

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS  
ESCOLA POLITÉCNICA  
GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO



**DETECÇÃO E MEDIÇÃO AUTOMÁTICA DE HALOS PELO MÉTODO  
DISCO-DIFUSÃO (KIRBY-BAUER) UTILIZANDO APRENDIZADO  
PROFUNDO E VISÃO COMPUTACIONAL**

WILLGNNER FERREIRA SANTOS

GOIÂNIA  
2022

WILLGNNER FERREIRA SANTOS

**DETECÇÃO E MEDIÇÃO AUTOMÁTICA DE HALOS PELO MÉTODO  
DISCO-DIFUSÃO (KIRBY-BAUER) UTILIZANDO APRENDIZADO  
PROFUNDO E VISÃO COMPUTACIONAL**

Trabalho de Conclusão de Curso apresentado à Escola Politécnica, da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para a obtenção do título de Bacharel em Engenharia de Computação.

Orientador(a): Prof. Dr. Clarimar José Coelho.

Banca examinadora:

Eng. Douglas Vieira do Nascimento.

Eng. Diogo Fernandes Costa Silva.

Eng. Arthur Galdino Dangoni.

GOIÂNIA  
2022

WILLGNNER FERREIRA SANTOS

**DETECÇÃO E MEDIÇÃO AUTOMÁTICA DE HALOS PELO MÉTODO  
DISCO-DIFUSÃO (KIRBY-BAUER) UTILIZANDO APRENDIZADO  
PROFUNDO E VISÃO COMPUTACIONAL**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de Bacharel em Engenharia de Computação, e aprovado em sua forma final pela Escola Politécnica, da Pontifícia Universidade Católica de Goiás, em 13/06/2022.

---

Profa. Ma. Ludmilla Reis Pinheiro dos Santos.  
Coordenador(a) de Trabalho de Conclusão de Curso.

Banca examinadora:

---

Orientador(a): Prof. Dr. Clarimar José Coelho.

---

Coorientador(a): Eng. Douglas Vieira do Nascimento.

---

Eng. Diogo Fernandes Costa Silva.

---

Eng. Arthur Galdino Dangoni.

GOIÂNIA  
2022

“A Deus pela minha vida e oportunidades.”

“Aos meus familiares e amigos.”

## **AGRADECIMENTOS**

Aos meus pais Hilda Santos e Nirto Ferreira por todo apoio prestado e incentivo durante ao longo do curso.

Ao Prof. Orientador Dr. Clarimar, pela confiança e por ter aceitado este desafio.

Ao Prof. Dr. Arlindo Galvão, por ter ajudado neste projeto, além ter conseguido uma bolsa de Iniciação Científica (IC), com o auxílio financeiro do Centro de Excelência em Inteligência Artificial (CEIA) do Estado de Goiás.

Ao Laboratório de Biotecnologia de Microrganismos (LBMic) e à Profa. Dra. Lilian Carla, Coordenadora do LBMic e parceira do Laboratório de Computação Científica (LCC) da Pontifícia Universidade Católica de Goiás (PUC Goiás), juntamente com à Profa. Dra. Raylane Gomes, por ter fornecido a base de imagens deste projeto e por todo o auxílio prestado.

Aos meus queridos amigos Giulianni Santos, Amanda Kelly e Vinícius Biasi por estarem sempre ao meu lado durante a graduação.

Aos meus colegas do LCC, por toda a ajuda e conhecimento adquirido ao longo desta jornada. Em específico meu amigo Walcy Rios, por ter sanado minhas dúvidas deste projeto quando precisei e por sempre estar disposto em ajudar.

Para esta universidade PUC Goiás, seu corpo docente, direção e administração me forneceram uma janela na qual agora vejo uma visão mais elevada, cheia de forte confiança nas virtudes e na ética que aqui presentes.

Ao Eng. Douglas Vieira, por sempre confiar em meu trabalho, me proporcionando um excelente estágio, abrindo oportunidades no mercado de trabalho. Pela parceria em várias matérias da faculdade, pela amizade incondicional, por sempre estar ao meu lado, pelo companheirismo, pela cumplicidade, por toda sua inestimável ajuda, suas dicas preciosas, discussões técnicas e incalculáveis sugestões. Você desempenhou um papel muito significativo no meu crescimento acadêmico e profissional, e deve ser recompensado com minha eterna gratidão. Conte comigo sempre no que precisar!

A todos que direta ou indiretamente fizeram parte de minha formação, o meu muito obrigado!

“A inteligência artificial só será um problema se não reconhecermos o valor do ser humano.”

Aldo Novak.

## RESUMO

A detecção e identificação bacteriana em nível patogênico é essencial para a tomada de decisões de ações corretivas em casos de contaminação, de modo a não levar ao agravamento do estado de saúde do paciente. A identificação fenotípica também é importante na detecção bacteriana, e um de seus métodos é verificar a resistência das bactérias aos antimicrobianos. Uma metodologia clássica para essa verificação é o teste disco-difusão (Kirby-Bauer), onde os discos de antimicrobianos são dispensados em uma placa de ágar após a aplicação do inóculo bacteriano, causando a geração de halos. Halos são áreas claras formadas ao redor dos discos antimicrobianos fixados a placa de Petri, chamados de zona de inibição de crescimento, tal medida é feita por uma régua, um trabalho demorado e sujeito ao erro humano. Visando melhorar os resultados do disco-difusão, este trabalho propõe detectar e medir automaticamente os diâmetros dos halos para todos os discos, a partir de 74 imagens digitais em escala RGB. Uma RNC foi treinada para segmentar área dos discos e dos halos, posteriormente a visão computacional foi utilizada para realizar as medições dos halos com as imagens segmentadas. Os resultados são comparados com valores reais e preditos pelo modelo em visão computacional, apresentando valores de EQM EMA e  $R^2$  para os dados de treinamento e teste, sendo: 0,0816, 0,0440, 0,9053; 0,0895, 0,0533, 0,9114 respectivamente. Desta forma, conclui-se que o modelo é uma técnica promissora para ser realizada em dispositivos móveis, em prol de facilitar a rotina laboratorial.

**Palavras-chave:** Halos de inibição. Segmentação semântica. Rede neural convolucional.

## ABSTRACT

Bacterial detection and identification at the pathogenic level is essential for making decisions about corrective actions in cases of contamination, so as not to lead to a worsening of the patient's health status. Phenotypic identification is also important in bacterial detection, and one of its methods is to verify the resistance of bacteria to antimicrobials. A classic methodology for this verification is the disk-diffusion test (Kirby-Bauer), where antimicrobial disks are dispensed onto an agar plate after application of the bacterial inoculum, causing the generation of halos. Halos are clear areas formed around antimicrobial disks attached to the Petri dish, called the growth inhibition zone, this measurement is made by a ruler, a time consuming job and subject to human error. Aiming to improve the disk-diffusion results, this work proposes to automatically detect and measure the diameters of the halos for all disks, from 74 digital images in RGB scale. A RNC was trained to segment the area of the disks and the halos, later computer vision was used to perform the measurements of the halos with the segmented images. The results are compared with real values and predicted by the model in computer vision, presenting values of EQM EMA and  $R^2$  for the training and test data, being: 0.0816, 0.0440, 0.9053; 0.0895, 0.0533, 0.9114 respectively. In this way, it is concluded that the model is a promising technique to be performed on mobile devices, in order to facilitate the laboratory routine.

**Keywords:** Inhibition halos. Semantic segmentation. Convolutional neural network.



## LISTA DE FIGURAS

Figura 1 - Estrutura de uma rede neural convolucional.....	16
Figura 2 - Exemplo de uma convolução em execução. ....	17
Figura 3 - Função ReLU.....	18
Figura 4 - Representação da camada de max-pooling com stride 2. ....	18
Figura 5 - Representação da fully connected.....	19
Figura 6 - Formação de imagem pelo olho humano.....	20
Figura 7 - Demonstração da imagem digital do TSA.....	21
Figura 8 - Exemplo da aplicação de segmentação semântica.....	22
Figura 9 - Cubo da saída do segmentador semântico. ....	22
Figura 10 - TSA fundo branco (a) e fundo preto (b): .....	23
Figura 11 - Exemplo de anotação da imagem do TSA – fundo branco.....	24
Figura 12 - Bloco de construção residual. ....	25
Figura 13 - Arquitetura ResNet 34. ....	27
Figura 14 - Estrutura codificador-decodificador. ....	28
Figura 15 - Imagem original (a) e sua máscara (b): .....	31
Figura 16 -Fluxograma do algoritmo para medir os halos. ....	32
Figura 17 - Imagem original (a), predição da área dos halos segmentados (b), predição da área dos discos segmentados (c): .....	34
Figura 18 - Limiar de ativação de 0.5 para a classe dos halos (a) e discos (b): .....	35
Figura 19 - Limiar de ativação de 0.9 para a classe dos halos (a) e discos (b): .....	35
Figura 20 - Halos medidos (a) e zoom aplicado na medição (b):.....	36

## LISTA DE TABELAS

Tabela 1 - Estrutura da ResNet 34.....	27
Tabela 2 - Parâmetros da taxa de aprendizado. ....	30
Tabela 3 - Parâmetros de limiar de ativação.....	30
Tabela 4 - Resultados dos modelos de segmentação. ....	34
Tabela 5 - Erros de medidas dos halos. ....	36

## LISTA DE ABREVIATURAS E SIGLAS

TSA	Teste de Sensibilidade aos Antimicrobianos
TCIE	Transformada de Característica Invariante de Escala
IAA	Identificação Automática do Antibiograma
RDI	Região de Interesse
PDI	Processamento Digital de Imagens
RNC	Rede Neural Convolucional
IA	Inteligência Artificial
RNA	Rede Neural Artificial
MLP	Perceptron Multicamada
CTC	Camada Totalmente Conectada
RGB	<i>Red - Green - Blue</i>
IPTSP	Instituto de Patologia Tropical e Saúde Pública
UFG	Universidade Federal de Goiás
NL	Normalização em Lote
ReLU	Conformidade Linear Retificada
ISU	Interseção Sobre União
MM	Milímetro
EQM	Erro Quadrado Médio
EMA	Erro Médio Absoluto
JSON	<i>JavaScript Object Notation</i>
ROC	Reconhecimento Óptico de Caracteres
Conv	Camadas Convolucionais
<i>Max Pool</i>	<i>Max-Pooling</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>13</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>14</b>
2.1	Inteligência Artificial.....	14
2.2	Rede Neural Artificial.....	15
2.3	Rede Neural Convolucional.....	15
2.3.1	Camada de Convolução.....	16
2.3.2	Camada de Ativação.....	17
2.3.3	Camada de Agrupamento.....	18
2.3.4	Camada Totalmente Conectada.....	19
2.3.5	Classificador <i>Softmax</i> e Entropia Cruzada.....	20
2.4	Processamento de Imagens.....	20
2.4.1	Imagem Digital.....	20
2.5	Segmentação Semântica.....	21
<b>3</b>	<b>MATERIAIS E MÉTODOS.....</b>	<b>23</b>
3.1	Coleta de Amostra.....	23
3.1.1	Anotação das Imagens.....	23
3.2	Rede Neural Residual.....	24
3.2.1	ResNet.....	25
3.2.2	ResNet 34.....	25
3.2.3	Arquitetura ResNet 34.....	26
3.3	Codificador-Decodificador.....	28
3.4	Função de Custo.....	29
3.5	Intersecção Sobre União.....	29
3.6	Treinamento do Modelo RNC.....	30
3.7	Pré-Processamento.....	31
3.8	Pós-Processamento em PDI.....	31
3.9	Critério de Avaliação do Pós-Processamento em PDI.....	33
<b>4</b>	<b>RESULTADOS.....</b>	<b>33</b>
4.1.1	Resultados da Segmentação Semântica.....	33
4.1.2	Resultados do Pós-Processamento da Segmentação.....	34
4.1.3	Resultados do Pós-Processamento em PDI.....	35
<b>5</b>	<b>CONCLUSÃO.....</b>	<b>36</b>
<b>6</b>	<b>REFERÊNCIAS.....</b>	<b>37</b>

## 1 INTRODUÇÃO

A principal causa de doença e morte em humanos são doenças infecciosas provocadas por bactérias. As doenças infecciosas podem ser definidas como qualquer doença causada por bactérias e que produza a existência e reprodução de agentes biológicos (JESUS et al., 2017; RAMIREZ, 2017). Existem alguns testes para classificação de bactérias, testes fenotípicos *etc.* Um desses processos de identificação fenotípica é o Teste de Sensibilidade aos Antimicrobianos (TSA), que é uma tarefa muito importante na área da saúde. Dentre as técnicas utilizadas, uma é comumente manipulada, sendo a diluição em ágar (BHAT et al., 2021; SICURO et al., 2020).

