



**INSTITUTO
FEDERAL**
Paraíba

INSTITUTO FEDERAL DA PARAÍBA - CAMPUS CAMPINA GRANDE
DEPARTAMENTO DE ENGENHARIA DE COMPUTAÇÃO

TRABALHO DE CONCLUSÃO DE CURSO

**TÉCNICAS DE SEGMENTAÇÃO DE IMAGENS ATRAVÉS DE REDES
NEURAI CONVOLUCIONAIS**

EVERSON BATISTA MARIANO
MARCOS VINÍCIUS LINS DE OLIVEIRA
TALISON KENNEDY OLIVEIRA COSTA

ORIENTADOR: PROF. DR. CÉSAR ROCHA VASCONCELOS

CAMPINA GRANDE – PB

DEZEMBRO, 2022

TRABALHO DE CONCLUSÃO DE CURSO

**TÉCNICAS DE SEGMENTAÇÃO DE IMAGENS ATRAVÉS DE REDES
NEURAS CONVOLUCIONAIS**

EVERSON BATISTA MARIANO
MARCOS VINÍCIUS LINS DE OLIVEIRA
TALISON KENNEDY OLIVEIRA COSTA

ORIENTADOR: PROF. DR. CÉSAR ROCHA VASCONCELOS

CAMPINA GRANDE – PB

DEZEMBRO, 2022

M333t Mariano, Everson Batista.

Técnicas de segmentação de imagens através de redes neurais convolucionais. - Campina Grande, 2022.

51 f. : il.

Trabalho de Conclusão de Curso (Curso de Graduação em Engenharia de Computação) - Instituto Federal da Paraíba, 2022.

Orientador: Prof. Dr. César Rocha Vasconcelos.

1.Redes neurais 2. Limiarização 3. Segmentação semântica I.Oliveira, Marcos Vinícius Lins de II. Costa, Talison Kennedy Oliveira III. Vasconcelos, César Rocha IV. Título.

CDU 004.032.26

EVERSON BATISTA MARIANO
MARCOS VINÍCIUS LINS DE OLIVEIRA
TALISON KENNEDY OLIVEIRA COSTA

TÉCNICAS DE SEGMENTAÇÃO DE IMAGENS ATRAVÉS DE REDES NEURAS CONVOLUCIONAIS

Trabalho de Conclusão de Curso apresentado à
Disciplina de Projeto em Engenharia de
Computação II do Curso em Engenharia de
Computação do Instituto Federal da Paraíba -
Campus Campina Grande, seguindo às
exigências para obtenção do Título de
Bacharel em Engenharia de Computação.

CAMPINA GRANDE – PB

DEZEMBRO, 2022

"Eu acredito que às vezes são as pessoas que ninguém espera nada que fazem as coisas que ninguém consegue imaginar."

Alan Turing

AGRADECIMENTOS

Agradecemos aos nossos familiares, cônjuges e amigos - Os de verdade sabem quem são.

Agradecemos imensamente ao nosso orientador Prof. Dr. César pelas suas ricas contribuições e por ter nos aceitado, tão logo foi solicitado, para esta orientação.

Aos membros da banca examinadora: Prof. Dr. César Rocha Vasconcelos, Prof. Dr. David Candeia Medeiros Maia, Prof. Dr. Petrônio Carlos Bezerra, Prof. Dr. George Sobral Silveira por terem aceitado este processo de avaliação e aprimoramento deste trabalho.

Agradecimento a todos os professores que nos proporcionaram, através de suas experiências, técnicas preciosas para nossa formação profissional.

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO BIBLIOGRÁFICA	17
2.1	Limiarização	17
2.2	Segmentação de Instância	20
2.3	Segmentação Semântica	22
3	MATERIAL E MÉTODOS	25
3.1	Limiarização	25
3.1.1	Estudos de Caso	25
3.1.2	Configuração de Ambiente de Desenvolvimento	25
3.2	Segmentação de Instância	26
3.2.1	Base de Dados e Etapas do Processamento	28
3.2.2	Bibliotecas Utilizadas	28
3.3	Segmentação Semântica	29
3.3.1	Plataforma e Bibliotecas	29
3.3.2	Estrutura da Base de Dados	29
3.3.3	Carregamento das Imagens e Pré-Processamento	30
3.3.4	Construção da Rede Neural	30
3.3.5	Treinamento, Teste e Validação	30
4	RESULTADOS E DISCUSSÃO	31
4.1	Limiarização de Imagens	31
4.2	Testes com os Principais Métodos de Limiarização Biblioteca OPENCV	32
4.3	Implementação de Limiarização na Tomografia Computadorizada (TC)	33
4.4	Estudo de Aplicabilidade em Imagem Infravermelha	34
4.5	Método de Otsu	35
4.6	Segmentação de instância com Mask RCNN	37
4.7	Segmentação Semântica - Carregamento das Imagens	41
4.8	Pré Processamento	41
4.9	Construção da rede neural	42
4.10	Treinamento da Rede Neural	42
4.11	Acurácia	43
4.12	Imagens de Testes	44

5	CONCLUSÕES.....	47
	REFERÊNCIAS BIBLIOGRÁFICAS.....	49

RESUMO

O objetivo dessa pesquisa foi apresentar referências e aplicações das Redes Neurais Convolucionais no campo da *Deep Learning* e *Machine Learning* e analisar estudos de casos existentes em cada uma das suas vertentes. O ambiente de desenvolvimento utilizado foi Google Collab que já é bastante utilizado para pesquisas científicas. Aqui foram abordados processos de Limiarização, Segmentação de Instâncias e Semântica através das Redes Neurais Convolucionais e suas respectivas arquiteturas existentes. Para limiarização, os resultados no estudo realizado com imagens infravermelhas, observando que no primeiro teste e em todas as diversas limiarização, separando assim as pessoas do fundo da imagem. Na proposição da segmentação de instância, os resultados apresentados condizem com que a rede neural se propôs a analisar, dando uma pontuação que foi capaz de distinguir os objetos na imagem, mediante a rede neural de outros exemplos existente. Por fim, a Segmentação Semântica, a arquitetura U-Net no final de seu treinamento, apresentou um processamento de um total de 31.055.297 parâmetros e seus resultados foram salvos em arquivos na extensão JSON. Com o estudo sobre a segmentação de imagens adquiriu-se valiosos conhecimentos sobre métodos de processamento de imagens, o que mostra que o processo de aprendizagem de uma abordagem inicialmente desconhecida para o tratamento de imagens para variados fins.

Palavras-chave: Segmentação; Redes Neurais; Limiarização; Instâncias; Semântica.

ABSTRACT

The objective of this research was to present references and applications of Convolutional Neural Networks in the field of Deep Learning and Machine Learning and to analyze existing case studies in each of its aspects. The development environment used was Google Collab, which is already widely used for scientific research. Here, processes of Thresholding, Segmentation of Instances and Semantics through Convolutional Neural Networks and their respective existing architectures were approached. For thresholding, the results in the study performed with infrared images, noting that in the first test and in all the various thresholding, the objective was achieved, thus separating people from the background of the image. In the proposition of instance segmentation, the presented results match what the neural network proposed to analyze, giving a score that was able to distinguish the objects in the image, through the neural network from other existing examples. Finally, the Semantic Segmentation, the U-Net architecture at the end of its training, presented a processing of a total of 31,055,297 parameters and its results were saved in files with the JSON extension. With the study on image segmentation, valuable knowledge on image processing methods was acquired, which shows that the learning process of an initially unknown approach to image processing for various purposes.

Keywords: Segmentation; Neural networks; Threshold; Instances; Semantics.

Lista de Figuras

Figura 1 - Etapas da R-CNN.....	27
Figura 2 - Etapas da Masck R-CNN.....	27
Figura 3 - Imagem original x imagem com a limiarização Binária.....	31
Figura 4 - Imagem com limiar 100 x imagem com limiar 140.....	31
Figura 5 - Imagem original com suas respectivas limiarizações.....	32
Figura 6 - Imagem original de uma tomografia computadorizada (TC) com suas respectivas limiarizações.....	33
Figura 8 - Imagem original de uma imagem infravermelha com suas respectivas limiarizações.....	34
Figura 7 - Imagem original de uma câmera infravermelha com suas respectivas limiarizações.....	34
Figura 9 - Método de Otsu.....	35
Figura 10 - Histograma.....	36
Figura 11 - Imagem original.....	38
Figura 12 - Imagem segmentada.....	38
Figura 13 - Aplicação da máscara.....	39
Figura 14 - Processo da mask RCNN.....	39
Figura 15 - Kernel aplicado a uma imagem de plantas.....	39
Figura 16 - Representa diversos objetos segmentados.....	40
Figura 17 - Imagem original e suas respectivas máscaras segmentadas.....	41
Figura 18 - Parte do resumo da construção da rede neural.....	42
Figura 19 - Épocas representativas no processo de treinamento e validação da rede neural ...	43
Figura 20 - Imagem original - Máscara - Predição.....	44
Figura 21 - Representa imagens do banco da base de testes.....	45
Figura 22 - Máscara real x Predição.....	45

Lista de Gráficos

Gráfico 1 - Contagem de pixels na imagem	36
Gráfico 2 - Gráfico com preenchimento.....	37
Gráfico 3 - Gráficos de acurácia (<i>Accuracy</i>) e de Perdas (<i>Loss</i>) representando a base de treinamento e validação	43

Lista de Tabelas

Tabela 1 - IoU das 67 imagens da base de testes.....	46
--	----

1 INTRODUÇÃO

Atualmente é observado uma grande adesão de implementações em inteligências artificiais atuando nas aplicações consumidas no cotidiano vivido por todos os seres humanos. O tratamento segmentado de imagens se fez necessário para integrar esse campo da visão computacional e toda sua evolução. Com uma breve reflexão sobre suas aplicações, relatados pela literatura, é possível perceber que com o processamento digital de imagens, nem sempre um determinado ponto da imagem se manterá estático, ou o objeto pode estar de uma posição diferente entre imagens, fazendo com que seja necessário um método que possa perceber e contornar contratempos presentes citados (GOODFELLOW et. al., 2016).

Em meados de 1970 pesquisadores da *Stanford University* deram início a grandes estudos sobre visão computacional. No princípio, seus modelos preditivos eram baseados em imagens capturadas de câmeras com sensores precários, que tinha como objetivo capturar o objeto e vetorizá-lo a fim de obter um refinamento (TENENBAUM, 1970).

Na mesma linha de pesquisa, Baumgart (1974) em sua tese, estudou modelos preditivos de imagens em 3-D que utilizavam algoritmos chamados de GEOMED e CRE. O autor teve como objetivo prever características dos objetos em ambientes com pouca iluminação. Essas imagens eram capturadas através de sensores de câmeras dispostas para que pudessem obter um ângulo dessas imagens. Posteriormente, os modelos eram alimentados para assim obter a previsão e aplicadas às devidas correções. Os resultados dessas imagens eram utilizados em sequências de imagens dinâmicas formando assim as primeiras imagens televisivas. Os resultados eram muito incertos visto que o modelo não possuía grandes refinamentos (BAUMGART, 1974). Entretanto, mediante esses estudos, iniciou-se uma corrida para aprimoramentos através das primeiras redes neurais aplicadas a essa área de estudo, a qual atualmente chamamos de Deep Learning (DL) dentro da visão computacional moderna.

Então, dentro do campo da DL surgem as *Convolutional Neural Network* (CNN) e com isso uma técnica denominada de limiarização de imagens. Esta é uma das mais utilizadas no âmbito de detecção de borda de imagens e que consiste em converter imagem de cor cinza para uma imagem binária, assim tornando a imagem mais simples, ignorando detalhes irrelevantes da mesma, para detectar os limiares de uma imagem geralmente é analisado o histograma das imagens (DURASAMY et al., 2010) (KAUR et al., 2012).

Observa-se que a técnica de limiarização de imagens possui vários tipos de aplicabilidades e que atua em diferentes áreas como a da saúde, aeroespacial e arquitetônica.

Na literatura tem-se uma ampla quantidade de trabalhos em que se aplicam as técnicas na área da saúde, como por exemplo, utilizando imagens de tomografia computadorizada (HANGARTNER, 2007), detecção de tumor no cérebro (KIRANMAYEE *et al.*, 2016) (VERMA *et al.*, 2013) e detecção de câncer de mama (HARRABI *et al.*, 2012).

A problemática da técnica de limiarização seria definir um valor de limiares adequados para um conjunto de imagens para uma determinada aplicação em que tenham resultados de imagens vindas de bases de dados diferentes, e assim não apresentando os mesmos resultados e com isso podendo causar uma divergência (SILVA, 2011).

As CNN's vêm nas últimas décadas obtendo resultados e incrementando com o passar do tempo suas funcionalidades e complexidades em diversas áreas da segmentação (DUMOULIN & VISIN, 2016), inclusive na segmentação de imagens, a qual desempenha um grande papel em diferentes aplicações de visão computacional de alto nível, empregando um desempenho acima da média comparada às demais soluções (LECUN *et al.*, 2015).

