

UNIVERSIDADE FEDERAL DO PARANÁ

GEISLA DE ALBUQUERQUE MELO LASKOSKI

MICROAER_UFPR: BASE DE IMAGENS DE MICROCYSTIS AERUGINOSA EM
ÁGUA BRUTA PARA CLASSIFICAÇÃO/DETECÇÃO DE CIANOBACTÉRIAS

CURITIBA

2021

GEISLA DE ALBUQUERQUE MELO LASKOSKI

MICROAER_UFPR: BASE DE IMAGENS DE MICROCYSTIS AERUGINOSA EM
ÁGUA BRUTA PARA CLASSIFICAÇÃO/DETECÇÃO DE CIANOBACTÉRIAS

Tese apresentada ao curso de Pós-Graduação em
Informática, Setor de Ciências Exatas,,
Universidade Federal do Paraná, como requisito
parcial à obtenção do título de Doutora em
Informática.

Orientador: Prof. Dr. Lucas Ferrari de Oliveira

Coorientador: Prof. Dr. Antonio Carlos Sobieranski

CURITIBA

2021

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
UNIVERSIDADE FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA CIÊNCIA E TECNOLOGIA

Laskosk, Geisla de Albuquerque Melo

Microaer_UFPR: base de imagens de *microcystis aeruginosa* em água bruta para classificação/detecção de cianobactérias. / Geisla de Albuquerque Melo Laskoski. – Curitiba, 2022.

1 recurso on-line : PDF.

Tese (Doutorado) – Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Informática

Orientador: Prof. Dr. Lucas Ferrari de Oliveira

Coorientador: Prof. Dr. Antonio Carlos Sobieranski

1. Processamento de Imagens. 2. Cianobactérias (*Microcystis aeruginosa*). I. Oliveira, Lucas Ferrari de. II. Sobieranski, Antonio Carlos. III. Universidade Federal do Paraná. Programa de Pós-Graduação em Informática. IV. Título.

Bibliotecária: Roseny Rivelini Morciani CRB-9/1585



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO INFORMÁTICA -
40001016034P5

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **GEISLA DE ALBUQUERQUE MELO LASKOSKI** intitulada: **MICROAER_UFPR: BASE DE IMAGENS DE MICROCYSTIS AERUGINOSA EM ÁGUA BRUTA PARA CLASSIFICAÇÃO/DETECÇÃO DE CIANOBACTÉRIAS**, sob orientação do Prof. Dr. LUCAS FERRARI DE OLIVEIRA, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de doutora está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 13 de Dezembro de 2021.

Assinatura Eletrônica

14/12/2021 09:28:03.0

LUCAS FERRARI DE OLIVEIRA
Presidente da Banca Examinadora

Assinatura Eletrônica

14/12/2021 08:34:17.0

GISELLE LOPES FERRARI RONQUE
Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

14/12/2021 13:34:04.0

EDUARDO TODT
Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

14/12/2021 11:15:00.0

RAZER ANTHOM NIZER ROJAS MONTAÑO
Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

Dedico este trabalho ao meu amado esposo. Seu amor e apoio incondicional tornam minha vida mais leve, e fazem tudo valer à pena.

AGRADECIMENTOS

Meu principal agradecimento é ao meu Deus, todo-poderoso. Não existem palavras ou sentimentos capazes de expressar minha gratidão por Seu amor e cuidado infinitos, em todos os aspectos de minha vida.

Agradeço ao meu amado esposo Ricardo, pois sem seu conforto e apoio nos momentos mais difíceis, talvez eu não terminasse esta caminhada. Ele foi parte essencial em mais esta etapa de minha vida, e me sinto imensamente privilegiada por tê-lo como meu melhor amigo e esposo.

Agradeço também à minha amada mãe, por todas as suas orações, e por se importar e cuidar tanto de mim.

Agradeço ao meu orientador Lucas, que topou embarcar no meio desta jornada comigo, me apoiou, realmente me orientou, me ajudou a tomar as melhores decisões e a escolher os melhores caminhos a seguir ao longo da pesquisa. Obrigada por ter sido essencial. Meus mais sinceros agradecimentos.

Agradeço ao meu coorientador, Antonio, por me apoiar, me entender e me animar desde o início desta longa caminhada, que passou por três trocas de tema, uma troca de orientação e uma troca de foco de tema. Obrigada por estar sempre lá.

Agradeço à professora Thelma por me receber, por me apoiar, por compartilhar sua sala e seu conhecimento comigo, numa das etapas mais importantes e decisivas da pesquisa.

Agradeço ao Maurício que me apresentou a demanda das cianobactérias na empresa de saneamento, dando propósito à pesquisa. Muito obrigada pela solicitude.

Agradeço à Sônia, também sempre muito solícita para tirar minhas dúvidas e orientar como especialista no assunto.

Agradeço a UFPR por me dar a oportunidade de cursar algo que eu sempre almejei, e à CAPES pelo apoio financeiro ao longo da pesquisa, que também foi essencial para minha subsistência.

Meu coração transborda em alegria e gratidão. Apenas obrigada!

“Pesquisar é acordar para o mundo.”

(Marcelo Lamy)

RESUMO

As cianobactérias representam um problema real que pode ocorrer em reservatórios e mananciais de água utilizados para suprir a população. Sua ocorrência se deve principalmente à eutrofização dos ambientes aquáticos, isto é, o enriquecimento de nutrientes, que por sua vez, se deve especialmente a atividades como despejo de dejetos nos mananciais e próximo a lençóis. Algumas espécies de cianobactérias são potencialmente tóxicas, e podem causar danos ao meio ambiente e ao ser humano, ingerindo a água ou apenas por contato com a pele. Técnicas de *deep learning* (aprendizado profundo) vêm ganhando destaque em aplicações envolvendo visão computacional. Tais técnicas são consideradas o estado da arte no processamento de imagens. A literatura mostra estudos baseados em *deep learning* sendo desenvolvidos ao longo dos anos, objetivando encontrar soluções para problemas relacionados às cianobactérias, como reconhecimento de espécies e contagem de células e/ou filamentos. O uso de aprendizado profundo geralmente demanda grande quantidade de imagens para alcançar bons resultados. Assim, a necessidade de uma base de dados de espécies de cianobactérias para o desenvolvimento de pesquisas torna-se um fator consideravelmente importante, pois proporciona aos pesquisadores a concepção de trabalhos voltados para os problemas incidentes ao tema, sem se preocupar com a criação da base, focando no aprimoramento dos resultados. O monitoramento de florações de cianobactérias é obrigatório por lei. O processo de detecção e contagem é moroso, pode causar confusão na identificação entre espécies e toma horas para analisar uma única amostra. Logo, o objetivo deste trabalho é criar e disponibilizar uma base de imagens de lâminas de microscópio padronizada e rotulada de florações da cianobactéria *Microcystis aeruginosa*, e uma base de imagens de fitoplâncton, isto é, imagens dos organismos aquáticos, para que pesquisadores e interessados possam ter acesso e desenvolver estudos. A base *Microaer_ufpr* é composta por 2.265 imagens de floração de *Microcystis aeruginosa*, rotuladas célula a célula, e 1.657 imagens de organismos diversos no fitoplâncton, sendo um total de 3.922 imagens. O trabalho também aplica processos de classificação de imagens bem como identificação e contagem de células, apresentando as potencialidades da base criada. A *Microaer_ufpr* foi treinada nos modelos de redes neurais convolucionais de aprendizado profundo *Mobilenet*, *Inception* e *Faster-RCNN*, além da *YOLO*. Na tarefa de classificação de imagens, onde se objetivou distinguir imagens com e sem floração, os resultados foram excelentes, alcançando acurácia acima de 99% em todos os testes, tanto para *Mobilenet* quanto para *Inception*. Já nos testes de detecção de objetos, os resultados foram igualmente promissores, chegando a uma precisão média de 75,75% com a rede *Faster-RCNN* com *threshold* de 30% (células são consideradas positivas com probabilidade superior a 30%) e de 66,15% com *threshold* de 50%. A rede *Inception* trouxe uma precisão média de 69,56% (*threshold* de 30%). Os resultados denotam a potencialidade de utilização da base apresentada na utilização para processamento de imagens, e dão direcionamento para o ajuste fino das redes citadas, para as diversas aplicações que pesquisadores possam desenvolver. A base está disponibilizada gratuitamente, objetivando ser um facilitador no caminho para resolver problemas relacionados.

Palavras-chave: *Microcystis aeruginosa*. Processamento de Imagens. Deep Learning. Cianobactérias. Base de Imagens.

