido.

Vale ressaltar que a proposta inicial da restauração completa de um filme tem como base várias técnicas, algumas técnicas usam super-resolução, que não é este caso. Aqui veremos como funciona a remoção de ruídos usando o algoritmo DnCNN com os parâmetros definidos anteriormente, podendo melhorar dependendo da máquina em que utiliza.

Um dos objetivos do restauro deve ser a ética na reconstrução de uma cópia do filme, em que todo o conteúdo do filme deve ser preservado. A importância do uso de técnicas respeitadas, que não diferenciam tanto do conteúdo antigo, é primordial para um trabalho completo e preservado.

As imagens restauradas são retiradas do filme *Nosferatu* (1922) e, tendo em vista que as imagens estão danificadas, as necessidades mudam, para um restauro completo deve se ter um treinamento maior: é necessária a utilização de outros algoritmos, como o de super-resolução. Podemos perceber que a remoção de ruído foi aplicada nas Figuras 6, 8, 10 e 12, entendendo que as mudanças na utilização do algoritmo de remoção de ruídos são visíveis.

Figura 6: Imagem 1 restaurada do filme *Nosferatu* (1922), antes e depois



Fonte: Elaborado pelo autor

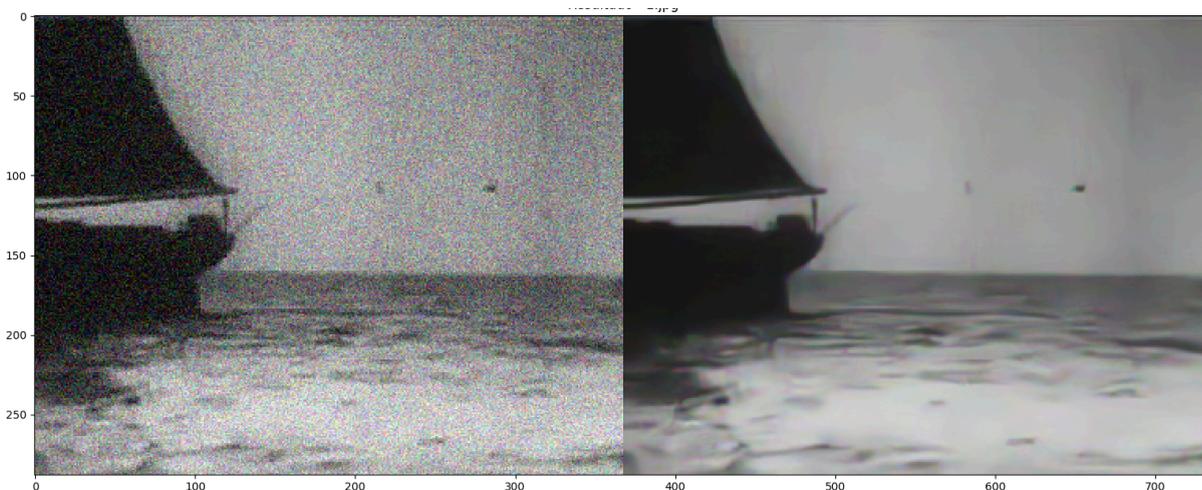
Após alguns testes com imagens retiradas de filmes preto e branco e ruidosas, foi visto que o algoritmo funciona de forma eficiente, podendo ser melhorado nos parâmetros e combinado com outras técnicas. Na Figura 6 está disponível o antes, com uma imagem cheia de ruídos no mar, no céu e no barco.

Vale ressaltar que a imagem, após o teste, fica mais limpa. Porém, como a imagem original está bem danificada, a cópia melhorada irá perder alguns detalhes. Para resolver esse caso seria necessário a utilização de algum algoritmo de super resolução.

Para compreensão e estudo, o DnCNN é um algoritmo de remoção de ruídos. A melhoria completa de um filme segue diversas etapas, e, para utilizar IA nessas etapas, é necessário um grande estudo de diversos algoritmos e distribuir muito tempo para executá-los da melhor forma possível.

É importante destacar que o tempo para treinamento, parâmetros, imagens para treino são essenciais para uma imagem ainda mais limpa. No treinamento deste algoritmo, foram necessárias 50 épocas, em que cada uma vai ser treinada com o parâmetro de 400 imagens de um conjunto de dados comumente usados em diversos algoritmos que utilizam imagens como base.

