

CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC		
(X) PRÉ-PROJETO	() PROJETO	ANO/SEMESTRE: 2021/2

RECONSTRUÇÃO DE SUPERFÍCIES 3D A PARTIR DE IMAGENS DERMATOLÓGICAS 2D

Gabriel Castellani de Oliveira

Prof. Aurélio Faustino Hoppe - Orientador

1 INTRODUÇÃO

O câncer é um termo dado para doenças malignas que possuem em comum o crescimento desordenado de células (INSTITUTO NACIONAL DE CâNCER, 2020). Segundo o Instituto Nacional de Câncer (INCA), “pessoas de pele clara, sensíveis à ação dos raios solares, com história pessoal ou familiar deste câncer ou com doenças cutâneas prévias são as mais atingidas”. No Brasil, o câncer de pele é o mais frequente entre todos os tumores malignos registrados no país, estimando-se que cerca de 3% são do tipo melanoma e 30% para o não melanoma (INSTITUTO NACIONAL DE CâNCER, 2021).

Segundo Domingues *et al.* (2018), o melanoma surge de mutações genéticas nos melanócitos, células produtoras de pigmentos. Esse câncer é mais frequente em adultos brancos, sendo o mais grave e com alta probabilidade de disseminar-se para outros órgãos. No último ano, somente no Brasil, surgiram cerca de 8.450 novos casos, sendo 4.200 homens e 4.250 mulheres. Diferentemente do melanoma, o câncer de pele não melanoma possui altos percentuais de cura e uma mortalidade baixa entre os tumores de pele (INSTITUTO NACIONAL DE CâNCER, 2021).

Ainda de acordo com Instituto Nacional de Câncer (2021), o diagnóstico desta doença é feito por um dermatologista, através de exames clínicos, laboratoriais, endoscópios ou radiológicos. Em certas situações, o especialista utiliza um dermatoscópio para visualizar algumas camadas de pele que não podem ser vistas a olho nu. Além disso, os dermatologistas usam uma regra adotada internacionalmente para apontar sinais sugestivos de câncer de pele melanoma, conhecida como regra ABCDE, que representa assimetria, bordas irregulares, cor variável, diâmetro e evolução (INSTITUTO NACIONAL DE CâNCER, 2021).

Segundo Póvoa e Andrade (2006, p. 1555), “a distribuição geográfica dos médicos influencia o bem-estar social uma vez que estes são os principais provedores dos serviços de saúde”.

Particularmente no Brasil, a distribuição dos dermatologistas é bastante irregular, visto que 63,5% deles se concentram na Região Sudeste, que tem apenas 41,6% da população brasileira. Há 5.058 dermatologistas em atividade, que assistem mais de 190 milhões de pessoas numa área de 8.514.876,599 km (YARAK; RUIZ; PISA, 2017, p. 348).

Diante desse cenário, o uso da tele dermatologia torna-se importante para o Brasil, visto que sua proporção ocasiona em certas regiões, a falta de acesso à consulta especializada em dermatologia (YARAK; RUIZ; PISA, 2017, p. 348). Segundo Miot, Paixão e Wen (2005), “tele dermatologia representa o envio de informações médica dermatológica entre dois ou mais pontos separados fisicamente, utilizando tecnologia de telecomunicação e informática, visando à promoção de saúde e educação de pacientes, paramédicos ou profissionais médicos”.

A tele dermatologia proporciona o acesso a consultas dermatológicas em áreas que antes era impossível, além de reduzir os custos (YARAK; RUIZ; PISA, 2017, p. 348). Embora a tele dermatologia apresente muitos benefícios na sua utilização, existem as limitações no uso dessa abordagem. Entre as limitações enfrentadas, encontra-se a precisão, a qualidade das fotos capturadas e a representação adequada da lesão cutânea (RISMILLER; CARTRON; TRINIDAD, 2020). A qualidade das fotos influencia decisivamente no diagnóstico, uma das queixas mais frequentes nas imagens é a identificação de regiões como o pavilhão auricular, cujo local é de difícil visualização (SILVA *et al.*, 2009).

Um recurso comum da medicina é a Tomografia computadorizada, utilizada para aplicações clínicas no início da década de 70 para examinar o encéfalo, o sistema ventricular e as partes ósseas do crânio. Essa abordagem é baseada em raio-X, onde se obtém o crânio em fatias (AMARO JÚNIOR; YAMASHITA, 2001). Porém, tal recurso não é aplicado em lesões de pele. Segundo Benalcazar *et al.* (2020), para realizar uma reconstrução 3D é necessário: (i) um conjunto de imagens da região a ser reconstruída em diferentes perspectivas (ii) reconstruir uma nuvem de pontos 3D a partir das imagens obtidas e (iii) converter a nuvem de pontos em uma superfície.

Neste contexto, a reconstrução tridimensional da superfície possibilita a interação do profissional com a estrutura reconstruída, permitindo a rotação e a escala do objeto reconstruído tridimensionalmente (KLEMT;

INFANTOSI, 2011). Segundo Benalcazar *et al.* (2020), o resultado de uma reconstrução 3D contém informações de profundidade e cor, possibilitando uma visualização mais detalhada, por exemplo, de uma lesão de pele. Meschender *et al.* (2019) mencionam que as abordagens baseadas em aprendizagem para reconstrução 3D se popularizaram mais do que os algoritmos estéreos *multi-view* tradicionais, pois conseguem codificar informações prévias sobre o espaço das formas 3D, assim, resolvendo problemas de ambiguidades nas imagens de entrada. Além disso, os autores comentam que os modelos generativos alcançaram sucessos notáveis na geração de imagens de alta resolução para um domínio 2D, porém, para um domínio 3D ainda não foi alcançado.

Diante do cenário exposto, este trabalho propõe a construção de um método que seja capaz de auxiliar no diagnóstico de doenças de pele, visando melhorar a precisão e a qualidade das informações necessárias para um diagnóstico mais certo.