ABSTRACT

Cyanobacteria represent a real problem that can occur in reservoirs and water sources used to supply the population. Its occurrence is mainly due to the eutrophication of aquatic environments, that is, the enrichment of nutrients, which, in turn, is mainly due to activities such as dumping of waste into water sources and close to groundwater. Some species of cyanobacteria are potentially toxic and can cause damage to the environment and to humans, ingesting the water or just by contact with the skin. Deep learning techniques have been gaining prominence in applications involving computer vision. Such techniques are considered the state of the art in image processing. Literature shows studies based on deep learning being developed over the years, aiming to find solutions to problems related to cyanobacteria, such as species recognition and cell and/or filament count. The use of deep learning often requires a large amount of images to achieve good results. Thus, the need for a database of cyanobacteria species for the development of research becomes a considerably important factor, as it provides researchers to conceive works aimed at the problems related to the theme, without worrying about the creation of the basis, focusing on improving results. Monitoring cyanobacterial blooms is mandatory by law. The detection and counting process is time-consuming, can cause confusion in species identification, and takes hours to analyze a single sample. Therefore, the objective of this work is to create and make available a standardized and labeled microscope slide image base of *Microcystis aeruginosa* cyanobacterial blooms, and a phytoplankton image base, that is, images of aquatic organisms, so that researchers and interested parties can have access and develop studies. The database *Microaer_ufr* is composed of 2.265 images of *Microcystis aeruginosa* blooming, labeled cell by cell, and 1.657 images of diverse organisms in phytoplankton, for a total of 3.922 images. The work also applies image classification processes as well as cell identification and counting, presenting the potential of the created base. The *Microaer_ufr* was trained in the deep learning convolutional neural network models Mobilenet, Inception and Faster-RCNN, in addition to YOLO. In the image classification task, which aimed to distinguish images with and without blooms, the results were excellent, reaching an accuracy above 99% in all tests, both for Mobilenet and for Inception. In the object detection tests, the results were equally promising, reaching an average accuracy of 75.75% with the Faster-RCNN network with a threshold of 30% (cells are considered positive with a probability greater than 30%) and 66.15% with a threshold of 50%. The Inception network brought an average accuracy of 69.56% (30% threshold). The results denote the potential use of the base presented in the use for image processing, and provide guidance for the fine-tuning of the aforementioned networks, for the various applications that researchers can develop. The base is available free of charge, aiming to be a facilitator on the way to solving related problems.

Keywords: *Microcystis aeruginosa*. Image processing. Deep learning. Cyanobacteria.

Image database.

LISTA DE FIGURAS

FIGURA 1 – Câmara de <i>utermohl</i>	20
FIGURA 2 – Floração no Reservatório de Salto Grande – SP	26
FIGURA 3 – <i>Microcystis aeruginosa</i>	30
FIGURA 4 – <i>Microcystis wesenbergii</i>	31
FIGURA 5 – Neurônio biológico simplificado.....	35
FIGURA 6 – Neurônio artificial	35
FIGURA 7 – Exemplo de RNA multicamadas típicas	37
FIGURA 8 – Modelo de CNN	40
FIGURA 9 – YOLO	45
FIGURA 10 – Arquitetura da YOLO	46
FIGURA 11 – Fórmula <i>Intersection Over Union</i>	47
FIGURA 12 – Vidro de cor âmbar com amostra de água bruta manipulada e identificada	61
FIGURA 13 – Preparação de lâminas de microscópio	62
FIGURA 14 – a) Classe <i>Microcystis aeruginosa</i> b) Classe Outros.....	63
FIGURA 15 – Rotulação de Colônia de <i>Microcystis</i> 1	64
FIGURA 16 – Rotulação de Colônia de <i>Microcystis</i> 2	65
FIGURA 17 – Processo de Rotulação.....	66
FIGURA 18 – Curva de aprendizado de <i>Inception_v3</i>	69
FIGURA 19 – Curva de aprendizado de <i>Mobilenet</i>	70
FIGURA 20 – Variedade de <i>Microcystis aeruginosa</i> 1	73
FIGURA 21 – Variedade de <i>Microcystis aeruginosa</i> 2	74
FIGURA 22 – Variedade de <i>Microcystis aeruginosa</i> 3	74
FIGURA 23 – Variedade de <i>Microcystis aeruginosa</i> 4	75
FIGURA 24 – Variedade de <i>Microcystis aeruginosa</i> 5	75
FIGURA 25 – Variedade de <i>Microcystis aeruginosa</i> 6	76
FIGURA 26 – Variedade de Fitoplâncton 1	76
FIGURA 27 – Variedade de Fitoplâncton 2	77
FIGURA 28 – Variedade de Fitoplâncton 3	77
FIGURA 29 – Variedade de Fitoplâncton 4	77
FIGURA 30 – Variedade de Fitoplâncton 5	78
FIGURA 31 – Exemplo de Imagem Rotulada 1	79

FIGURA 32 – Exemplo de Imagem Rotulada 2.....	80
FIGURA 33 – Exemplo de Imagem Rotulada 3.....	81
FIGURA 34 – Exemplo de Imagem Rotulada 4.....	81
FIGURA 35 – Exemplo de Imagem Rotulada 5.....	82
FIGURA 36 – Arquivo XML	83
FIGURA 37 – Imagem em a) Original, em b) Espelhamento Horizontal e em c) Espelhamento Vertical	83
FIGURA 38 – Detecção com <i>Mobilenet</i> em a), com <i>Inception</i> em b), com <i>Faster-RCNN</i> em c) e com YOLO em d).....	89
FIGURA 39 – <i>Ground-truth boxes x Predicted Boxes</i>	90
FIGURA 40 – Curva PR <i>Faster-RCNN</i> 3000 épocas (IoU 30%)	94
FIGURA 41 – Curva PR <i>Inception</i> 3000 épocas (IoU 30%)	95
FIGURA 42 – Curva PR <i>Faster-RCNN</i> 3000 épocas (IoU 50%)	95

LISTA DE QUADROS

QUADRO 1 – EVOLUÇÃO DO CONSUMO PER CAPITA DE ÁGUA PELO HOMEM	16
QUADRO 2 – CIANOTOXINAS, EFEITOS EM RATOS, NÚMERO DE VARIAÇÕES ESTRUTURAIS, TOXICODINÂMICAS E OBSERVAÇÕES CLÍNICAS	28

LISTA DE TABELAS

TABELA 1 – Matriz de Confusão.....	48
TABELA 2 – Acurácias dos Modelos de Classificação.....	84
TABELA 3 – Matriz de Confusão dos Testes com Modelo Simples	85
TABELA 4 – Matriz de Confusão dos Testes com Validação Cruzada	85
TABELA 5 – Acurácia nos Testes Simples	86
TABELA 6 – Acurácia nos Testes com Validação Cruzada	86
TABELA 7 – Perda dos Modelos.....	91
TABELA 8 – Resultados 500 épocas	92
TABELA 9 – Resultados 3000 épocas	92
TABELA 10 – Resultados YOLOv3	93

LISTA DE ABREVIATURAS OU SIGLAS

AM	- Aprendizado de Máquina
AP	- <i>Average Precision</i>
API	- <i>Application Programming Interface</i>
Casan	- Companhia Catarinense de Águas e Saneamento
CETESB	- Companhia Ambiental do Estado de São Paulo
CNN	- <i>Convolutional Neural Network</i>
Copasa	- Companhia de Saneamento e Minas Gerais
CPU	- <i>Central Processing Unit</i>
EFDC	- <i>Environmental Fluid Dynamics Code</i>
ELA	- Esclerose Lateral Amiotrófica
FN	- <i>False Negative</i>
FP	- <i>False Positive</i>
GPU	- <i>Graphics Processing Unit</i>
IA	- Inteligência Artificial
IoU	- <i>Intersection Over Union</i>
MLP	- <i>Multilayer Perceptron</i>
NSE	- <i>Nash-Sutcliffe Efficiency</i>
PCA	- <i>Principal Component Analysis</i>
RCNN	- <i>Region-based Convolutional Neural Network</i>
RMSE	- <i>Root Mean Squared Error</i>
RNA	- Rede Neural Artificial
TN	- <i>True Negative</i>
USAV	- Unidade de Serviços de Avaliações de Conformidade
VP	- <i>True Positive</i>
YOLO	- <i>You Only Look Once</i>

SUMÁRIO

1 INTRODUÇÃO	16
1.1 JUSTIFICATIVA	19
1.2 OBJETIVOS	21
1.2.1 Objetivo geral	21
1.2.2 Objetivos específicos.....	21
1.3 METODOLOGIA.....	22
1.4 ESTRUTURA DA TESE	22
2 FUNDAMENTAÇÃO TEÓRICA	24
2.1 O PROBLEMA DA EUTROFIZAÇÃO.....	24
2.2 AS CIANOBACTÉRIAS	25
2.3 CASOS RELATADOS	32
2.4 INTELIGÊNCIA ARTIFICIAL E REDES NEURAIS.....	33
2.4.1 REDES NEURAIS CONVOLUCIONAIS.....	39
2.5 DETECÇÃO DE OBJETOS.....	47
3 REVISÃO DE LITERATURA	51
3.1 CRIAÇÃO E DISPONIBILIZAÇÃO DE BASES DE DADOS.....	51
3.2 PROCESSAMENTO DE IMAGENS APLICADO À CIANOBACTÉRIAS	56
4 MATERIAIS E MÉTODOS	60
4.1 FORMULAÇÃO DA BASE.....	60
4.1.1 Aquisição das Imagens	60
4.1.2 Classificação e Rotulação da Base	63
4.1.3 Processo de <i>Data Augmentation</i>	66
4.2 APLICAÇÃO DE DEEP LEARNING	67
4.2.1 O processo de Classificação	67
4.2.2 O Processo de Detecção de Objeto e Contagem de Células.....	70
5 RESULTADOS	73
5.1 RESULTADOS DA MONTAGEM DA BASE.....	73
5.1.1 Imagens Originais	73
5.1.2 Imagens Rotuladas	78
5.2 RESULTADOS DA APLICAÇÃO DE DEEP LEARNING.....	84
5.2.1 Classificação de Imagens.....	84
5.2.2 Identificação de Objeto e Contagem	88

5.2.3 Disponibilização da base.....	97
6 CONSIDERAÇÕES FINAIS	98
6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS	99
REFERÊNCIAS.....	101
APÊNDICE 1 – GRÁFICOS CURVA PR	107

1 INTRODUÇÃO

A água é um constituinte fundamental, que influencia de maneira irrefutável todas as formas de vida em todos os ecossistemas. A água possibilita a integração entre os ecossistemas natural, urbano e rural, atua como regulador de temperatura climática, permite a fotossíntese de organismos, age como solvente e matéria prima na indústria, entre muitas outras atuações e benefícios para a humanidade. Neste cenário, qualquer alteração na qualidade e/ou quantidade deste recurso, que seja capaz de desequilibrar o ambiente, tem impacto direto na disponibilidade hídrica, produtividade e saúde dos ecossistemas (ANDREOLI, 2003).

