

**Universidade Estadual Paulista “Júlio de Mesquita Filho”
FCT — Faculdade de Ciências e Tecnologia
DMC — Departamento de Matemática e Computação
Bacharelado em Ciência da Computação**

**Trabalho de Conclusão de Curso
Revisão bibliográfica**

Aplicação de métodos de inpainting em imagens digitais

Gustavo Becelli do Nascimento

Orientador: Prof. Dr. Almir Olivette Artero

**Presidente Prudente
2023**

LISTA DE FIGURAS

1	Exemplo do inpainting de imagens (COMMONS, 2022). A imagem da esquerda mostra uma imagem danificada. A imagem da direita mostra a imagem após o processo de inpainting.	5
2	Exemplo da seleção de regiões a serem preenchidas (OPENCV, 2022). A imagem no canto superior esquerdo é a imagem danificada. No canto superior direito há a máscara (região) que selecionada pelo usuário para ser preenchida. A imagem no canto inferior esquerdo mostra a imagem após o processo de inpainting utilizando (BERTALMIO; BERTOZZI; SAPIRO, 2001), e a imagem no canto inferior direito mostra a imagem após o processo de inpainting utilizando (TELEA, 2004).	7

LISTA DE TABELAS

SUMÁRIO

1	Introdução	2
2	Fundamentação teórica	3
2.1	Métodos clássicos	4
2.2	Métodos baseados em aprendizado	4
2.2.1	Redes neurais convolucionais	5
2.2.2	Redes generativas adversárias	5
2.2.3	Depth Inpainting	6
2.3	Otimizações dos resultados	6
2.3.1	Patch-based inpainting	6
2.3.2	Preenchimentos intra e inter imagem	6
2.3.3	Inpainting guiado	6
2.3.4	Context-aware inpainting	7
2.3.5	Representação esparsa	7
2.4	Otimizações computacionais	7
2.4.1	Profiling	7
2.4.2	Aceleração de hardware	7
2.4.3	Crop	7
2.4.4	Redução de memória	7
2.4.5	Paralelismo	7
2.4.6	Estruturas de dados	7
2.4.7	Redução de redundância	7
2.4.8	Otimização de acesso a memória	7
2.4.9	Vetorialização	7
2.4.10	Aproximações	7
2.4.11	Cálculos incrementais	7
2.5	Aplicações	7
3	Trabalhos Relacionados	7
4	Conclusão	7
5	Conclusão	8
6	Referências bibliográficas	9

1 Introdução

O Inpainting de Imagens é o processo de preencher regiões faltantes ou danificadas de uma imagem com o intuito de restaurá-la ou melhorar sua aparência. Este é um problema amplamente estudado em visão computacional e possui uma vasta gama de oportunidades de aplicações, incluindo o restauro de obras artísticas e fotografias danificadas, a remoção de objetos de imagens e síntese de texturas (CRIMINISI; PÉREZ; TOYAMA, 2004).

Esta área de pesquisa tem sido estudada há várias décadas. Os primeiros trabalhos sobre o tema surgiram na década de 1950, mas foi somente a partir dos anos 1990 que o inpainting começou a ser utilizado de maneira mais ampla.

(ELHARROUSS et al., 2019) apresenta uma revisão sobre o inpainting de imagens, descrevendo os principais métodos e aplicações. Neste campo, existem várias técnicas para o inpainting de imagens, cada tal com suas próprias vantagens e desvantagens. Uma abordagem popular é apresentada em (BERTALMIO et al., 2000), a qual introduz uma técnica de inpainting baseada no conceito de complementar uma região de uma imagem existente. Neste trabalho é utilizado equações parciais diferenciais para propagar informações das bordas da região danificada para o centro da região, permitindo que o algoritmo estime como os pixels devem ser restaurados.