A diluição em ágar, também conhecida como método disco-difusão (Kirby-Bauer), envolve a cultura de microrganismos apropriadamente para cada cultura. Usando substrato, ele é espalhado em toda placa de Petri, contendo vários discos de papel, cada um contendo um antimicrobiano. Depois de um tempo, a reação produzida com um antibiótico contra as bactérias aparece em várias áreas circulares de textura diferente que crescem ao redor do disco antimicrobiano. Atualmente, a avaliação e mensuração dessas áreas denominadas halos de inibição, são realizadas em laboratórios por profissionais responsáveis pelo teste, utilizando uma régua e sujeito a diversos erros, pois é um trabalho demorado e delicado (BOTH et al., 2017; NELSON et al., 2017).

Três classificações podem ser obtidas visualmente do TSA, através da presença de halos ao redor dos discos: sem halo, halo com diâmetro dentro da faixa de referência ou halo com diâmetro maior que o valor de referência. Assim, essas medidas são indicadores de resistência total, resistência intermediária ou suscetibilidade total da bactéria (BAUER et al, 1966). Por meio dos resultados do TSA, o profissional de saúde irá analisar qual antibiótico é mais adequado para tratar a bactéria no paciente, evitando assim o uso de antibióticos desnecessários que não são resistentes à infecção e prevenindo o surgimento de resistência aos medicamentos (BAUER et al, 1966). Essa técnica é usada para medir qualitativamente a atividade *in vitro* de agentes antibacterianos em termos de suscetibilidade bacteriana e resistência a drogas (MIFTAHUSSURUR et al., 2020; MORRIS et al., 2020). Em alguns tratamentos, o tempo adicional de 24 horas após a incubação pode não ser suficiente para os resultados do teste (HU et al., 2021), isso abre possibilidades para novos projetos que visam automatizar esse processo ou reduzir o tempo de incubação (DIXIT et al, 2021).

Com o objetivo de melhorar o TSA, a visão computacional faz parte de uma ampla série, baseada na representação de dados de aprendizado. As observações podem ser expressas de várias maneiras, como o vetor de valor de intensidade de cada elemento da imagem, traduzido para o inglês: *pixel*, ou de uma forma mais abstrata, como um conjunto de arestas, uma área com uma forma específica e assim por diante (MINAEE et al, 2022; QIU; NAGL, 2021). Alguns trabalhos sugerem o uso de visão computacional e aprendizado profundo para melhorar a precisão e automatizar o processo do TSA. Costa (DA COSTA, 2015) utilizou a técnica do descritor baseada em mecanismos invariantes de momento em comparação com o uso da Transformada de Característica Invariante de Escala (TCIE), um procedimento comercial utilizado pela Identificação Automática do Antibiograma (IAA).

Priya et al., (PRIYA et al., 2021) propôs um método baseado na segmentação semântica de placa de ágar como Região de Interesse (RDI) (ZHANG; LI; DU, 2018), com avaliação feita por kappa e índices de correlação. No trabalho de Costa (COSTA, 2016), foram desenvolvidas técnicas de Processamento Digital de Imagens (PDI) para automatizar o resultado do antibiograma do disco-difusão. O ambiente utilizado foi controlado por um protótipo, que recebe placas de Petri e realiza diversos processos de PDI para medir halos e informar o que está escrito nos discos. As limitações dos trabalhos citados são que eles foram desenvolvidos em ambientes controlados.

Este trabalho propõe realizar o TSA de forma ágil em um ambiente não controlado, que pode ser automatizado, menos laboratorial, podendo ser ainda mais preciso que a medição de halos manualmente. O estudo utiliza uma Rede Neural Convolutiva (RNC) para segmentar imagens digitais do TSA, com a área de discos e dos halos como RDI. Posteriormente, utilizar PDI para realizar as medições dos halos.

## **2 REFERENCIAL TEÓRICO**

### **2.1 Inteligência Artificial**

O conceito de Inteligência Artificial (IA) foi registrada em 1943, pelos cientistas McCulloch e Pitts, com base nos princípios simples da fisiologia e com as funções dos neurônios. Tais pesquisadores criaram um neurônio artificial que possuiu na época apenas dois estados, um ligado e o outro desligado (MCCULLOCH; PITTS, 1943). As pesquisas em IA começaram a aparecer com mais frequência em academias, influenciadas

principalmente por Turing em 1950 que propôs o teste de Turing que consiste em verificar se uma máquina poderia imitar um humano em conversas de texto (TURING, 1950).

IA é um campo da ciência que analisa a composição de agentes inteligentes, visando buscar um jeito de permanência para as máquinas solucionarem problemas complexos. Fornece soluções para que as máquinas consigam resolver de forma intuitiva, visto que os seres humanos resolveriam de forma inconsciente (ROSEBROCK, 2017).

## **2.2 Rede Neural Artificial**

Uma Rede Neural Artificial (RNA) é um método de aprendizado profundo capaz de reconhecer padrões complexos ou abundantes e ter aprendizado dentro da própria rede. As RNAs são baseadas na arquitetura dos neurônios humanos e visam replicar o aprendizado multiplicando sistemas que aprendem com amostras de treinamento. Uma RNA consiste em uma camada de entrada, uma camada de saída e pelo menos uma camada oculta, que converte os dados recebidos pela entrada em dados que podem ser usados pela saída. Essas camadas são compostas de neurônios (KROGH, 2008).

Frank Rosenblatt foi o primeiro cientista a usar um neurônio artificial, inspirado pelo neurônio dos seres humanos em 1958, denominado de Perceptron, visto que tem uma simples arquitetura está ultrapassado para solucionar desafios complexos, por falta de rigor e estabilidade, juntamente com sua função de ativação que é o que determina se o neurônio será ativado ou não (ROSENBLATT, 1958).

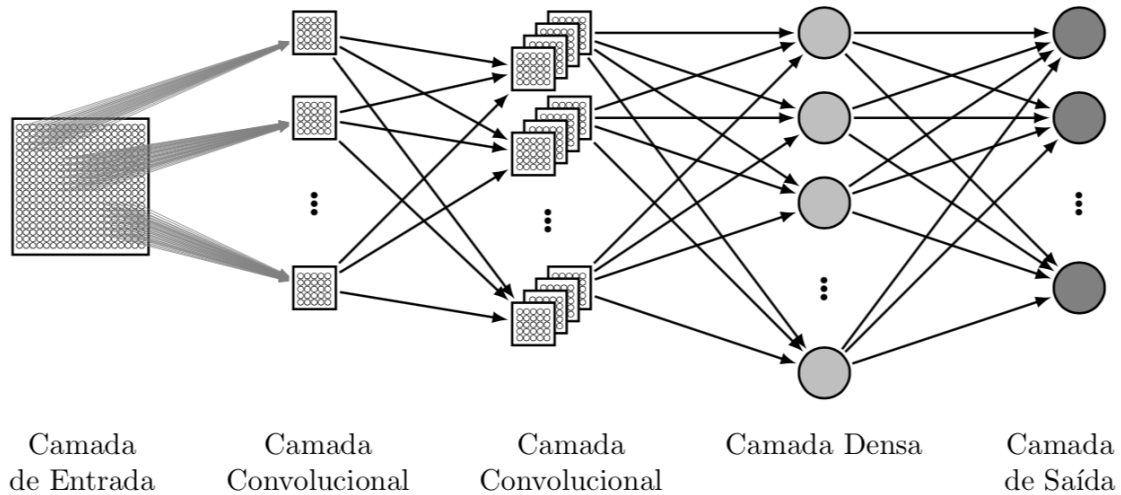
## **2.3 Rede Neural Convolutacional**

A aplicação de camadas de convolução em conjunto de dados pode ser utilizada como extratora de características implícitas nos dados (ALBAWI; MOHAMMED; AL-ZAWI, 2017). Uma RNC é uma arquitetura composta de quatro camadas distintas: uma camada de entrada, uma camada convolutacional com sua respectiva função de ativação, uma camada de agrupamento e uma camada densa responsável por fazer a regressão ou classificação. O parâmetro  $k$  especifica o número de mapas de recursos na camada convolutacional (ALBAWI; MOHAMMED; AL-ZAWI, 2017).

A Figura 1 representa que as RNCs são diferenciadas porque possuem muitas camadas convolucionais, que após aplicada uma função matemática de convolução nos conjuntos de dados de entrada e por fim realizado o agrupamento. A saída convolutacional é passada para a próxima camada convolutacional até atingir a última camada, chamada de

camada densa, que geralmente é representada por uma Perceptron Multicamada (MLP) (SAKURAI, 2017).

Figura 1 - Estrutura de uma rede neural convolucional.



Fonte: (SAKURAI, 2017).

### 2.3.1 Camada de Convolução

Na operação básica de um modelo (neurônio artificial), a convolução é a operação necessária. É responsável por identificar as características de uma imagem e gerar uma matriz de pesos representando estas características, ou seja, um mapa de características (UCHIDA; TANAKA; OKUTOMI, 2018).

Na operação de convolução, uma multiplicação de dois elementos da matriz é realizada. Os elementos dessa matriz de *pixels* da imagem são multiplicados pelos elementos da matriz de filtros, ou núcleo, traduzido para o inglês: (*kernel*), e por fim os produtos são somados (UCHIDA; TANAKA; OKUTOMI, 2018).

A matriz do filtro é menor que a matriz da imagem, a matriz do filtro vai se movendo da esquerda para a direita e de cima para baixo na matriz da imagem, representado na Figura 2, e após cada movimento, a convolução é realizada e o resultado é alocado.



Figura 2 - Exemplo de uma convolução em execução.

131	162	232	84	91	207
104	-1	0	+1	237	109
243	-2	0	+2	135	26
185	-1	0	+1	61	225
157	124	25	14	102	108
5	155	16	218	232	249

Fonte: (ROSEBROCK, 2017, p. 172).

A convolução é representada pela seguinte equação:

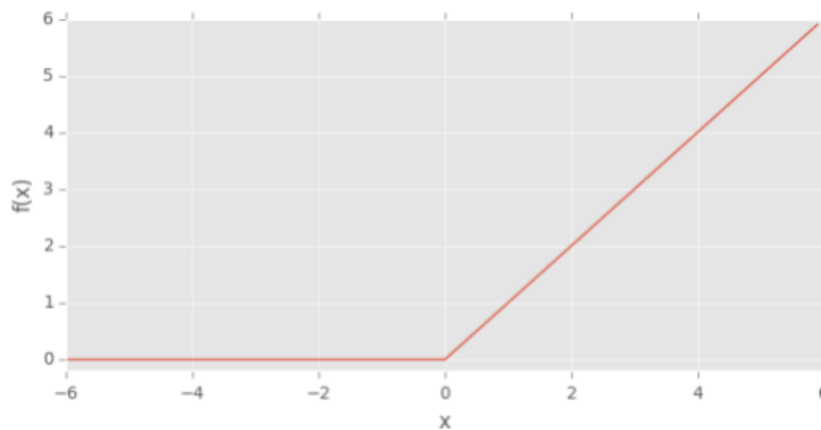
$$z(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) \cdot f(x - s, y - t) \quad (1)$$

onde cada etapa tem com resultado o somatório da multiplicação dos elementos do *kernel*, corresponde a cada *pixel* da imagem, levando em conta a região do passo que está sendo executado. O consequente dos passos fica com o tamanho menor, onde  $z$  representa a imagem convolucionada, com base na imagem  $f$  em função do *kernel*  $w$  (GONZALEZ; WOODS, 2009).