1.1 OBJETIVOS

Este trabalho tem como objetivo disponibilizar um método capaz de realizar a reconstrução de superfície 3D a partir de imagens dermatológicas 2D.

Os objetivos específicos são:

- a) detectar as lesões cutâneas utilizando técnicas de aprendizado de máquina;
- b) estimar a profundidade das estruturas internas das lesões cutâneas a partir de algoritmos de visão computacional;
- c) disponibilizar um mecanismo para visualização de lesão de forma volumétrica.

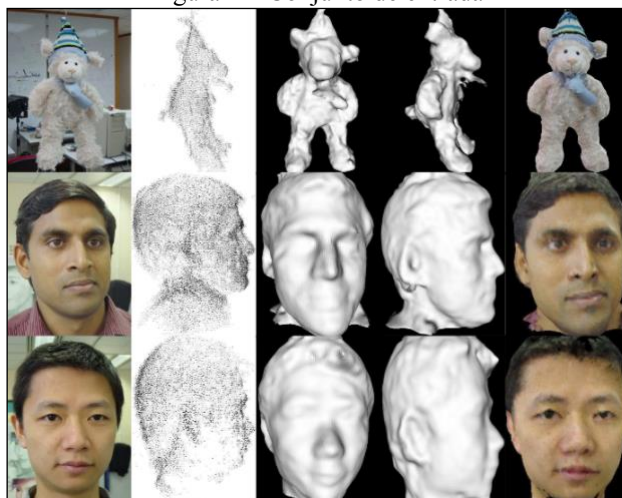
2 TRABALHOS CORRELATOS

Nesse capítulo serão apresentados trabalhos com características semelhantes aos principais objetivos do estudo proposto. A seção 2.1 apresenta o desenvolvimento de uma abordagem para realizar a reconstrução de superfícies, através da propagação de dados estéreos 3D e imagens 2D (ZENG *et al.*, 2004). A seção 2.2 aborda uma rede de ocupação para a reconstrução de objeto 3D (MESCHENDER *et al.*, 2019). Por fim, na seção 2.3 será apresentado uma abordagem para reconstruir objetos com base em uma única imagem de entrada (FAN; SU; GUIBAS, 2017).

2.1 SURFACE RECONSTRUCTION BY PROPAGATING 3D STEREO DATA IN MULTIPLE 2D IMAGES

Zeng *et al.* (2004) construíram um *framework* para reconstrução de superfície através da propagação de dados estéreo 3D utilizando várias imagens 2D. O *framework* analisa o ciclo de propagação da superfície, explorando a integração dos pontos estéreo e as imagens calibradas. Segundo os autores, inicialmente instancia-se uma lista de pontos estéreos 3D que serão utilizados para realizar a propagação da superfície. Estes pontos são computados a partir de um conjunto de imagens de entrada, onde são coletadas em torno de 30 imagens ao redor de cada objeto, conforme exemplificado na Figura 1.

Figura 1 – Conjunto de entrada

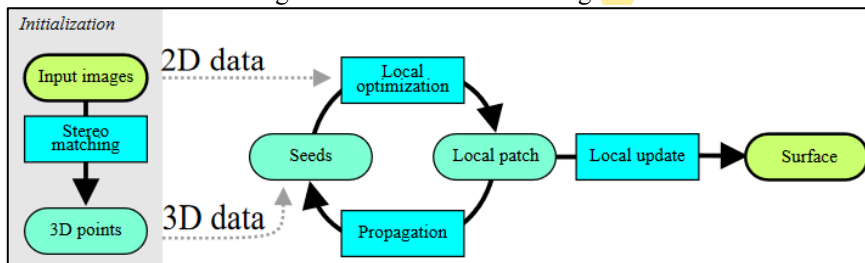


Fonte: Zeng *et al.* (2004).

Zeng *et al.* (2004) destacam que o ciclo de propagação da superfície realiza correções da superfície com base nos pontos 3D obtidos. A Figura 2 demonstra que este ciclo possui três etapas: (i) seleção do ponto 3D para

a propagação (ii) geração de um remendo a partir do ponto escolhido e (iii) geração de novos pontos a partir do remendo criado. Segundo os autores, essas etapas serão executadas para cada ponto da lista, fazendo com que a superfície cresça até atingir o resultado desejado. O ciclo de propagação termina quando não existirem mais pontos disponíveis a serem propagados.

Figura 2 – Estrutura da abordagem.

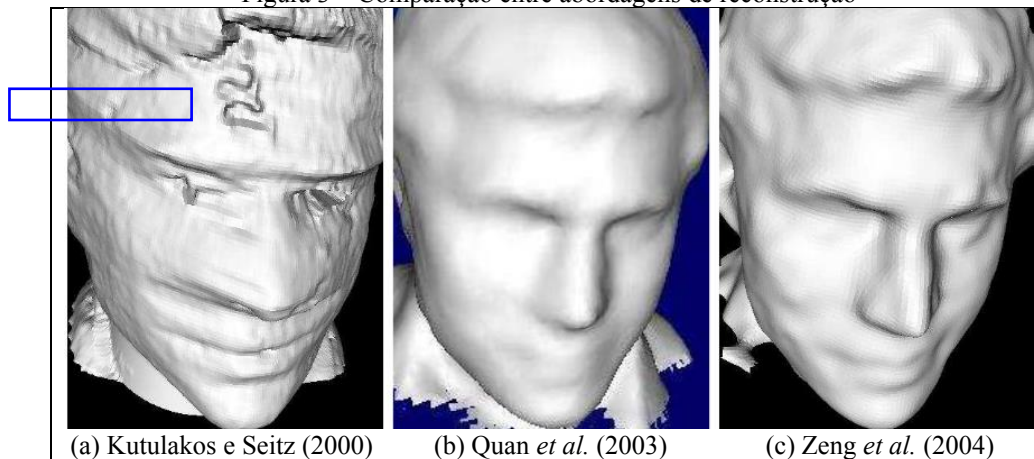


Fonte: Zeng *et al.* (2004)