Uma das contribuições mais significativas para o campo de inpainting de imagem é o desenvolvimento do "Método de Marcha Rápida" (TELEA, 2004), um algoritmo eficiente que usa equações diferenciais parciais para propagar informações de pixels conhecidos para os pixels desconhecidos na região de inpainting. Este algoritmo tem sido amplamente utilizado em várias aplicações e foi implementado em muitas bibliotecas de software, incluindo a biblioteca de código aberto de visão computacional OpenCV (OPENCV, 2000). Este método tem demonstrado produzir resultados de alta qualidade para uma ampla variedade de tipos de imagem e cenários, e se tornou uma das referências para avaliar o desempenho de outros algoritmos de inpainting.

Nos últimos anos, houve um aumento de interesse em utilizar métodos baseados em aprendizado de máquina, como redes adversárias gerativas (GANs) e redes neurais convolucionais (CNNs) (PATHAK et al., 2016), para resolver o problema de inpainting. Esses métodos têm o potencial de aprender padrões e estruturas complexas a partir de grandes conjuntos de dados, o que pode ser usado para gerar resultados de inpainting de alta qualidade. No entanto, esses métodos podem ser computacionalmente caros e podem exigir uma abundância de dados de treinamento para obter bons resultados.

Hoje, o inpainting é utilizado em uma variedade de aplicações. Dentre elas, está incluída o restauro de obras de arte danificadas, a remoção de objetos de imagens (como nuvens em imagens de sensoriamento remoto) e até mesmo a remoção de imperfeições de fotografias. Além disso, pode-se utilizar os mesmos conceitos de restauração para modificar as perspectivas de imagens, como é apresentado em (HUANG et al., 2014).

O principal problema encontrado neste campo de pesquisa são a escolha do método de inpainting mais adequado para cada tipo de imagem e cenário. Este fator se deve às desvantagens que os métodos existentes proporcionam (SALEM, 2021). Em geral, os maiores desafios são:

- **Preservação da consistência e estrutura da imagem de entrada:** muitos modelos geradores podem gerar imagens visualmente agradáveis, mas que não correspondem perfeitamente com a vizinhança ou à textura da imagem original.
- **Preservação de pequenos detalhes:** este problema é particularmente desafiador para imagens com texturas complexas, como imagens de paisagens naturais, ou a área que

está sendo restaurada é muito grande.

- **Reconstruir objetos sobrepostos:** quando um objeto é parcial ou completamente sobreposto, como um objeto que está coberto por uma nuvem, o modelo gerador pode não ser capaz de reconstruir o objeto original.
- **Preenchimento de grandes regiões:** a geração de imagens de alta qualidade quando há uma grande região a ser reconstruída ainda é um problema desafiador.
- **Desempenho computacional:** a maioria dos métodos de inpainting de imagens que geram imagens de alta qualidade são computacionalmente caros, dificultando sua aplicação em tempo real.

A etapa posterior a este trabalho, iniciada com as disciplinas de Projeto Científico II e Trabalho de Conclusão de Curso, será a implementação de algumas técnicas de preenchimento de imagens (inpainting) e sua aplicação em exemplos reais. O objetivo será comparar os resultados obtidos e avaliar a eficiência computacional e os resultados de cada um deles. É esperado que, com a aplicação desta sequência de trabalhos, sejam obtidos bons resultados, com um tempo de processamento aceitável. Além disso, caso oportuno, serão realizadas otimizações no código-fonte, para aumentar a qualidade dos resultados ou diminuir o tempo de processamento, a fim de tornar os métodos mais adequados para aplicações em tempo real.