A questão da qualidade da água se caracteriza como uma situação no âmbito de saúde pública mundial, pois está diretamente ligada à qualidade de vida diária da população. A humanidade consome hoje mais água do que todos os outros bens e matérias-primas combinados. Ao longo do desenvolvimento das civilizações, o consumo de água per capita aumentou significativamente, como se pode observar no QUADRO 1.

QUADRO 1 – EVOLUÇÃO DO CONSUMO PER CAPITA DE ÁGUA PELO HOMEM

Homem	Volume (litros/dia)
100 a.C.	12
Império Romano	20
Século XIX	
Pequenas Cidades	40
Grandes Cidades	60

FONTE: MACÊDO *apud* ANDREOLI (2003).

A água é um recurso natural renovável, sendo que seu volume permanece constante, independentemente de seu tratamento inadequado ou má utilização. A água flui entre os diversos compartimentos ambientais da Terra, passando por processos de transferência entre a atmosfera, lençóis subterrâneos, rios, oceanos, geleiras e novamente para a atmosfera. Através deste ciclo hidrológico a água é purificada e renovada constantemente. A questão da disponibilidade da água é que vem gerando problemas. O fato é que a água disponível se concentra basicamente na parcela que flui pelos rios e lagos e lençóis próximos à superfície. O restante da

água apresenta disponibilidade restrita, devido à profundidade superior à 800m, ou por estarem em oceanos e geleiras, sendo de exploração a custos elevados (ANDREOLI e CARNEIRO, 2005). Estes fatores levam ao que é considerado escassez.

Segundo Andreoli (2003), os principais problemas de abastecimento no Brasil não são devido à escassez de água, pois o país não apresenta tal situação, pelo volume de água, mas sim devido a dois principais fatores: a redução do potencial hídrico (pelo desmatamento, degradação estrutural do solo e aumento da evapotranspiração da água) e a perda da qualidade da água (pela acumulação de sedimentos agrícolas e esgoto nos rios e mananciais).

Deste modo, faz-se necessário o constante monitoramento e tratamento de água nos reservatórios utilizados para o abastecimento da população, no intuito de garantir a disponibilidade de água de qualidade para o consumo. O frequente monitoramento é exigido por lei, segundo a Portaria N° 2.914, de 12 de dezembro de 2011 (BRASIL, 2011), e abrange a análise de uma variedade de fatores relacionados à qualidade da água, incluindo a presença e a medição de quantidade de cianobactérias.

As cianobactérias são microrganismos que surgem nos corpos d'água devido a diferentes fatores, incluindo o enriquecimento de nutrientes presentes na água, por ação natural ou humana (como descargas de esgotos ou erosão por exemplo). Desta forma, sua presença em reservatórios e mananciais é relativamente comum. As cianobactérias são capazes de alterar a cor, o odor e o sabor da água, e algumas espécies (pelo menos sete) possuem a capacidade de produzir toxinas potencialmente nocivas tanto para seres humanos quanto para animais que passam por algum tipo de contato com a água contaminada. Dezenas de pessoas já morreram devido a exposição às cianotoxinas, incluindo uma toxina produzida pela espécie de cianobactéria *Microcystis aeruginosa* (JOCHIMSEN *et al*, 1998). Por isto, o monitoramento visando detectar e quantificar a presença de cianobactérias é obrigatório por lei.

O monitoramento das cianobactérias é feito de acordo com as normas dos órgãos reguladores, de forma manual. As fases incluem coletar água dos corpos d'água, levá-las ao laboratório para gerar amostras menores, colocá-las em câmaras cilíndricas de 10ml, e então analisar uma parte desta amostra no microscópio. Quando uma cianobactéria é encontrada, é necessário identificar seu gênero e

efetuar a contagem de suas células manualmente, e os valores podem chegar facilmente a centenas de células por ml em uma floração. O valor encontrado pelo profissional é multiplicado por três, pois considera que a câmara tem três camadas de profundidade.

A Inteligência Artificial (IA) vem sendo utilizada ao longo dos anos na resolução de problemas reais, incluindo o processamento de imagens para identificação de objetos de interesse. As técnicas de inteligência artificial visam facilitar o trabalho humano, geralmente com algum tipo de automação. Dentro da IA, existe uma área chamada *Deep Learning*, também conhecida por sua tradução do inglês, aprendizado profundo. A *Deep Learning* possui uma arquitetura conhecida como Redes Neurais Convolucionais, que é empregada também no reconhecimento de imagens, e é atualmente considerada como o estado da arte na área (LECUN *et al*, 2015).

Quando se fala em processamento digital de imagens e redes neurais convolucionais, um fato conhecido entre pesquisadores é a necessidade de grandes quantidades de imagens, independente do problema a ser resolvido.

Uma das bases mais conhecidas e utilizadas foi apresentada por Deng *et al* (2009), batizada de *ImageNet*. No momento de escrita desta tese, a base já contava com mais de 14 milhões de imagens anotadas, isto é, pertencentes à algum tipo de categoria. Segundo os autores, a base criada é inspirada pelo sentimento de crescente pesquisas no campo de imagem e visão computacional. A ferramenta de aprendizado de máquina *TensorFlow*¹ faz uso de bases para seus modelos de treinamento, e a *ImageNet* é uma delas.

Neste cenário, a proposta deste trabalho está direcionada ao desenvolvimento de uma base de dados intitulada *Microaer_ufpr*, de imagens de cianobactérias da espécie *Microcystis aeruginosa*, proveniente de lâminas de microscópio preparadas por meio de amostras de água bruta de reservatórios com floração da referida cianobactéria. Ressalta-se que não existe nenhuma base pública rotulada disponível até o momento do desenvolvimento tese para a execução do trabalho. Assim, objetiva-se suprir esta lacuna disponibilizando a base em sua forma original, bem como com suas células rotuladas, de modo que estudos envolvendo o processamento das imagens sejam facilitados e encorajados. A

¹ Tensorflow. <https://www.tensorflow.org/>

proposta envolve também a utilização de métodos computacionais para a classificação e contagem das células de *Microcystis aeruginosa*, para exemplificação do potencial de uso da base.

1.1 JUSTIFICATIVA

A pesquisa torna-se justificável pelo fato do impacto ambiental e de saúde pública que permeia o contexto das cianobactérias. Atualmente, de acordo com os procedimentos sanitários, as análises da água são feitas do seguinte modo: são coletadas as amostras de água nos pontos de captação dos mananciais e reservatórios; a água é colocada em câmaras cilíndricas chamadas de *utermöhl* com capacidade de 10ml (FIGURA 1); a câmara é colocada para análise no microscópio visando identificar a presença de organismos do fitoplâncton, incluindo as cianobactérias; o profissional escolhe um transecto, isto é, uma linha imaginária em formato de cruz, e então faz a leitura da amostra nesta cruz. Quando é encontrado um gênero de cianobactéria, o profissional precisa fazer a identificação e a contagem das células desta. Sabe-se que é possível passar facilmente da escala de 1.000 células na leitura considerando uma floração, onde as pequenas células, que no caso da cianobactéria *Microcystis aeruginosa* medem de 4-6 μ m de diâmetro ficam adjacentes, umas sobre as outras. Esta contagem é feita manualmente, o que em se tratando de seres humanos, pode causar resultados imprecisos.

FIGURA 1 – Câmara de *utermohl*

FONTE: LIMNOTECH, 2018.

Assim, percebe-se dois fatos importantes: primeiro, quanto à identificação nas normas legais seguidas, só é necessário identificar a cianobactéria a nível de gênero. Além disso, a identificação a nível de espécie é considerada uma tarefa de alto grau de dificuldade entre os profissionais da área, devido à similaridade de espécies de um mesmo gênero, e mesmo de gêneros diferentes; segundo, quanto à contagem, é um processo fisicamente fadigoso e custoso para os olhos e a visão de modo geral do profissional responsável, podendo esta tarefa causar desconfortos aos olhos, e até mesmo complicações na acuidade visual pela exposição prolongada (horas para a contagem, e anos realizando a tarefa). Pode-se citar ainda que só é lida uma pequena parcela da amostra (a “cruz”), o que pode não representar com fidelidade o que contém na amostra, para o caso de haver organismos importantes fora do limite estabelecido.