O primeiro estopim das CNN's na segmentação foi demonstrado por Krizhevsky *et al.*, (2012) que observaram que técnicas de rede neural poderiam aprimorar o desempenho e capacidade de processamento além da melhor precisão em detecção de objetos em cenários de larga escala. Portanto, demonstrando que o limiar da capacidade da CNN ficando apenas nos limites do *hardware* utilizado e da diversidade do treinamento da rede (KRIZHEVSKY *et al.*, 2012).

Atualmente as CNN's são umas das principais ferramentas utilizadas para identificação e análise de objetos na segmentação de imagem, sendo aplicada em diversas áreas como, por exemplo, na agricultura para detecção de ervas daninhas em meio a plantação de soja (FERREIRA, 2017); Reconhecimento de palavras e números em documentos manuscrito (LECUN *et al.*, 1998); Reconhecimento de tumores em imagens histopatológicas aumentando consideravelmente as chances de acertos na detecção de tumores (VRIESMANN *et al.*, 2018).

A ideia de Convoluções aplicada à segmentação de imagens tem como objetivo reduzir a invariância dos valores dos pixels da imagem facilitando o reconhecimento de padrões pelo computador, e assim, tornando mais eficiente a identificação dos objetos na imagem (LECUN *et al.*, 1998). No processo de convolução da imagem, três ideais de arquitetura são combinados para alcançar uma melhor precisão ao fim do processo inicial, são elas: utilização de um peso unificado entre os processos, interferência local entre os demais campos e amostras temporais entre as camadas de convoluções. Criando uma imagem consideravelmente menor, porém mantendo o máximo de informações possível da imagem original (LECUN *et al.*, 1998). Após as camadas de convoluções, os resultados são repassados

para a rede neural para que a mesma possa seguir com os pesos relativos para cada fatia da imagem e chegando a uma conclusão de qual a maior probabilidade de resposta seria a ideal (SOLICHIN, 2019).

Um dos progressos feitos pelas CNN's foi à segmentação semântica. Esta se dá pela aplicação em imagens 2D, vídeos e dados 3D. Alguns dos desafios para a segmentação semântica são os dados volumétricos se tornando um desafio para a área da visão computacional. A importância da percepção do cenário como um problema central de visão computacional é destacada pelo fato de que um número expressivo em aplicativos se retroalimenta à inferência de conhecimento a partir de imagens (GARCIA-GARCIA *et al.*, 2017). A DL traz aplicabilidades em todas as áreas da automação, como por exemplo, nos carros e direções autômatos (ESS, 2009), mecanismos de busca de imagens (YOON *et al.*, 2015), fotografia computacional, entre outras.

A DL ainda enfrenta muitos desafios e trilha um longo caminho até obter sua maturidade, em comparação com a visão computacional e *Machine Learning* (ML). Isso ocorre devido a este campo estar em constante modificação, o que dificulta o acompanhamento de seu progresso, fazendo com que as pesquisas sobre a segmentação semântica sejam consolidadas em longo prazo, refletindo na escassez da literatura existente publicada e outras que virão (GARCIA-GARCIA *et al.*, 2017).

Por fim, dentro do campo das CNN's se estende um tipo de segmentação chamada de panóptica que visa juntar o resultado que é feito no processo da segmentação semântica (que se propõe em rotular à classe de cada pixel) e da segmentação de instância (identificar e fragmentar cada instância de objeto) (KIRILLOV *et al.*, 2019). A segmentação panóptica faz uso de uma arquitetura chamada de "*DEtection TRansformer*" (DETR) que condensa o cenário de detecção, sem o uso de camadas personalizadas, podendo ser usado por qualquer padrão de CNN (CARION *et al.*, 2020; WANG *et al.*, 2021).

Nesta pesquisa, se abordou no âmbito da segmentação semântica uma aplicação da rede neural denominada de U-net, onde sua definição encontra-se em (RONNEBERGER *et al.*, 2015). A U-net possui arquitetura desta rede neural resultante pois possui uma simetria que lhe dá uma forma de "U". A mesma baseia-se em dois processos: no primeiro ocorre uma codificação, que também é chamada de o caminho de análise. No segundo ocorre no processo de decodificação ou expansão. Neste processo ocorre essa expansão admite que a rede passe pelo processo de aprendizagem ou treinamento e posteriormente devolva a imagem segmentada (SIDDIQUE *et al.*, 2021).

O objetivo dessa pesquisa é apresentar referências e aplicações das Redes Neurais Convolucionais no campo da Deep Learning e analisar estudos de casos existentes em cada uma das suas vertentes.

Objetivo Geral:

Desenvolver uma abordagem sobre técnicas de segmentação de imagens através de Redes Neurais Convolucionais no campo da *Machine Learning* e *Deep Learning*.

Objetivos Específicos:

1. Verificar a segmentação de imagens é capaz de identificar objetos em um plano RGB;
2. Investigar uma rede neural consegue distinguir campos da imagem como objetos diferentes entre si;
3. Examinar a segmentação de imagens pode auxiliar na tomada de 'decisão em exames através de imagens;
4. Observar a segmentação de imagens atua no contexto de um corpo microscópico e no campo celular e/ou molecular.

Disposição deste Documento

Esta pesquisa está organizada em 5 partes. Na primeira parte apresentou-se **Introdução** ao problema e os **Objetivos (Geral e Específicos)**; Na segunda parte foram apresentados uma **Revisão da Literatura** das CNN's e aplicações de suas arquiteturas; Na terceira é descrito no **Material e Métodos** como foram aplicadas às técnicas de segmentação de imagens para cada estudo de caso; Na quarta parte foram analisados os respectivos resultados obtidos; E por fim, a quinta parte às conclusões finais desta pesquisa e sugestões para futuros trabalhos.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados trabalhos relevantes no que diz respeito às técnicas de segmentação de imagens. Para facilitar o entendimento e tornar a leitura mais fluida, este tema foi dividido em quatro partes de acordo com a metodologia utilizada pelos autores. Inicialmente serão revisados os trabalhos nos quais às técnicas de **Limiarização (2.1)**, apresentam-se abordagens sobre às **Segmentações de Instância (2.2)** e **Semântica (2.3)**, respectivamente.

2.1 Limiarização

Duraisamy & Kayalvizhi (2010) conceituaram a técnica de Limiarização como um método que é bastante utilizado para segmentação de imagens em que ele converte a imagem de cinza para uma imagem binária. Um desafio que permaneceu por várias épocas foi determinar a seleção de ótimos limiares, sua maioria é analisada a partir de um histograma de uma imagem. Os limites ótimos podem ser observados a partir da maximização ou minimização de uma função em relação aos seus limites. O artigo deles apresenta um novo algoritmo que se utiliza de inteligência artificial, conhecida como "*Particle Swarm Optimization*" (PSO). Este algoritmo é utilizado para melhorar o desempenho de duas funções existentes da Limiarização vistos em Kapur et., al (1985) e Otsu (1979), em que foi testado o desempenho em algumas imagens e foi visto um resultado superior ao algoritmo genético. O desempenho dessa abordagem tem sido comparado com o método de algoritmo genético, verifica-se que o PSO supera a abordagem algoritmo genético em termos de qualidade da solução, convergência e robustez. Comparado com todos os casos, o Kapur-PSO forneceu um valor de desvio padrão mais baixo. Embora o Kapur-PSO dê um desvio padrão mais baixo, o método Otsu-PSO converge mais rápido do que o método Kapur. Portanto, a abordagem Otsu-PSO é uma ferramenta eficiente para encontrar valores de limiar otimizados.

Jaskirat Kaur *et al.*, (2012) citaram os algoritmos de limiarização e detecção de borda, demonstrando que eles são recursos utilizados antes de ser feita a extração de reconhecimento de uma imagem para análise, assim ignorando mínimos detalhes ou ruídos que seriam desnecessários. Foram demonstrados em quais as aplicações os métodos são utilizados e testados com imagens de satélite, imagens médicas e imagens de obras arquitetônicas. Em seguida, foram aplicados os algoritmos e testados para assim analisar a consistência de erros, medindo seus resultados, quantificá-los. Como característica usada para medir a consistência

desses detectores de borda, foi utilizada a raiz do erro quadrático médio entre a imagem de entrada e a imagem de saída. Esta medida permite uma comparação de princípios entre diferentes resultados gerados pelos detectores de borda. Nos resultados, fora analisado que dependendo da aplicação, um algoritmo apresentou melhor desempenho dependendo do tipo de aplicação feita no teste.

Hangartner (2007) aplicou a técnica de limiarização para a tomografia computadorizada (TC) no objetivo de fazer uma segmentação do osso do tecido para ter uma análise quantitativa. Seu principal objetivo é destacar que dependendo do parâmetro ósseo, o uso dos limiares pode mudar assim tendo limiares recomendados dependendo do tipo de parâmetro ósseo, assim tendo duas informações obtidas através da análise que são de fundamental interesse para a análise óssea, parâmetros geométricos e parâmetros de densidade e isso é obtido a partir da limiarização. O autor demonstra que existe certa correção que deve ser feita, seria em relação à geometria e análise de densidades das imagens da tomografia computadorizada quantitativa periférica (pQCT), assim identificando quais são os limiares corretos para uma justa avaliação dos parâmetros de uma imagem, que seria estabelecer os limites geométricos corretos e um limiar mais alto para extrair valores corretos de densidade das imagens de TC.

Kiranmayee *et al.*, (2016) visa desenvolver um aplicativo para demonstrar a utilidade dos algoritmos propostos na detecção de tumores, eles aplicam a mineração de dados no campo médico e estudam as mais variadas técnicas na detecção de tumores usando tecnologia, no estudo da técnica de limiarização eles citam pesquisadores que testaram ela na detecção de tumores, Atkins & Mackiewich (1998) que fez experimentos para segmentação automática de imagens e assume que Cheng *et al.*, (2000) usou de forma efetiva o conceito de limiarização, baseando-se em histograma, em que eles sempre assumem que há um fundo uniforme quando objetos estão espalhados. Por fim, eles concluem que a aplicação de algoritmos utilizando aplicativos para classificação de tumores cerebrais seria de importante contribuição para indústria de profissionais de saúde, assim melhorando na qualidade do serviço e na tomada de decisão de forma mais eficiente, tendo ótimos resultados nos experimentos e que poderia ser integrado a softwares de saúde do mundo real.

Verma *et al.*, (2013) também estuda a aplicabilidade em detecção em tumores cerebrais, aprimorando imagens usando equalização de histograma e detecção de borda para obter padrões de tumores. Eles fazem todo um estudo de base literária, dividindo em etapas, para chegar ao resultado final esperado, uma dessas etapas primordiais é o estudo da limiarização, onde ele conclui que os resultados revelam vantagens e limitações, mas que eles

forneem uma compreensão adicional e alguns novos procedimentos a serem adotados, dependendo do tipo de situação.

Harrabi & Braiek (2012) testaram uma nova estratégia para detecção de células de câncer de mama, que apesar de ser complexa e exigir um grande tempo de processamento, ela melhora resultados de segmentação, em termos de sensibilidade da segmentação, que seria o processo de limiarização multinível e técnicas de fusão de dados para uma imagem colorida, assim esperando obter resultados precisos e confiáveis e ajudando no caso de diferentes médicos, no caso de terem diferentes interpretações, a partir de uma imagem de um câncer. A ideia consiste em fundir os pixels provenientes de diferentes fontes, assim melhorando a comparação de dados em um histograma e estudando os seus melhores limiares. Os resultados esperados foram obtidos, demonstraram melhor desempenho na segmentação. Esse método também pode ser aplicado em várias aplicações de visão computacional e no ramo de segmentação e aprimoramento de imagens médicas no geral.

Haddon (1988) desenvolveu uma técnica que visa selecionar o limite para detecção de borda em imagens. Para atingir o tal feito ele derivou um limiar para todos os operadores de borda, com base nas estatísticas de ruídos em uma imagem e usa uma técnica para selecionar limites de mapas de resistência de bordas. O uso da técnica de limiar foi aplicado a qualquer operador de borda digital e estendendo para uso de combinações, para desenvolver a técnica foram utilizados dois exemplos de mapas de bordas, com o intuito de remover ruídos que atrapalham na interpretação da imagem. A técnica é robusta e pode ser utilizada com qualquer operador de limite digital ou mistura de operadores. Para sistemas com estatísticas de ruído conhecidas e constantes, a limitação pode ser selecionada trazendo estatísticas do sistema em vez de estatísticas de imagens.

Silva (2001) apresenta uma solução para o problema de segmentação de imagens de profundidade. Foi desenvolvido um método de segmentação baseado em detecção de bordas. Para ter certeza da efetividade do método foram comparados os resultados com outros quatro segmentadores avaliados e com isso, foi possível comprovar que foram percebidas boas contribuições para o desenvolvimento de detecção de bordas, visto que, na remoção de ruídos e eliminação de limiares fixos, contribuiu para que diversos tipos de imagens pudessem ser usados como entradas de testes. Claro que com a análise, também destacam alguns problemas e apontadas possíveis soluções. Os resultados mostram que o método utilizado por ele, se torna mais efetivos em segmentação de profundidade, depois de comparar com os testes com outras imagens, em outros casos os resultados mais relevantes são em imagens com relação à topologia das superfícies e eliminação de ruídos. Com os testes foi visto que a detecção de

bordas na segmentação de imagens é efetiva para resolver problemas na segmentação de imagens de profundidade.