### 2.3.2 Camada de Ativação

Depois de aplicar a convolução, é ativado uma função de ativação não linear. Essa função é encarregada por diminuir a saída da camada convolucional a uma faixa de valores admissíveis. A saída dessa camada tem o mesmo número de mapas de recursos que a entrada e suas dimensões não mudam. KRIG (KRIG, 2016) reforça a importância da camada de ativação ReLU ao mostrar que ao projetar a convolução em um espaço não linear, ignorando o espaço vazio (valores negativos na matriz de características), obtém-se melhor o resultado e aumenta a convergência de rede durante o treinamento usando a retropropagação. A Figura 3 representa o gráfico da função ReLU onde os valores negativos de  $x$  são zerados e os valores não negativos permanecem os mesmos.

Figura 3 - Função ReLU.

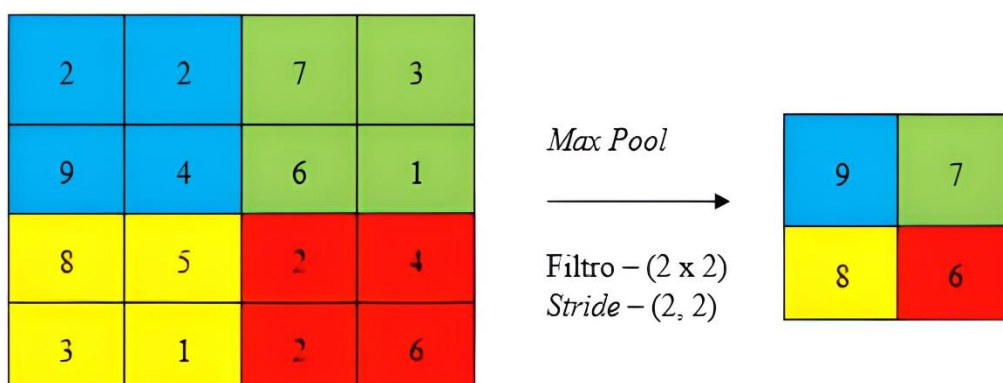


Fonte: (ROSEBROCK, 2017, p. 125).

### 2.3.3 Camada de Agrupamento

As camadas de *pooling*, geralmente usadas após as camadas convolucionais, são semelhantes às convoluções em termos de operações de deslizamento usando janelas e passos largos, traduzido para o inglês: (*strides*). Seu papel é reduzir a dimensionalidade, reduzindo o tamanho do mapa de características gerado pela camada convolucional (ou camada anterior), reduzindo assim a quantidade de informações a serem processadas pela próxima camada. Uma técnica amplamente utilizada é o agrupamento máximo, traduzido para o inglês: (*max pooling*), em que os dados são reduzidos selecionando o maior valor em uma determinada região do filtro (YU et al., 2014), aplicada na Figura 4. Demais procedimentos também podem ser realizados, como a média dos valores das janelas (LACERDA, 2019).

Figura 4 - Representação da camada de *max-pooling* com *stride 2*.



Fonte: (Autoria Própria, 2022).

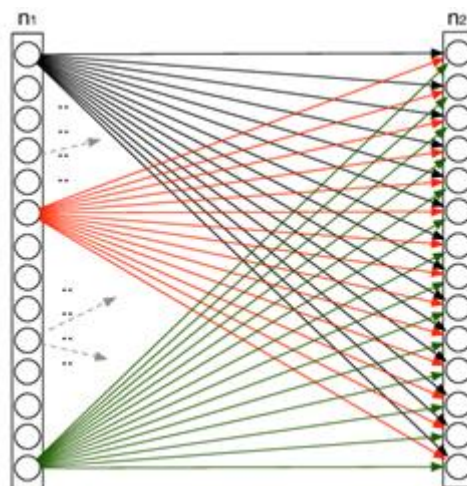
### 2.3.4 Camada Totalmente Conectada

A Camada Totalmente Conectada (CTC), traduzido para o inglês: (*fully connected*), é iniciado o processo de classificação das informações extraídas pela camada anterior. A camada totalmente conectada nivela os sub-blocos contendo os dados extraídos, ou seja, converte o bloco em uma única linha contendo todas as informações extraídas (MA; LU, 2017).

A CTC sempre aparece no final de uma arquitetura RNC. Krig (KRIG, 2016) apontou alguns usos para a CTC, seja como uma conexão completa de todos os pesos de recursos em um vetor de uma dimensão para classificação linear, ou como um redutor, convertendo-o em uma pequena camada, reduzindo assim o número de parâmetros. Krig (KRIG, 2016) destacou que a CTC trabalhada como intermediação entre as RNC, faz uma associação da camada de convolução e função *softmax*.

Visto que a saída da camada anterior é um vetor linha com as características obtidas pelas camadas convolucionais, todo esse vetor é passado para a CTC (pode ter mais de 1 camada oculta) e todos os neurônios entre as camadas são interligados (Figura 5). Calcula-se os *logits* e em seguida usa-se a função *softmax* para calcular a probabilidade, associada a cada classe. Esta saída equivale ao tamanho ao número de classes e para cada classe é gerado valores entre 0 e 1 para demonstrar a probabilidade (*score*), sendo 0 a probabilidade mínima e 1 a probabilidade máxima (MA; LU, 2017).

Figura 5 - Representação da *fully connected*.



Fonte: (MA; LU, 2017).

### 2.3.5 Classificador *Softmax* e Entropia Cruzada

O classificador *softmax* é uma função de ativação, podendo ser logística ou exponencial. Essa função aparece logo depois da camada CTC, ou seja, no final da RNC, seu objetivo é classificar as imagens, possibilitando probabilidades para todas as classes de saída, normalizadas entre valores de 0 e 1 (KRIG, 2016, p.327).

Rosebrock (ROSEBROCK, 2017) afirma que a entropia cruzada permite calcular o desempenho de uma rede neural em um determinado conjunto de treinamento. Ou seja, é uma função para verificar e minimizar o erro, tendo como base a probabilidade, pelo *softmax*.

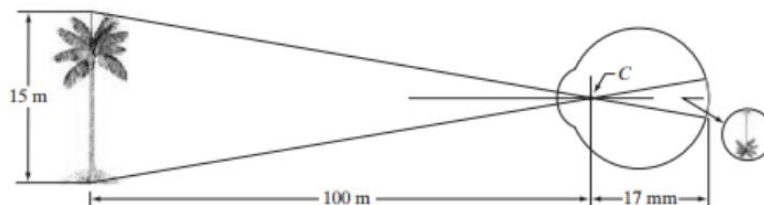
Os modelos de aprendizado supervisionado tais como RNCs, durante o treinamento, se a previsão estiver incorreta, os pesos e tendência são atualizados. O algoritmo usado para correção de erros é o *backpropagation* (O'SHEA; NASH, 2015).

## 2.4 Processamento de Imagens

O PDI é um conjunto de técnicas que utilizam computadores para capturar, representar e transformar imagens. (PEDRINI; SCHWARTZ, 2007). O processo de aquisição de uma imagem digital, envolve a captura da mesa, usando um dispositivo que converte sinais analógicos em dados digitais. O dispositivo pode ser uma câmera, um tomógrafo médico, um scanner, um satélite ou similar (PEDRINI; SCHWARTZ, 2007).

No olho humano, o cristalino é acomodado pelo músculo ciliar, que afina ou engrossa o cristalino para que possa ajustar o foco a objetos localizados a diferentes distâncias. A Figura 6 representa a imagem produzida pelo olho humano, onde a luz que entra no olho é projetada na retina, onde estão localizados os cones e bastonetes.

Figura 6 - Formação de imagem pelo olho humano.



Fonte: (GONZALEZ & WOODS, 2008, p. 39).

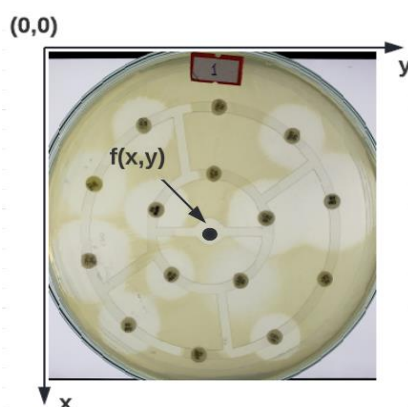
### 2.4.1 Imagem Digital

Uma imagem digital é uma representação de uma cena por um grupo de tamanho finito de elementos discretos denominados de *pixels*, acoplados em uma matriz

bidimensional. A cada *pixel* é integrado, um valor, no caso de imagens em escala de cinza, ou um conjunto de três valores, vermelho – verde-azul, traduzido para o inglês: (*Red - Green - Blue*) - RGB para representar uma cor (GONZALEZ; WOODS, 2008).

A representação de imagens em formato digital é feita por meio de matrizes, que mapeiam objetos e cores do mundo real para o mundo digital. A matriz é composta por elementos de imagem, também chamados de *pixels*. Um *pixel* representa um ponto amostral no eixo  $f(x, y)$ , contendo um valor que representa a cor naquela coordenada, conforme representado na Figura 7 (GONZALEZ; WOODS, 2008).

Figura 7 - Demonstração da imagem digital do TSA.



Fonte: (Autoria Própria, 2022).

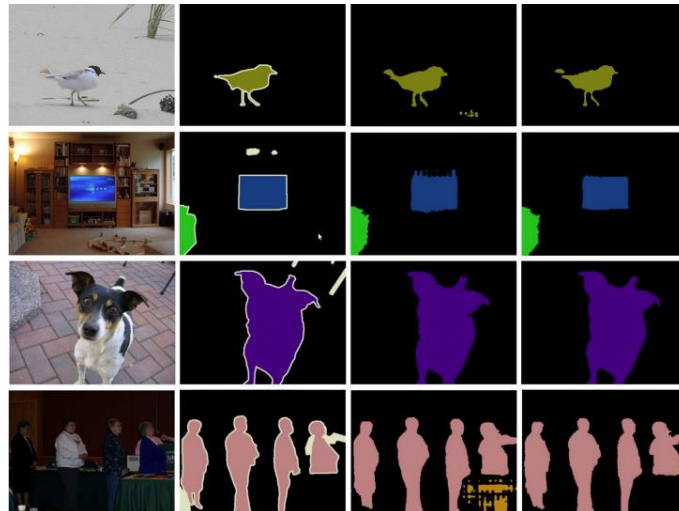
## 2.5 Segmentação Semântica

A segmentação semântica tem como objetivo classificar cada *pixel* de uma imagem. A segmentação difere da classificação, pois a segmentação atribui uma classe a cada *pixel* da imagem, já a classificação concede uma classe à imagem inteira. Algumas redes parecidas com segmentação semântica podem ser trabalhadas para imagens em tons de cinza (WANG et al., 2018).

Em específico, o foco da segmentação semântica de imagens é rotular *pixel a pixel* na imagem, de acordo com a classe equivalente que está sendo representada (WANG et al., 2018). Neste trabalho, a segmentação semântica foi utilizada para prever cada *pixel* nas imagens digitais das placas de Petri, sendo o RDI da área dos halos e dos discos.

Na Figura 8, da esquerda para a direita: imagem original, imagem anotada e o resultado antes e depois pela segmentação semântica.

Figura 8 - Exemplo da aplicação de segmentação semântica.



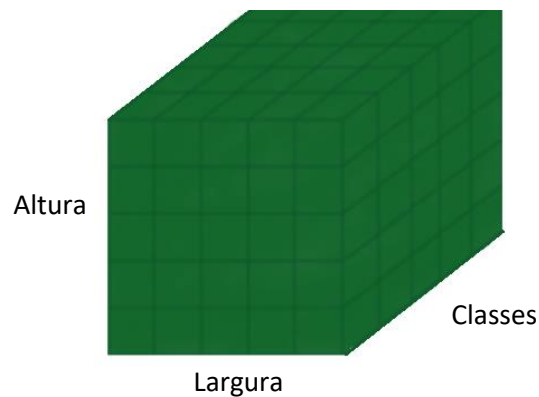
Fonte: (WANG et al., 2018).