Segundo Zeng *et al.* (2004), a primeira etapa da propagação seleciona o melhor ponto da lista para servir como referência para o remendo. Para isso, os autores utilizam o critério π para avaliar se o ponto escolhido é a melhor opção para a propagação, no qual os pontos iniciais da lista são considerados como pontos confiáveis. A partir do ponto escolhido, a segunda etapa da propagação é executada, onde o algoritmo gera um remendo ao redor do ponto escolhido para estender a superfície existente. Por fim, a última etapa gera novos pontos para garantir a propagação da superfície, onde são criados a partir do remendo anterior. Os autores comentam que a localização de um novo ponto é determinada pelos seguintes aspectos: (i) qualidade do remendo (ii) qualidade de correspondência (iii) regularidade da superfície e (iv) eficiência da propagação. Além disso, os autores comentam que a quantidade de novos pontos é inspirada na configuração de malha triangular, onde a distância angular entre dois pontos vizinhos fica entre $\left[\frac{2\pi}{5}, \frac{2\pi}{7}\right]$.

Zeng *et al.* (2004) aplicaram o *framework* em três conjuntos de entrada. No primeiro conjunto tem-se a representação de um brinquedo, no qual os autores examinaram a exatidão e a robustez da reconstrução. Os dois últimos conjuntos são representações de rostos humanos. Percebe-se a partir da Figura 1 que o primeiro rosto possui mais textura do que o segundo. Os autores comentam que o resultado da reconstrução da superfície obteve para os dois conjuntos praticamente a mesma qualidade.

Figura 3 – Comparação entre abordagens de reconstrução



Fonte: Zeng *et al.* (2004).

Zeng *et al.* (2004) também realizaram um comparativo entre o método de escultura espacial (KUTULAKOS; SEITZ, 2000) que utiliza apenas informações de imagens 2D para a reconstrução da superfície e o método de conjunto de níveis (QUAN *et al.*, 2003) e a abordagem desenvolvida, que ambas utilizam informações 3D e 2D para a reconstrução, o resultado dessa comparação pode ser observado na Figura 3. Os autores destacam que o método de escultura espacial (KUTULAKOS; SEITZ, 2000) perde muitos detalhes do rosto, por conta de usar apenas dados 2D para a reconstrução. Já o resultado do método de conjunto de níveis (QUAN *et al.*, 2003), utiliza as mesmas informações de entrada que a abordagem desenvolvida pelos autores. Porém, (QUAN *et al.*, 2003) aplicam um método de definição de nível que suaviza a superfície, perdendo detalhes geométricos. Contudo, o método desenvolvido por Zeng *et al.* (2004) conseguiu gerar um resultado sem perder os detalhes e as características do objeto de entrada, assim, gerando a superfície com precisão e robustez.

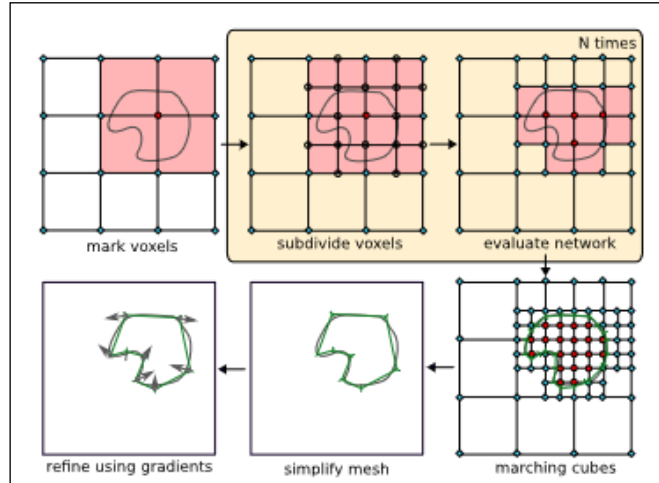
2.2 OCCUPANCY NETWORKS: LEARNING 3D RECONSTRUCTION IN FUNCTION SPACE

Meschender *et al.* (2019) propuseram uma abordagem que não necessita de um modelo do objeto de entrada como referência. Tal abordagem utiliza aprendizado profundo para obter uma representação mais expressiva, podendo ser integrado com um *pipeline* de aprendizado de ponta a ponta. Além disso, os autores desenvolveram uma função de ocupação 3D contínua para reconstrução de objetos 3D, sendo equivalente a uma rede neural para classificação binária, mas com o objetivo de decidir implicitamente as fronteiras das superfícies de um objeto.

Segundo Meschender *et al.* (2019), a função de ocupação do objeto 3D é denotada pelo resultado da função $o : R^3 \rightarrow \{0, 1\}$, no qual os autores aproximam essa função com uma rede neural que atribui para cada local $p \in R^3$ uma probabilidade de ocupação entre 0 e 1. Além disso, os autores enfatizam a necessidade de condicioná-la na entrada quando forem aplicar para reconstruções de objetos 3D com base em observações (imagem, nuvem de pontos etc.). Para realizar esse processo de condicionamento, Meschender *et al.* (2019) utilizaram uma equivalência funcional simples: onde é representada por uma função que recebe um par $(p, x) \in R^3 \times X$ como entrada e produz um número real como saída. Essa representação pode ser parametrizada por uma rede neural f_θ , onde recebe um par (p, x) como entrada e gera uma probabilidade de ocupação. Os autores chamam a rede $f_\theta : R^3 \times X \rightarrow [0, 1]$ de rede de ocupação.