Outrossim, durante a pesquisa, em contato com profissionais diretamente envolvidos no assunto e visita à outras cidades, também foi constatada esta necessidade nas empresas de saneamento dos estados de Santa Catarina (Casan – Companhia Catarinense de Águas e Saneamento) e de Minas Gerais (Copasa – Companhia de Saneamento de Minas Gerais). Portanto, entende-se que pesquisar sobre cianobactérias é uma questão de preocupação a nível nacional, no que diz respeito à qualidade de água consumida pela população.

1.2 OBJETIVOS

Com o estudo em questão, pretende-se alcançar alguns objetivos, destacados a seguir como gerais e específicos.

1.2.1 Objetivo geral

Como objetivo geral, busca-se criar e disponibilizar para a comunidade acadêmica e científica, uma base de dados rotulada da cianobactéria *Microcystis aeruginosa*. Para isso imagens de água bruta (sem tratamento) foram adquiridas utilizando microscópio, além disso um especialista auxiliou na rotulação das imagens e nas marcações das cianobactérias. Assim, foram testados métodos de processamento de imagens, a fim de obter resultados iniciais de identificação dos organismos na base, e para demonstrar as potencialidades dela.

Diante da escassez de bases de imagens na área, sua necessidade e importância dentro da visão computacional, a principal ideia é oferecer um conjunto de imagens que possam ser úteis para quaisquer interessados em desenvolver pesquisa sobre a cianobactéria utilizando técnicas de processamento de imagens.

1.2.2 Objetivos específicos

Como objetivos específicos têm-se:

- Criar e disponibilizar uma base de imagens de *Microcystis aeruginosa* com suas imagens originais, isto é, não rotulada;
- Rotular as células da cianobactéria, e disponibilizar a base rotulada;
- Criar e disponibilizar uma base de imagens de organismos do fitoplâncton não rotulada;
- Aplicar técnicas de *Deep Learning* para a classificação das imagens que contenham a cianobactéria *Microcystis aeruginosa* e testar técnicas de *Deep Learning* para a identificação e contagem das células de *Microcystis aeruginosa*.

1.3 METODOLOGIA

Nesta seção, serão brevemente descritos os passos seguidos para se alcançar o objetivo proposto, sendo eles explicados detalhadamente no capítulo 4, Materiais e Métodos:

- Realizou-se a coleta de material, isto é, de água de mananciais/reservatórios da região de Curitiba;
- Foram preparadas as amostras menores, e então colocadas em lâminas para a leitura no microscópio;
- Adquiriu-se as imagens da cianobactéria, varrendo toda a extensão de todas as lâminas, e fotografando-as no microscópio;
- As imagens foram separadas entre as que continham a cianobactéria *Microcystis aeruginosa* e as que não continham, isto é, que representavam os demais organismos do fitoplâncton;
- As imagens que continham a cianobactéria tiveram todas as suas células rotuladas individualmente;
- Foram utilizadas técnicas de *data augmentation*, para aumentar a quantidade de imagens;
- Utilizou-se conceitos e ferramentas de aprendizado profundo, mais especificamente a ferramenta *TensorFlow* para a classificação das imagens;
- Utilizou-se a técnica de *Object Detection* para detectar as células nas imagens e então contá-las;
- Por fim, a base de imagens de *Microcystis aeruginosa* original bem como a base anotada, mais a base contendo os demais organismos do fitoplâncton, foram disponibilizadas gratuitamente para a comunidade acadêmica, em arquivo na nuvem.

1.4 ESTRUTURA DA TESE

No primeiro capítulo, é apresentada uma breve introdução sobre os assuntos tratados na tese, a hipótese da pesquisa, bem como os objetivos e a motivação para

a realização do estudo. São apresentados breves conceitos sobre as técnicas e estruturas a serem utilizadas na pesquisa.

O segundo capítulo aborda a fundamentação teórica do trabalho, onde serão aprofundados assuntos importantes para o entendimento do projeto realizado, como cianobactérias e inteligência artificial. Já no terceiro capítulo, é feita uma revisão de literatura, que busca apresentar trabalhos desenvolvidos relacionados à criação e à disponibilização de bases de imagens para estudos computacionais, bem como mostra trabalhos direcionados ao reconhecimento automático de cianobactérias por meio de processamento digital de imagens. Ambas as questões servem para balizar e indicar os rumos desta pesquisa.

O quarto capítulo relata os materiais e os métodos utilizados ao longo do trabalho, de forma aprofundada. O capítulo seguinte, apresenta os resultados alcançados com os experimentos, na busca para a resolução do problema apresentado. No sexto capítulo, são apresentadas as considerações finais à que se chegaram após a conclusão a pesquisa, seguido das referências utilizadas no estudo.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, alguns conceitos importantes para o entendimento da problemática e da proposta são explanados.

2.1 O PROBLEMA DA EUTROFIZAÇÃO

A eutrofização é o enriquecimento dos ecossistemas aquáticos, com nutrientes, especialmente fósforo e nitrogênio (ESTEVES, 1998). Este aumento gera o aumento da produtividade do corpo hídrico, assim, um ambiente eutrófico possui altas quantidades de nutrientes, tornando a água com transparência limitada, e fazendo com que comumente ocorram florações de algas.

A eutrofização se dá tanto por processos naturais quanto por processos artificiais. Quando naturais, são lentos e contínuos, resultados das chuvas que lavam a superfície terrestre, trazendo nutrientes. Quando artificial, isto é, por atividade humana, os nutrientes são oriundos de atividades domésticas, urbanas e agrícolas. Incluem descargas orgânicas de resíduos tratados ou não, de águas de fossas sépticas e ligações clandestinas de esgoto em redes pluviais. Isto constitui um ambiente rico em nitrogênio e fósforo (SIQUEIRA e OLIVEIRA-FILHO, 2005).

Outro processo artificial ocorre quando existe a retirada da vegetação de uma bacia para a ocupação por agricultura, e a posterior colheita das culturas plantadas, reflete na remoção de nutrientes que não são repostos naturalmente. Assim, para aumentar a produtividade, o solo é fertilizado com químicos que possuem altas quantidades de fósforo e nitrogênio, o que pode causar redução na capacidade de infiltração do solo, levando ao escoamento do excesso de nutrientes para os corpos d'água, causando aumento de algas e organismos consumidores.

Um outro caso de processo artificial é o derivado de efluentes domésticos. O despejo de detergentes sintéticos na água, que são ricos em polifosfatos, carbonatos e silicatos, bem como os dejetos humanos, também causam a médio e longo prazo a eutrofização de corpos d'água. Isto se dá pelas consideráveis concentrações de nitrogênio e fósforo e pela composição orgânica dos esgotos. As indústrias contribuem também para a eutrofização, devido ao manuseio, descarga e

uso inadequado de compostos à base de nitrogênio, utilizados em seus processos industriais (SIQUEIRA e OLIVEIRA-FILHO, 2005).

Já quanto aos processos naturais, as chuvas também têm importante papel em regiões de intensa poluição atmosférica. Lagos e outros corpos d'água próximos a rodovias tendem a receber aporte adicional de fosfato e nitrogênio, em decorrência do tráfego de veículos (ESTEVES, 1998).

O florescimento algal causado pelo aumento de nutrientes pode atingir níveis que prejudiquem a qualidade da água, e em casos em que as algas sejam tóxicas, podem causar graves problemas para o ecossistema onde estão inseridas, bem como graves problemas de saúde para o ser humano, além de prejudicar todo o sistema de tratamento e distribuição de água.

2.2 AS CIANOBACTÉRIAS

As cianobactérias conhecidas também como algas azuis ou ainda cianofíceas, surgem, entre outros fatores, também como uma das consequências da eutrofização. São microrganismos que não são nem algas e nem bactérias comuns. Elas não têm membrana nuclear (características celulares de procariontes) e apresentam sistema fotossintetizante semelhante ao das algas. Não possuem estrutura de locomoção, como cílios ou flagelos, mas algumas podem se locomover combinando deslizamento com rotação em seu próprio eixo. Também possuem movimentos espasmódicos (SIQUEIRA e OLIVEIRA-FILHO, 2005).

Elas podem se apresentar de forma unicelular ou filamentosa, formando colônias de filamentos ramificados ou colônias irregulares. As células individuais são capazes de se dividir quando isoladas, formando uma nova colônia. As cianobactérias podem ainda colonizar áreas nuas sobre rochas e solo.

Quando há floração, isto é, aumento da quantidade das cianobactérias, ela altera a coloração da água produzindo uma espuma esverdeada, bem como altera o odor e sabor. A FIGURA 2 mostra um exemplo da condição visível de locais onde há florações de cianobactérias.

FIGURA 2 – Floração no Reservatório de Salto Grande – SP



FONTE: FAPESP, 2018.

LEGENDA: A espuma de coloração esverdeada na margem representa a alta concentração de cianobactérias no corpo d'água.

As cianobactérias de água doce possuem a capacidade de desenvolver toxinas em larga escala. Segundo Duy *et al* (2000), dos cerca de 50 gêneros das cianobactérias de água doce, pelo menos sete (*Anabaena*, *Aphanizomenon*, *Coelosphaerium*, *Gloetrichia*, *Microcystis*, *Nodulária* e *Nostoc*) contém espécies tóxicas.