A segmentação semântica não separa instâncias da uma mesma classe. O objetivo maior é a classe de cada *pixel* (WANG et al., 2018). Por exemplo, se tiver dois objetos da mesma classe (halos/discos) na imagem de entrada, o esquema de segmentação não os difere essencialmente como objetos separados.

Um das tarefas da segmentação semântica, está relacionado com uma arquitetura codificador-decodificador, onde a etapa codificadora extrai características dominantes para que a etapa decodificadora possa construir semanticamente atributos discriminativos que são aprendidos. (ZHOU et al., 2018).

A Figura 9 representa o segmentador semântico, tendo três tamanhos, no caso um cubo de imagens, variando a altura, largura e classes. As classes são os quantitativos dos conjuntos de dados, tendo a classe incluída como fundo, altura e largura, são os atributos da imagem.

Figura 9 - Cubo da saída do segmentador semântico.



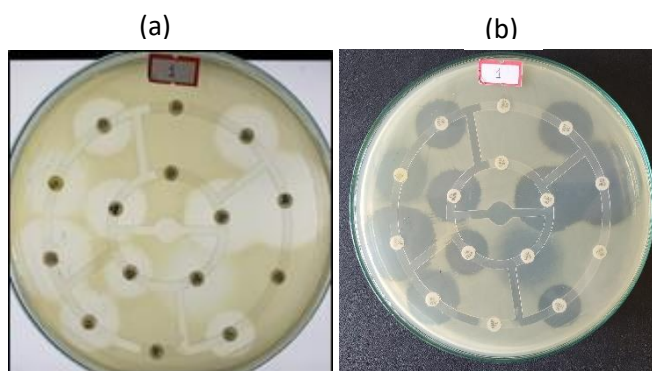
Fonte: (Autoria Própria, 2022).

### 3 MATERIAIS E MÉTODOS

#### 3.1 Coleta de Amostra

Este trabalho utilizou dados do TSA, fornecidos cordialmente pelo Instituto de Patologia Tropical e Saúde Pública (IPTSP) da Universidade Federal de Goiás (UFG), localizado em Goiânia - Goiás. O IPTSP (EDITORES, 2007) é considerado uma unidade multidisciplinar para o estudo de doenças tropicais e tem como objetivo realizar análises médicas, microbiológicas, biotecnológicas de microrganismos e disponibilizou um conjunto de dados de 74 imagens originais do TSA, totalizando 37 imagens com o fundo branco e 37 imagens com o fundo preto, tendo 555 discos e 452 halos para treinamento e propostas de testes. A Figura 10 representa uma imagem digital original que o IPTSP forneceu, contendo os halos e os discos.

Figura 10 - TSA fundo branco (a) e fundo preto (b):



Fonte: (IPTSP, 2022).

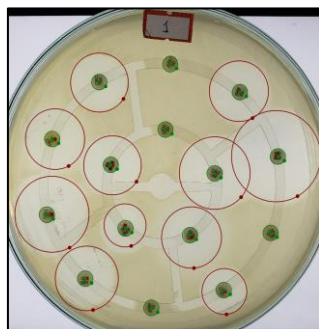
Para realização do experimento, essas imagens foram capturadas por um *smartphone* Samsung, câmera modelo SM-G965U1. As imagens do TSA foram separadas empiricamente em 80% para treinamento e 20% para teste. As imagens foram salvas em formato com 256 níveis de cada componente RGB de dimensões 2500 x 2500 *pixels*.

##### 3.1.1 Anotação das Imagens

Para que o modelo aprenda a reconhecer os padrões das imagens do TSA, é necessário utilizar exemplos e, para identificar corretamente os objetos em uma imagem, é preciso capturar 100% dos *pixels* das imagens. Anotar e rotular objetos com base em seu nível de detalhes leva muito tempo, o que piora dependendo da complexidade da imagem. Quando se trata de agilidade, economizar tempo e custo por meio da anotação de imagem é especialmente importante (BERGER; SEUNG; LICHTMAN, 2018).

Na Figura 11, o círculo em vermelho é classe halo e o verde é a classe do disco. Todas as imagens foram anotadas com a ferramenta *LabelMe* (WADA, K., 2021), com as RDIs devidamente preenchidas com suas respectivas cores, para que todos os *pixels* fossem capturados e segmentados utilizando aprendizado profundo.

Figura 11 - Exemplo de anotação da imagem do TSA – fundo branco.



Fonte: (Autoria Própria, 2022).

Todas as rotinas computacionais foram programadas em MATLAB e Python.

### 3.2 Rede Neural Residual

O bloco de composição residual consiste em muitas camadas convolucionais (Conv), normalização em lote (NL), funções de ativação de Conformidade Linear Retificada (ReLU) e caminhos (GAO et al., 2021). A saída do bloco residual pode ser representada por:

$$y = F(x) + x \quad (1)$$

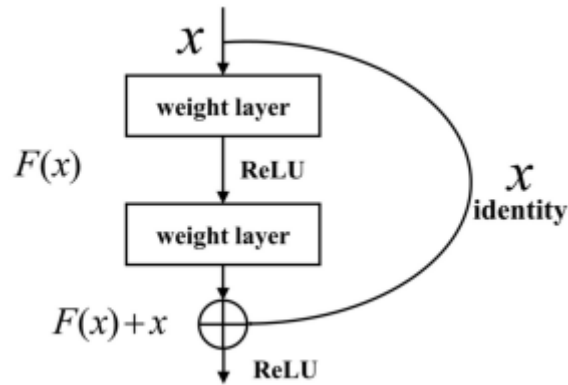
onde  $F$  é a função residual e  $x$  e  $y$  as entradas/saídas da função residual, relativamente. A rede residual consiste na precedente camada de rede convolucional, juntamente com vários blocos de construção (GAO et al., 2021).

Um exemplo é o cenário para prever o preço de 10 antibióticos contra uma bactéria, apenas observando para eles. Se o preço real dos antibióticos é de R\$ 3.000,00 e foi previsto R\$ 4.000,00, então o resíduo é -1.000,00 R\$. Caso contrário, se fosse previsto os antibióticos como R\$ 1.000,00, então o residual neste caso seria mais R\$ 2.000,00. Neste caso, o residual é o valor ou número pelo qual se deve alterar sua previsão para atingir o valor real (GAO et al., 2021).



Este cenário pode ser representado na Figura 12:

Figura 12 - Bloco de construção residual.



Fonte: (GAO et al., 2021).

### 3.2.1 ResNet

A ResNet é importante e muito utilizada em casos de visão computacional (GAO et al., 2021). A ResNet lida com uma condição de aprendizado residual para descomplexificar a geração de redes profundas e sua arquitetura equivale a blocos residuais (PENG et al., 2021). Redes profundas são compostas por muitas funções não lineares, onde as dependências entre camadas podem ser muito complexas, tornando os cálculos de gradientes instáveis (TOMITA; CHEUNG; HASSANPOUR, 2018).

Para resolver este problema, a ResNet dá início a conexões de salto de combinação para contornar as camadas residuais, de forma que os dados passem sem desvios por quaisquer camadas seguidamente a outras. Desta forma, eles se intercalam na representação residual em vez de camadas se encaixando no mapeamento implícito (TOMITA; CHEUNG; HASSANPOUR, 2018).

### 3.2.2 ResNet 34

É considerado a entrada  $x$  uma demonstração subentendida  $H(x)$  a ser ajustado por 34 camadas que são empilhadas (KOONCE, 2021), a representação residual é definida como:

$$F(x) = H(x) - x \quad (3)$$

A arquitetura residual única é visualizada em:

$$\hat{x} = \sigma(F(x, W)) + H(x) \quad (4)$$

já que  $\hat{x}$  é a saída de apenas um resíduo,  $H(x)$  é uma apresentação de ajuste,  $W$  é uma classe de peso, a função  $\sigma$  é a funciona como intensificações do nó ReLU e  $F(x, W)$  tem como performance uma residual a ser aprendida. A ResNet utilizado neste trabalho é chamado de ResNet34, equilibrado por 34 camadas, que são compostas por blocos convolucionais e residuais (KOONCE, 2021).

Logo depois, uma camada de filtro convolucional ( $7 \times 7$ ) é empregada, contínua por um bloco residual, juntamente com uma série de camadas residuais filtradas ( $3 \times 3$ ). A última camada é conectada com uma camada totalmente conectada seguida por uma camada de regressão (KOONCE, 2021).

### 3.2.3 Arquitetura ResNet 34

No início da arquitetura da ResNet 34, é composta por uma camada convolucional, tendo 64 filtros com um tamanho de *kernel* de  $7 \times 7$ , no caso é a primeira convolução, contínua por uma camada de *max pooling*. Nos dois casos, é detalhada a etapa 2. Então, em *conv2\_x* tem a camada de *pooling*, seguida das camadas convolucionais. Devido à forma como os resíduos estão conectados, essas camadas geralmente são agrupadas em pares. As setas indicadoras manifestam que elas são saltadas a cada duas camadas (KUMAR et al., 2020).

Nas duas camadas com o tamanho do *kernel* de  $3 \times 3$ , número de filtros em 64, e todos consecutivos vezes 3, correspondendo às camadas entre *pool*, dividido por 2 com o filtro de 128, para um total de 6 camadas (um par multiplicado por 3). Tais camadas têm o tamanho do *kernel* de  $3 \times 3$  com os números dos filtros de 128, elas também se repetem, mas com o valor em 4. Na medida que o número de filtros é dobrado, é visto que a primeira camada especifica os números dos filtros divididos por 2 (KUMAR et al., 2020), a arquitetura da ResNet 34 é representada na Figura 13.

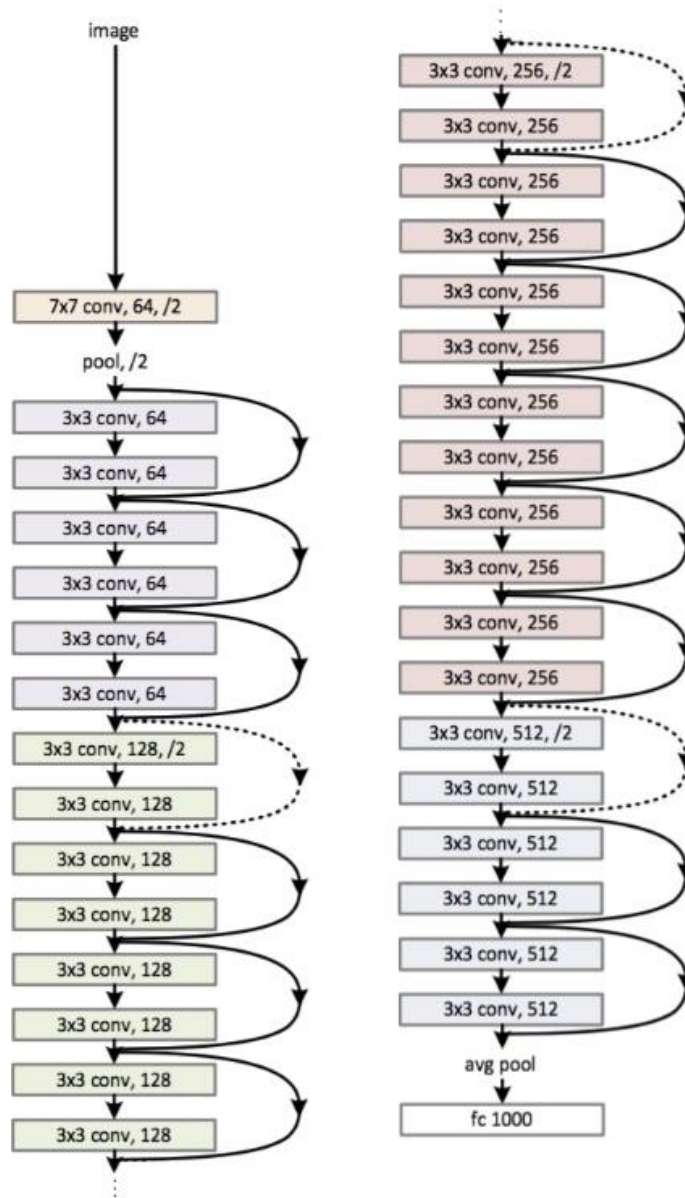
Ao implementar qualquer arquitetura da ResNet em 5 blocos, representado na Tabela 1, o primeiro bloco conta com 64 filtros com *stride* de 2, constando com a *max-pooling* de *stride* 2 também. Na arquitetura, é utilizado um preenchimento de 3. A ativação ReLu é acionada no final. No primeiro bloco, é igual para qualquer arquitetura, seguido pelos outros blocos que são modificados para cada camada, contínuos em um mesmo padrão (KORFIATIS et al., 2017).