Figura 7: Representação da Imagem 1 com ruído gaussiano e a imagem restaurada



Fonte: Elaborado pelo autor

Na Figura 6, o PSNR foi de 34.8765, e o SSIM foi de 0.8644. O PSNR mede a qualidade de uma imagem reconstruída, comparando-a com a imagem original, em decibéis (dB).

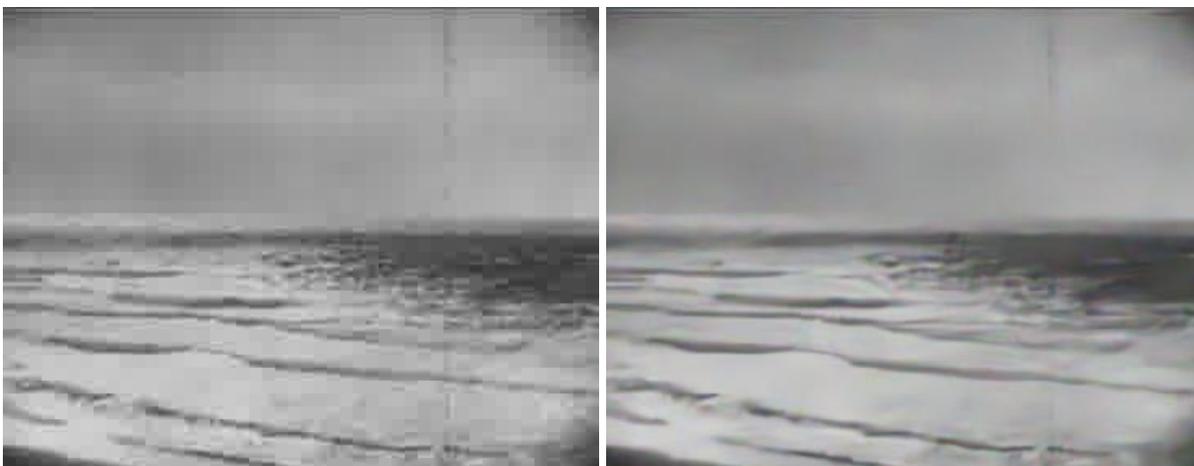
Para Wang et al. (2004), valores mais altos de PSNR indicam uma melhor qualidade, ou seja, que a imagem restaurada está mais próxima da original. PSNR acima de 30 dB são considerados bons em tarefas de restauração de imagem. O valor de PSNR 34.8764 é um bom resultado, sugerindo que a imagem *denoised* (sem ruído) tem uma qualidade bastante próxima à da imagem original, com pouca

diferença perceptível. Já em relação ao SSIM, valores próximos ou acima de 0.9 indicam que a qualidade da imagem restaurada é muito boa em termos de estrutura, preservando as características visuais mais importantes.

Assim como a Figura 6, na Figura 8 foram utilizados os mesmos parâmetros e o PSNR foi de 34.3566 com valor bom, conforme foi definido anteriormente. Visualmente a imagem aparenta ter mais mudanças. Vale ressaltar que a existência de algo que não tem tanta variação de elementos, como na Figura 6, tendo um barco no mar, pode não ter o mesmo impacto, devido às imagens utilizadas para treino. Nessa nova imagem podemos entender que a existência de apenas um mar pode diminuir o PSNR, fazendo com que a imagem tenha mais adição de ruído e diferenciando um pouco mais da imagem original.

O SSIM da figura 8 foi definido como 0.8212, indicando que a qualidade da imagem em sua estrutura também está boa e a remoção do ruído foi concluída com sucesso.