Meschender *et al.* (2019) também desenvolveram um algoritmo hierárquico *Multiresolution IsoSurface Extraction* (MISE), que utiliza a rede de ocupação para extrair uma isosuperfície correspondente a uma observação. Inicialmente esse algoritmo discretiza o espaço volumétrico e avalia a rede de ocupação $f_\theta(p, x)$ para todos os p da grade, marcando todos os pontos da grade como ocupados se o resultado da ocupação for maior ou igual a 0 ou 1, os demais pontos que não entraram nessa condição são marcados como desocupados. Em seguida, o algoritmo marca como ativo todos os *voxels* que possuem pelo menos dois pontos adjacentes com diferentes previsões de ocupação. Neste mesmo momento, os *voxels* ativos são subdivididos em 8 *subvoxels*, gerando novos pontos na grade de ocupação que posteriormente serão avaliados. Por fim, o algoritmo repete essas etapas até alcançar a resolução desejada. Essa estrutura pode ser observada na Figura 4, onde os pontos ocupados estão marcados com círculos vermelhos, os pontos desocupados em diamantes cianos, os *voxels* ativos em vermelho claro e os novos pontos marcados com círculos vazios.

Figura 4 – Algoritmo MISE.



Fonte: Meschender *et al.* (2019).

Após obter a malha resultante das operações descritas acima, Meschender *et al.* (2019) apontam que a malha passa por um processo de extração, onde inicialmente é aplicado o algoritmo *Marching cubes* (LORENSEN; CLINE, 1987) que cria modelos triangulares de superfícies constantes a partir de dados 3D e, em seguida, a malha extraída passa pelo algoritmo *Fast-Quadric-Mesh-Simplification* (GARLAND; HECKBERT, 2000) que simplifica a superfícies poligonais. Além disso, os autores também utilizaram informações de gradientes de primeira e segunda ordem para realizar um refinamento da malha de saída. Para isso, os autores aplicaram pontos p_k aleatórios em cada face da malha de saída e minimizaram a perda através da equação descrita na Figura 5, onde $n(p_k)$ denota o vetor normal da malha em p_k e λ é definido como $\lambda = 0.01$. Para a minimização do segundo termo da equação, Meschender *et al.* (2019) utilizaram informações de gradiente de segunda ordem que pode ser aplicada usando *Double-Backpropagation* (DRUCKER; CUN, 1992). Os autores ressaltam que essa última etapa remove os artefatos da discretização do algoritmo *Marching cubes* (LORENSEN; CLINE, 1987) e que não seria possível apenas com uma representação baseada em *voxel*.

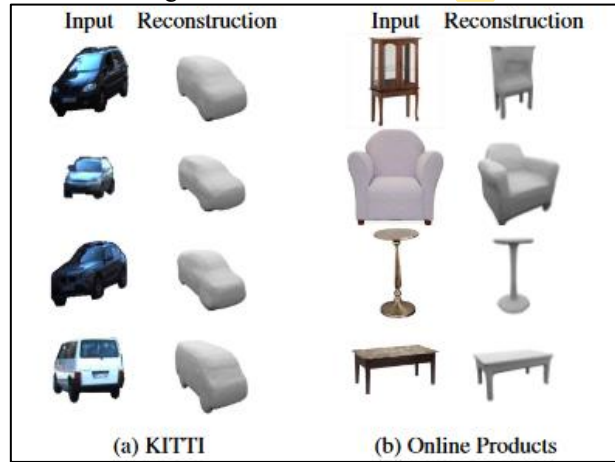
Figura 5 – Operação de minimização

$$\sum_{k=1}^K (f_{\theta}(p_k, x) - \tau)^2 + \lambda \left\| \frac{\nabla_p f_{\theta}(p_k, x)}{\|\nabla_p f_{\theta}(p_k, x)\|} - n(p_k) \right\|^2$$

Fonte: Meschender *et al.* (2019).

Meschender *et al.* (2019) aplicaram a rede de ocupação nos conjuntos de dados KITTI (GEIGER *et al.*, 2013) e *Online Products* (SONG *et al.*, 2015), para verificar o comportamento da rede com dados reais. Os autores ressaltam que para os dados KITTI (GEIGER *et al.*, 2013) foi necessário aplicar a máscara de instância (ALHAIJA *et al.*, 2018) para cortar as regiões pertencentes a um carro, aplicando em seguida a rede de ocupação sobre as imagens. Para o conjunto de dados do *Online Products* (SONG *et al.*, 2015), os autores aplicaram o mesmo modelo pré-treinado. Na Figura 6 pode-se ver o resultado da rede de ocupação para os conjuntos de dados KITTI (GEIGER *et al.*, 2013) e *Online Products* (SONG *et al.*, 2015).

Figura 6 – Resultados dos testes.



Fonte: Meschender *et al.* (2019).

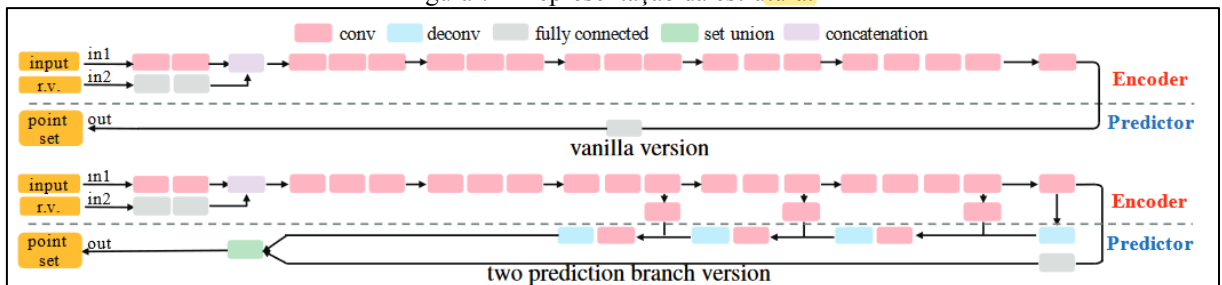
Meschender *et al.* (2019) concluem que a rede de ocupação consegue gerar reconstruções semelhantes ao objeto de entrada, gerando malhas realistas de alta resolução. Além disso, os autores comentam que a partir dos experimentos realizados, foi possível observar que a rede de ocupação pode ser utilizada para os aprendizados supervisionados e os não supervisionados.