Em relação às toxinas, os fatores que favorecem sua produção são variáveis e ainda pouco entendidos pela comunidade científica. Segundo Siqueira e Oliveira (2005), estudos apontaram que a produção e acumulação de toxinas está relacionado com o crescimento das colônias, a temperatura, oferta de nutrientes, presença de metais, entre outros. Por exemplo, o gênero *Microcystis* apresenta aumento e diminuição de toxinas proporcionais a aumento e redução de temperatura ótima. Apresenta também redução de toxicidade após a retirada do nitrogênio e carbono inorgânico da cultura, e aumento relacionado ao fósforo e ao metal zinco.

As cianotoxinas se diferem tanto em aspectos químicos quanto toxicológicos. Seus efeitos conhecidos variam de hepatotóxicos, neurotóxicos até a inibição de síntese de proteínas. Elas podem ser classificadas em peptídeos cíclicos, alcalóides e lipossacarídeos. O primeiro grupo é caracterizado pela presença de toxinas

hepatotóxicas, como a microcystina e a nodularina, e geralmente encontradas em água doce ou salgada. Em ensaios com ratos, as hepatotoxinas causam morte por hemorragia no fígado em poucas horas (CHORUS e BARTRAM, 1999). Em ambientes aquáticos, estas hepatotoxinas permanecem no interior das cianobactérias, e representam grande risco de contaminação do meio ambiente e possível contaminação de seres humanos. O segundo grupo, os alcalóides, são produzidos por plantas e algumas bactérias, geralmente tóxicos e bioativos. As toxinas alcalóides de cianobactérias de água doce (anatoxina-a e saxitoxina) são todas neurotóxicas. A cilindrospermopsina é um exemplo de toxinas citotóxicas. Em bioensaios com ratos, a morte ocorre por rápida parada respiratória (de 2 a 30 minutos), e causam também patologias em órgãos internos (DUY *et al*, 2000).

O último grupo, dos lipossacarídeos, são encontrados na parede celular de bactérias e cianobactérias, e também são tóxicas em geral (CHORUS e BARTRAM, 1999).

As cianobactérias apresentam efeitos em organismos aquáticos e em seres humanos. No zooplâncton, a suscetibilidade é altamente variável entre gênero e espécies. De fato, as cianotoxinas podem influenciar a estrutura do zooplâncton, especialmente quando as cianobactérias são grupos dominantes no ambiente. Os efeitos estudados incluem alteração e redução da capacidade de filtração e a sobrevivência de filhotes de crustáceos, paralisia de organismos, necrose hepática e morte de peixes, entre outros (CHORUS e BARTRAM, 1999).

Já em seres humanos, vários são também os efeitos. Os sintomas causados pela intoxicação também dependem da dose ingerida e via de contaminação combinados com o prévio estado de saúde da pessoa. Por fim, os efeitos também dependem diretamente do tipo de toxina, já que cada uma possui seu próprio mecanismo de toxicidade. O QUADRO 2 apresenta observações clínicas, a toxicodinâmica, e o número de variações estruturais que cada uma das cianotoxinas possui.

QUADRO 2 – CIANOTOXINAS, EFEITOS EM RATOS, NÚMERO DE VARIAÇÕES ESTRUTURAIS, TOXICODINÂMICAS E OBSERVAÇÕES CLÍNICAS

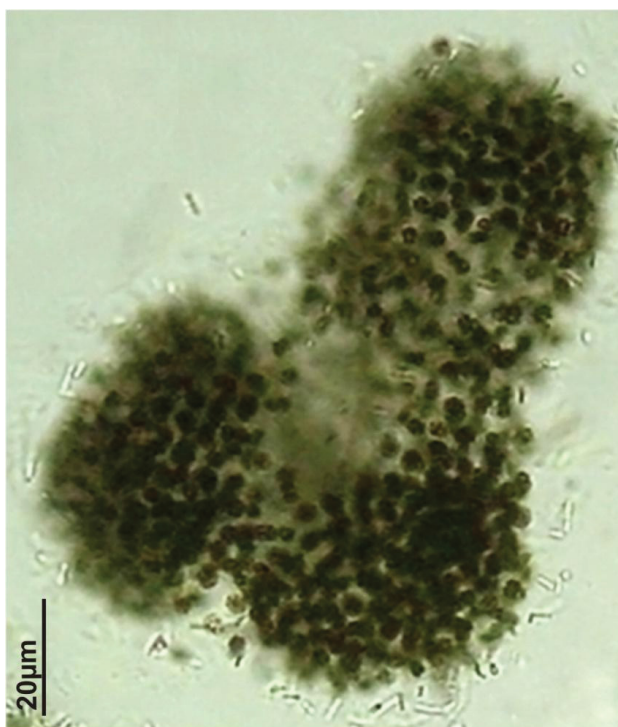
CIANOTOXINA	EFEITO	VARIAÇÕES ESTRUTURAIS	TOXICODINÂMICA	OBSERVAÇÕES CLÍNICAS
Anatoxina – a	Neurotóxico	Uma	Agonista nicotínico. Ação pós-sináptica de despolarização e bloqueio da atividade neuromuscular.	Cambaleamento, convulsões, salivação viscosa, lacrimação (em ratos), incontinência urinária, fraqueza muscular e morte por parada respiratória em poucos minutos.
Anatoxina – a (s)	Neurotóxico	Uma	Inibição irreversível da colinesterase (ChE).	Lacrimação, dispnéia, mucosa oral cianótica, defecação e crises clônicas antes da morte, que pode ocorrer por parada respiratória.
Saxitoxina	Neurotóxico	Cerca de seis	Bloqueio dos canais de sódio. Inibição da transmissão nervosa.	Não documentado.
Neosaxitoxina	Neurotóxico	-	Bloqueio dos canais de sódio. Inibição da transmissão nervosa.	Não documentado.
Microcistinas	Hepatotóxico	Mais de 50	Em hepatócitos, inibição da fosfatase protéica e ativação da fosfolipase A2 e da ciclooxigenase. Inibição da fosfatase protéica 1ª e 2ª, causando hiperfosforilação das proteínas associadas com o citoesqueleto, levando à deformidade celular. Em macrófagos, indução da produção de FNT e IL-1, culminando com a ativação da ciclooxigenase.	Taquicardia, colapso, exagerada respiração abdominal, cianose, convulsões e morte.

CIANOTOXINA	EFEITO	VARIAÇÕES ESTRUTURAIS	TOXICODINÂMICA	OBSERVAÇÕES CLÍNICAS
Nodularina	Hepatotóxico	Cinco	Indução de mudanças nos microfilamentos de actina e outros elementos do citoesqueleto, consequentemente destruição das células da camada sinusóide endotelial.	Taquicardia, dano hepático, carcinógenese, colapso, exagerada respiração abdominal, cianose, convulsões e morte.
Cilindrospermopsina	Hepatotóxico	-	Ação citotóxica. Deslocamento dos ribossomos para o centro da célula, diminuição do P-450 dos microsossomos hepáticos, acumulação de gotículas de gordura na porção central dos lóbulos hepáticos.	Anorexia, leve diarreia, respiração ofegante e possível morte entre 6 e 9 horas.
Lipopolissacarídeos – Endotoxina	Choque tóxico	Mais de três	Distúrbios gastrintestinais, choque tóxico e inflamação.	Gastroenterite, irritação cutânea, irritação ocular, reações alérgicas e asma.

FONTE: Adaptado de DUY *et al*, 2000.

A FIGURA 3 mostra a cianobactéria do gênero *Microcystis*, espécie *aeruginosa* vista pelo microscópio. A espécie possui colônias micro ou macroscópicas, alongadas, irregulares, possui mucilagem transparente (substância gelatinosa que as envolve) e suas células têm de 4-6 μ m de diâmetro (SANT'ANNA *et al*, 2012). A cor do conteúdo celular original é verde-escuro, entretanto, varia conforme as substâncias utilizadas na manipulação da amostra. Determinar espécies do gênero *Microcystis* não é tarefa trivial, pois as espécies são facilmente confundíveis entre si, devido à semelhança morfológica, variando minimamente tamanho, cor e disposição das células. A FIGURA 4 mostra a espécie *Microcystis wesenbergii*.

FIGURA 3 – *Microcystis aeruginosa*



FONTE: Sant'Anna *et al.*, 2012.

FIGURA 4 – *Microcystis wesenbergii*FONTE: Sant'Anna *et al.*, 2012

As cianotoxinas produzidas pelas cianobactérias afetam negativamente tanto os organismos aquáticos quanto seres humanos. Os sintomas da intoxicação dependem da via de contaminação, da dose e do estado de saúde em que a pessoa se encontra. O contato direto da pele no banho em duchas naturais, na natação, no esqui aquático pode irritar a pele do indivíduo, causar erupções, inchar os lábios, irritação dos olhos e ouvidos, dor de garganta, inflamações nos seios da face e asma. Já se ingerida a água contaminada, podem ocorrer náuseas, vômitos, dores abdominais, diarreias, complicações no fígado e fraqueza muscular, e um médico deve ser consultado o mais rápido possível. Também não é indicado utilizar a água para cozimento ou manuseio de alimentos, bem como não se deve comer pescados de águas contaminadas (COPASA, 2018).

O controle das cianobactérias pode ser feito por meio de algicidas, que são comparados à agrotóxicos. Eles são capazes de matar as cianobactérias, entretanto, liberam as toxinas existentes nelas. Devido a isto, no Brasil, o uso de algicidas para o controle do crescimento das cianobactérias em mananciais de abastecimento foi proibido pela Portaria 2914/11. Para o controle das florações, é necessário fazer a

regulação dos nutrientes internos, através de meios físicos, químicos e biológicos (CETESB, 2013).