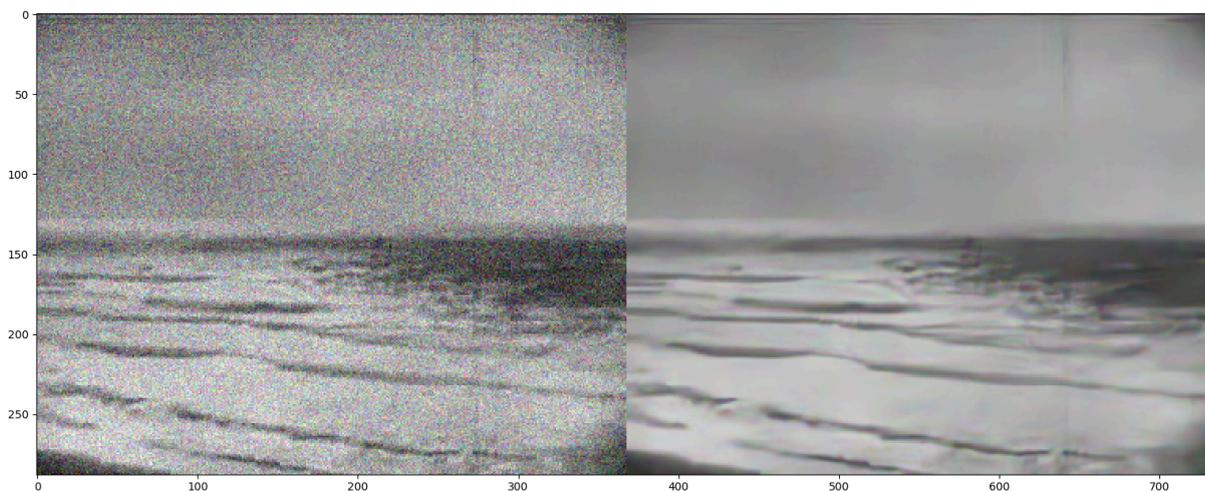
Figura 8: Imagem 2 restaurada do filme *Nosferatu* (1922), antes e depois



Fonte: Elaborado pelo autor

Na Figura 9 podemos ver como a imagem do mar está com o ruído gaussiano, esse ruído é feito com o nível *sigma* 25, para definir como o modelo vai ser adaptado no momento da restauração.

Figura 9: Representação da Imagem 2 com ruído gaussiano e a imagem restaurada



Fonte: Elaborado pelo autor

Na Figura 10 podemos destacar que está trabalhando com imagens muito escuras e como o modelo conseguiu adaptar a restauração e fazer de forma eficiente. O PSNR dessa imagem foi de 36.3328 e o SSIM foi de 0.9007. Nesse caso, pode-se visualizar uma melhoria bastante significativa, principalmente no seu SSIM, fazendo com que a imagem permaneça com as características necessárias da ambientação. A remoção de ruído foi aplicada com sucesso nela.

Figura 10: Imagem 3 restaurada do filme *Nosferatu* (1922), antes e depois



Fonte: Elaborado pelo autor

Na figura 11 pode-se ver o nível do ruído aplicado na imagem.

Figura 11: Representação da Imagem 3 com ruído gaussiano e a imagem restaurada



Fonte: Elaborado pelo autor

Em destaque, a Figura 12, é importante entender que é a única imagem com a representação da imagem de um ser vivo. Foi importante a utilização de uma imagem assim para descobrir se o modelo iria compreender bem. Assim, percebe-se que a remoção de ruído foi bem aplicada e não deve-se deixar de notar algumas falhas que permaneceram como rasuras permanentes na imagem restaurada.

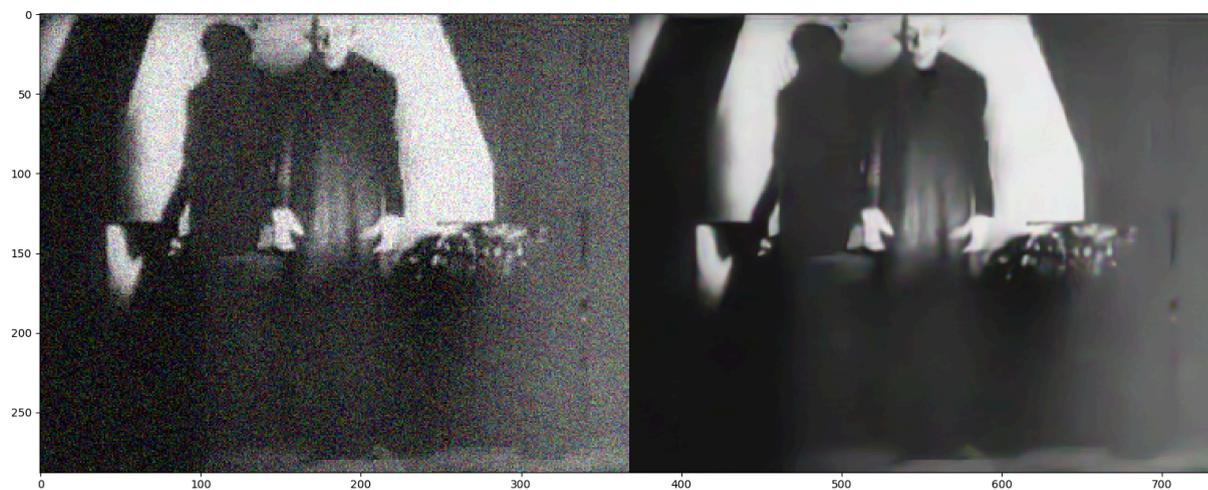
Figura 12: Imagem 4 restaurada do filme *Nosferatu* (1922), antes e depois