2.2 Segmentação de Instância

Sobre a segmentação de instâncias podemos constatar que a idealização do conceito de separação dos objetos além dos seus tipos e semelhantes se mostra um desafio real e que está sendo frequentemente abordado e melhorado, pois como citado por Kaiming (2017), tanto a visão computacional quanto a sociedade de análise computacional de imagens vem crescendo desenfreadamente, e um dos tópicos mais tocados seriam o fato da utilização de redes neurais, mais especificamente, a métodos mais rápidos e eficazes das RCNN o *Fast-RCNN* descritos em Girshick (2015), e redes neurais convolucionais completas Long (2015) e os demais *frameworks* de detecção de objetos e segmentação semântica em imagens. Criando assim um ritmo de desenvolvimento surpreendente para as tecnologias do meio, como o Mask RCNN, que seria um método de RCNN baseado no *Fast-RCNN* mais rápido que os encontrados no momento, capaz de apresentar mais precisão na detecção semântica dos objetos como foi mostrado em suas conclusões do desempenho dos métodos, mostrando que a Mask RCNN é capaz de dar uma precisão melhor na identificação dos objetos, mesmo eles estando sobrepostos ou em posições desfavoráveis.

Bolya (2019) realizou uma revisão das dificuldades enfrentadas na segmentação de instâncias e a detecção de objetos em si, criando assim um produto capaz de utilizar a abordagem dos demais métodos conhecidos, apresentando um modelo simplificado para segmentação de instância utilizando uma rede neural convolucional completa, e dividindo a segmentação em dois processos paralelos, gerando a máscara de protótipos e prevendo os coeficientes das máscaras, conseguindo um desempenho muito superior na criação das máscaras de instância definitivas, combinando os resultados. Além de utilizar o método de repooling, o que produz uma máscara de alta qualidade. Assim foi desenvolvido o procedimento de segmentação YOLACT, que divide a segmentação da instância em duas tarefas paralelas e realiza a segmentação em tempo real atingindo marcas muito altas em testes padronizados na área de processamento de imagem.

Liu (2018) realizou uma análise na propagação da informações relevantes nas redes neurais de maior importância no meio, no artigo foi proposto a abordagem "Path Aggregation Network", com o intuito de escalonar o fluxo estrutural da segmentação de instância baseada

em proposta do aprendizado anterior da rede neural, especificamente aprimorando toda a hierarquia dos recursos com sinais da localização do núcleo precisa nas camadas inferiores da rede neural, até as mais ao topo, diminuindo assim a quantidade de passos da rede neural e conseqüentemente diminuindo o caminho da informação até a apresentação do resultado da análise. Além da citação do método de *pool* de recursos adaptativos, que vincula as grades de recursos e todos os demais níveis de recursos para assim, fazer com que as informações úteis para a análise dos pixels em cada nível da rede neural se propaguem diretamente para as seguintes sub-redes, criando uma ramificação complementar que captura diferentes visões para a proposta são criados para melhorar a definição da máscara final.

Milanez (2022) sugeriu a implantação da segmentação de instâncias na indústria do aço, pois a mesma apresenta uma grande necessidade de auxílios para melhor realizar seus processos, os quais abrangem todos os campos e nichos da fábrica, foi proposto o uso de redes neurais convolucionais para segmentar e detectar as pás dos raspadores de escória em painéis de ferro em um reator de uma siderúrgica, com o objetivo de detectar o desgaste das pás, fazendo assim o controle de uso e se necessário demarcando-as para substituição, utilizando a abordagem da *mask RCNN* segmentando as pás e prevendo o desgaste e com essa métrica, realizando uma previsão de quando seria necessário realizar uma troca.

O câncer do colo do útero é o quarto tipo de câncer mais comum e mortal nas mulheres nos últimos anos. Assim o Amorim (2022) realizou uma análise no método de diagnóstico deste tipo de câncer que é através da citologia, podendo detectar o câncer precocemente, mas a análise desses diagnósticos é feita manualmente por profissionais treinados, surgindo então a possibilidade do erro humano na análise do problema, por diversos fatores. Pois na citologia é utilizado um corante que destaca as células e suas características. No estudo, foram utilizadas amostras cervicais que tiveram a aplicação de corantes, as quais após a aplicação são fotografadas e utilizando redes neurais convolucionais para identificação dos NORs que são estruturas celulares muito pequenas presentes nos núcleos das células. Assim é possível identificar os focos de NORs com a segmentação de instância, e classificá-los caso esteja apresentando algo que não seja o esperado.

Sacadura (2021) realizou uma avaliação das capacidades de processamento de imagem das redes neurais convolucionais em identificação e classificação de entidades geoespaciais em imagens de satélite, como telhas, vias, edifícios industriais e áreas agrícolas. Utilizando diferentes abordagens para melhor avaliação das técnicas de otimização como o aumento de dados, junção de modelos e transferência de aprendizagem entre processos, as quais têm o intuito de potencializar a capacidade do modelo de classificar novos dados, ou acelerar o

reconhecimento dos já conhecidos. Foram utilizadas para o teste imagens de alta resolução de áreas com diversos tipos de topografia e objetos, utilizando a estrutura de tensor Flow e a biblioteca Keras para construção da arquitetura da rede neural de treino dos modelos, junto com a arquitetura da U-Net, que no momento da análise era a técnica de segmentação e processamento mais competente para a área de observação de imagens de satélite. Por fim, foi possível constatar que a arquitetura U-Net teve grande sucesso em identificar grande parte dos objetos na imagem, tendo um sucesso de até 87% na identificação de áreas industriais.

Souza (2019) investigou algumas técnicas de detecção de objetos utilizando redes neurais aplicadas à detecção de frutos em viticultoras, e uma base de dados para treinamento da CNN aplicada para segmentação de instâncias. Assim, foram treinadas duas redes neurais convolucionais e v3 Peiyuan (2022), para detecção e localização dos cachos de uvas nas imagens, resultando em até 92% de precisão dos locais onde estão localizados os cachos.

Ferreira (2017) teve como objetivo utilizar as redes neurais convolucionais para realizar a detecção de ervas daninhas em imagens de lavouras de soja e classificar essas ervas daninhas entre gramíneas e folhas largas, com o intuito de localizar e alertar para o agricultor os locais que apresentam ervas daninhas indesejadas que podem prejudicar a produção e suas classes, ajudando assim o agricultor a escolher qual herbicida o mesmo deve aplicar em suas lavouras para melhor efetividade do controle das plantas indesejadas. Os testes foram realizados com plantações de soja no mato grosso do Sul, com drones, e nas imagens capturadas foi aplicado uma cnn com o intuito de segmentar as instâncias das plantas, e utilizando a rede neural identificar as ervas daninhas e classificá-las.

2.3 Segmentação Semântica

Sabe-se que a segmentação semântica vai muito além de apenas relacionar e/ou classificar uma imagem a algo, então Arnab *et al.*, (2018) constataram que na alteração estrutural as CNN's atuam na identificação de objetos no seu formato mais natural possível, assim como na segmentação de instâncias focalizando as imagens com detalhes mais nítidos. Assim, esta segmentação se propõe a definir melhor cada pixel estrutural de uma imagem, categorizando todo e qualquer objeto em diversos ambientes. A partir disso, se pode ter como base de como agem os sensores na captação, possibilitando assim uma maior sensibilidade na abstração das perspectivas do mundo real para virtual.

Garcia-Garcia *et al.*, (2017) fizeram uma revisão sobre métodos de Deep Learning para a segmentação semântica com aplicabilidade em muitos campos de aplicação. Foram abordadas definições básicas, assim como as terminológicas utilizadas. Em seguida, são expostos os principais conjuntos de dados e desafios para ajudar os pesquisadores a decidir quais são os que melhor atendem às suas necessidades e aos seus objetivos. Esta pesquisa proporcionou uma série de contribuições que darão base para projetos de segmentação semântica futuros, usando técnicas de Deep Learning. Também foram abordados testes de desempenho que combinaram medidas quantitativas, como o tempo consome de memória e precisão. Por fim, apresentaram sugestões para trabalhos futuros e análises sobre o estado da arte da segmentação semântica.

Ess *et al.*, (2009) sugeriram um sistema baseado em imagens capaz de reconhecer objetos em vias urbanas. O entendimento se deu em um conjunto de características obtidos através de mapas de probabilidade da segmentação, analisando as repetições, curvaturas e estrutura de modo geral. Este conjunto é então usado novamente com um treinamento um contra todos para inferir tanto o tipo de segmento de estrada à frente, quanto à presença e travessia de pedestres e carros. As informações obtidas para os estudos poderiam ser implementadas para somar com mapas existentes nos GPS locais. Com isso beneficiaria os motoristas de aplicativos, e, de forma geral contribuiria positivamente no rastreamento de objetos e corpos na pista, sejam eles estáticos ou em movimento. Então, empregaram uma arquitetura formada por dois processos denominados de "*Segmentation-Based Urban Traffic Scene Understanding*" (SUTSU) mediante uma sequência de vídeos captados através de um carro em movimento por vias urbanas. Esse processo formou a base de representação intermediária da imagem em uma visão computacional baseada na profundidade do campo de uma imagem corrigida. Um de seus resultados apresentados mostrou que a utilidade de uma abordagem baseada na segmentação semântica sobre um descritor global, com um maior desempenho de fácil adaptação levado a outros dados de imagens urbanas como, por exemplo, o entardecer ou tempo chuvoso, sem a necessidade de modificar algum classificador outro alto nível, fazendo com que houvesse redução no tamanho dessas imagens.

A rede U-Net é amplamente utilizada para todos os tipos de problemas de biomedicina, sendo essa a estrutura da segmentação semântica mais importante das CNN. Então, diante disso, Du *et al.*, (2020) com o objetivo de solucionar problemas clínicos e otimizar seus resultados, fez uso de uma rede neural baseada em U-net, utilizando um amplo banco de imagens médicas como dados de entrada a ser analisados e classificados. Concluíram ainda que as técnicas de DL forneçam um bom desempenho no processamento de imagens médicas,

não são suficientes para dar um bom diagnóstico, pois surgiram algumas particularidades ainda não resolvidas por esses sistemas nas operações de convolução. Isso limita suas aplicações a outras áreas da medicina clínica manual. Uma rede DL necessita um grande volume de dados e poder de computacional. Se esses forem insuficientes, mais tempo de treinamento da rede se fará necessário.

Chang *et al.*, (2021) com o objetivo de minimizar as limitações das análises e classificação das CNN em imagens médicas, sugeriram uma configuração de redes denominada de *TransClaw U-Net*, na qual em processamento combinam dois algoritmos: a operação de convolução juntamente com a operação de transformador na parte de codificação. A *transClaw U-Net* consiste na junção de três processos: parte de convolução é aplicada para obter as peculiaridades espaciais rasas para favorecer e a restauração da resolução da imagem após a expansão e filtragem; O processo da modificação é usada para codificar às correções e o mecanismo de auto-observação é utilizada para obter informações globais entre às sequências de imagens. E finalmente, a parte de decodificação guarda a estrutura de expansão inferior para melhor desempenho de segmentação dos detalhes. As consequências dos dados de segmentação de múltiplas camadas e as sinapses das redes demonstraram que a execução do *TransClaw U-Net* foi superior do que outras estruturas CNN. Os testes de remoção comprovaram o desempenho de generalização do *TransClaw U-Net*.

Zhang *et al.*, (2019) fizeram uma revisão bibliográfica com base em técnicas de DL em aplicações da segmentação semântica em pontos existentes na nuvem. Para isso, usaram dois agrupamentos de dados classificados como forma direta (ordenação de pontos, multi-escala, fusão de recursos e fusão de rede neural convolucional de grafos); e, indireta (multi-views e grades de voxel) dispostos em pontos na nuvem com o objetivo de uma investigação nas suas características de acordo com cada resultado obtido. Aqui alguns deles: RGCNN, GAPNet, A-CNN, SEGCloud, MVCNN, VoxNet, SpiderCNN, S3DIS, entre outros. Em um de seus resultados exaltou a importância da segmentação semântica atrelada à visão computacional e que a segmentação em pontos de nuvem é utilizada para discernir uma categoria de objeto de um mesmo tipo de objetos, e, ocasionalmente, usada para dividi-los em uma mesma categoria num conjunto de tipos diferentes. Por fim concluíram que no geral técnicas de segmentação semântica aplicada na nuvem apresentaram resultados superiores, entretanto ainda ocorrem problemas e deixando assim um apanhado de sugestões e discussões para pesquisas futuras.