2.3 A POINT SET GENERATION NETWORK FOR 3D OBJECT RECONSTRUCTION FROM A SINGLE IMAGE

Fan, Su e Guibas (2017) utilizaram redes generativas para realizar a reconstrução 3D com base em uma única imagem de entrada. Os autores destacam que a rede estima a profundidade das partes visíveis do objeto na imagem. Fan, Su e Guibas (2017) também ressaltam que a representação por nuvem de pontos não é a ideal para a reconstrução 3D, porém, possui uma estrutura simples e uniforme para o aprendizado.

De acordo com Fan, Su e Guibas (2017), construir uma rede generativa para uma representação em pontos é um processo desafiador, pois a representação é desordenada e possui ambiguidades. Para isso, os autores desenvolveram uma arquitetura composta por um gerador de conjunto de pontos, função de perda para comparação do conjunto de pontos e a modelagem da incerteza da verdade fundamental (partes invisíveis do objeto), conforme exibe a Figura 7.

Figura 7 – Representação da estrutura.



Fonte: Fan, Su e Guibas (2017).

Segundo Fan, Su e Guibas (2017), a primeira parte da arquitetura consiste em uma rede para representar estruturas e estatísticas dos dados geométricos, tendo dois estágios:

- a) codificador: mapeia a imagem de entrada e um vetor aleatório em um espaço de incorporação;
- b) preditor: emite uma matriz $N \times 3$, onde cada linha contém as coordenadas de um ponto.

Fan, Su e Guibas (2017) apontam que o *design* do ramo preditor foi modificado para suportar superfícies grandes e lisas, pois a estrutura do *vanilla version* não faz uso completo das estatísticas geométricas. Esse novo *design* possui dois ramos preditores paralelos, sendo um totalmente conectado e o outro de deconvolução. Segundo os autores, o ramo totalmente conectado continua prevendo pontos N_1 e o ramo de deconvolução prevê uma imagem com três canais de tamanho $H \times W$, onde os valores de cada pixel representam a coordenada de um ponto.

Fan, Su e Guibas (2017) descrevem que a segunda parte da arquitetura é uma função de perda, que serve para comparar a nuvem de pontos prevista com a parte invisível do objeto gerado. Além disso, os autores listaram três condições para definir uma distância adequada para conectar a uma rede neural: (i) diferenciável em relação às localizações dos pontos (ii) eficiente para computar e (iii) robusto para um número pequeno de pontos discrepantes. Com base nessas condições, os autores propuseram duas funções:

- a) *chamfer distance*: para cada ponto, a função localiza o vizinho mais próximo no outro conjunto e soma as distâncias ao quadrado;
- b) *earth Mover's distance*: o cálculo da função em si é muito custoso para um aprendizado profundo. Para isso, os autores implementaram um esquema de aproximação desenvolvida originalmente por Bertsekas (1985).

De acordo com Fan, Su e Guibas (2017), a terceira parte da arquitetura é responsável por realizar a reconstrução da parte invisível do objeto, onde os autores encontraram problemas de regressão e ambiguidades. Para modelar a melhor representação do objeto oculto, os autores aplicaram o método *Min-of-N* que calcula o mínimo de n distancias.

Nos experimentos, Fan, Su e Guibas (2017) utilizaram o conjunto de dados *ShapeNet* (CHANG *et al.*, 2015) que possui um grande volume de modelos de objetos 3D, ao qual foram escolhidas imagens que possuem iluminação simples. Em cada modelo, os autores normalizam o hemisfério delimitador e alinham seu plano base. Em seguida, o modelo da imagem 2D é renderizado utilizando a fórmula de sombreado de *Blinn-Phong*. Fan, Su e Guibas (2017) compararam a abordagem desenvolvida com o 3D-R2N2 (CHOY *et al.*, 2016) que utiliza aprendizado profundo para reconstruir representações 3D a partir de imagens com visualização única. As comparações podem ser visualizadas na Figura 8.

Figura 8 – Resultados da comparação.



Fonte: Fan, Su e Guibas (2017).

Com base nos resultados obtidos, Fan, Su e Guibas (2017) relatam que a abordagem 3D-R2N2 (CHOY *et al.*, 2016) geralmente não consegue reconstruir as características finas de um objeto, gerando uma reconstrução

incompleta do objeto de entrada. Já a abordagem dos autores, conseguiu reconstruir o objeto com suas características finas e as partes ocultas da imagem a partir de uma visão única.

3 PROPOSTA DO MÉTODO

Esse capítulo tem como objetivo apresentar a proposta e a justificativa para a elaboração desse trabalho, assim como os requisitos e metodologias que serão adotadas. Também serão apresentadas breves revisões bibliográficas das principais áreas de estudo que serão explorados.

3.1 JUSTIFICATIVA

No Quadro 1 é possível observar um comparativo das abordagens escolhidas para a resolução dos problemas propostos pelos respectivos trabalhos. As linhas representam as características e as colunas os trabalhos.

Quadro 1 - Comparativo dos trabalhos correlatos

Trabalhos Correlatos Características	Zeng <i>et al.</i> (2004)	Meschender <i>et al.</i> (2019)	Fan, Su e Guibas (2017)
Objetivo	Reconstruir superfícies através da propagação dos pontos estéreo 3D e imagens 2D.	Reconstruir superfícies utilizando uma rede neural.	Reconstruir um objeto a partir de uma única imagem.
Base de dados pública	Não	Sim	Sim
Informações de entrada	Pontos estéreos 3D e imagens 2D	Imagens 2D	Imagem 2D
Algoritmos/técnicas utilizadas	Ciclo de propagação da superfície	Multiresolution IsoSurface Extraction, Fast-Quadric-Mesh-Simplification, Double-Backpropagation	Redes generativas
Cenário de teste	Conjunto de pontos (brinquedo e rostos)	KITTI e Online Products	ShapeNet

Fonte: elaborado pelo autor.