Tabela 1 - Estrutura da ResNet 34.

Nome da camada	Tamanho da saída	34-Camadas
Conv1	112 x 112	7 x 7, 64, stride 2 3 x 3 max pooling, stride 2
Conv2_x	56 x 56	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 3$
Conv3_x	28 x 28	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 4$
Conv4_x	14 x 14	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 6$
Conv5_x	7 x 7	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 3$

Fonte: (Adaptado de GAO et al., 2022).

Figura 13 - Arquitetura ResNet 34.



Fonte: (KUMAR et al., 2020).

O segundo bloco é desenvolvido com um *pool* máximo com início em 1 x 1, o tamanho do *kernel* é de 3 x 3, com dois blocos residuais, tais blocos recebem 64 de canais de entrada e saída, são constantes 3 vezes, conforme a figura 11. Ao fim, o último bloco é ligado ao próximo bloco, contendo um *stride* de 2 (KORFIATIS et al., 2017)

No terceiro bloco, o tamanho do *kernel* é de 3 x 3, contendo um canal de acesso, chegando do último bloco de 64, tendo um canal de saída em 128. Os 128 ficam como entrada, o mesmo valor para o canal de saída. Do mesmo jeito, o quarto e quinto blocos são desenvolvidos, com 256 e 512 no tamanho do *kernel* (KORFIATIS et al., 2017).

### 3.3 Codificador-Decodificador

Este trabalho propõe uma estrutura de rede codificadora (ResNet 34) e decodificadora (Unetplusplus), (ZHOU et al., 2018) na tarefa de segmentação semântica RGB interior. Os blocos residuais são empregados como blocos de composição para barrar o problema de degradação do modelo (ZHOU et al., 2018). Permitindo que a atuação da rede se expanda à medida que a estrutura se aprofunda. A estrutura de salto é utilizada para circundar a cognição espacial do codificador para o decodificador. No decorrer do treinamento da rede, ele efetua diversas camadas no decodificador para uma melhor otimização (ZHOU et al., 2018).

Na Figura 14, é representado a arquitetura do codificador-decodificador, o codificador atribui um segmento da extensão da variável como entrada e a modifica no estado com uma configuração que não muda. O decodificador representa o estado codificado da configuração fixa, para posteriormente ter uma continuidade da extensão da variável (ZHOU et al., 2018).

Figura 14 - Estrutura codificador-decodificador.



Fonte: (Autoria Própria, 2022).

Uma rede genérica representa três componentes nomeados, codificador, decodificador e regressor. Em uma determinada resolução espacial, o codificador é completamente convolucional. Desta forma, o decodificador amplia aos poucos a resolução dos mapas de características, que por fim são conectados a um regressor que consiste em três camadas, sendo elas totalmente conectadas (ZHOU et al., 2018).

### 3.4 Função de Custo

A função de custo trabalhada no modelo é denominada Perda Focal, onde tem relação com um desequilíbrio de uma camada no decorrer do treinamento, focando na detecção de objetos. A perda leva a um termo de modulação de perda de entropia cruzada, focando no aprendizado em amostras difíceis e mal categorizadas. Essa perda é uma entropia cruzada calculada ativamente, onde o fator de escala tende a cair a zero, na medida em que a segurança na classe certa vai aumentando. Espontaneamente, esse fator de escala pode diminuir involuntariamente a contribuição de amostras fáceis ao longo do treinamento e salientar agilmente o modelo em exemplos mais complexos (LIN et al., 2017).

É definido a equação da Perda Focal desta maneira:

$$(1 - p_t)^\gamma \quad (5)$$

onde as amostras são mal classificadas e  $p_t$  é reduzido, o fator de modulação tende a ficar mais próximo de 1, desta maneira, a perda não é prejudicada. Visto que o fator  $p_t$  transforma em 0 e a perda para exemplos bons classificados vai diminuindo. O parâmetro de foco  $\gamma$  vai calibrando devagar a taxa de diminuição de modelo único. À medida que  $\gamma$  cresce, ocorre o mesmo com o efeito do fator de modulação (LIN et al., 2017).

A Perda Focal foi utilizada pelo fato as classes halos, discos e contendo o fundo também, sendo que é esperado ter muito mais classes fundo em relação as outras duas, ou seja, elas são mal classificadas.

### 3.5 Intersecção Sobre União

A Intersecção sobre União (ISU) é uma métrica de avaliação bastante popular para tarefas como segmentação, detecção de objetos e rastreamento (REZATOFIGHI et al., 2019).

ISU é um termo empregado na detecção de objetos, tendo duas tarefas primordiais: localização e classificação. A localização informa onde o objeto está na imagem, já a classificação, concede uma classe para o objeto. A ISU é bastante aplicada em tarefas de detecção de objetos, onde cada modelo treinado gera caixas, que fazem a combinação ideal com os objetos. Neste trabalho, os objetos detectados são os halos e os discos, da mesma a forma as classes classificadas com halos e discos (REZATOFIGHI et al., 2019).

O ISU é representado da seguinte forma:

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{|I|}{|U|} \quad (6)$$

onde  $A$  e  $B$  representam as caixas delimitadoras de previsão e verdade, mutuamente (REZATOFIGHI et al., 2019).

### 3.6 Treinamento do Modelo RNC

Para o treinamento do modelo, foi utilizada uma janela de busca de taxa de aprendizado e é necessário um limiar de ativação para servir de filtro para a confiança do modelo (limiar), para cada combinação foi treinado um modelo com 100 épocas. Os valores dos parâmetros e o número de épocas foram escolhidos empiricamente, tendo em vista que foi treinado até 500 épocas e o pico da acurácia foi encontrado em 100 épocas. Com todos os modelos treinados, apenas aquele com maior valor de ISU foi testado no pós-processamento. Os parâmetros são descritos na Tabela 2 e Tabela 3:

Tabela 2 - Parâmetros da taxa de aprendizado.

<b>Taxa de aprendizado</b>
0.001
0.003
0.005
0.007
0.009
0.01
0.07
0.09

Fonte: (Autoria Própria, 2022).

Tabela 3 - Parâmetros de limiar de ativação.

<b>Limiar de ativação</b>
0.5
0.6
0.7
0.8
0.9

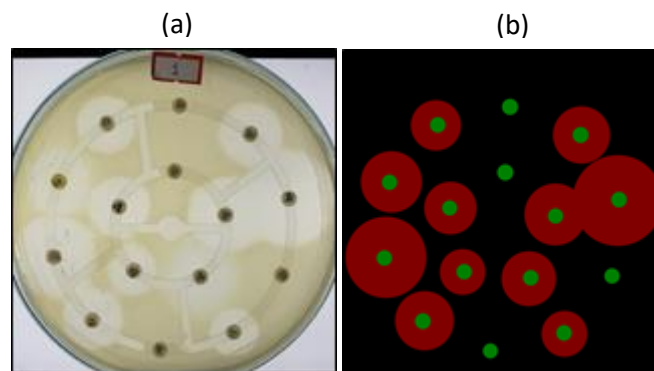
Fonte: (Autoria Própria, 2022).

### 3.7 Pré-Processamento

As técnicas de pré-processamento têm o efeito de melhorar a qualidade da imagem. Essas técnicas se enquadram em duas categorias principais: métodos que operam no domínio espacial e métodos que operam no domínio da frequência. As técnicas de processamento no domínio espacial são baseadas em filtros que manipulam o plano da imagem, enquanto as técnicas de processamento no domínio da frequência são baseadas em filtros que atuam no espectro da imagem. É comum aprimorar determinados recursos de uma imagem, combinando vários métodos baseados nessas duas categorias (ZHANG; ZHAO, 2018). Os tamanhos das imagens do TSA de 2500 x 2500 *pixels* foram redimensionados no pré-processamento para tamanhos de 512 x 512 *pixels*, esse processo foi realizado para melhorar o pós-processamento.

O método proposto é baseado na segmentação semântica supervisionada. O modelo nesta tarefa é outra imagem do mesmo tamanho, o valor da classe que contém a classe *pixel a pixel*, chamada de máscara (VEZHNEVETS; FERRARI; BUHMANN, 2011). Após o processamento da anotação, um arquivo Notação de Objeto Java Script (JSON), traduzido para o inglês: (*JavaScript Object Notation*) contendo todos os pontos dos polígonos rotulados, que são atribuídos para cada imagem, e uma imagem do mesmo tamanho é gerada através do código de programação Python (Figura 15), com apenas um canal de cor e valores entre 0 e N -1, onde N é a quantidade da classe.

Figura 15 - Imagem original (a) e sua máscara (b):



Fonte: (Autoria Própria, 2022).

### 3.8 Pós-Processamento em PDI

O pós-processamento foi trabalhado em mensurar os halos com base na previsão do modelo de segmentação, este trabalho aplicou o pós-processamento somente no modelo de valor maior de ISU. As previsões do modelo foram contidas por uma imagem

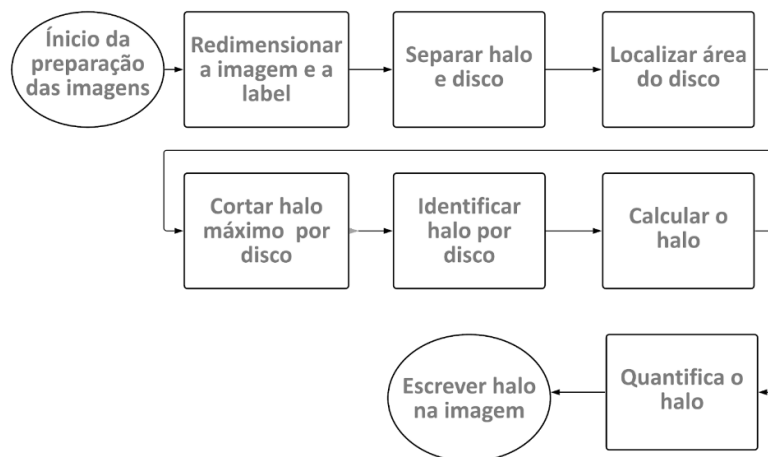
em que cada canal de cor conferiu a uma classe, classe 1 (primeiro canal) é área dos halos e classe 2 (segundo canal), é área de discos. Os discos foram separados em forma de círculos, igualmente para os halos, então a transformada de Hough foi usada (RIZON et al., 2005), para detectar a localização dos discos nas imagens para conseguir obter o raio dos discos.

Para medir os halos automaticamente é essencial ajustar a distância da borda do halo com o centro do disco. Para encontrar a borda do halo, foi trabalhado a binarização com limiar, calibrado pelo algoritmo de Otsu, com o objetivo de alterar os valores dos *pixels* de acordo com a área do halo tendo 0 e o fundo correspondente a 1, buscando todo o canal da imagem com a classe halos limiarizada pelos contornos da borda de cada disco, alternando a linha superior e inferior em relação a linha do meio do disco e transformando a coluna para a direita e esquerda, obtendo quatro valores e a média deles retratadas com a distância média absoluta do raio entre a borda do halo e o meio do disco, desta forma é calculado o diâmetro com a seguinte função:

$$v = 2 * \left( \frac{\psi * \omega}{\theta} \right) \quad (7)$$

onde  $\psi$  é a distância média absoluta entre o meio do disco e a borda do halo com previsão do modelo,  $\theta$  representa a distância entre o meio e a borda do disco com base na previsão do modelo,  $\omega$  é o valor verdadeiro do raio dos discos, os discos trabalhados neste trabalho possuem 7 mm, com um raio de 3,5 mm. Por fim,  $v$  é o diâmetro de halo. A Figura 16 representa todo pós-processamento da mensuração dos halos.