3 MATERIAL E MÉTODOS

3.1 Limiarização

3.1.1 Estudos de Caso

A coleta de dados foi realizada a partir de imagens de tomografias encontradas em artigos mencionados nas Revisão Bibliográfica. O critério de inclusão das imagens para efetuar os testes são artigos que abordaram a aplicações da técnica de limiarização com imagens de tumores, onde foram comparados os resultados obtidos com a aplicação dos principais algoritmos de limiarização.

3.1.2 Configuração de ambiente de desenvolvimento

O ambiente de desenvolvimento a ser utilizado será o Google Collab o qual é bastante utilizado em pesquisas científicas, foi escolhido por conta de ser possível codificar, demonstrar gráficos de forma simples e satisfatória, assim tornando a aprendizagem e análise de dados mais compreensível.

Foi utilizada a linguagem Python com a biblioteca OpenCV, que serve para visão computacional em Python. Para auxiliar nos cálculos científicos com vetores e matrizes de funções matemáticas e tornar as operações mais rápidas, será utilizada a biblioteca numpy. Para demonstrar de forma significativa os resultados e melhor interpretação dos dados de forma gráfica, será utilizada a biblioteca Matplotlib.

Serão utilizados algoritmos de limiarização disponíveis no openCV sendo eles:

- **THRESH_BINARY:**
 - Pega todos os pixels que estiverem acima de um limiar (valor 255) e os transformam em brancos e os que estiverem abaixo do valor ficarão pretos;
- **THRESH_BINARY_INV:**
 - Faz o inverso da THRESH_BINARY. Os que estiverem acima ficarão pretos, e os que tiverem abaixo do limiar ficarão brancos. Seus resultados são

literalmente uma imagem limiarizada invertida em relação a limiarização binária normal;

- **THRESH_TOZERO:**

- Tudo o que for abaixo do limiar fica preto e tudo o que está acima ou igual ao limiar fica normal, assim conservando o fundo da imagem;

- **TRESH_TRUNC:**

- Tudo abaixo do limiar permanece como está, o que está acima, fica com o valor de limiar fixo. Por isso que leva o nome de *trunc* ("truncar"). Estabelece um teto e ao chegar nele o valor fica fixo nesse limiar, não havendo nenhum valor de pixel de intensidade superior ao limiar definido;

- **THRESH_TOZERO_INC:**

- É o contrário do anterior: Todos os pixels abaixo do limiar permanecem iguais. Tudo que está igual ou maior que o limiar ficará com a cor preta.

Cada um desses métodos foi analisado na sua aplicação em imagens de tomografia computadorizada, para simular o destaque de um tumor para se efetuar uma análise melhor e mais detalhada, observando assim qual método obteve um melhor resultado para essa aplicação.

O método de OTSU foi visto com o objetivo de obter o valor do limiar automaticamente em imagens com ruídos. Para melhor entendimento da técnica de OTSU na geração de um limiar, também será feita a geração dos histogramas utilizando da biblioteca Numpy. Em seguida, foram feitos gráficos para melhoramento da análise na divisão dos pixels, assim analisando o padrão dos picos dos gráficos e seus valores.

3.2 Segmentação de Instância

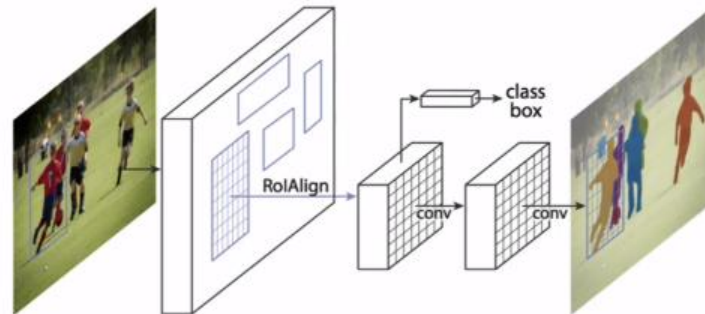
A arquitetura utilizada para este estudo de caso é a Mask R-CNN que é um filtro ou máscara usada na segmentação de instância descrito por He *et al.*, (2017). O ciclo desta técnica será feito em duas fases:

1 - O escaneamento da imagem, gerando áreas prováveis de contenção do objeto (proposals);

2 - A classificação do proposals e geração de caixas delimitadoras da máscara.

Na Figura 1 é apresentado um exemplo de como esse ciclo é feito.

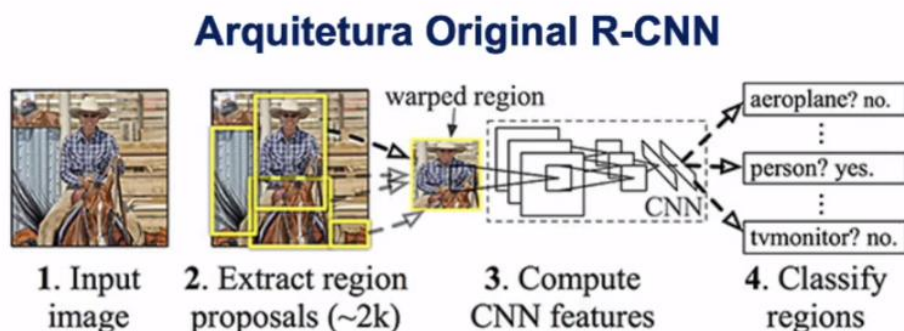
Figura 1 - Etapas da Mask R-CNN



Fonte: He *et al.*, (2017)

Para entendimento desta arquitetura (Mask-R CNN) é preciso entender a que antecedeu a ela: uma arquitetura que é chamada da R-CNN. É observada na imagem abaixo vista em Girshick & Malik (2013) que é um processo inicialmente com quatro etapas, onde a primeira é feita a inserção de uma imagem na rede neural. Na fase dois são extraídos os *proposals* e logo na sequência, cada uma dessas regiões, que potencialmente contém um objeto é classificada para saber se de fato trata se ou não de um objeto. E para isso é utilizada uma CNN uma rede neural convencional que vai receber cada um desses quadradinhos desses possíveis boxes e fará a classificação. Note que no final é apresentada uma indicação de que somente essa parte da imagem nesse caso o objeto Pessoa é classificada como uma ‘pessoa’ e a não detecção da classificação do que seria um avião ou uma televisão e assim por diante.

Figura 2 - Etapas da R-CNN



Fonte: Girshick & Malik (2013)

Em resumo é utilizado um algoritmo de pesquisa seletiva para extrair possíveis regiões, sendo cada uma dessas regiões enviada para a rede neural com o objetivo de verificar se de fato trata-se de um determinado objeto alvo.

Buscando melhorar a arquitetura original o mesmo autor construiu a arquitetura chamada Teste RCNN que ainda utiliza a busca seletiva para obter os proposals da região descritos em Girshick & Malik (2013), onde será mais detalhada nos resultados.

3.2.1 Base de dados e Etapas do Processamento

Foi feito o clone do Github, cujo link é “https://github.com/matterport/Mask_RCNN” e que dá acesso a todos os arquivos do banco de imagens chamado COCO que é melhor descrito em Lin *et al.*, (2014).

Após a estruturação do banco de dados, foram criadas classes e carregamento dos pesos para o uso de uma rede neural pré treinada.

Diante disso, foram feitas avaliações através de inferências, criação de modelos e respectivos carregamento de pesos, mediante as etapas do processo.

Posteriormente foram nomeadas as classes, onde ajudou no processo de predição, visualização e remoção do fundo dos objetos, evidenciado apenas o alvo.

Por fim, os testes levaram em conta todos esses fatores gerados e serão dadas classificações mediante as instâncias que foram obtidas nas imagens, como também a contagem dos objetos.

3.2.2 Bibliotecas Python Utilizadas

- OS
- SYS;
- Random;
- Math;
- CV2;
- Numpy;
- Skimage;
- Matplotlib;
- Tensorflow 1.5;

- Keras.

A plataforma para desenvolvimento utilizado foi o Google Colab.

3.3 Segmentação Semântica

3.3.1 Plataforma e Bibliotecas

A linguagem de programação utilizada na pesquisa foi a linguagem Python.

As bibliotecas que se farão usos são:

- Numpy;
- OpenCV;
- Matplotlib;
- TensorFlow;
- Sklearn.

3.3.2 Estrutura da base de dados

Para este experimento será utilizado a base de dados de imagens microscópicas, obtidos no site do Kaggle (<https://www.kaggle.com/competitions/data-science-bowl-2018>), sendo 603 imagens para a fase de treinamento, 67 para a fase de validação (cerca de aproximadamente 10% do treinamento) e novamente 603 para as máscaras.

O primeiro banco será utilizado para o treinamento da rede neural. Nesta etapa, a rede neural irá captar o peso de cada pixel da imagem através de uma varredura. Nesse banco de imagens, serão criadas duas pastas: uma da imagem real e outra com várias imagens segmentadas, observando cada núcleo individual identificado (máscaras).

O banco de testes será utilizado posteriormente, para efeitos de comparação dos resultados, logo após da rede neural terminar o treinamento. Nessa fase não haverá aplicação de máscaras.

E por fim, a fase de validação. Nesta etapa, serão aplicadas métricas para analisar o que se propõe o algoritmo utilizado. Para isso será utilizado o índice de Jaccard ou *Intersection over Union* – IoU. Este índice é uma métrica avaliativa usado para medir a precisão de modelos de

segmentação de imagens. Além disso, é uma métrica frequentemente usada na avaliação de desafios de detecção de objetos, como o desafio PACAL VOC.

Nas RCNN, a aplicação deste índice vem através das medições como uma forma de conceituar a precisão da detecção de objetos. Quanto mais o IoU próximo de 1, melhor será o resultado.

3.3.3 Carregamento das Imagens e Pré-Processamento

Depois de estruturado o banco de dados, nesta etapa será dado o ajuste nas imagens colocando-as, todas, em uma mesma dimensão, como também serão unidas em uma só. Isso será feito nas pastas das segmentadas, nas máscaras e na de testes (sem máscara).

As dimensões que serão dadas para estas imagens são de 256 x 256 Mega pixel, sendo aplicados à aproximadamente 64 filtros.

3.3.4 Construção da Rede Neural

A construção da rede neural se dará pelo uso da biblioteca de Deep Learning TensorFlow e a API Keras para concepção da sua arquitetura da classe e todas suas camadas. A arquitetura que será implementada neste momento é a U-NET. Suas camadas são compostas em:

- Camada de Convolução;
- Camada de Concatenação;
- Camada de Max-Pooling – Reduz a dimensionalidade;
- Up-Sampling – Aumenta a dimensionalidade;

Dentre os processos, são criados módulos que vão determinar as referidas camadas. Esses módulos são os que fazem o processo de Convolução, Encoder e Decoder, até ser criado o modelo final.

3.3.5 Treinamento, Teste e Validação

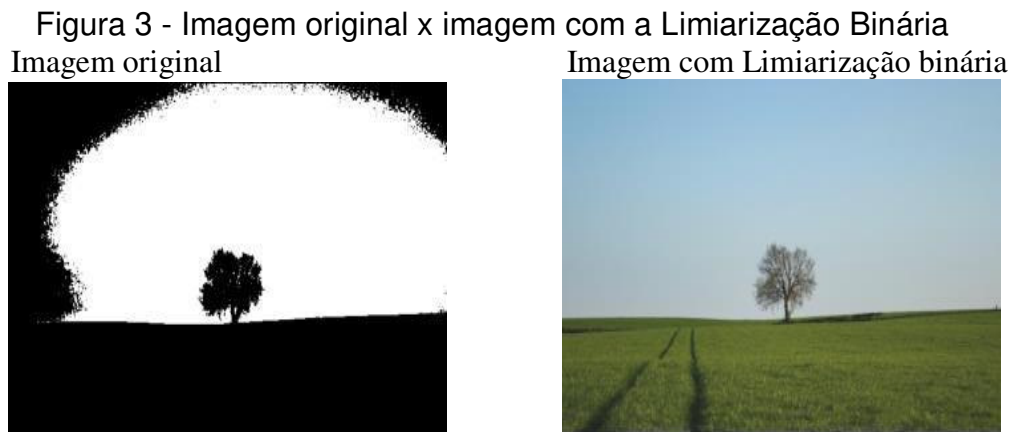
Neste processo vão ser definidas quantas etapas ou épocas de treinamento e o ajuste dos pesos.

Ao final do treinamento da rede neural foram produzidas figura já no seu estado comparativo.

4 RESULTADOS E DISCUSSÃO

4.1 Limiarização de Imagens

A imagem com limiarização binária é o resultado da limiarização da primeira imagem original utilizando como parâmetro o limiar com o valor de 180 e o método de limiarização utilizado foi a limiarização binária (Todos os pixels que estiverem acima do limiar irão ficar brancos e os que estiverem abaixo ficarão pretos).



Pode-se observar que não foi possível dissociar o solo da árvore e o céu na imagem processada, pelo fato dos cantos da imagem não terem um limiar menor do que 180, portantoo resultado adequado não foi obtido com esse limiar.