A partir do Quadro 1 é possível observar que Meschender *et al.* (2019) e Fan, Su e Guibas (2017) mantiveram seu foco na reconstrução do objeto utilizando apenas dados 2D, tendo o trabalho de Fan, Su e Guibas (2017) o diferencial de receber apenas uma imagem 2D como entrada, e prever o restante do objeto. Já Zeng *et al.* (2004) utilizaram tanto dados 3D quanto dados 2D para a reconstrução.

Todos os trabalhos possuem uma abordagem diferente para realizar a reconstrução 3D. Zeng *et al.* (2004) desenvolveram um *framework* que realiza uma propagação no espaço 3D orientada por informações 2D, criando remendos da superfície a partir de um ponto estéreo 3D escolhido na propagação. Esse *framework* foi testado em três conjuntos de dados, sendo dois rostos humanos e um brinquedo. Todos os objetos foram reconstruídos por completo com exatidão e robustez. Porém, é possível perceber que algumas partes não ficaram com uma boa qualidade de reconstrução.

Já a abordagem de Meschender *et al.* (2019) utiliza uma rede de ocupação em conjunto com o algoritmo *Multiresolution IsoSurface Extraction*, que verifica todos os pontos ocupados da superfície e retornam uma malha resultante, aplicando em seguida os algoritmos *Marching cubes* (LORENSEN; CLINE, 1987), *Fast-Quadric-Mesh-Simplification* (GARLAND; HECKBERT, 2000) e *Double-Backpropagation* (DRUCKER; CUN, 1992), para extrair, simplificar e refinar a malha de saída. Essa abordagem foi utilizada para reconstruir objetos reais obtidos através da base de dados KITTI (GEIGER *et al.*, 2013) e *Online Products* (SONG *et al.*, 2015), onde a reconstrução ficou bem parecido com o objeto de entrada.

Por fim, o trabalho de Fan, Su e Guibas (2017) utilizou-se de redes generativas para reconstruir superfícies a partir de uma imagem 2D de entrada. Inicialmente, essa abordagem reconstrói a primeira parte do objeto com as informações presentes na imagem de entrada, em seguida, a parte do objeto oculta na imagem. Essa abordagem é aplicada sobre os conjuntos de dados *ShapeNet* (CHANG *et al.*, 2015). Percebe-se nos resultados que a rede conseguiu reconstruir objetos com características finas e as partes ocultas apenas com uma única imagem.

Diante deste contexto, esse trabalho se torna relevante pelo fato de reconstruir superfícies 3D para imagens dermatológicas, visto que uma imagem 2D não oferece a qualidade necessária para um diagnóstico certo. Essa reconstrução possui alta complexidade, devido a sensibilidade das informações para a tridimensionalidade. No entanto, uma representação 3D da lesão possibilita aos dermatologistas visualizar os relevos e outras características que não podem ser observadas através de uma imagem 2D. Além disso, por ser uma representação 3D vista através de um computador, os atendimentos aos pacientes não precisam ser realizados presencialmente, possibilitando o

uso do atendimento remoto (tele dermatologia). Dessa forma, agiliza as consultas, reduz os custos e permite a cobertura da saúde para regiões com carência em médicos especializados em dermatologia ou regiões de difícil acesso pelos profissionais.

3.2 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

O método para a reconstrução de superfícies em 3D deverá:

- a) permitir ao usuário carregar imagens dermatológicas a serem analisadas (Requisito Funcional – RF);
- b) realizar o realce e melhoramento das imagens (ruídos, distorções e problemas de iluminação) utilizando técnicas de processamento de imagens (RF);
- c) utilizar técnicas de aprendizado de máquina para a detecção das lesões cutâneas (RF);
- d) utilizar métricas de cortes para obter a superfície a ser reconstruída (RF);
- e) permitir ao usuário visualizar a reconstrução 3D considerando elevação e textura (RF);
- f) ser desenvolvida na linguagem de programação Python (Requisito Não Funcional – RNF);
- g) utilizar as bibliotecas OpenCV para o processamento de imagens e TensorFlow para o aprendizado de máquina (RNF);
- h) ser capaz de realizar a análise em um tempo máximo de 1 minuto (RNF);
- i) reconstruir superfícies a partir de imagens de baixa resolução (RNF);
- j) utilizar a base de dados ShapeNet para obtenção das imagens de teste (RNF).

3.3 METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

- a) levantamento bibliográfico: buscar fontes bibliográficas sobre dermatologia, lesões cutâneas, detecção de objetos, reconstrução de superfícies em 3D e trabalhos correlatos;
- b) montagem da base de imagens: buscar em projetos *open sources* (código livre) ou privados, uma base de dados com imagens dermatológicas, como por exemplo, a ShapeNet. Estas imagens serão utilizadas para o treinamento e validação do método. Quanto mais imagens forem utilizadas, melhores devem ser os resultados, portanto, a busca por estas imagens pode se estender durante todo o desenvolvimento do trabalho;
- c) rotulação das imagens: rotular as imagens que serão utilizadas, permitindo assim que o algoritmo possa ter sua assertividade testada;
- d) pesquisa e escolha do algoritmo de segmentação: pesquisar os principais algoritmos de segmentação e detecção de objetos, escolhendo o adequado para o desenvolvimento do trabalho;
- e) **implementação**: implementar o algoritmo de detecção de lesões cutâneas utilizando a linguagem de programação Python e a biblioteca TensorFlow juntamente com o OpenCV;
- f) validação do algoritmo de segmentação: realizar testes em conjunto com especialistas para validar a taxa de acerto em relação ao processo manual;
- g) pesquisa e escolha do algoritmo de reconstrução: pesquisar os principais algoritmos de reconstrução 3D a partir de imagens 2D, escolhendo o mais adequado para o desenvolvimento do trabalho;
- h) **implementação**: a partir dos item (f) e (g) implementar o algoritmo de superfície 3D, utilizando a linguagem de programação Python no ambiente de desenvolvimento Visual Studio Code, representando e disponibilizando o resultado de maneira volumétrica;
- i) validação da reconstrução de superfícies em 3D: em paralelo ao desenvolvimento, verificar a coerência da reconstrução 3D em conjunto com especialistas, e caso necessário, alterar os requisitos para atender o problema a ser resolvido.