2 Fundamentação teórica

A gama de aplicações processo de inpainting em imagens digitais é ampla, uma vez que não se limita a um único domínio. A técnica pode ser aplicada em diversas áreas, tais como a medicina, engenharia, arquitetura, entre outras, as quais serão mais exploradas mais a fundo no tópico 2.5. Além disso, o conceito de inpainting, inicialmente apresentado em 1, não agrega apenas as ideias de restauração de imagens ou remoção de objetos, mas também a construção e reconstrução de imagens (YOU; YOU; PAN, 2019), a preservação da privacidade (por exemplo, borramento de rostos e placas de veículos automotivos) em imagens públicas (GOOGLE, 2022), ao preenchimento de lacunas em imagens de satélites (MAALOUF et al., 2009), entre outros possíveis usos, aos quais alguns serão mais explorados no prosseguir desta revisão. Por fim, o tópico 3 apresenta os trabalhos relacionados ao tema.

Neste trabalho, o escopo de inpainting de imagens pode ser compreendido como uma técnica utilizada para preencher regiões vacantes ou danificadas em uma imagem digital. O objetivo do inpainting é reconstruir as regiões afetadas de forma que o resultado final seja o mais visualmente plausível possível, ou seja, que a imagem resultante seja natural à ótica humana e não apresente artefatos visuais, enquanto preserva a coerência e consistência da imagem original (LEVIN; ZOMET; WEISS, 2003).

Há diversas formas de se implementar o processo de inpainting, sendo que a escolha do método a ser utilizado depende do domínio de aplicação, das características da imagem, do tipo de máscara, entre outros fatores (BLACK; KESHAVARZ; SOUVENIR, 2020). Contudo, é comum distinguir duas variedades de métodos amplamente utilizados: os métodos "clássicos", os quais geralmente utilizam informações da estrutura ou da imagem ou outras regiões para preencher lacunas, e os métodos baseados em aprendizado, os quais comumente utilizam grandes quantidades de dados para serem treinados, sendo que os métodos baseados em aprendizado são os mais recentes e apresentam geralmente resultados mais satisfatórios.

2.1 Métodos clássicos

Os métodos clássicos utilizam informações dos pixels e da região vizinha das regiões que estão vacantes ou danificadas para preencher as lacunas. Estes métodos são baseados no princípio de que regiões próximas e similares ao alvo são os mais adequados para preencher a região (BARNES et al., 2009; BERTALMIO; BERTOZZI; SAPIRO, 2001). É possível ainda, dentro desta estratégia, distinguir dois tipos de métodos: os métodos baseados em patches, tais que utilizam amostras próximas da região-alvo para preencher a lacuna, e os métodos baseados em síntese de textura que utilizam a textura das regiões conhecidas para preencher a lacuna.

Métodos baseados em difusão: Segundo (BLACK; KESHAVARZ; SOUVENIR, 2020), estes métodos transmitem os valores dos pixels adjacentes para as regiões a serem preenchidas e suas implementações são variadas. Métodos como (TELEA, 2004) utilizam uma o Método da Marcha Rápida, que consiste em propagar médias ponderadas das intensidades dos pixels adjacentes para a região-alvo, mantendo as altas e baixas frequências. Há outras implementações como (BERTALMIO; BERTOZZI; SAPIRO, 2001), que utilizam as equações de Navier-Stokes para a difusão de valores de pixels adjacentes ou até os isótopos das imagens. Estes métodos são simples e eficientes computacionalmente, porém, geralmente apresentam resultados com artefatos visuais, como bordas irregulares, texturas deformadas, etc, em especial, para imagens com texturas complexas.

Métodos baseados em patches: estes métodos utilizam amostras de regiões conhecidas e similares a região-alvo para reconstruí-la. Assume-se que a região corrompida possui pixels semelhantes a alguma área próxima, então o método seleciona uma amostra da área próxima e a transfere para a região-alvo, mantendo a coerência da imagem e a consistência da textura. Essa estratégia é simples e eficiente computacionalmente, em especial para imagens com texturas e estruturas simples (BARNES et al., 2009), apesar de apresentarem resultados superiores aos de difusão. Contudo, não produz resultados satisfatórios para lidar com imagens mais compostas, como imagens com texturas complexas, imagens com bordas irregulares, imagens com objetos com texturas semelhantes, etc.