Figura 4 - Imagem com limiar 100 x Imagem com Limiar 140

Imagem com Limiar 100



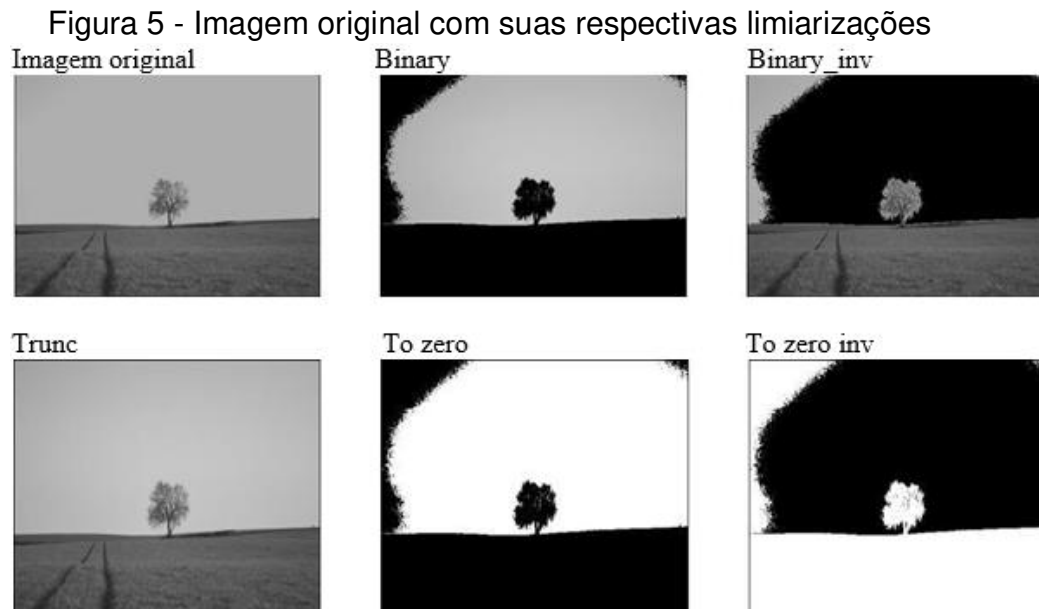
Imagem com Limiar 140



A imagem foi reprocessada com o limiar de 100 utilizando novamente da limiarização binária, assim se observa que vários pixels no solo e na árvore, que estão sendo considerados como fundoda imagem, portanto esse não seria também um valor interessante, por ser muito baixo e ter uma perda de informação da imagem.

Um melhor resultado foi obtido com o limiar de 140, assim conseguindo obter uma separação da árvore e do solo, ou seja, aqui foi conseguido obter realmente a segmentação do terreno e do céu.

4.2 Testes com os principais métodos de limiarização biblioteca OPENCV



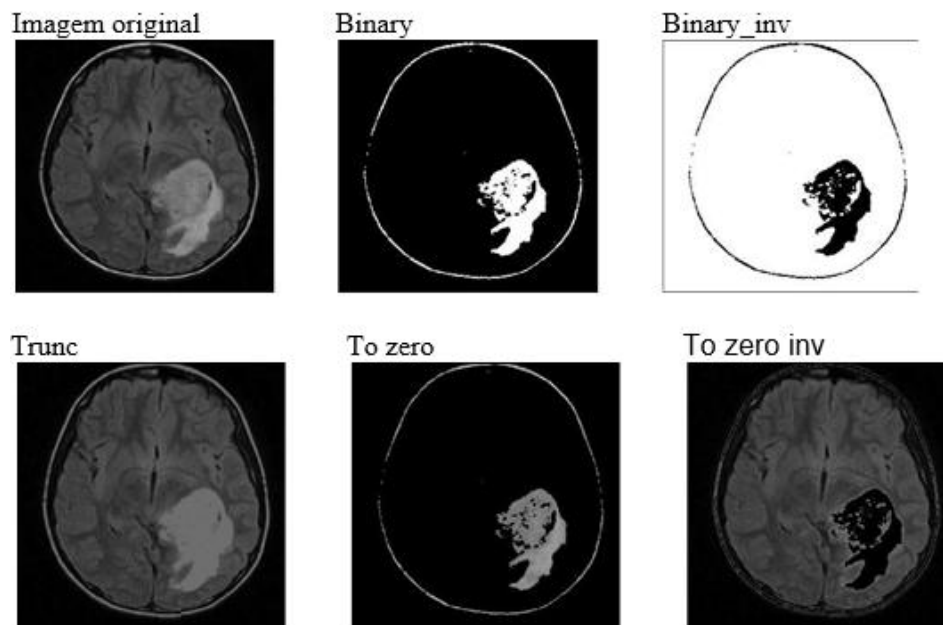
Vemos que na Figura 5. na imagem Binary_inv faz o inverso do primeiro método explicado anteriormente o fundo fica preto e o conteúdo fica branco, o resultado é uma imagem invertida em relação a uma imagem limiarizada binária.

Na imagem trunc, em que tudo abaixo do limiar permanece como está e o que está acima fica com o valor de limiar fixo. Ele estabelece um teto e ao chegar nele o valor fica fixo, podemos ver que o céu da imagem truncada ele possui uma única cor, enquanto o restante da imagem, permanece original.

O método “to zero” consiste em deixar o que for abaixo do limiar preto e o que está igual ou acima do limiar da mesma forma que a imagem original é visto que o céu está exatamente da mesma cor que a imagem original. Por fim, o thresh to zero inv que funciona justamente ao contrário do anterior, deixando o céu de cor escura.

4.3 Implementação de limiarização na Tomografia Computadorizada (TC)

Figura 6 - Imagem original de uma Tomografia computadorizada (TC) com suas respectivas limiarizações



A Figura 6 representa uma imagem de uma TC de um cérebro humano com tumore aplicados os algoritmos de limiarização no objetivo de conseguir destacar melhor o tumor do cérebro para melhor interpretação do mesmo .

Foi utilizado o limiar com o valor de 110 no processamento de todas as imagens para melhor análise dos resultados. É observado que o método trunc contém uma semelhança maior em relação a imagem original. No to zero é notável que as cores originais da parte segmentada (o tumor) foi mantida já na última imagem. Também é visível que apenas a parte segmentada ficou preta e o restante com a imagem original, vemos aqui que o uso dos métodos de limiarização vão apresentar melhores resultados a depender da aplicabilidade, caso o objetivo seja uma análise mais apurada do tumor, seria interessante utilizar do trunc e o to zero, pois no trunc, daria para comparar a parte segmentada do restante e no to zero possui apenas informações acerca do tumor, que poderia ser enviada para outro algoritmo de

interpretação de imagem com aprendizagem de máquina para definir o tipo do tumor.

4.4 Estudo de aplicabilidade em imagem infravermelha

Figura 7 - Imagem original de uma câmera infravermelha com suas respectivas limiarizações.

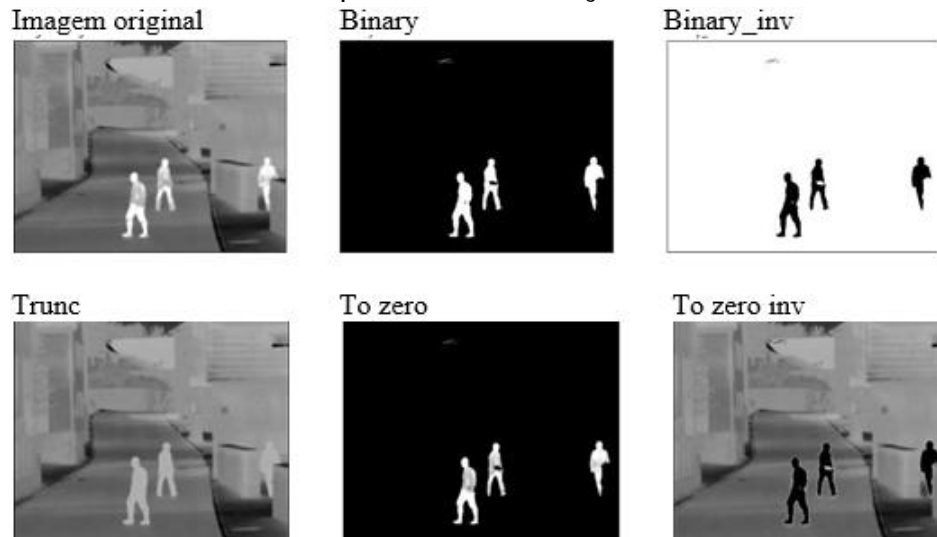
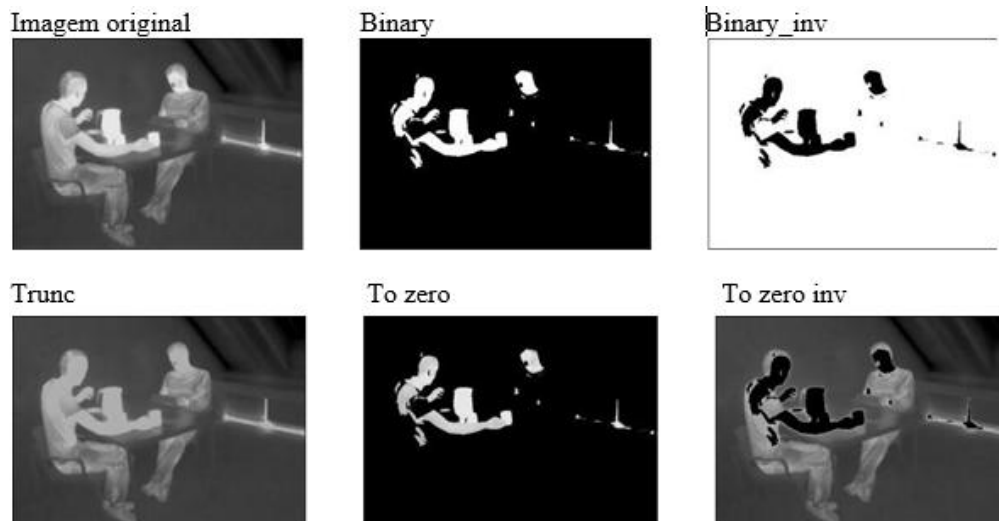


Figura 8 - Imagem original de uma imagem infravermelha com suas respectivas limiarizações



Esse estudo foi realizado com imagens infravermelhas para testar o desempenho do algoritmo de segmentação nas mesmas, podemos observar que no primeiro teste, em todas as limiarização conseguimos separar as pessoas do fundo da imagem.

Com a realização dos testes fica claro que com técnicas diferentes podem mostrar resultados diferentes, e que com a Figura 8. as imagens Binary e Binary_inv é possível identificar apenas a parte superior do corpo, enquanto com a truncada, se identifica mais detalhes da parte inferior do corpo. As imagens de infravermelho levam em consideração o calor emanado pelo objeto, nisso é observado mais detalhes das partes do corpo que são mais claras, ou seja, partes do corpo que estão amostra, com essas interpretações, esses testes demonstram que seria possível uma aplicabilidade associando imagens infravermelhas a limiarização, para detectar regiões específicas do corpo humano ou até mesmo usar o calor como maneira para fazer uma análise de temperatura de diferentes partes do corpo humano.

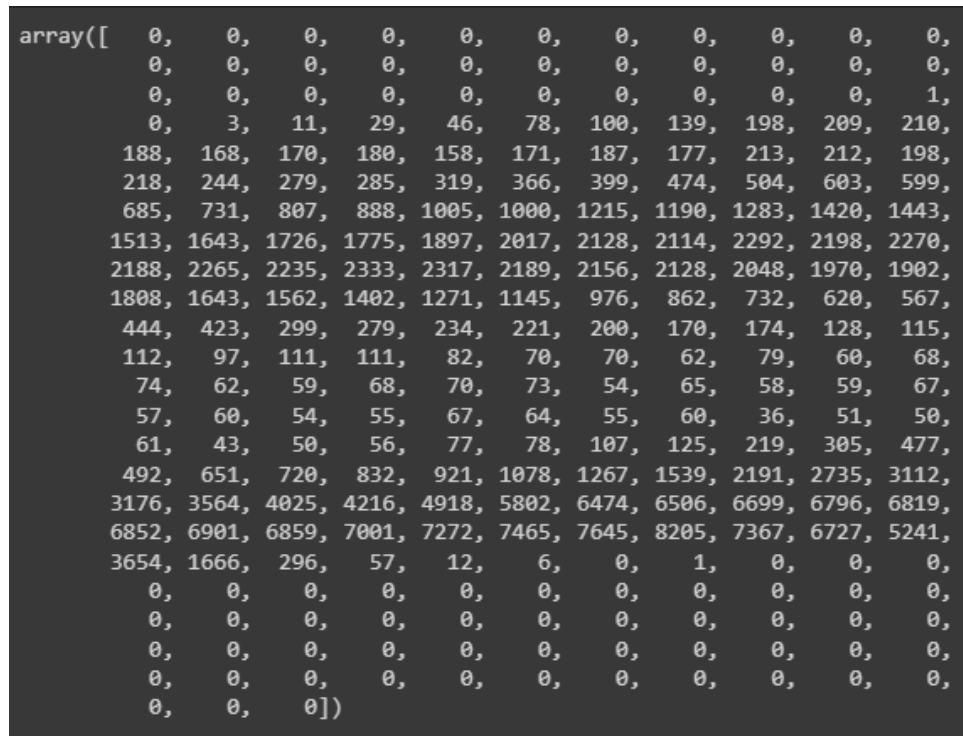
Observa-se diferentes padrões de limiares a depender da aplicabilidade, para o primeiro caso utilizamos 140, para a do tumor utilizamos 110 já no segundo e terceiro caso utilizamos 175, com isso percebe-se que o limiar varia de acordo com a imagem a ser aplicada o algoritmo, assim podendo entregar melhores detalhes e informações.