O estado do Paraná segue as normas da Cetesb (Companhia Ambiental do Estado de São Paulo) no que diz respeito ao monitoramento das cianobactérias, que tem por base a portaria 2914/2011 do Ministério da Saúde. Não basta a identificação, é necessário realizar a contagem das células presentes nas amostras de água. De acordo com as normas da Cetesb (2013), é considerada ocorrência de floração quando se ultrapassa 10.000 células por ml. A portaria obriga o monitoramento mensal nos pontos de captação de água nos mananciais de abastecimento, quando o número de células de cianobactérias não exceder 10.000 células por ml, e semanalmente quando exceder esse valor. Também exige a análise semanal de cianotoxinas quando o valor exceder 20.000 células por ml no ponto de captação (CETESB, 2013).

2.3 CASOS RELATADOS

As cianobactérias se tornaram um problema mundial devido à sua toxicidade. Durante as últimas décadas, as estruturas químicas de uma série de toxinas cianobacterianas (“cianotoxinas”) foram identificadas e estudadas. Chorus *et al* (2000) conduziram uma pesquisa de casos atribuídos a cianotoxinas relatados ao longo do tempo. A lista descreve 14 casos notificados nas quatro formas de exposição (oral (ingestão), pulmonar (inalação, aspiração), dérmico (pele, contato mucoso) e hemodiálise (água usada para hemodiálise)), entre 1931 e 1996. Em 1878, foi publicado o primeiro caso documentado de intoxicação letal de animais causada pela ingestão de água de um lago infestado por cianobactérias (FRANCIS, 1878). Desde então, casos registrados incluem porcos, cães, cavalos, ovelhas, roedores, morcegos e outros (Codd *et al*, 1989). Há também relatos de mortes humanas.

Em 1996, 61 pessoas morreram na cidade de Caruaru, em Pernambuco, no caso que ficou conhecido como a “Tragédia da Hemodiálise”. Os pacientes de uma clínica de hemodiálise foram submetidos às máquinas que filtram o sangue, processo que utiliza água para esta filtragem. A água utilizada nos procedimentos estava contaminada com cianobactérias tóxicas. A toxina responsável foi a

microcystina LR. Estudos indicaram que o nível de água do reservatório estava baixo, então a cianobactéria se multiplicou gerando altos níveis de toxina. Os pacientes começaram a passar mal logo após a sessão de hemodiálise (JOCHIMSEN *et al*, 1998).

Outros casos também relacionam cianobactérias a mortes massivas de pássaros selvagens em uma lagoa no Japão (MATSUNAGA *et al*, 1999), mortes de flamingo-pequeno no Kenya (KRIENITZ *et al*, 2003), e ainda vários outros animais ao redor do mundo.

2.4 INTELIGÊNCIA ARTIFICIAL E REDES NEURAIS

A Inteligência Artificial (IA) é uma área da computação já consolidada que com o passar do tempo vem atuando de modo a auxiliar na resolução de problemas e casos reais. De modo geral, utilizam-se técnicas de inteligência artificial para que a partir de dados e informações do problema, a IA possa criar um modelo capaz de auxiliar no problema em questão. A este processo de indução de uma hipótese por experiência passada dá-se o nome de Aprendizado de Máquina (AM) (FACELI *et al*, 2011). Em conjunto com processamento de imagens e sistemas de visão computacional, torna-se uma ferramenta capaz de solucionar questões nas mais diversas áreas, como na ciência dos materiais (ALBUQUERQUE *et al*, 2007), no urbanismo (BOTELHO e CENTENO, 2005), na saúde (VERONEZI *et al*, 2011), mesmo na medicina, como por exemplo no monitoramento online de parâmetros respiratórios de pacientes com Esclerose Lateral Amiotrófica (ELA) (SILVA *et al*, 2013), entre outras áreas.

Dentro do aprendizado de máquina uma das técnicas amplamente utilizadas são as Redes Neurais Artificiais.

Tarefas cotidianas, como caminhar ou pegar um objeto, desenvolvidas com facilidade por seres humanos, ainda se apresentam como tarefas difíceis de serem ensinadas a robôs. A realização destas e de outras atividades aparentemente simples é possível graças à complexa estrutura biológica do cérebro humano.

As redes neurais artificiais, inspiradas no cérebro humano, visam simular a capacidade cognitiva e de aquisição de conhecimento na realização de tarefas.

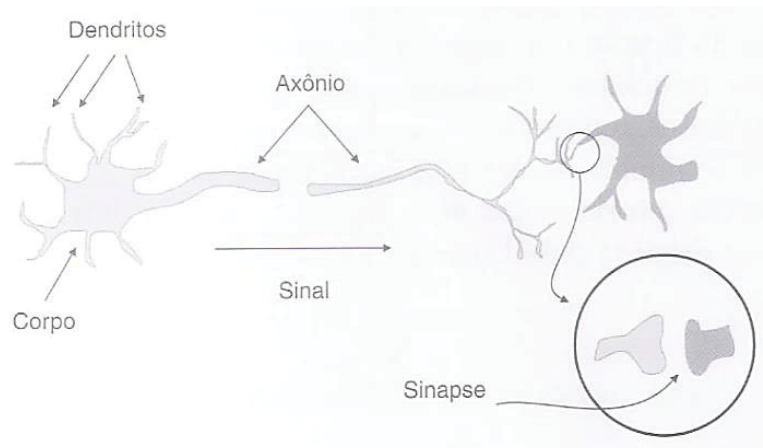
A procura por modelos computacionais ou matemáticos do sistema nervoso teve início da década de 1940. McCulloch e Pitts realizaram alguns dos primeiros estudos na área, propondo um modelo matemático de neurônio artificial em que os neurônios executavam funções lógicas simples. Eles mostraram que a combinação de vários neurônios artificiais em sistemas neurais produz um elevado poder computacional, já que qualquer função representada por uma combinação de funções lógicas poderia ser modelada por uma rede formada por estes neurônios, que inicialmente não possuíam capacidade de aprendizado (FACELI *et al*, 2011).

Outros autores ao longo dos anos trouxeram mais contribuições, ao passo que mais poder de processamento com máquinas mais potentes eram disponibilizados, até que em 1956 nasceram os dois paradigmas da Inteligência Artificial: Simbólica e Conexionista. A Inteligência Artificial Simbólica objetiva a simulação do comportamento inteligente humano sem uma inspiração biológica. A Inteligência Artificial Conexionista baseia-se na simulação da estrutura do cérebro construindo um sistema que o simule, de modo que este sistema apresentará inteligência, será capaz de aprender, assimilar, errar e ainda aprender com seus erros (FURTADO, 2019). Os avanços nos estudos de redes neurais continuam.

Uma rede neural pode ser considerada uma estrutura conexionista, um processador paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso (HAYKIN, 2001). Estas redes se caracterizam por um funcionamento similar ao da troca de informações dos neurônios do cérebro humano, possuindo também os neurônios artificiais e conexões entre eles, e assim são capazes de “aprender”, se treinadas adequadamente.

Um neurônio biológico, representado na FIGURA 5, é formado por dendritos, corpo celular e axônio. Os dendritos são receptores de estímulos nervosos advindos de outros neurônios ou mesmo do ambiente. Estes estímulos são enviados para o corpo celular ou soma, que os coleta, combina e processa. Por sua vez, o soma gera um novo impulso, enviando-o para o axônio. A função do axônio é conduzir os impulsos elétricos produzidos no corpo celular até outro local mais distante. O sinal do neurônio flui da esquerda para a direita, isto é, dos dendritos para o corpo celular, então para o axônio. Sinapse é o nome que se dá para o contato existente entre a terminação de um axônio e o dendrito de outro neurônio (HAYKIN, 1999).

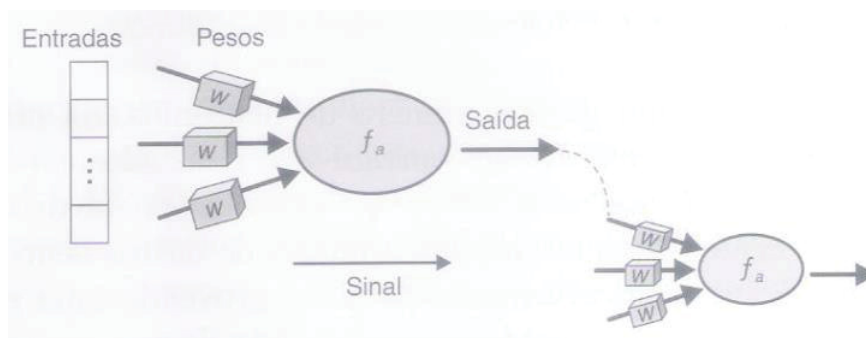
FIGURA 5 – Neurônio biológico simplificado



FORTE: FACELLI *et al*, 2011.