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro 2 – Cronograma de atividades a serem realizadas

etapas / quinzenas	2021				2022											
	nov.		dez.		jan.		fev.		mar.		abr.		maio		jun.	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
levantamento bibliográfico																
montagem da base de imagens																
rotulação das imagens																
pesquisa e escolha do algoritmo de segmentação																
implementação																
validação do algoritmo de segmentação																
pesquisa e escolha do algoritmo de reconstrução																
implementação																
validação da reconstrução de superfícies em 3D																

Fonte: elaborado pelo autor.

4 REVISÃO BIBLIOGRÁFICA

Este capítulo descreve brevemente sobre os assuntos que fundamentarão o estudo a ser realizado: dermatologia, lesões cutâneas, detecção de **objetos, reconstrução** de superfícies em 3D.

Inicialmente a palavra dermatologia era entendida como referência à anatomia da pele. Mais tarde, adquiriu o significado de uma área médica concentrada no diagnóstico, prevenções e tratamentos de doenças relacionadas à pele (HOENIG; PARISH, 2021). Segundo Miot, Paixão e Wen (2005), “a dermatologia lida com o diagnóstico de mais de duas mil condições que afetam a pele e colabora com as demais especialidades por meio do reconhecimento das manifestações cutâneas das doenças sistêmicas”. De acordo com o Instituto Nacional de Câncer (2021), o diagnóstico das doenças é feito por dermatologistas, através de exames clínicos. Porém, em alguns momentos é necessário que o profissional utilize a dermatoscopia, exame no qual é possível visualizar as camadas de pele que não podem ser vistas a olho nu. Além disso, em certos casos o uso da biópsia é necessário para o diagnóstico da doença.

Conforme o Instituto Nacional de Câncer (2021), a detecção precoce de um câncer, é uma estratégia que visa encontrar o tumor na fase inicial, assim, possibilitando um tratamento bem-sucedido. A regra ABCDE (assimetria, bordas irregulares, cor variável, diâmetro e evolução) é adotada internacionalmente para identificar possíveis tumores do tipo melanoma, onde cada letra representa uma informação que é avaliada para o diagnóstico.

De acordo com Szeliski (2019), a área de visão computacional vem sendo desenvolvida por pesquisadores, para permitir a recuperação da forma e aparência dos objetos em imagens. Os seres humanos conseguem diferenciar formas, adivinhar emoções ou até mesmo reconhecer pessoas a partir de uma imagem com muita facilidade. Com o avanço do desenvolvimento pelos pesquisadores, essas tarefas já podem ser realizadas por máquinas (SZELESKI, 2019, p. 3).

Diante deste contexto, surgiram as Redes Neurais Convolucionais, que vem do inglês **Convolutional Neural Network** (CNN). Essa rede foi projetada para processar dados em formato de *array*, como por exemplo, uma imagem 2D que possui três canais de cores, e foi utilizada no início de 1990 para identificar objetos em uma imagem (LECUN; BENGIO; HINTON, 2015). A Mask R-CNN é um exemplo de redes convolucionais que realiza a detecção de objetos em uma imagem 2D. Essa rede consegue distinguir os **pixels** pertencentes ao objeto detectado dos **pixels** não pertencentes ao objeto. Dessa forma, é possível detectar com precisão os objetos em imagens 2D (HE *et al.*, 2017). Além dela, também existe o modelo conhecido como U-Net que é comumente utilizada em reconhecimento de padrões de imagens médicas. O objetivo desse modelo é reduzir os erros humanos e obter um diagnóstico melhor e mais rápido, assim, garantindo o melhor tratamento para o paciente (ALOM *et al.*, 2019).

Szeliski (2019) ressalta que utilizando técnicas confiáveis, é possível calcular com precisão um modelo 3D de um ambiente a partir de várias imagens sobrepostas, ou até mesmo criar modelos de superfície 3D densas e precisas. Assim como nomear as pessoas em uma imagem utilizando combinações de detecção de rosto. Paris, Sillion e Quan (2006) desenvolveram uma abordagem para reconstruir superfícies usando um algoritmo de corte global, que realiza o recorte de superfície em um gráfico 3D, produzindo uma geometria de superfície detalhada.

De acordo com Meschender *et al.* (2019), a representação de uma saída 3D pode ser categorizada em (a) representações baseadas em *voxel*: são as representações mais comuns para tarefas 3D, porém, os autores afirmam que o consumo elevado da memória se torna um ponto negativo; (b) representações baseadas em pontos: são representações dadas por uma nuvem de pontos 3D, que inicialmente eram utilizadas para robótica e computação gráfica, (FAN; SU; GUIBAS, 2016) introduziram essa representação como forma de saída 3D para reconstruções de objetos a partir de uma imagem de entrada. Porém, os autores comentam que essa abordagem necessita de etapas não triviais de pós-processamento; (c) representações de malha: aplicadas inicialmente para classificações

3D discriminativa ou para tarefas de segmentação, porém, recentemente foram consideradas como representações de uma saída 3D.