4.5 Método de Otsu

Figura 9 - Método de Otsu

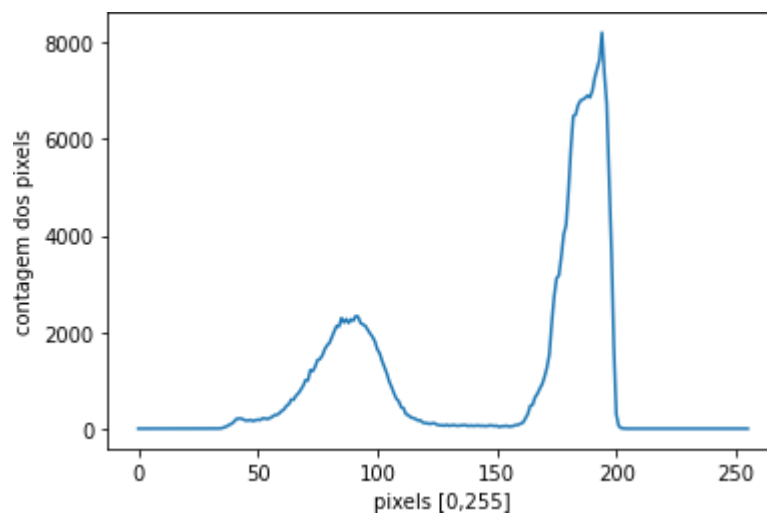


Figura 10 - Histograma



O número do Limiar gerado a partir do método de Otsu para a imagem acima é de 136 e foram feitos testes utilizando o limiar de 140 através de testes manuais, foi muito próximo do que obtivemos anteriormente, mas com o número correto, obteremos o melhor resultado na limiarização da imagem.

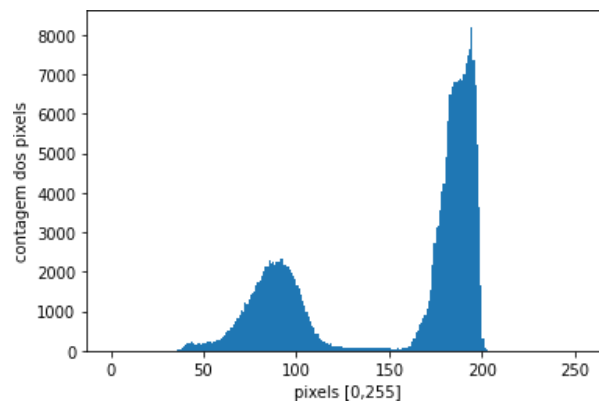
Gráfico 1 - Contagem de pixels na imagem



Aqui está sendo gerado o histograma a partir da imagem anterior, onde podemos ver a contagem de cada pixel da imagem, onde a contagem é efetuada nos pixels que estão de 0 (que seria o pixel mais escuro) ao pixel 255 (que seria o pixel mais claro).

O gráfico 1 está plotado o gráfico em relação ao histograma gerado anteriormente, podemos ver que entre 50 e 100 apresenta os pixels mais escuros, enquanto entre 160 e 200 estão os pixels mais claros, isso também pode ser observado nas contagens dos valores, onde os mais numerosos estão também nessa faixa. Se é observado que o número de Otsu é de 136 e está exatamente entre esses dois picos de contagem de pixels, mas esse método obteve melhor desempenho em imagens bimodais, onde possuem duas grandes distribuições de pixels na histograma, bem distinguíveis uma ao lado da outra.

Gráfico 2 - Gráfico com preenchimento



Aqui foi gerado o histograma, transformando a matriz em um vetor, assim obtendo o preenchimento das áreas do histograma.

4.6 Segmentação de instância com Mask RCNN

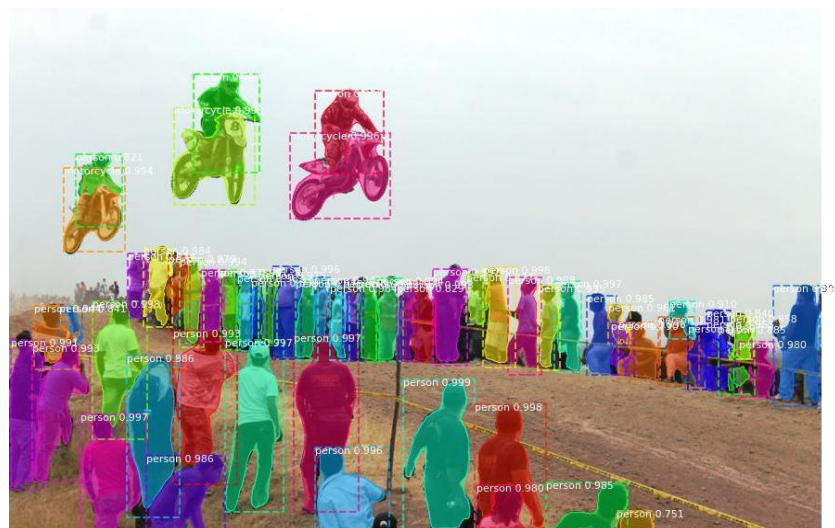
Com a implementação da segmentação de instância na análise das imagens, pode-se dar um passo a mais da segmentação semântica fazendo a distinção dos objetos além de suas classes, separando-os individualmente e assim tendo a possibilidade de criar máscaras específicas para cada objeto. Em quesito de testes utilizamos a figura 11, para aplicação das técnicas de segmentação de instância.

Figura 11 - Imagem original



A Figura 12 representa uma imagem com as segmentações aplicadas, com seus *bound-boxes* e suas respectivas porcentagens de assertividade do objeto com aproximadamente 60 segmentos detectados.

Figura 12 - Imagem segmentada



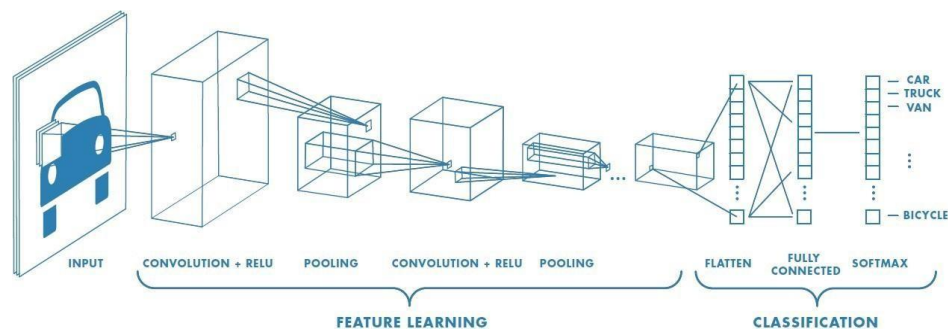
Assim, além de possuir os *bound-boxes* dos objetos, esses separados pertencem a uma mesma classe, sendo assim possível criar uma máscara para a imagem e assim aplicá-la removendo apenas os objetos em questão, como visto na Figura 13.

Figura 13 - Aplicação da máscara



A segmentação de instância se beneficia do uso da inteligência artificial a Mask RCNN Kaiming He(2017), que foi treinada especificamente para análise de imagens com o processo convolucional, analisando as camadas RGB da imagem e pontuando cada pixel em relação aos seus adjacentes, descritos na Figura 15.

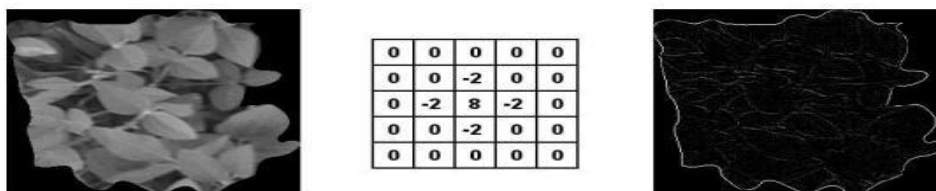
Figura 14 - Processo da mask RCNN



Fonte: “A Comprehensive Guide to Convolutional Neural Networks”

O algoritmo percorre conjuntos de pixels de cada camada da imagem que são conhecidos como *kernel*, comparando-os e dando pontuações de acordo com a coerência do pixel com os demais na região, como por exemplo, na figura 14 que representa um *kernel* aplicado a uma imagem de plantas.

Figura 15 - kernel aplicado a uma imagem de plantas.



Após isto, os dados criados pelos kernels passa por um outro processo denominado "polling", que realiza uma análise nos valores adquiridos e faz uma média dos dados comprimindo-os em um único pixel, que criará uma nova imagem com os resultados e a mesma voltará novamente para o processo de convolução, até o número que a imagem esteja rica o suficiente de informações para que a rede neural possa gerar uma resposta válida para os dados da imagem reduzida.

Assim a rede neural analisa essa pontuação e é capaz de distinguir os objetos na imagem, a rede neural do exemplo anterior foi treinada para detecção das pessoas e das motos.

Abaixo são apresentadas imagens nas quais foram utilizadas na rede neural para detecção de mais objetos nas fotos, detectando zebras e objetos do cotidiano na imagem.

Figura 16 – Imagens que apresentam diversos objetos segmentados



4.7 Segmentação Semântica - Carregamento das imagens

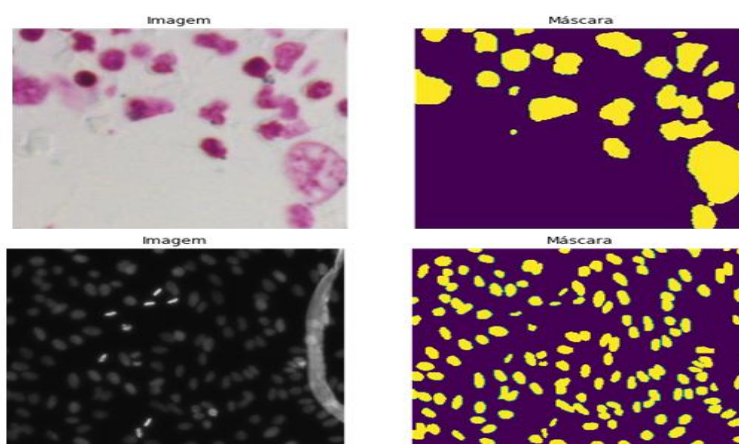
Mediante a importação das bibliotecas foram baixados dados do *Kaggle Data Science Bowl Challenge 2018* (DSB18) diretamente no site do Kaggle de imagens microscópicas de núcleos celulares. Após isso, ajustaram-se alguns parâmetros como as dimensões e canais RGB das imagens. Essa padronização é feita para os dados sejam utilizados como entrada no algoritmo da arquitetura U-NET. Num total de 670 imagens para a base de treinamento e outras 67 (cerca de 10% do total) para serem compostas na base de testes.

4.8 Pré Processamento

Após isso, foram feitos o pré-processamento das imagens, onde se utilizou como entrada um banco de imagens que deram origem as máscaras. Para cada imagem original foi gerada uma máscara através da concatenação de todas as imagens, ou seja, ocorreu uma união de imagens para que através desse pré-processamento fosse possível a identificação dos núcleos da imagem original através da máscara.

No final desta etapa ocorreu uma divisão total na base de dados: base de dados de treinamento e base de dados de validação. Na primeira, a rede neural obteve os pesos da imagem e na segunda foi utilizada para verificar o percentual de acerto em cada passo do processamento da fragmentação. A base de dados de teste foi usada posteriormente logo após o término desse treinamento. Todos os processo desse estudo de caso tomaram como base no trabalho de Pun & Agarwal (2020).

Figura 17 - Imagem original e suas respectivas máscaras segmentadas.



4.9 Construção da Rede Neural

Nessa fase é usada a biblioteca *Tensorflow* que é uma biblioteca mais utilizada para *Deep Learning* e redes neurais em geral, onde foi construída a arquitetura U-NET. Observou nesse processo também às implementações das camadas de convolução, *Maxpooling*, *Up-sampling* dentre os blocos Encoder e o Decoder.