As unidades de uma RNA (Rede Neural Artificial), conhecidas como neurônios artificiais, trabalham computando funções matemáticas. Pode-se dizer que em um neurônio é realizado o processamento sobre uma ou, geralmente, várias entradas, a fim de gerar uma saída. Um neurônio artificial é uma estrutura lógico matemática que procura simular o comportamento e as funções de um neurônio biológico. Em comparação ao neurônio biológico, observa-se que cada terminal de entrada simula um dendrito, que recebe um valor. Estes valores são ponderados e combinados por meio de uma função matemática f_a , o que equivale ao processamento do soma. Por fim, a saída da função é a resposta do neurônio para a entrada. Desta forma, pode-se considerar o neurônio como a unidade fundamental que processa a informação (FURTADO, 2019). A FIGURA 6 mostra um neurônio artificial simplificado.

FIGURA 6 – Neurônio artificial



FORTE: HAYKIN, 1999 apud FACELLI *et al* 2011.

Para melhor compreensão dos valores de entrada de um neurônio, considera-se um objeto \mathbf{x} com d atributos, representado na forma de vetor como $\mathbf{x} = [x_1, x_2, \dots, x_d]^t$ e um neurônio com d terminais de entradas cujos pesos são w_1, w_2, \dots, w_d , que podem ser representados na forma vetorial como $\mathbf{w} = [w_1, w_2, \dots, w_d]$. A entrada total recebida pelo neurônio, u , pode ser definida pela equação:

$$u = \sum_{j=1}^d x_j w_j$$

Os neurônios podem apresentar conexões de entrada negativas ($w_j < 0$) ou positivas ($w_j > 0$). Um valor peso 0 equivale a ausência da conexão associada.

Já a saída de um neurônio é dada pela definição da função de ativação à entrada total, como definida a seguir:

$$f_a(u)$$

Existem várias funções propostas na literatura, como por exemplo as funções linear, limiar e sigmoidal. Elas limitam a amplitude do sinal de saída do neurônio. A função linear identidade implica retornar como saída o valor de u . Na função limiar (empregada no modelo de neurônio artificial de McCulloch e Pitts de 1943), o valor do limiar define quando o resultado da função limiar será igual a 1 ou 0 (alternativamente pode-se usar -1). Quando a soma das entradas recebidas ultrapassa o limiar que foi estabelecido, o neurônio torna-se ativo (saída +1). Quanto maior o valor do limiar, maior tem que ser o valor da entrada total para que o valor de saída do neurônio seja igual a 1. Na função sigmoidal, diferentes inclinações podem ser empregadas, e esta representa uma aproximação contínua e diferenciável da função limiar (FACELLI *et al*, 2011).

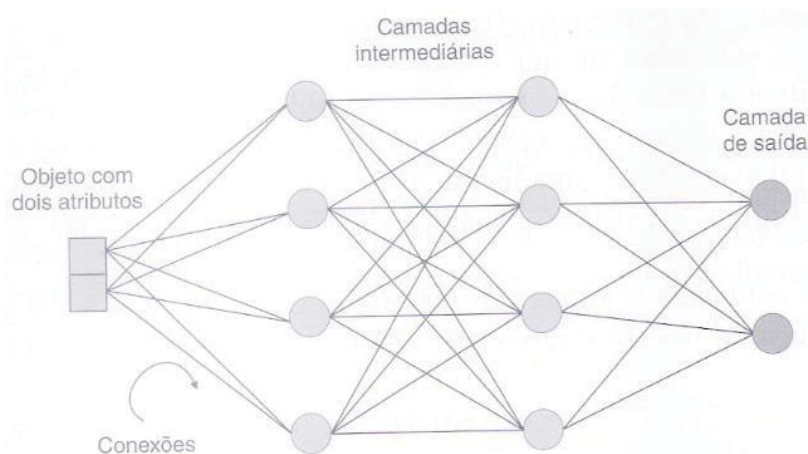
Os neurônios são organizados em uma ou mais camadas na arquitetura da rede, e interligadas por conexões, onde, na maioria dos modelos de arquitetura, simulam as sinapses dos neurônios biológicos do cérebro humano. Todas as conexões existentes possuem pesos associados, que ponderam a entrada recebida por cada neurônio da rede, o que simula as sinapses (FACELLI *et al*, 2011). Os

pesos têm seus valores ajustados em um processo de aprendizado e codificam o conhecimento adquirido pela rede (BRAGA *et al*, 2007).

Pode-se caracterizar uma RNA portanto, de forma simplificada, por dois aspectos: a arquitetura e o aprendizado. A arquitetura se refere ao tipo, número de unidades de processamento e forma de conexão. Já o aprendizado é relacionado às regras utilizadas para o ajuste dos pesos da rede e que informação é utilizada pelas regras.

Uma arquitetura de RNA pode ser organizada em uma ou mais camadas de neurônios. Quando há mais de uma, recebe o nome de multicamadas. Quando duas ou mais camadas de neurônios são utilizadas, um neurônio pode receber como entrada a saída de um neurônio da camada anterior, e também pode enviar suas saídas para a entrada de neurônios da camada posterior. A FIGURA 7 apresenta um exemplo de uma RNA com duas camadas intermediárias (escondidas ou ocultas) de neurônios, recebendo dois valores de entrada, e gerando dois valores de saída na camada de neurônios de saída (HAYKIN, 1999).

FIGURA 7 – Exemplo de RNA multicamadas típicas



FONTE: FACELLI *et al*, 2011.

Modelos conexionistas podem se modificar de várias formas, conforme a necessidade de aplicação. A arquitetura pode variar através das diferentes conexões entre as camadas: pelo número de camadas intermediárias; pelo número de unidades processadoras (neurônios ou nós); pela função de ativação; e pelo processo de aprendizado.

A forma como os neurônios de uma rede é estruturada está intimamente relacionada com o algoritmo de aprendizagem usado para treinar a rede.

Os tipos de arquiteturas geralmente são divididos em três grupos (FURTADO, 2019):

- *Feedforward* de uma camada, onde os neurônios da camada de entrada se conectam diretamente com a camada de saída. Os nós de entrada não processam nada, só apresentam os padrões à rede;
- *Feedforward* Multicamadas, onde também há uma ou mais camadas escondidas. Com a adição de camadas extras, a rede pode organizar problemas de mais alta complexidade de melhor forma do que uma rede de uma camada;
- Rede recorrente ou realimentada, que se distingue dos tipos *feedforward* ao permitir a realimentação de uma camada com as saídas geradas pela posterior, ou mesmo a realimentação do mesmo neurônio (*self-feedback*).

Através dos algoritmos de treinamento os pesos sinápticos são atribuídos, bem como o processo de aprendizagem é definido. Os algoritmos podem ser divididos em três classes basicamente (FURTADO, 2019):

- Aprendizado supervisionado: as saídas desejadas para o padrão apresentado à rede são conhecidas. Os parâmetros da rede são ajustados até que o erro seja minimizado ao máximo. Um algoritmo de treinamento supervisionado muito conhecido segundo Furtado (2019), é o *backpropagation*;
- Aprendizado não supervisionado: não há saída conhecida, desta forma, a própria rede deve ser capaz de extrair características relevantes, classificando as saídas em grupos pré-existentes. O algoritmo de Aprendizado Competitivo é um exemplo;
- Aprendizado por reforço: variante do aprendizado supervisionado, em que pode-se saber se as saídas estão corretas, mas não se dispõe de respostas corretas. É um método baseado em tentativa e erro, onde as conexões que levam a respostas satisfatórias são reforçadas.

Desenvolvida por Rosenblatt (1958), a rede Perceptron foi a primeira rede neural artificial a ser implementada. Ela apresenta apenas uma camada de neurônio, contendo um único neurônio. Embora simples, par Haykin (2001), demonstrou uma

boa acurácia preditiva em problemas de classificação de imagens. A rede perceptron é treinada por um algoritmo supervisionado e utiliza função de ativação limiar (HAYKIN, 2001).

Já as redes perceptron multicamadas MLP (*multilayer perceptron*) possuem uma ou mais camadas ocultas de neurônios e uma camada de saída. Nas camadas intermediárias utilizam funções de ativação não lineares (como sigmoideal), e cada neurônio realiza uma função específica, que é resultante das operações dos neurônios da camada anterior, que estão conectados a ele, de modo que à cada avanço, o processo vai se tornando mais complexo (HAYKIN, 2001; FACELLI, *et al*, 2011).

Portanto, pode-se dizer que três etapas são necessárias para aplicar redes neurais artificiais à resolução de um problema qualquer:

- Treinamento: ensina a rede a resolver um conjunto de padrões de saída associados a padrões de entrada;
- Teste: onde são apresentados padrões de entrada à rede, e as saídas obtidas são comparadas às saídas previamente desejadas;
- Aplicação: por fim, a rede sintetizada é utilizada na resolução do problema.

As redes neurais deram origem a redes neurais densas que são a base para o *Deep Learning*. Essas redes são altamente conectadas, capazes de gerar modelos de alto nível de abstração, atuando na visão computacional, no reconhecimento de sons e também processamento de linguagem natural. Sua atuação é muito variada em todos os campos do conhecimento, presente até mesmo em estudos para detecção do câncer (BROWN, 2017).