REFERÊNCIAS

- ALHAJJA, Hassan Abu *et al.* **Augmented Reality Meets Computer Vision: efficient data generation for urban driving scenes.** *International Journal Of Computer Vision*, [S.l.], v. 126, n. 9, p. 961-972, 7 mar. 2018.
- ALOM, Md Zahangir *et al.* **Recurrent residual U-Net for medical image segmentation.** *Journal Of Medical Imaging*, [S.l.], v. 6, n. 01, p. 1-16, 27 mar. 2019.
- AMARO JÚNIOR, Edson; YAMASHITA, Helio. **Aspectos básicos de tomografia computadorizada e ressonância magnética.** *Revista Brasileira de Psiquiatria*, [S.l.], v. 23, n. 1, p. 2-3, maio 2001.
- BENALCAZAR, Daniel P. *et al.* **A 3D Iris Scanner From a Single Image Using Convolutional Neural Networks.** *Ieee Access*, [S.l.], v. 8, p. 98584-98599, maio 2020.
- BERTSEKAS, Dimitri P. **A distributed asynchronous relaxation algorithm for the assignment problem.** In: 1985 24th IEEE Conference on Decision and Control. IEEE, 1985. p. 1703-1704
- CHANG, Angel X. *et al.* **ShapeNet: An information-rich 3d model repository.** *arXiv preprint arXiv:1512.03012*, 2015.
- CHOY, Christopher B. *et al.* **3D-R2N2: a unified approach for single and multi-view 3d object reconstruction.** *Computer Vision – Eeccv 2016*, [S.l.], v. 1, n. 0, p. 628-644, out. 2016.
- DOMINGUES, Beatriz *et al.* **Melanoma treatment in review.** *Immunotargets And Therapy*, [S.L.], v. 7, p. 35-49, jun. 2018. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5995433/>>. Acesso em: 01 out. 2021.
- DRUCKER, H.; CUN, Y. Le. **Improving generalization performance using double backpropagation.** *Ieee Transactions On Neural Networks*, [S.l.], v. 3, n. 6, p. 991-997, nov. 1992.
- FAN, Haoqiang; SU, Hao; GUIBAS, Leonidas J. **A Point Set Generation Network for 3D Object Reconstruction from a Single Image.** In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 605-613.
- GARLAND, Michael; HECKBERT, Paul S. **Simplifying surfaces with color and texture using quadric error metrics.** In: *Proceedings Visualization'98* (Cat. No. 98CB36276). IEEE, 1998. p. 263-269.
- GEIGER, Andreas *et al.* **Vision meets robotics: The KITTI dataset.** *The International Journal Of Robotics Research*, [S.l.], v. 32, n. 11, p. 1231-1237, 23 ago. 2013.
- HE, Kaiming *et al.* **Mask R-CNN.** *IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988, 2017.
- HOENIG, Leonard J.; PARISH, Lawrence Charles. **How dermatology got its name.** *Clinics In Dermatology*, [S.l.], v. 39, n. 2, p. 353-354, mar. 2021.
- INSTITUTO NACIONAL DE CÂNCER. **O que é câncer?** 2020. Disponível em: <<https://www.inca.gov.br/o-que-e-cancer/>>. Acesso em: 30 set. 2021.
- INSTITUTO NACIONAL DE CÂNCER. **Câncer de pele melanoma.** 2021. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-pele-melanoma/>>. Acesso em: 30 set. 2021.
- KLEMT, A.; INFANTOSI, Antonio Fernando Catelli. **Método de superfície na visualização 3D da dissecação do crânio humano.** *Research on Biomedical Engineering*, v. 16, n. 1, p. 21-37, 2011
- KUTULAKOS, Kiriakos N.; SEITZ, Steven M. **A Theory of Shape by Space Carving.** *International Journal Of Computer Vision*, [S.l.], v. 38, n. 3, p. 199-218, jan. 2000.
- LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. **Deep learning.** *Nature*, [S.l.], v. 521, n. 7553, p. 436-444, 27 maio 2015.
- LORENSEN, William E.; CLINE, Harvey E. **Marching cubes: a high resolution 3d surface construction algorithm.** *Acm Siggraph Computer Graphics*, [S.l.], v. 21, n. 4, p. 163-169, ago. 1987.
- MIOT, Hélio Amante; PAIXÃO, Maurício Pedreira; WEN, Chao Lung. **Tele dermatologia: passado, presente e futuro.** *Anais Brasileiros de Dermatologia*, [S.l.], v. 80, n. 5, p. 523-532, out. 2005.
- MESCHENDER, Lars *et al.* **Occupancy Networks: Learning 3d Reconstruction in Function Space.** *2019 Ieee/Cvf Conference On Computer Vision And Pattern Recognition (Cvpr)*, [S.l.], v. 2, p. 4460-4470, jun. 2019.
- PARIS, Sylvain; SILLION, François X.; QUAN, Long. **A Surface Reconstruction Method Using Global Graph Cut Optimization.** *International Journal Of Computer Vision*, [S.l.], v. 66, n. 2, p. 141-161, fev. 2006.
- PÓVOA, Luciano; ANDRADE, Mônica Viegas. **Distribuição geográfica dos médicos no Brasil: uma análise a partir de um modelo de escolha locacional.** *Cadernos de Saúde Pública*, [S.l.], v. 22, n. 8, p. 1555-1564, ago. 2006.
- QUAN, Long *et al.* **Surface reconstruction by integrating 3d and 2d data of multiple views.** In: *Proceedings Ninth IEEE International Conference on Computer Vision*. IEEE, 2003. p. 1313-1320.

RISMILLER, Kyle; CARTRON, Alexander M.; TRINIDAD, John C. L. **Inpatient teledermatology during the COVID-19 pandemic**. Journal Of Dermatological Treatment, [S.l.], v. 31, n. 5, p. 441-443, 13 maio 2020.

SILVA, Cristiana **Silveira** *et al.* **Teledermatologia**: correlação diagnóstica em serviço primário de saúde. Anais Brasileiros de Dermatologia, v. 84, p. 489-493, 2009.

SONG, Hyun O. *et al.* **Deep metric learning via lifted structured feature embedding**. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 4004-4012.

SZELISKI, Richard. **Computer Vision**: Algorithms and Applications (Texts in Computer Science). 2011. ed. Washington, D.C: Springer, 2019. 832 p.

YARAK, Samira; RUIZ, Evandro **Eduardo Seron**; PISA, Ivan Torres. **A Teledermatologia na Prática Médica**. Revista Brasileira de Educação Médica, [S.l.], v. 41, n. 2, p. 346-355, jun. 2017.

ZENG, Gang et al. **Surface Reconstruction by Propagating 3D Stereo Data in Multiple 2D Images**. Lecture Notes In Computer Science, [S.l.], p. 163-174, maio 2004.