Figura 18 - Parte do resumo da construção da Rede Neural

```

Model: "UNet"
-----
Layer (type)                Output Shape          Param #   Connected to
-----
input_1 (InputLayer)        [(None, 256, 256, 3  0           []
                          )]

conv2d (Conv2D)             (None, 256, 256, 64  1792        ['input_1[0][0]']
                          )

batch_normalization (BatchNorm (None, 256, 256, 64  256         ['conv2d[0][0]']
alization)                  )

activation (Activation)     (None, 256, 256, 64  0           ['batch_normalization[0][0]']
                          )

conv2d_1 (Conv2D)          (None, 256, 256, 64  36928       ['activation[0][0]']
                          )

batch_normalization_1 (BatchNo (None, 256, 256, 64  256         ['conv2d_1[0][0]']
rmalization)                  )

activation_1 (Activation)   (None, 256, 256, 64  0           ['batch_normalization_1[0][0]']
                          )

max_pooling2d (MaxPooling2D) (None, 128, 128, 64  0           ['activation_1[0][0]']
                          )

conv2d_2 (Conv2D)          (None, 128, 128, 12  73856       ['max_pooling2d[0][0]']
                          8)

batch_normalization_2 (BatchNo (None, 128, 128, 12  512         ['conv2d_2[0][0]']
rmalization)                  8)

activation_2 (Activation)   (None, 128, 128, 12  0           ['batch_normalization_2[0][0]']
                          8)

conv2d_3 (Conv2D)          (None, 128, 128, 12  147584      ['activation_2[0][0]']
                          8)

batch_normalization_3 (BatchNo (None, 128, 128, 12  512         ['conv2d_3[0][0]']
rmalization)                  8)

activation_3 (Activation)   (None, 128, 128, 12  0           ['batch_normalization_3[0][0]']
                          8)

```

Nesse treinamento houve um processamento de um total de 31.055.297 parâmetros e seus resultados foram salvos em arquivos na extensão JSON.

4.10 Treinamento da Rede Neural

O treinamento da rede é feito em no máximo de 100 épocas, sendo interrompido se a função de perda da validação não cair por 15 épocas. Nesta pesquisa foram requisitadas 15 épocas no total, com função de perdas de 5 épocas.

Figura 19 - Épocas representativas no processo de treinamento e validação da rede neural

```

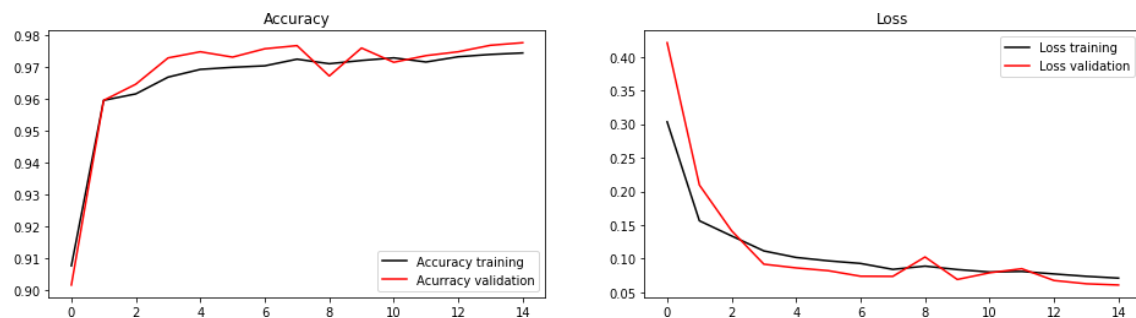
Epoch 1/15
76/76 [=====] - ETA: 0s - loss: 0.4947 - accuracy: 0.8238
Epoch 1: val_loss improved from inf to 0.53691, saving model to modelo_unet.h5
76/76 [=====] - 60s 569ms/step - loss: 0.4947 - accuracy: 0.8238 - val_loss: 0.5369 - val_accuracy: 0.8909
Epoch 2/15
76/76 [=====] - ETA: 0s - loss: 0.2172 - accuracy: 0.9580
Epoch 2: val_loss improved from 0.53691 to 0.35508, saving model to modelo_unet.h5
76/76 [=====] - 40s 526ms/step - loss: 0.2172 - accuracy: 0.9580 - val_loss: 0.3551 - val_accuracy: 0.9263
Epoch 3/15
76/76 [=====] - ETA: 0s - loss: 0.1951 - accuracy: 0.9577
Epoch 3: val_loss improved from 0.35508 to 0.23537, saving model to modelo_unet.h5
76/76 [=====] - 41s 535ms/step - loss: 0.1951 - accuracy: 0.9577 - val_loss: 0.2354 - val_accuracy: 0.9597
Epoch 4/15
76/76 [=====] - ETA: 0s - loss: 0.1632 - accuracy: 0.9657
Epoch 4: val_loss improved from 0.23537 to 0.15948, saving model to modelo_unet.h5
76/76 [=====] - 41s 543ms/step - loss: 0.1632 - accuracy: 0.9657 - val_loss: 0.1595 - val_accuracy: 0.9671
Epoch 5/15
76/76 [=====] - ETA: 0s - loss: 0.1515 - accuracy: 0.9654
Epoch 5: val_loss did not improve from 0.15948
76/76 [=====] - 40s 524ms/step - loss: 0.1515 - accuracy: 0.9654 - val_loss: 0.1782 - val_accuracy: 0.9680
Epoch 6/15
76/76 [=====] - ETA: 0s - loss: 0.1389 - accuracy: 0.9675
Epoch 6: val_loss improved from 0.15948 to 0.12176, saving model to modelo_unet.h5
76/76 [=====] - 43s 560ms/step - loss: 0.1389 - accuracy: 0.9675 - val_loss: 0.1218 - val_accuracy: 0.9715
Epoch 7/15
76/76 [=====] - ETA: 0s - loss: 0.1302 - accuracy: 0.9684
Epoch 7: val_loss improved from 0.12176 to 0.11678, saving model to modelo_unet.h5
76/76 [=====] - 41s 542ms/step - loss: 0.1302 - accuracy: 0.9684 - val_loss: 0.1168 - val_accuracy: 0.9746
Epoch 8/15
76/76 [=====] - ETA: 0s - loss: 0.1198 - accuracy: 0.9709
Epoch 8: val_loss improved from 0.11678 to 0.10228, saving model to modelo_unet.h5
76/76 [=====] - 41s 544ms/step - loss: 0.1198 - accuracy: 0.9709 - val_loss: 0.1023 - val_accuracy: 0.9762
Epoch 9/15
76/76 [=====] - ETA: 0s - loss: 0.1148 - accuracy: 0.9700
Epoch 9: val_loss improved from 0.10228 to 0.09350, saving model to modelo_unet.h5
76/76 [=====] - 41s 540ms/step - loss: 0.1148 - accuracy: 0.9700 - val_loss: 0.0935 - val_accuracy: 0.9746

```

4.11 Acurácia

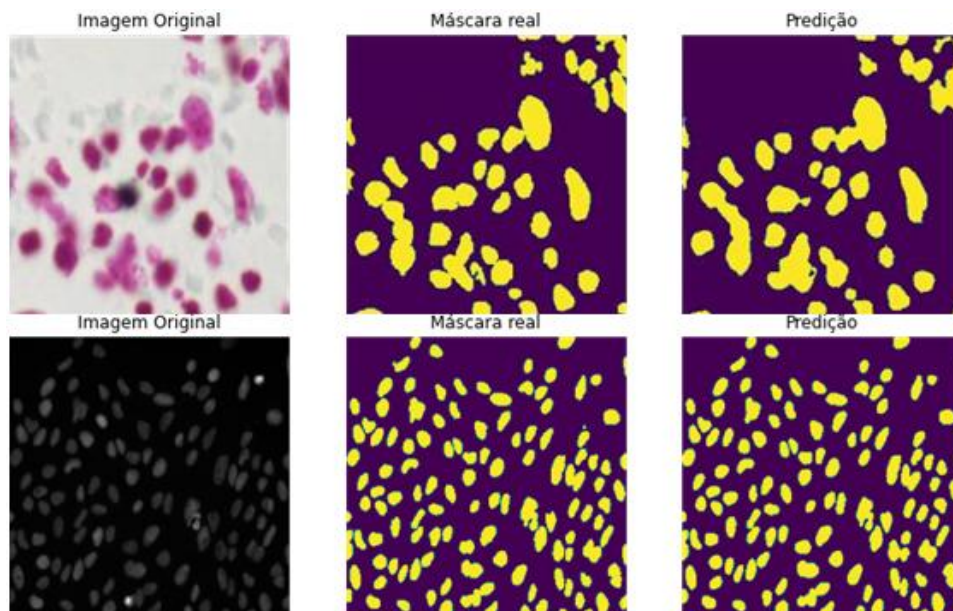
Os gráficos mostram as comparações do processo de treinamento e validação no entre a abordagem proposta para o modelo de U-Net e seus respectivos ciclos. É observado que a curva possui alguns índices menores de acurácia no início do processo, devido a rede neural estar se adequando aos dados e assim aprendendo no seu processo de treinamento. O contrário ocorre com a curva de perdas, visto que quanto mais a rede neural capta informações e se adéqua aos dados, menos perdas terão no seu processo de treinamento e validações.

Gráfico 3 - Gráficos de acurácia (Accuracy) e de Perdas (Loss) representando a base de treinamento e validação



A Figura 20 representa às imagens de validação da base de treinamento e serve para que se possa fazer comparações visuais da imagem original, máscara e predições, respectivamente. As imagens foram escolhidas de maneira aleatórias através dos índices. Como a base possui muitas imagens seria praticamente inviável de apresentá-las aqui. Entretanto, no que é observado, percebe-se que a rede neural está bem ajustada, pois a mesma conseguiu retratar a imagem real, realçando os núcleos das células através dos seus segmentos, assim como conseguiu prevê-los de forma positiva.

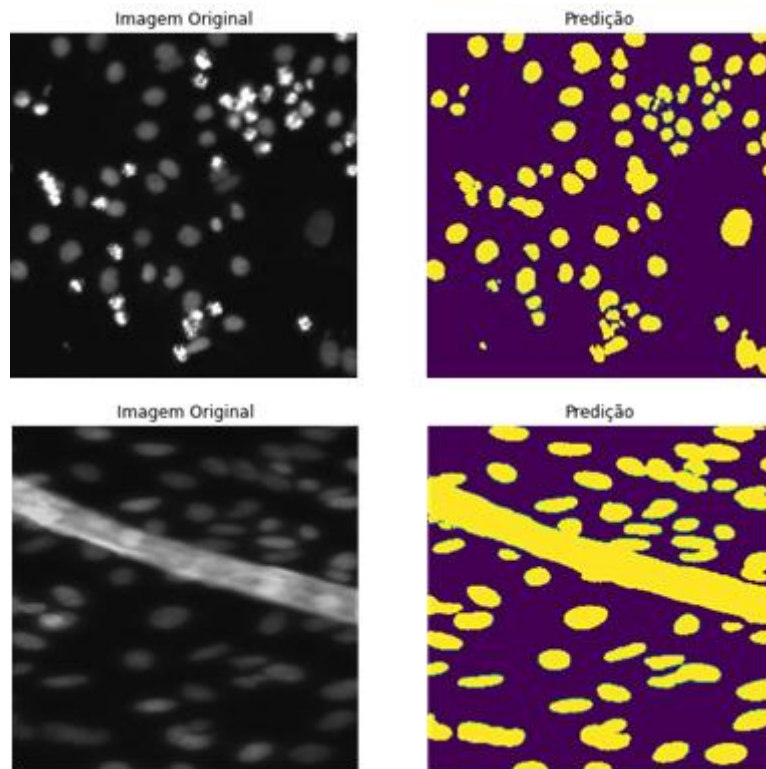
Figura 20 - Imagem original - Máscara - Predição



4.12 Imagens de Testes

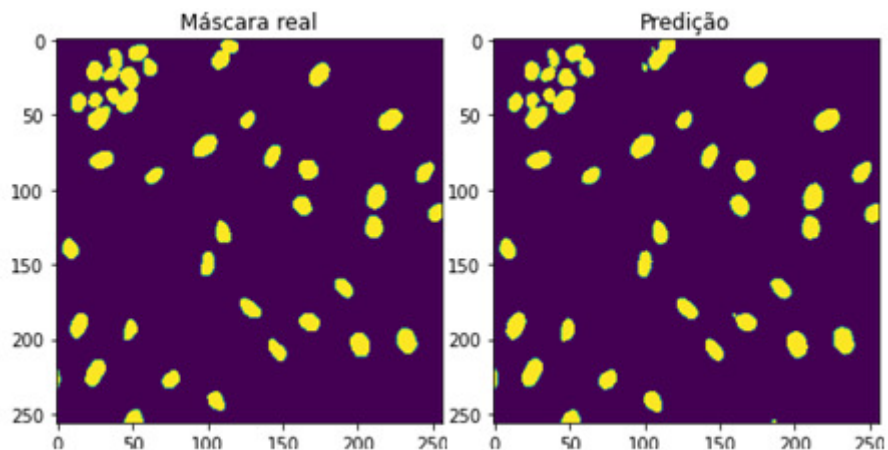
Na Figura 21 são representadas as imagens de testes, onde estas não possuem máscaras, pois dessa forma pode-se avaliar visualmente o desempenho da rede neural. Observa-se nas imagens que a rede neural foi capaz de identificar a maior parte dos núcleos celulares. Isso já era esperado, visto que os valores de perdas que foram apresentados são pequenos e o da acurácia valores em sua maior parte do processo, próximos a 1.