Fonte: Elaborado pelo autor

A Figura 13 mostra o nível de ruído aplicado e a imagem restaurada. O PSNR dessa imagem foi de 35.0708 e o SSIM foi de 0.9039. Sendo uma das melhores em relação ao seu PSNR das imagens e o melhor SSIM.

Figura 13: Representação da Imagem 4 com ruído gaussiano e a imagem restaurada



Fonte: Elaborado pelo autor

5 CONSIDERAÇÕES FINAIS

Considerando todo o contexto sociocultural que o cinema vem exercendo ao longo dos anos, se vê a necessidade de uma preservação maior sobre o nosso patrimônio. No caso dos filmes, alguns aspectos estruturais das películas foram se perdendo com o tempo em meio a desastres naturais ou falta de ajuda externa, alguns conseguindo se manter com muita preservação e restauro manual. Este trabalho mostra que os meios computacionais no século XXI vêm conseguindo ajudar diversas áreas do conhecimento e pode, por sua vez, também ajudar no meio cinematográfico.

Entendendo que a Inteligência Artificial é amplamente utilizada em âmbitos diversos no dia a dia. Desse modo, é possível inserir alguns métodos que possam ajudar na restauração de filmes que foram desgastados com o tempo. Os filmes que precisam de restauro podem se beneficiar com as tecnologias de hoje em dia, buscando um meio de respeito mútuo na reconstrução respeitosa e ética de uma cópia restaurada a partir de algoritmos amplamente utilizados para a restauração de imagens e vídeos.

Neste trabalho, foi apresentada uma técnica de remoção de ruído proposta pela equipe de Kai Zhang (2007). Esta técnica é uma forma segregatória em que separa o ruído da imagem para conseguir limpá-la. Desse modo, filmes com muitos ruídos ou manchas podem se beneficiar positivamente e, assim, facilitar os meios de restauro manual feitos antigamente. Portanto, o DnCNN vem como algoritmo proposto, demonstrando sucesso nos resultados obtidos. Com as quatro imagens disponíveis do filme *Nosferatu* (1922), é possível ver uma melhoria significativa nas imagens, podendo ser adaptadas para outras técnicas com outros algoritmos como os de super-resolução. Vale ressaltar que nesse trabalho foram utilizadas apenas imagens estáticas e não houve teste com imagens em movimento.

A proposta de uma técnica de restauro em que o filme é modificado de forma ética ocorreu bem e com particularidades interessantes, como apresentadas nas métricas de *PSNR* e *SSIM*. Algumas imagens saíram melhores que outras nos testes, o que pode ser diferenciado por sua parte inicial altamente danificada. Então, as imagens podem se beneficiar, mas, para ter um resultado 100% positivo, deve-se organizar diversos parâmetros e propor futuras técnicas ainda melhores.

Para trabalhos futuros, é possível identificar meios que restaurem o filme por completo. Alguns algoritmos podem ser estudados, de forma que melhorem a imagem em diferentes aspectos, como reconstrução do áudio, técnicas de super-resolução, correção de quadros faltantes considerando um filme sem cor e extremamente danificado com o tempo, e utilizando outras técnicas de IA. Deve-se usar essas técnicas também para outros meios externos ao cinema, como em vídeos e imagens antigas relacionadas a aspectos emocionais, como os patrimônios históricos de famílias.