Figura 16 -Fluxograma do algoritmo para medir os halos.



Fonte: (Autoria Própria, 2022).



### 3.9 Critério de Avaliação do Pós-Processamento em PDI

O método proposto no pós-processamento foi avaliado por três critérios distintos: Erro Quadrado Médio (EQM), expresso pela equação (8), Erro Médio Absoluto (EMA) na equação (9) e Coeficiente de Determinação ( $R^2$ ) na equação (10).

$$EQM = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (8)$$

$$EMA = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (10)$$

onde  $N$  são todos os dados,  $\hat{y}_i$  são os valores preditos,  $y_i$  são os valores reais e  $\bar{y}$  os valores medidos das mensurações dos halos. O EQM é a medida que calcula a raiz quadrada média dos erros entre valores observados (reais) e previsões (previsões). O EMA calcula a média sobre o conjunto de dados entre estimativa e medições de halos por imagens digitais. O  $R^2$  mostra a combinação da variância da estimativa e a variância total dos dados. Sobre os valores, EQM e EMA têm seu ótimo em 0 e o  $R^2$  em 1. Os valores de avaliação desses erros são mostrados na seção de resultados.

## 4 RESULTADOS

### 4.1.1 Resultados da Segmentação Semântica

Para avaliar o método proposto, os resultados do pós-processamento foram comparados com a medição dos halos fornecida pelo IPTSP. A segmentação semântica é treinada e testada em 40 modelos diferentes da arquitetura da ResNet 34, para segmentar as imagens digitais do TSA. Devido a muitos experimentos, a Tabela 4 representa os cinco melhores modelos, com seus limiares de ativação e os valores de ISU.

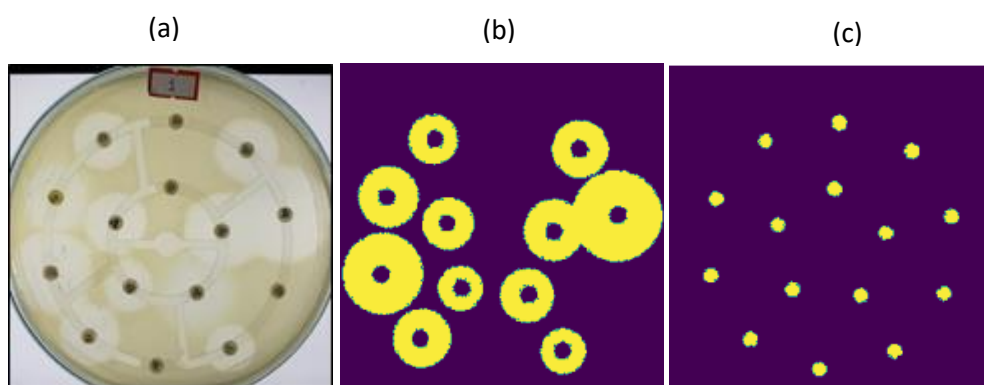
Tabela 4 - Resultados dos modelos de segmentação.

Treino		Teste	
Limiar de ativação	ISU	Limiar de ativação	ISU
0.5	0.9204	0.5	0.9284
0.6	0.9162	0.6	0.9283
0.7	0.9186	0.7	0.9203
0.8	0.9119	0.8	0.9164
0.9	0.9013	0.9	0.9117

Fonte: (Autoria Própria, 2022).

A Figura 17 representa o resultado da predição do modelo de segmentação semântica, sendo a área dos halos e discos segmentados, onde as anotações das áreas deles foram classificadas, encontrados na caixa delimitadora e por fim agrupado os *pixels* da imagem localizada, seguida pela máscara de segmentação.

Figura 17 - Imagem original (a), predição da área dos halos segmentados (b), predição da área dos discos segmentados (c):



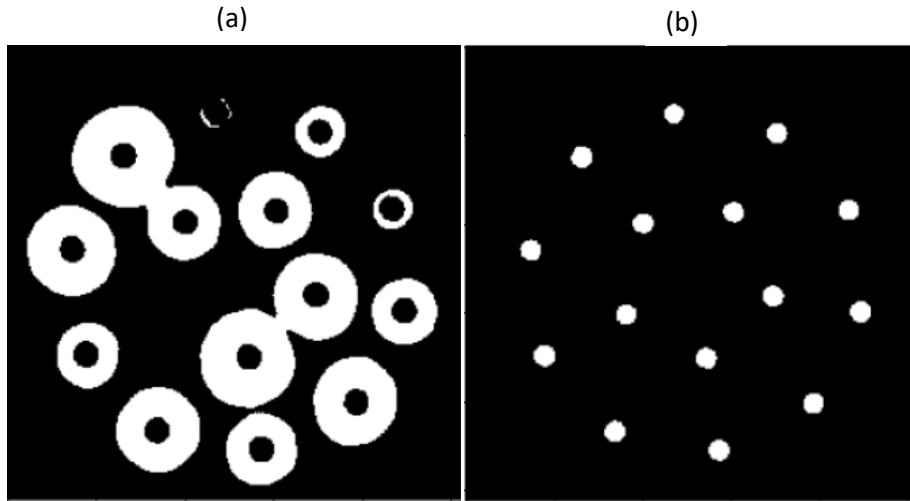
Fonte: (Autoria Própria, 2022).

#### 4.1.2 Resultados do Pós-Processamento da Segmentação

Na fase do pós-processamento, o limiar de ativação é um fator importante para avaliar a confiança do modelo, visto que na medida em que aumenta esse parâmetro, a segmentação das classes dos halos e discos sofrem alteração, essas alterações interferem no resultado do algoritmo baseado em PDI para quantificar a área dos halos, caso a classe dos discos sofra uma dilatação grande, irá impactar no valor original de 7 mm, com um raio de 3,5 mm nos discos. A Figura 18 representa o limiar de ativação de 0.5 do melhor modelo, juntamente com um limiar de ativação em 0.9, representando na Figura 19, tendo dilatação na área dos discos e halos. A diferença das imagens é justamente a dilatação da

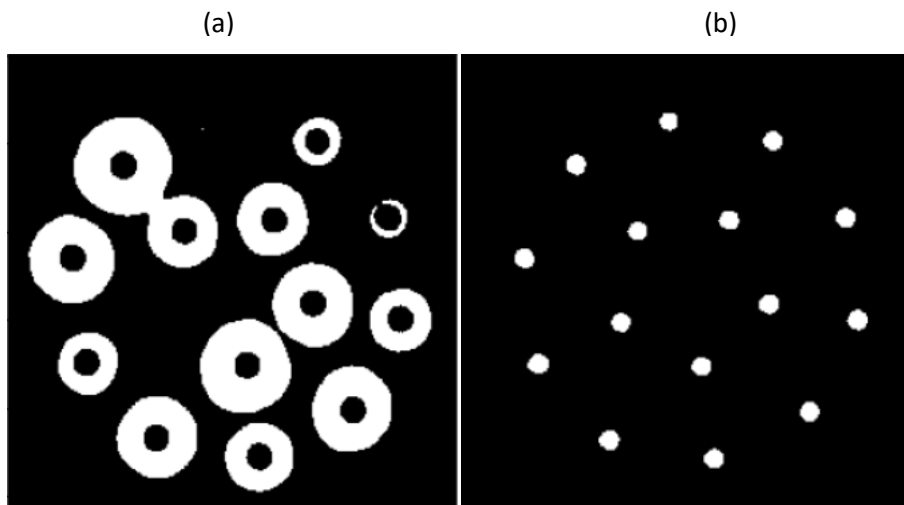
área dos halos e discos em mm, visto que impacta no resultado do algoritmo em PDI para quantificar os halos.

Figura 18 - Limiar de ativação de 0.5 para a classe dos halos (a) e discos (b):



Fonte: (Autoria Própria, 2022).

Figura 19 - Limiar de ativação de 0.9 para a classe dos halos (a) e discos (b):



Fonte: (Autoria Própria, 2022).

#### 4.1.3 Resultados do Pós-Processamento em PDI

A fim de comparar os resultados com a medição dos halos fornecidos pelo IPTSP, os resultados do melhor modelo para segmentação foram apresentados a um pós-processamento baseado em uma rotina PDI para quantificar a área dos halos. A Figura 20 a representa os halos medidos na placa de Petri, já a Figura 20 b é um *zoom* aplicado a medidas em mm.

Estes resultados experimentais são apresentados em termos de EQM, EMA e  $R^2$ , atingindo erros baixos para conjunto de treinamento e para o conjunto de teste. Todos os resultados antes de serem verificados os erros, foram normalizados entre valores de 0 e 1, representados na Tabela 5.

Figura 20 - Halos medidos (a) e *zoom* aplicado na medição (b):

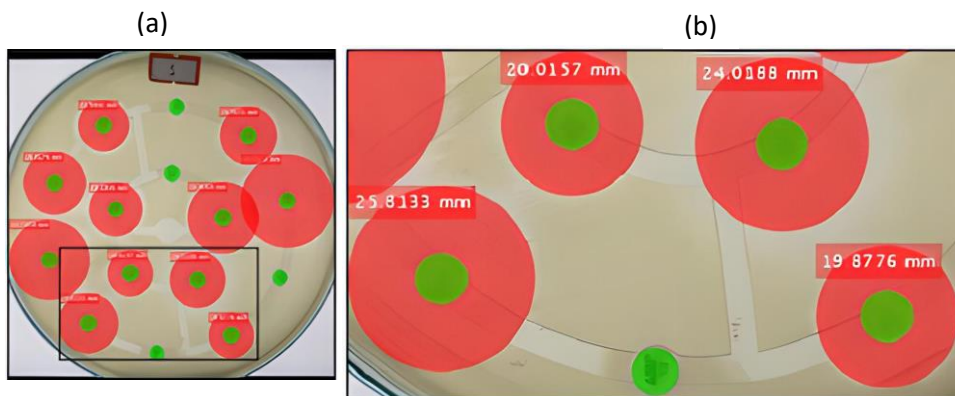


Tabela 5 - Erros de medidas dos halos.

Conjunto de dados					
Treino			Teste		
EQM	EMA	$R^2$	EQM	EMA	$R^2$
0.0816	0.0440	0.9053	0.0895	0.0533	0.9114

Fonte: (Autoria Própria, 2022).

## 5 CONCLUSÃO

Este trabalho apresenta um método para detecção e medição automática de halos em imagens digitais utilizando uma RNC e o método disco-difusão (Kirby-Bauer) para pós-processamento baseado em rotinas de PDI. Este modelo RNC é treinado e testado usando um conjunto de dados de imagem digital com classes segmentadas de halo e disco. O modelo é avaliado por ISU, resultando em altas pontuações do mesmo para conjuntos de treinamento e teste. Em seguida, uma rotina de PDI é aplicada aos resultados do melhor modelo RNC para quantificar a área do disco. Essa etapa é avaliada de acordo com EQM, EMA e  $R^2$ , resultando em umas baixas taxas de erro. Portanto, pode-se concluir que a RNC combinada com PDI é uma estratégia promissora para detectar e medir halos automaticamente.

Existem algumas dificuldades em relação ao modelo RNC, a segmentação erra em algumas classificações dos *pixels*, isso porque possui halos muito grandes que se cruzam

na placa de Petri, resultando em imagens processadas muito complexas. Se os dados forem aumentados para ajustar os parâmetros das redes profundas (ou seja, viés, outras taxas de aprendizado e limiar de ativação etc.), pode melhorar não apenas o modelo RNC proposto, mas também os modelos testados para fornecer uma estimativa mais rápida na detecção e medição automática de halos.