Métodos baseados em síntese de textura: estes métodos estão incluídos na categoria de métodos baseados em patches, pois também utilizam amostras de regiões conhecidas e similares a região-alvo para reconstruí-la. Contudo, estes métodos utilizam amostras de texturas para sintetizar novos pixels para a região-alvo. Estes métodos são baseados no princípio de que a textura de uma região é uma função de sua vizinhança, e assume que a área vizinha ao alvo possui texturas semelhantes, o que viabiliza propagar as bordas e frequências para preenchimento. Utilizar esta estratégia pode produzir resultados satisfatórios para imagens com texturas complexas, embora ainda encontrem dificuldades para lidar com imagens com bordas ou padrões irregulares, imagens com objetos com texturas semelhantes, entre outras (BERTALMIO et al., 2003).

2.2 Métodos baseados em aprendizado

Os métodos baseados em aprendizado, no que lhe concernem, utilizam técnicas de aprendizado de máquina, como redes neurais convolucionais (CNNs) e redes generativas adversárias (GANs) para preencher a região danificada. Estas técnicas "aprendem" a reconstruir a região danificada a partir de um conjunto de dados de imagens com e sem máscara. Estes métodos são mais recentes e apresentam resultados mais satisfatórios, em especial para imagens com

texturas complexas, imagens com bordas irregulares, imagens com objetos com texturas semelhantes, etc. Contudo, estes métodos são mais complexos e computacionalmente mais custosos, devido ao fato que eles necessitam de um grande conjunto de dados para serem treinados (GOODFELLOW; BENGIO; COURVILLE, 2016).

Redes neurais convolucionais: utilizam uma CNN para reconstruir as regiões faltantes. Estes métodos se beneficiam da habilidade das CNNs de aprender as características de uma imagem e conseguem produzir bons resultados para imagens com texturas e estrutura complexas. A rede é treinada em um vasto conjunto de dados de imagens e aprende a reconhecer padrões e as características das imagens. Após o treinamento, a rede utiliza o conhecimento adquirido para gerar novos pixels para a região danificada, ainda preservando a região que a cerca.

Redes generativas adversárias: utilizam uma GAN para gerar novos pixels para preencher a região danificada. As GANs consistem em duas partes: uma rede geradora que gera novos pixels, e um discriminador que avalia a qualidade dos pixels obtidos (BLACK; KESHAVARZ; SOUVENIR, 2020). Estes métodos podem gerar resultados mais realísticos do que os baseados em CNNs, mas eles são mais custosos computacionalmente e requerem um conjunto de dados maior para treinamento. Consequentemente, estes métodos são mais complexos e costumam levar mais tempo para treinamento (PATHAK et al., 2016).

Em suma, o inpainting de imagem é uma técnica utilizada para preencher regiões faltantes ou corrompidas de uma imagem, como uma nuvem em uma imagem de sensoriamento remoto, ou danos causados pela degradação em fotografias digitalizadas, como na Figura 1.



Figura 1: Exemplo do inpainting de imagens (COMMONS, 2022). A imagem da esquerda mostra uma imagem danificada. A imagem da direita mostra a imagem após o processo de inpainting.

2.2.1 Redes neurais convolucionais

Definição Arquitetura Treinamento Teste Erros Aplicações Limitações Melhorias

2.2.2 Redes generativas adversárias

Definição Arquitetura Treinamento Teste Erros Aplicações Limitações Melhorias

2.2.3 Depth Inpainting

O Depth Inpainting é uma técnica utilizada para realizar o inpainting de imagens 3D, como é o caso de imagens obtidas por câmeras fotográficas ou sistemas de visão computacional. Esta técnica é útil em situações onde a câmera é incapaz de capturar corretamente as informações de profundidade, como é o caso de objetos translúcidos ou afetados pela luz especular (reflexos) (SHIH et al., 2020).