REFERÊNCIAS

- ...E o vento levou; Direção: Victor Fleming. Produção: Selznick International Pictures. Estados Unidos: Metro-Goldwyn-Mayer, 1939. 1 DVD.
- A LISTA de Schindler; Direção: Steven Spielberg. Produção: Amblin Entertainment. Estados Unidos: Universal Pictures, 1993. 1 DVD.
- BATCHEN, G. **Forget Me Not: Photography and Remembrance**. In: New York: Princeton Architectural Press, 2004.
- BERTALMIO, M. *et al.* **Image Inpainting**. In: Proceedings of the 27th annual conference on Computer Graphics and Interactive Techniques (SIGGRAPH), 2000.
- BONS Companheiros; Direção: Martin Scorsese. Produção: Warner Bros. Estados Unidos: Metro-Goldwyn-Mayer, 1990. 1 DVD.
- FERRO, M. **Cinema e História**. In: Paz e Terra, Rio de Janeiro, 1992.
- GONZALEZ, Rafael C.; WOODS, Richard E. **Digital Image Processing**. Global Edition. Pearson Education, 2018.
- GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. MIT Press, 2016.
- GURNEY, Kevin. **An Introduction to Neural Networks**. UCL Press, 1997.
- KUIPER, J. B. Silent Films for Contemporary Audiences. **Journal of Film Preservation**, George Eastman House, jun. 1982.
- LIM, Bee; SON, Sanghyun; KIM, Heewon; NAH, Seungjun; LEE, Kyoung Mu. **Enhanced Deep Residual Networks for Single Image Super-Resolution**. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- MARKHAM, Katie. **Two-dimensional engagements: photography, empathy and interpretation at District Six Museum**. In: International Journal of Heritage Studies, Routledge, 2019.
- MAYER, Richard E. **Multimedia Learning**. In: Cambridge University Press, 2001.
- NOSFERATU; Direção: F. W. Murnau. Produção: Prana Film. Alemanha: UFA, 1922. 1 DVD.
- NVIDIA. **O que é Cuda?**. In: Nvidia.com.br. 25 jun. 2024. Disponível em: <https://blog.nvidia.com.br/blog/o-que-e-cuda/>. Acesso em 2 out.2024.
- O GLOBO; El País; NYT. Scorsese faz apelo pela preservação dos acervos cinematográficos dos EUA. In: Globo.com. **O GLOBO** . [S.l.], 2 abr. 2013. Disponível em: <https://oglobo.globo.com/cultura/scorsese-faz-apelo-pela-preservacao-dos-acervos-cinematograficos-dos-eua-8005845>. Acesso em: 2 jun.2024.
- PRESSE, F. Mais de dois terços dos filmes mudos feitos nos EUA se perderam. In: Globo.com. **G1**. [S.l.], 5 dez. 2013. Disponível em: <https://g1.globo.com/pop-arte/cinema/noticia/2013/12/mais-de-dois-tercos-dos-filmes-mudos-feitos-nos-eua-se-perderam.html>. Acesso em: 1 jun.2024.

PSICOSE; Direção: Alfred Hitchcock. Produção: Paramount Pictures. Estados Unidos: Universal Pictures, 1960. 1 DVD.

RUSSELL, S. J.; NORVIG, P. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013 p. 23.

SZELISKI, Richard. **Computer Vision: Algorithms and Applications**. Springer. Disponível em: <http://szeliski.org/Book/>. Acesso em: 10 out.2024.

SMITH, Steven W. **The Scientist and Engineer's Guide to Digital Signal Processing**. California Technical Publishing, 1997. p. 15.

TURING, Alan M. **Computing Machinery and Intelligence**. *Mind*, v. 59, n. 236, p. 433-460, 1950.

TAXI Driver; Direção: Martin Scorsese. Produção: Columbia Pictures. Estados Unidos: Sony Pictures, 1976. 1 DVD.

USAI, Paolo C. **Silent Cinema: An Introduction**. *In*: British Film Institute (BFI), 2000.

VINCENT, P. *et al.* **Denosing autoencoders: unsupervised learning of representations**. *In*: Proceedings of the 25th International Conference on Machine Learning (ICML). 2010.

WANG, Zhou. *et al.* **Image Quality Assessment: From Error Visibility to Structural Similarity**. *In*: IEEE Transactions on Image Processing, 2004.

WAZLAWICK, Raul S. **Metodologia de Pesquisa para Ciência da Computação**. 6. ed. Rio de Janeiro: Elsevier, 2009.

ZHANG, K. *et al.* Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. **IEEE**, National Natural Scientific Foundation of China, v. 26, n. 7, p. 3142-3155, fev. 2017.

ZHANG, Lei. *et al.* **Image Restoration: From Sparse and Low-rank Priors to Deep Priors**. *In*: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.