Em colaboração com o desempenho do disco-difusão de maneira ágil, a detecção dos halos pode ser automatizada e medições mais precisas do diâmetro do halo podem ser realizadas, levando a um impacto positivo nas atividades de pesquisa biológica e laboratorial. Para trabalhos futuros, o conjunto de dados das imagens do TSA serão aumentados para fornecer maiores conjuntos de treinamento e teste para o modelo. Diferentes calibrações e configurações de redes neurais profundas possuem potencial para serem exploradas para adaptar o modelo ao problema. O modelo poderá ser executado em um dispositivo móvel e treinado em Reconhecimento Óptico de Caracteres (ROC) para identificar o agente antimicrobiano em cada disco por suas iniciais impressas nele.

## 6 REFERÊNCIAS

- ALBAWI, Saad; MOHAMMED, Tareq Abed; AL-ZAWI, Saad. *Understanding of a convolutional neural network*. 2017 *International Conference On Engineering And Technology (Icet)*, [S.L.], p. 1-6, ago. 2017. *IEEE*. Disponível em: <http://dx.doi.org/10.1109/icengtechnol.2017.8308186>. Acesso em: 20 abr. 2022.
- BAUER, A. W.; KIRBY, W. M. M.; SHERRIS, J. C.; TURCK, M. *Antibiotic Susceptibility Testing by a Standardized Single Disk Method*. *American Journal Of Clinical Pathology*, [S.L.], v. 45, n. 4, p. 493-496, 1 abr. 1966. *Oxford University Press (OUP)*. Disponível em: [http://dx.doi.org/10.1093/ajcp/45.4\\_ts.493](http://dx.doi.org/10.1093/ajcp/45.4_ts.493). Acesso em: 05 fev. 2022.
- BERGER, Daniel R.; SEUNG, H. Sebastian; LICHTMAN, Jeff W. *VAST (Volume Annotation and Segmentation Tool): efficient manual and semi-automatic labeling of large 3d image stacks*. *Frontiers In Neural Circuits*, [S.L.], v. 12, p. 0-0, 16 out. 2018. *Frontiers Media SA*. Disponível em: <http://dx.doi.org/10.3389/fncir.2018.00088>. Acesso em: 11 maio. 2022.
- BHAT, Sevitha; MUTHUNATARAJAN, Shruthi; MULKI, Shalini Shenoy; BHAT, K. Archana; KOTIAN, K. Himani. *Bacterial Infection among Cancer Patients: analysis of isolates and antibiotic sensitivity pattern*. *International Journal Of Microbiology*, [S.L.], v. 2021, p. 1-7, 7 jan. 2021. *Hindawi Limited*. Disponível em: <http://dx.doi.org/10.1155/2021/8883700>. Acesso em: 15 jan. 2022.

BOTH, Anna; BÜTTNER, Henning; HUANG, Jiabin; PERBANDT, Markus; CAMPOS, Cristina Belmar; CHRISTNER, Martin; MAURER, Florian P.; KLUGE, Stefan; KÖNIG, Christina; AEPFELBACHER, Martin; WICHMANN, Dominic; ROHDE, Holger. *Emergence of ceftazidime/avibactam non-susceptibility in an MDR Klebsiella pneumoniae isolate*. *Journal Of Antimicrobial Chemotherapy*, [S.L.], v. 72, n. 9, p. 2483-2488, 16 jun. 2017. *Oxford University Press (OUP)*. Disponível em <http://dx.doi.org/10.1093/jac/dkx179>. Acesso em: 25 jan. 2022.

COSTA, L. F. R. (2016). **Sistema de automatização do antibiograma por disco-difusão em aplicação clínica e ambiental**. Dissertação de Mestrado em Engenharia Biomédica, Publicação 053A/2016, Programa de Pós-Graduação em Engenharia Biomédica, Faculdade Gama, Universidade de Brasília, Brasília, DF, no.p.82. Disponível em: <https://repositorio.unb.br/handle/10482/22126>. Acesso em 25 mar. 2022.

DA COSTA, L. A. (2015). **IDENTIFICAÇÃO AUTOMÁTICA DE CARACTERES EM ANTIBIOGRAMAS COM USO DE MOMENTOS INVARIANTES**. Dissertação de Mestrado em Engenharia Biomédica, Publicação 34A/2015. Programa de Pós-Graduação em Engenharia Biomédica, Faculdade UnB Gama, Universidade de Brasília, Brasília, DF, no.p.9. Disponível em: <https://repositorio.unb.br/handle/10482/19883>. Acesso em: 10 mar. 2022.

DIXIT, Snehil; AHMAD, Irfan; GULAR, Kumar; EID, Refaat A.; REDDY, Ravi Shankar; RIBEIRO, Ivana Leão; ABOHASHRH, Mohammed; ALSHAHRANI, Mastour Saeed; TEDLA, Jaya Shanker; DIXIT, Nitin Arun. *Efficacy of single versus multiple exposure by electromagnetic modalities on gram-negative and positive bacterial strains in an in-vitro model*. *Saudi Journal Of Biological Sciences*, [S.L.], v. 28, n. 3, p. 1678-1686, mar. 2021. *Elsevier BV*. Disponível em: <http://dx.doi.org/10.1016/j.sjbs.2020.12.004> . Acesso em: 25 fev. 2022.

EDITORES, R. P. T. **INAUGURAÇÃO DA NOVA SEDE DO INSTITUTO DE PATOLOGIA TROPICAL E SAÚDE PÚBLICA DA UNIVERSIDADE FEDERAL DE GOIÁS**. *Revista de Patologia Tropical / Journal of Tropical Pathology*, Goiânia, v. 18, n. 2, 2007. DOI: 10.5216/rpt.v18i2.20227. Disponível em: <https://www.revistas.ufg.br/iptsp/article/view/20227>. Acesso em: 10 maio. 2022.

GAO, Mingyu; QI, Dawei; MU, Hongbo; CHEN, Jianfeng. *A Transfer Residual Neural Network Based on ResNet-34 for Detection of Wood Knot Defects*. *Forests*, [S.L.], v. 12, n. 2, p. 212, 11 fev. 2021. MDPI AG. Disponível em: <http://dx.doi.org/10.3390/f12020212>. Acesso em: 12 maio. 2022.

GONZALEZ, Rafael C.; WOODS, Richard E. **PROCESSAMENTO DIGITAL DE IMAGENS**. 3. ed. São Paulo: Pearson Universidades, 2009. (ISNL: 9788576054016). Disponível em: <https://www.travessa.com.br/processamento-digital-de-imagens-3-ed-2009/artigo/26cce117-de8a-4eea-aCTC9-93c8da20e9d0>. Acesso em: 02 maio 2022.

HU, Yujie; ANES, João; DEVINEAU, Stéphanie; FANNING, Séamus. *Klebsiella pneumoniae: prevalence, reservoirs, antimicrobial resistance, pathogenicity, and infection*. *Foodborne Pathogens And Disease*, [S.L.], v. 18, n. 2, p. 63-84, 1 fev. 2021. Mary Ann Liebert Inc. Disponível em: <http://dx.doi.org/10.1089/fpd.2020.2847>. Acesso em: 20 fev. 2022.

JESUS, Claudemir Santos de; PINTO, Izabella dos Santos; REIS, Joyce Leandro Benevenuto dos; FERNANDES, Joyce; SANTOS, Rafael dos; SOARES, Edvânia Santos. **ATUAÇÃO DO ENFERMEIRO DO TRABALHO NA REDUÇÃO DE RISCOS BIOLÓGICOS NO ÂMBITO HOSPITALAR**. Revista Brasileira Ciências da Saúde - Uscs, [S.L.], v. 15, n. 54, p. 100-107, out. 2017. USCS Universidade Municipal de Sao Caetano do Sul. <http://dx.doi.org/10.13037/ras.vol15n54.4825>. Disponível em: [https://seer.uscs.edu.br/index.php/revista\\_ciencias\\_saude/article/view/4825](https://seer.uscs.edu.br/index.php/revista_ciencias_saude/article/view/4825). Acesso em: 05 jan. 2022.

KOONCE, Brett. ResNet 34. **Convolutional Neural Networks With Swift For Tensorflow**, [S.L.], p. 51-61, 2021. Apress. [http://dx.doi.org/10.1007/978-1-4842-6168-2\\_5](http://dx.doi.org/10.1007/978-1-4842-6168-2_5). Acesso em: 15 maio. 2022.

KORFIATIS, Panagiotis; KLINE, Timothy L.; LACHANCE, Daniel H.; PARNEY, Ian F.; BUCKNER, Jan C.; ERICKSON, Bradley J. **Residual Deep Convolutional Neural Network Predicts MGMT Methylation Status**. *Journal Of Digital Imaging*, [S.L.], v. 30, n. 5, p. 622-628, 7 ago. 2017. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10278-017-0009-z>. Acesso em: 17 maio. 2022.

KRIG, Scott. **Computer Vision Metrics**. Ibook, [S.L.], p. 1-652, 2016. Springer International Publishing. <http://dx.doi.org/10.1007/978-3-319-33762-3>. Disponível em: <https://ibook.pub/computer-vision-metrics-textbook-edition.html>. Acesso em: 31 maio 2022.

KROGH, Anders. **What are artificial neural networks?** *Nature Biotechnology*, [S.L.], v. 26, n. 2, p. 195-197, fev. 2008. Springer Science and Business Media LLC. Disponível em: <http://dx.doi.org/10.1038/nbt1386>. Acesso em: 10 abr. 2022.

KUMAR, Vinod; ARORA, Hritik; HARSH; SISODIA, Jatin. **ResNet-based approach for Detection and Classification of Plant Leaf Diseases**. 2020 *International Conference On Electronics And Sustainable Communication Systems (Icesc)*, [S.L.], p. 495-502, jul. 2020. IEEE. Disponível em: <http://dx.doi.org/10.1109/icesc48915.2020.9155585>. Acesso em: 16 maio. 2022.

LACERDA, Lucas. **Deep Learning & Visão Computacional — REDES NEURAIS CONVOLUCIONAIS**. 2019. Disponível em: <https://medium.com/@lucaaslb/deep-learning-vis%C3%A3o-computacional-redes-neurais-convolucionais-c21f19f5ec34>. Acesso em: 04 maio 2022.

LIN, Tsung-Yi; GOYAL, Priya; GIRSHICK, Ross; HE, Kaiming; DOLLAR, Piotr. **Focal Loss for Dense Object Detection**. 2017 *Ieee International Conference On Computer Vision (Iccv)*, [S.L.], p. 2999-3007, out. 2017. IEEE. Disponível em: <http://dx.doi.org/10.1109/iccv.2017.324>. Acesso em: 20 maio. 2022.

MA, Wei; LU, jun. **An Equivalence of Fully Connected Layer and Convolutional Layer**. *Cornell University*, [S.L.], p. 1-9, 2017. ArXiv. Disponível em: <http://dx.doi.org/10.48550/ARXIV.1712.01252>. Acesso em: 05 maio 2022.

MCCULLOCH, Warren S.; PITTS, Walter. **A logical calculus of the ideas immanent in nervous activity**. *The Bulletin Of Mathematical Biophysics*, [S.L.], v. 5, n. 4, p. 115-133, dez. 1943. Springer Science and Business Media LLC. Disponível em: <http://dx.doi.org/10.1007/bf02478259>. Acesso em: 30 mar. 2022.

MIFTAHUSSURUR, Muhammad; FAUZIA, Kartika Afrida; NUSI, Iswan Abbas; SETIAWAN, Poernomo Boedi; SYAM, Ari Fahrial; WASKITO, Langgeng Agung; DOOHAN, Dalla; RATNASARI, Neneng; KHOMSAN, Ali; ADNYANA, I. Ketut; AKADA, Junko; YAMAOKA, Yoshio. ***E-test versus agar dilution for antibiotic susceptibility testing of Helicobacter pylori: a comparison study.*** *Bmc Research Notes*, [S.L.], v. 13, n. 1, p. 1-6, 10 jan. 2020. Springer Science and Business Media LLC. Disponível em: <http://dx.doi.org/10.1186/s13104-019-4877-9>. Acesso em: 10 fev. 2022.