A técnica de Depth Inpainting consiste em utilizar uma rede neural convolucional para estimar a profundidade de uma imagem, treinada com imagens com valores conhecidos, e então utilizar esta estimativa para realizar o inpainting da imagem-alvo. Outra abordagem é utilizar técnicas de processamento de imagens, como a interpolação de valores, para aproximar os valores da região a ser preenchida. (SHIH et al., 2020) Isto pode ser obtido analisando os padrões e estruturas dos pixels vizinhos. Em ambos casos, o algoritmo precisa ser treinado com uma base de dados de imagens com valores de profundidade conhecidos para aprender a realizar previsões.

2.3 Otimizações dos resultados

2.3.1 Patch-based inpainting

2.3.2 Preenchimentos intra e inter imagem

2.3.3 Inpainting guiado

A escolha da região a ser preenchida é parte fundamental na qualidade do resultado do processo de inpainting, uma vez que ela afeta diretamente a qualidade da imagem gerada. A região selecionada deve abranger a quantidade necessária para prover contexto suficiente para o algoritmo de inpainting, mas não deve ser muito grande, pois isso pode levar a resultados com texturas e estruturas não realistas ou com maior tempo de processamento. Ademais, a região escolhida deve ser bem definida e conter uma delimitação clara, para que o algoritmo de inpainting possa identificar com facilidade a região e estruturas a serem preenchidas (HUANG et al., 2014). Alguns métodos realizam este processo de forma automática, sendo conhecidos como Blind Inpainting (WANG et al., 2020).

Por seguinte, é importante ponderar a situação da região a ser preenchida no contexto da imagem, pois o processo de correção pode afetar a coerência do resultado, isto é, o preenchimento de uma região no conteúdo principal de uma imagem possui mais impacto ao comparar-se com o preenchimento de uma região de fundo.

A Figura 2 mostra um exemplo da seleção de regiões para o processo de inpainting.



Figura 2: Exemplo da seleção de regiões a serem preenchidas (OPENCV, 2022). A imagem no canto superior esquerdo é a imagem danificada. No canto superior direito há a máscara (região) que selecionada pelo usuário para ser preenchida. A imagem no canto inferior esquerdo mostra a imagem após o processo de inpainting utilizando (BERTALMIO; BERTOZZI; SAPIRO, 2001), e a imagem no canto inferior direito mostra a imagem após o processo de inpainting utilizando (TELEA, 2004).

2.3.4 Context-aware inpainting

2.3.5 Representação esparsa

2.4 Otimizações computacionais

2.4.1 Profiling

2.4.2 Aceleração de hardware

2.4.3 Crop

2.4.4 Redução de memória

2.4.5 Paralelismo

2.4.6 Estruturas de dados

2.4.7 Redução de redundância

2.4.8 Otimização de acesso a memória

2.4.9 Vetorialização

2.4.10 Aproximações

2.4.11 Cálculos incrementais

2.5 Aplicações

3 Trabalhos Relacionados

entes computacionalmente e podem produzir excelentes resultados para imagens com texturas e estruturas simples. Os métodos baseados em aprendizado, por sua vez, utilizam técnicas de aprendizado de máquina para preencher a região alvo e podem produzir resultados mais realistas, incluindo aquelas com texturas e estruturas mais complexas. Contudo, estes métodos são mais complexos de serem implementados e treinados, e computacionalmente mais custosos.