Figura 21 - Imagens de do banco da base de testes



Na Figura 21 foram geradas imagens da base de validação, onde a predição é feita através não da imagem original, mas da máscara. Então, como nos outros casos, aqui também é observado um resultado satisfatório na predição provando-se mais uma vez que a rede neural apresenta uma parametrização inicial bem ajustada.

Figura 22 - Máscara real x Predição



Muitos critérios de avaliação foram propostos e são frequentemente usado para avaliar a precisão de qualquer tipo de técnica para segmentação semântica. Porém o mais utilizado é o *Intersection over Union (IoU)*. Ela é uma medida padrão para identificar contratempos no ato de segmentar imagem e é definida pela razão entre a interseção e a união dos conjuntos de pontos previstos como as organelas celulares na mascara da previsão e na imagem original, sendo o valor máximo 1. A IoU da imagem de previsão é 0,9461 aproximadamente. Já para a média das 67 imagens da base de testes apresentou uma média de 0,8932 aproximadamente. Isso significa que os dados processados através da rede neural U-Net cumpriu o seu papel na obtenção os núcleos celulares, afim que se possam identificar anomalias nas imagens microscópicas, e, posteriormente o laudo médico, ser aplicados medidas assertivas para resolução do problema que estiver a ser verificado. Resultados semelhantes foram obtidos em Punn & Agarwal (2020).

Tabela 1 - IoU das 67 imagens da base de testes.

IoU	
0	0.948165
1	0.964474
2	0.925081
3	0.828640
4	0.909868
...	...
62	0.961002
63	0.954314
64	0.877766
65	0.857334
66	0.881489

67 rows x 1 columns

5 CONCLUSÕES

Diante dos resultados verificados, conclui-se que:

1. Com o estudo sobre a segmentação de imagens adquiriu-se valiosos conhecimentos sobre métodos de processamento de imagens.
2. Além disso, alavancar no entendimento do comportamento de uma rede neural, e como é possível utilizá-las quando ajustados seus parâmetros corretamente.
3. Na Segmentação de instância em um de seus processos observou-se que é possível a identificação de objetos numa vetorização dos pixels, e, assim possam ser diferenciados, mesmo que possuam diversos objetos da mesma natureza.
4. Na Segmentação Semântica os objetos identificados com formas similares são agrupados num mesmo grupo. O contrário pode ser observado na Segmentação de Instância.
5. A Arquitetura U-Net apresentou bons resultados em seus processos, desde a captação do banco de imagens utilizado até os processos finais de decodificação das mesmas. Para fins avaliativos, a média do índice ficou IoU em torno de 89,32% o que significa um modelo ajustado para futuras pesquisas.
6. Além disto, foi compreendido como se dá o estudo no campo de análise e processamento de imagens com Segmentação de imagem é muito relevante, pois com o uso de tecnologias aplicadas, pode-se além de analisar a imagem, identificar corpos específicos na mesma e realizar operações no mesmo. Isso ao longo dos anos vem evoluindo e mostra útil para resolução de problemas na sociedade em geral de macro e micro escalas, como por exemplos, a identificação facial, participação direta na identificação de doenças, identificação de pragas em lavouras através da captação de imagens de drones e satélites, entre outros.
7. Como sugestões para trabalhos futuros o estudo mais aprofundados de outras arquiteturas de redes neurais utilizadas para o mesmo fim como, por exemplo, as Redes Neurais Recorrentes, *Long Short-Term Memory* (LSTM), Redes de

Hopfield, Máquinas de Boltzmann, entre outras.

REFERÊNCIAS BIBLIOGRÁFICAS

Amorim, J. G. A. Aprendizagem profunda em imagens citológicas: identificação e caracterização de objetos corados com a técnica AgNOR, 2022.

ARNAB, A. et al. Conditional Random Fields Meet Deep Neural Networks for Semantic Segmentation: Combining Probabilistic Graphical Models with Deep Learning for Structured Prediction. *IEEE Signal Processing Magazine*. v. 35. n. 1 p. 37-52, jan. 2018.

Atkins MS, M BT. Fully automatic segmentation of the brain in MRI. *IEEE Trans Med Imaging*. 1998 Feb;17(1):98-107. doi: 10.1109/42.668699. PMID: 9617911.

BAUMGART, B. G. Geometric modeling for computer vision. Stanford Univ Ca Dept Of Computer Science, 1974.

CARION, N. et al. End-to-End Object Detection with Transformers. In: *Computer Vision – ECCV 2020*. Cham: Springer International Publishing, 2020. p. 213–229.

Chang, Y.; Menghan, H.; Guangtao, Z., & Xiao-Ping, Z. Transclaw u-net: Claw u-net with transformers for medical image segmentation. *arXiv preprint arXiv:2107.05188*, 2021.

Cheng HD, Chen YH, Jiang XH. Thresholding using two-dimensional histogram and fuzzy entropy principle. *IEEE Trans Image Process*. 2000;9(4):732-5. doi: 10.1109/83.841949. PMID: 18255445.

Daniel Bolya, Chong Zhou, Fanyi Xiao, Yong Jae Lee. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019.

DU, G. et al. Medical image segmentation based on u-net: A review. *Journal of Imaging Science and Technology*, v. 64, p. 1-12, 2020.

DUMOULIN, Vincent; VISIN, Francesco. A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*. 2016.

DURASAMY, S. P.; R. KAYALVIZHI, "A New Multilevel Thresholding Method Using Swarm Intelligence Algorithm for Image Segmentation," *Journal of Intelligent Learning Systems and Applications*, Vol. 2 No. 3, 2010, pp. 126-138. doi: 10.4236/jilsa.2010.23016.

ESS, A. et al. Segmentation-Based Urban Traffic Scene Understanding. Proceedings of the British Machine Vision Conference 2009. Anais...British Machine Vision Association, 2009. Disponível em: <<http://dx.doi.org/10.5244/c.23.84>>. Acesso em: 24 nov. 2022

FERREIRA, S. ‘‘Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja’’, In: UFMG. 2017.

GARCIA-GARCIA, A. et al. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*, 2017.

GIRSHICK, R; MALIK, J. Training deformable part models with decorrelated features. In: Proceedings of the IEEE International Conference on Computer Vision. 2013. p. 3016-3023.

Haddon, J. F. "Generalised threshold selection for edge detection." *Pattern Recognition* 21.3 (1988): 195-203.

Hangartner, T. N. "Thresholding technique for accurate analysis of density and geometry in QCT, pQCT and muCT images." *Journal of Musculoskeletal and Neuronal Interactions* 7.1 (2007): 9.

HARRABI, R., et al. Color image segmentation using multi-level thresholding approach and data fusion techniques: application in the breast cancer cells images. *J Image Video Proc* 2012, 11 (2012). <https://doi.org/10.1186/1687-5281-2012-11>

HE, L. et al. Deep semantic role labeling: What works and what's next. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017. p. 473-483.

Long, j. E. S, and T. Darrell. Fully convolutional networks for semantic segmentation. In *CVPR*, 2015.

J. N. Kapur, P. K. Sahoo and A. K. C. Wong, "A New Method for Gray-Level Picture Thresholding Using the Entropy of The Histogram," *Computer Vision, Graphics and Image Processing*, Vol. 29, No. 3, 1985, pp. 273-285

Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.

KAUR, J.; SUNIL, A.; RENU, V.. "A Comparative Analysis of Thresholding and Edge Detection Segmentation Techniques." *International Journal of Computer Applications* 39.15. 2012. p. 29–34. Web.

KIRANMAYEE, B. V.; RAJINIKANTH, T. V.; NAGINI, S. "A novel data mining approach for brain tumour detection," 2nd International Conference on Contemporary Computing and Informatics (IC3I), 2016, pp. 46-50, doi: 10.1109/IC3I.2016.7917933.

Kirillov, A., He, K., Girshick, R., Rother, C., Dollár, P.. Panoptic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019. p. 9404-9413.

KRIZHEVSKY, Alex & SUTSKEVER, Ilya & Hinton, Geoffrey.. Image Net Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*. 2012. doi: 25. 10.1145/3065386.

LECUN, Y.; BENGIO, Y.; Hinton, G. Deep learning. *Nature* 521. 2015.

LECUN, Y.; BOTTOU, L.; Bengio, Y.; HAFFNER, P. Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*. 86. 2278 - 2324. 1998. doi:10.1109/5.726791.

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. Microsoft coco: Common objects in context. In: *European conference on computer vision*. Springer, Cham, 2014. p. 740-755.

Milanez, Carlos Eduardo Oliveira; DETECÇÃO E SEGMENTAÇÃO DE RASPADORES DE ESCÓRIA DE FERRO GUSA UTILIZANDO MASK R-CNN PARA CONTROLE DE DESGASTE E PREDIÇÃO DE TROCA, (2022).

Otsu, N. A Threshold Selection Method from Gray-Level Histogram. *IEEE Transactions on System Man Cybernetics*, 9, 62-66. <http://dx.doi.org/10.1109/TSMC.1979.4310076> (1979).

PUNN, Narinder Singh; AGARWAL, Sonali. Inception u-net architecture for semantic segmentation to identify nuclei in microscopy cell images. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, v. 16, n. 1, p. 1-15, 2020.

R. Girshick; Fast R-CNN. In *ICCV*, 2015.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2015. p. 234–241.

Sacadura, João Pedro Ventura de Cabral;Segmentação de imagens multiespectrais de alta resolução utilizando o modelo U-Net para cartografia de uso do solo, 2021.

Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, Jiaya Jia; *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

SIDDIQUE, Nahian et al. U-net and its variants for medical image segmentation: A review of theory and applications. *Ieee Access*, v. 9, p. 82031-82057, 2021.

SILVA, L. Segmentação De Imagens De Profundidade Por Detecção De Bordas. 2011. Silva, Luciano. "Segmentação de imagens de profundidade por detecção de bordas." (2001).

SOLICHIN, A. Seminário: "Computer Vision di Era Industri 4.0"; Jakarta, 22 Maret. 2019.

SOUZA, L. L. de AVILA, S.SANTOS, T. T; Detecção de frutos em viticultura utilizando redes neurais profundas, 2019.

TENENBAUM, Jay Martin. *Accommodation in computer vision*. Stanford Univ Ca Dept of Computer Science, 1970.

VERMA, K. et al. IMAGE PROCESSING TECHNIQUES FOR THE ENHANCEMENT OF BRAIN TUMOR PATTERNS. Disponível em: <<https://www.semanticscholar.org/paper/IMAGE-PROCESSING-TECHNIQUES-FOR-THE->

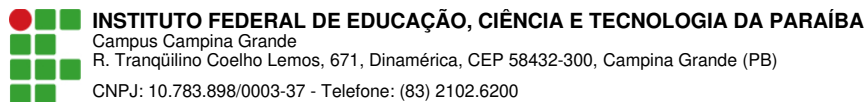
ENHANCEMENT-OF-Verma-Mehrotra/f603c030f6f6dd82d2e8371384c37016d9e41b46>.
Acesso em: 25 nov. 2022.

VRIESMANN, L. M.; BRITTO JR., A. de S. Método FEATORA na classificação de câncer de mama utilizando imagens histopatológicas. *Revista Saúde e Desenvolvimento*, [S. l.], v. 12, n. 11, p. 5–23, 2018. Disponível em: <https://www.revistasuninter.com/revistasauade/index.php/saudeDesenvolvimento/article/view/947>. Acesso em: 25 nov. 2022

WANG, W. et al. Fp-detr: Detection transformer advanced by fully pre-training. In: *International Conference on Learning Representations*. 2021.

YOON, Y. et al. Learning a deep convolutional network for light-field image super-resolution. In: *Proceedings of the IEEE international conference on computer vision workshops*. 2015. p. 24-32.

Zhang, J., Zhao, X., Chen, Z., & Lu, Z. A Review of Deep Learning-Based Semantic Segmentation for Point Cloud. *IEEE Access*, 7, 179118–179133. doi:10.1109/access.2019.2958671 (2019).



Documento Digitalizado Ostensivo (Público)

TCC

Assunto: TCC
Assinado por: Everson Mariano
Tipo do Documento: Anexo
Situação: Finalizado
Nível de Acesso: Ostensivo (Público)
Tipo do Conferência: Cópia Simples

Documento assinado eletronicamente por:

- **Everson Batista Mariano, ALUNO (201721250005) DE BACHARELADO EM ENGENHARIA DE COMPUTAÇÃO - CAMPINA GRANDE**, em 17/06/2023 20:10:30.

Este documento foi armazenado no SUAP em 20/06/2023. Para comprovar sua integridade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifpb.edu.br/verificar-documento-externo/> e forneça os dados abaixo:

Código Verificador: 859402
Código de Autenticação: a826ad6616