MINAEE, Shervin; KALCHBRENNER, Nal; CAMBRIA, Erik; NIKZAD, Narjes; CHENAGHLU, Meysam; GAO, Jianfeng. ***Deep Learning--based Text Classification.*** *Acm Computing Surveys*, [S.L.], v. 54, n. 3, p. 1-40, 30 abr. 2022. Association for Computing Machinery (ACM). Disponível em: <http://dx.doi.org/10.1145/3439726>. Acesso em: 01 maio. 2022.

MORRIS, C. Paul; BERGMAN, Yehudit; TEKLE, Tsigedera; FISSEL, John A.; TAMMA, Pranita D.; SIMNER, Patricia J. Cefiderocol. ***Antimicrobial Susceptibility Testing against Multidrug-Resistant Gram-Negative Bacilli: a comparison of disk diffusion to broth microdilution.*** *Journal Of Clinical Microbiology*, [S.L.], v. 59, n. 1, p. 1-12, 17 dez. 2020. American Society for Microbiology. Disponível em: <http://dx.doi.org/10.1128/jcm.01649-20>. Acesso em: 15 fev. 2022.

NELSON, Kirk; HEMARAJATA, Peera; SUN, Dongxu; RUBIO-APARICIO, Debora; TSIVKOVSKI, Ruslan; YANG, Shangxin; SEBRA, Robert; KASARSKIS, Andrew; NGUYEN, Hoan; HANSON, Blake M; LEOPOLD, Shana; WEINSTOCK, George; LOMOVSKEYA, Olga; HUMPHRIES, Romney M. ***Resistance to Ceftazidime-Avibactam Is Due to Transposition of KPC in a Porin-Deficient Strain of Klebsiella pneumoniae with Increased Efflux Activity.*** *Antimicrobial Agents And Chemotherapy*, [S.L.], v. 61, n. 10, p. 1-13, out. 2017. American Society for Microbiology. Disponível em: <http://dx.doi.org/10.1128/aac.00989-17>. Acesso em: 30 jan. 2022.

O'SHEA, Keiron; NASH, Ryan. ***An Introduction to Convolutional Neural Networks.*** Cornell University, [S.L.], p. 1-11, 2015. ArXiv. Disponível em: <http://dx.doi.org/10.48550/ARXIV.1511.08458>. Acesso em: 06 maio 2022.

PEDRINI, Hélio; SCHWARTZ, Willian Robson. **Análise De Imagens Digitais: Princípios, Algoritmos E Aplicações.** São Paulo: Cengage Learning, 2007. 528 p. (ISNL-10: 8522105952 - ISNL-13: 978-8522105953). Disponível em: <https://www.amazon.com.br/An%C3%A1lise-Imagens-Digitais-Princ%C3%ADpios-Algoritmos/dp/8522105952>. Acesso em: 07 maio 2022.

PENG, Li; NA, Yi; CHANGSONG, Ding; SHENG, L.I.; HUI, Min. ***Research on classification diagnosis model of psoriasis based on deep residual network.*** *Digital Chinese Medicine*, [S.L.], v. 4, n. 2, p. 92-101, jun. 2021. Elsevier BV. Disponível em: <http://dx.doi.org/10.1016/j.dcm.2021.06.003>. Acesso em: 13 maio. 2022.

PRIYA, B. Keerthi; REDDY, D. Akhila; RANI, A. Daisy; KALAHASTHI, Neelima; SOLIMAN, Wasim Ghder; REDDY, D. V. Rama Koti. ***Automatic Inhibition Zone Diameter Measurement for Disc Diffusion Test Using Image Segmentation.*** *Iete Journal Of Research*, [S.L.], p. 1-18, 7 out. 2021. Informa UK Limited. Disponível em: <http://dx.doi.org/10.1080/03772063.2021.1969293>. Acesso em: 15 mar. 2022.



QIU, Wenting; NAGL, Stefan. **Automated Miniaturized Digital Microfluidic Antimicrobial Susceptibility Test Using a Chip-Integrated Optical Oxygen Sensor**. *Acs Sensors*, [S.L.], v. 6, n. 3, p. 1147-1156, 15 mar. 2021. *American Chemical Society (ACS)*. <http://dx.doi.org/10.1021/acssensors.0c02399>. Acesso em: 05 mar. 2022.

RAMIREZ, Alejandro Rafael Garcia. **Pesquisa e prática docente no curso de Engenharia de Controle e Automação: automação da medição de halos de antibiogramas**. Campus Curitibanos, Santa Catarina: Biblioteca Central da Universidade do Contestado, 2017. 123 p. (9 788563 671523). Edição de Josiane Liebl Miranda CRB 14/1023 - ISNL: 978-85-63671-52-3. Disponível em: [https://uni-contestado-site.s3.amazonaws.com/site/pesquisa/anais+e+ebooks/e-book\\_Engenharia\\_Control\\_Automacao%20\(1\).pdf](https://uni-contestado-site.s3.amazonaws.com/site/pesquisa/anais+e+ebooks/e-book_Engenharia_Control_Automacao%20(1).pdf). Acesso em: 10 jan. 2022.

REZATOFIGHI, Hamid; TSOI, Nathan; GWAK, Junyoung; SADEGHIAN, Amir; REID, Ian; SAVARESE, Silvio. **Generalized Intersection Over Union: a metric and a loss for bounding box regression**. 2019 *Ieee/Cvf Conference On Computer Vision And Pattern Recognition (Cvpr)*, [S.L.], p. 658-666, jun. 2019. *IEEE*. Disponível em: <http://dx.doi.org/10.1109/cvpr.2019.00075>. Acesso em: 21 maio. 2022.

RIZON, Mohamed; YAZID, Haniza; SAAD, Puteh; SHAKAFF, Ali Yeon Md; SAAD, Abdul Rahman; SUGISAKA, Masanori; YAACOB, Sazali; MAMAT, M.Rozailan; KARTHIGAYA, M.. **Object Detection using Circular Hough Transform**. *American Journal Of Applied Sciences*, [S.L.], v. 2, n. 12, p. 1606-1609, 1 dez. 2005. *Science Publications*. Disponível em: <http://dx.doi.org/10.3844/ajassp.2005.1606.1609>. Acesso em: 22 maio. 2022.

ROSEBROCK, Adrian. **Deep Learning for Computer Vision with Python: starter bundle**. *Pyimagesearch*, 2017. 330 p. Disponível em: <https://books.google.com.br/books?id=9Ul-tgEACAAJ>. Acesso em: 05 abr. 2022.

ROSENBLATT, F. **The Perceptron: a probabilistic model for information storage and organization in the brain**. *Psychological Review*, [S.L.], v. 65, n. 6, p. 386-408, 1958. *American Psychological Association (APA)*. <http://dx.doi.org/10.1037/h0042519>. Acesso em: 15 abr. 2022.

SAKURAI, Rafael. **Implementando a estrutura de uma Rede Neural Convolutacional utilizando o MapReduce do Spark**. 2017. Disponível em: <https://www.sakurai.dev.br/RNC-mapreduce/>. Acesso em: 30 abr. 2022.

SICURO, Benedetto; PASTORINO, Paolo; BARBERO, Raffaella; BARISONE, Stefano; DELLERBA, Davide; MENCONI, Vasco; RIGHETTI, Marzia; VITA, Vito de; PREARO, Marino. **Prevalence and antibiotic sensitivity of bacteria isolated from imported ornamental fish in Italy: a translocation of resistant strains?** *Preventive Veterinary Medicine*, [S.L.], v. 175, p. 104880, fev. 2020. *Elsevier BV*. Disponível em: <http://dx.doi.org/10.1016/j.prevetmed.2019.104880>. Acesso em: 20 jan. 2022.

TOMITA, Naofumi; CHEUNG, Yvonne Y.; HASSANPOUR, Saeed. **Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans**. *Computers In Biology And Medicine*, [S.L.], v. 98, p. 8-15, jul. 2018. *Elsevier BV*. Disponível em: <http://dx.doi.org/10.1016/j.compbiomed.2018.05.011>. Acesso em: 14 maio. 2022.

TURING, A. M. I.—**COMPUTING MACHINERY AND INTELLIGENCE**. *Mind*, [S.L.], v., n. 236, p. 433-460, 1 out. 1950. *Oxford University Press (OUP)*. Disponível em: <http://dx.doi.org/10.1093/mind/lix.236.433>. Acesso em: 01 abr. 2022.

UCHIDA, Kazutaka; TANAKA, Masayuki; OKUTOMI, Masatoshi. **Coupled convolution layer for convolutional neural network**. *Neural Networks*, [S.L.], v. 105, p. 197-205, set. 2018. *Elsevier BV*. Disponível em: <http://dx.doi.org/10.1016/j.neunet.2018.05.002>. Acesso em: 01 maio 2022.

WADA, K. **Labelme: image polygonal annotation with python**. *Image Polygonal Annotation with Python*. 2021. Disponível em: <https://github.com/wkentaro/labelme>. Acesso em: 04 out. 2021.

WANG, Panqu; CHEN, Pengfei; YUAN, Ye; LIU, Ding; HUANG, Zehua; HOU, Xiaodi; COTTRELL, Garrison. **Understanding Convolution for Semantic Segmentation**. 2018 *Ieee Winter Conference On Applications Of Computer Vision (Wacv)*, [S.L.], p. 1451-1460, mar. 2018. *IEEE*. Disponível em: <http://dx.doi.org/10.1109/wacv.2018.00163>. Acesso em: 08 maio 2022.

VEZHNEVETS, Alexander; FERRARI, Vittorio; BUHMANN, Joachim M. **Weakly supervised semantic segmentation with a multi-image model**. 2011 *International Conference On Computer Vision*, [S.L.], p. 643-650, nov. 2011. *IEEE*. Disponível em: <http://dx.doi.org/10.1109/iccv.2011.6126299>. Acesso em: 01 fev. 2022.

YAMASHITA, Rikiya; NISHIO, Mizuho; DO, Richard Kinh Gian; TOGASHI, Kaori. **Convolutional neural networks: an overview and application in radiology**. *Insights Into Imaging*, [S.L.], v. 9, n. 4, p. 611-629, 22 jun. 2018. *Springer Science and Business Media LLC*. Disponível em: <http://dx.doi.org/10.1007/s13244-018-0639-9>. Acesso em: 25 abr. 2022.

YU, Dingjun; WANG, Hanli; CHEN, Peiqiu; WEI, Zhihua. **Mixed Pooling for Convolutional Neural Networks**. *Rough Sets And Knowledge Technology*, [S.L.], p. 364-375, 2014. *Springer International Publishing*. Disponível em: [http://dx.doi.org/10.1007/978-3-319-11740-9\\_34](http://dx.doi.org/10.1007/978-3-319-11740-9_34). Acesso em: 03 maio 2022.

ZHANG, Mengmeng; LI, Wei; DU, Qian. **Diverse Region-Based RNC for Hyperspectral Image Classification**. *Ieee Transactions On Image Processing*, [S.L.], v. 27, n. 6, p. 2623-2634, jun. 2018. *Institute of Electrical and Electronics Engineers (IEEE)*. Disponível em: <http://dx.doi.org/10.1109/tip.2018.2809606>. Acesso em: 20 mar. 2022.

ZHANG, Xiaoqing; ZHAO, Shu-Guang. **Cervical image classification based on image segmentation preprocessing and a CapsNet network model**. *International Journal Of Imaging Systems And Technology*, [S.L.], v. 29, n. 1, p. 19-28, 14 set. 2018. *Wiley*. Disponível em: <http://dx.doi.org/10.1002/ima.22291>. Acesso em: 10 maio 2022.

ZHOU, Zongwei; SIDDIQUEE, Md Mahfuzur Rahman; TAJBAKSH, Nima; LIANG, Jianming. **UNet++: a nested u-net architecture for medical image segmentation**. *Cornell University*, [S.L.], p. 1-8, 2018. *ArXiv*. Disponível em: <http://dx.doi.org/10.48550/ARXIV.1807.10165>. Acesso em: 19 maio. 2022.