5 Conclusão

6 Referências bibliográficas

- BARNES, C.; SHECHTMAN, E.; FINKELSTEIN, A.; GOLDMAN, D. B. PatchMatch: A randomized correspondence algorithm for structural image editing. **ACM Transactions on Graphics (Proc. SIGGRAPH)**, v. 28, n. 3, ago. 2009.
- BERTALMIO, M.; BERTOZZI, A.; SAPIRO, G. Navier-stokes, fluid dynamics, and image and video inpainting. In: **Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001**. [S.l.: s.n.], 2001. v. 1, p. I–I.
- BERTALMIO, M.; SAPIRO, G.; CASELLES, V.; BALLESTER, C. Image inpainting. In: **Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH 2000)**. [S.l.: s.n.], 2000.
- BERTALMIO, M.; VESE, L.; SAPIRO, G.; OSHER, S. Simultaneous structure and texture image inpainting. **IEEE Transactions on Image Processing**, v. 12, n. 8, p. 882–889, 2003.
- BLACK, S.; KESHAVER, S.; SOUVENIR, R. Evaluation of image inpainting for classification and retrieval. In: **2020 IEEE Winter Conference on Applications of Computer Vision (WACV)**. [S.l.: s.n.], 2020. p. 1049–1058.
- COMMONS, W. **File:Restoration.jpg — Wikimedia Commons, the free media repository**. 2022. [Online; accessed 18-January-2023]. Disponível em: <https://commons.wikimedia.org/w/index.php?title=File:Restoration.jpg&oldid=652328898>.
- CRIMINISI, A.; PÉREZ, P.; TOYAMA, K. Region filling and object removal by exemplar-based image inpainting. **IEEE Transactions on image processing**, IEEE, v. 13, n. 9, p. 1200–1212, 2004.
- ELHARROUSS, O.; ALMAADEED, N.; AL-MAADEED, S.; AKBARI, Y. Image inpainting: A review. **Neural Processing Letters**, Springer Science and Business Media LLC, v. 51, n. 2, p. 2007–2028, dec 2019. ISSN 1573-773X.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.
- GOOGLE. **Mediapipe-based library to redact faces from videos and images**. 2022. Disponível em: <https://google.github.io/magritte/>.
- HUANG, J.; KANG, S.; AHUJA, N.; KOPF, J. Image completion using planar structure guidance. **ACM Transactions on Computer Systems**, Association for Computing Machinery (ACM), v. 33, n. 4, 2014. ISSN 0730-0301.
- LEVIN; ZOMET; WEISS. Learning how to inpaint from global image statistics. In: **Proceedings Ninth IEEE International Conference on Computer Vision**. [S.l.: s.n.], 2003. p. 305–312 vol.1.
- MAALOUF, A.; CARRE, P.; AUGEREAU, B.; FERNANDEZ-MALOIGNE, C. A bandelet-based inpainting technique for clouds removal from remotely sensed images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 47, n. 7, p. 2363–2371, 2009.

OPENCV. **Open Source Computer Vision Library**. 2000. [⟨https://github.com/opencv/opencv⟩](https://github.com/opencv/opencv).

_____. **OpenCV: Image inpainting**. 2022. [⟨https://docs.opencv.org/4.7.0/df/d3d/tutorial_py_inpainting.html⟩](https://docs.opencv.org/4.7.0/df/d3d/tutorial_py_inpainting.html).

PATHAK, D.; KRÄHENBÜHL, P.; DONAHUE, J.; DARRELL, T.; EFROS, A. Context encoders: Feature learning by inpainting. In: **Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2016.

SALEM, N. A survey on various image inpainting techniques. 09 2021.

SHIH, M.-L.; SU, S.-Y.; KOPE, J.; HUANG, J.-B. **3D Photography using Context-aware Layered Depth Inpainting**. arXiv, 2020. Disponível em: [⟨https://arxiv.org/abs/2004.04727⟩](https://arxiv.org/abs/2004.04727).

TELEA, A. An image inpainting technique based on the fast marching method. **Journal of Graphics Tools**, v. 9, 01 2004.

WANG, Y.; CHEN, Y.-C.; TAO, X.; JIA, J. **VCNet: A Robust Approach to Blind Image Inpainting**. arXiv, 2020. Disponível em: [⟨https://arxiv.org/abs/2003.06816⟩](https://arxiv.org/abs/2003.06816).

YOU, S.; YOU, N.; PAN, M. Pi-rec: Progressive image reconstruction network with edge and color domain. **arXiv preprint arXiv:1903.10146**, 2019.