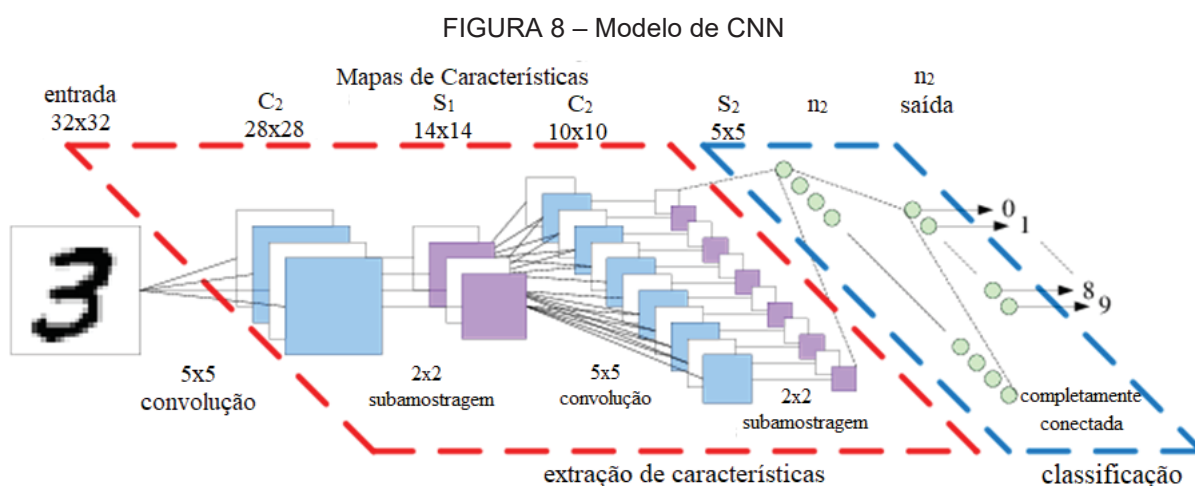
2.4.1 REDES NEURAS CONVOLUCIONAIS

As Redes Neurais Convolucionais são uma arquitetura de *Deep Learning* que vem sendo amplamente desenvolvidas nos últimos anos. Redes Neurais Convolucionais (CNN, sigla em inglês) foram inspiradas no modelo biológico da visão de animais. São redes neurais artificiais que aplicam a convolução (operação entre matrizes) em pelo menos uma de suas camadas de neurônios. Têm seu aprendizado baseado em camadas densas, fazendo multiplicações entre matrizes,

ou seja, convoluções de matrizes de acordo com um kernel (máscara) e tamanho determinado, semelhante aos filtros aplicados em processamento de imagens, tendo a representação dos dados em diferentes níveis de abstração, com a extração de características implícita. O fator chave do aprendizado profundo é que sua extração de características é aprendida automaticamente a partir do dado bruto (LECUN *et al*, 2015). Tem apresentado bons resultados em várias áreas, incluindo a visão computacional.

As CNNs são redes neurais especializadas em estruturas como texto, imagens, sons e vídeos. Atualmente, as redes neurais convolucionais representam o estado da arte na detecção de objetos (BOCHKOVSKIY *et al*, 2020)

A extração de características para a predição é realizada nas camadas convolucionais, como se pode observar na FIGURA 8.



FONTE: Adaptado de PEEMEN *et al*, 2011.

LEGENDA: Modelo de extração de características de uma CNN dentro da linha pontilhada vermelha e modelo de classificação dentro da linha pontilhada azul.

Assim, uma CNN consiste em múltiplas camadas, cada uma com sua função. Os dados de entrada sofrem convoluções, através destas camadas. Uma camada de convolução é composta por neurônios, que aplicam um filtro em uma parte específica da imagem. Cada neurônio é conectado a um conjunto de pixels da camada anterior, e um peso é atribuído a cada conexão. Desta forma, combinando-se as entradas deste neurônio com os pesos, produz-se uma saída, que será entrada da próxima camada. Os pesos podem ser interpretados como uma matriz

que representa o filtro de uma convolução de imagens no domínio espacial (também conhecido como máscara ou kernel).

Na clássica arquitetura de perceptrons de redes neurais, um neurônio é conectado a todos os neurônios da camada anterior, sendo esta arquitetura então chamada de completamente conectada. Já nas CNNs, apenas um subconjunto de entradas é conectado a cada neurônio. Esta mudança representa a capacidade que as redes neurais convolucionais possuem de análise de campos receptivos locais (*local receptive fields*). Os neurônios da mesma camada são agrupados em mapas, que são produzidos pelo agrupamento das saídas de neurônios que juntos cobrem uma parte da imagem processada. O mapa aplica o mesmo filtro em diferentes posições na imagem, devido ao compartilhamento de pesos durante o processo de treinamento, o que reduz significativamente a quantidade de parâmetros a serem aprendidos, e consequentemente seu tempo.

O tamanho do filtro define o tamanho da vizinhança que cada neurônio da camada vai processar. O passo representa quantos pixels serão pulados entre cada janela, de modo a saber qual o tamanho da camada seguinte. O processo de aprendizado da rede altera os pesos ao longo do treinamento, até encontrar os melhores valores dos filtros para o conjunto de dados utilizado.

Um ganho notável das CNNs em relação às técnicas clássicas é a possibilidade de as CNNs conseguirem misturar múltiplos mapas de características, o que permite extrair características cada vez mais complexas.

Após a convolução, aplica-se uma função de ativação, que basicamente serve para trazer a não-linearidade ao sistema, de modo que a rede consiga aprender qualquer tipo de funcionalidade. Há várias funções de ativação, como sigmoid, tanh, softmax, Relu e outras. Neste trabalho é utilizada a função *softmax* para predição das probabilidades na última camada. Ela define a probabilidade de uma classe acontecer em um problema multiclases. Esta função transforma a saída de cada unidade para um valor entre 0 e 1, e divide o valor de cada entrada de forma que a soma total das entradas seja igual a 1. Assim, o valor da saída revela a probabilidade de o objeto pertencer a cada uma das classes. Sua fórmula é dada como segue:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

Onde:

- \vec{z} = a entrada do vetor;
- z_i = elementos de entrada do vetor, podendo ser qualquer valor real;
- e^{z_i} = a função exponencial padrão aplicada a cada elemento de entrada do vetor, que resulta em um valor positivo;
- $\sum_{j=1}^K e^{z_j}$ = O termo abaixo da fórmula é o termo de normalização, que garante que todas as saídas da função estarão entre 0 e 1;
- K = O número de classes.

Seguindo na estrutura da rede, há também uma camada de *pooling*, uma camada de agrupamento que serve para simplificar a informação da camada anterior, reduzindo a dimensionalidade dos dados na rede. Da mesma forma que na camada de convolução, uma unidade de área é escolhida, e aplicada na saída da camada anterior. Por exemplo, se a saída da camada anterior for 12x12, a saída do *pooling* será 6x6, dependendo do tamanho do *pooling*. Um método comumente utilizado para a sumarização é o *maxpooling*, no qual apenas o maior número da unidade é passado para a saída. Com o *pooling*, há a diminuição da quantidade de pesos a serem aprendidos, e é também uma forma de se evitar o *overfitting*, isto é, a especialização da rede.

Ao final, uma camada completamente conectada (*fully connected*) é colocada na rede, onde sua entrada é a saída da camada anterior e sua saída são N neurônios, e N representa a quantidade de classes do modelo.

As arquiteturas de CNN repetem e combinam estas e outras funções em suas camadas. A LeNet-5 (LECUN *et al*, 1998), por exemplo, possui duas camadas de convolução seguidas de *pooling* e mais uma de convolução. Já a GoogLeNet (SZEGEDY *et al*, 2015) possui cinco camadas de convolução todas seguidas de uma de *pooling*.

O aprendizado profundo já atua com aplicações no mercado, como por exemplo aumentando a resolução de imagens, gerando sons para filmes mudos, dirigindo carros automatizados, predizendo de doenças e suporte para diagnósticos (UDACITY, 2018), atua ainda no reconhecimento de fala como por exemplo em dispositivos que realizam atividades de automação em residências, no reconhecimento facial como por exemplo para desbloqueio de *smartphones*, em

recomendações personalizadas na área de *marketing*, entre outras atividades cotidianas (PRÓXIMO NÍVEL EMBRATEL, 2020).

A empresa de tecnologia *Google* disponibilizou uma plataforma de ML, o *TensorFlow*, para que usuários possam trabalhar em seus próprios projetos de reconhecimento. Ele possui duas principais variações, importantes para o projeto em questão: uma técnica de transferência de aprendizado (a rede já é pré-treinada com um grupo de imagens, e então utiliza este treino para auxiliar no aprendizado de algo diferente) para classificação de imagens, e outra de detecção de objetos nas imagens.

TensorFlow não é a única ferramenta de aprendizado de máquina, há outras como o Keras², que consiste numa API de rede neural de alto nível, escrita na linguagem Python³. A biblioteca do Keras foi projetada para fornecer rápida experimentação com redes neurais profundas, e além de operar independentemente, também foi integrada ao *TensorFlow* em 2017 (BT ATSIT, 2021). Há também o Pytorch⁴, uma ferramenta também de código aberto que propõe flexibilidade e uso eficiente de memória nas operações de aprendizado profundo. *TensorFlow* foi a ferramenta escolhida para o desenvolvimento dos treinamentos efetuados neste trabalho, especialmente pelo fato de obter ampla documentação disponível.

O *TensorFlow* possui alguns modelos, como o *Inception* e o *Mobilenet*. A arquitetura *Mobilenet* é uma família de modelos de visão computacional especialmente desenvolvida para dispositivos móveis, criada pela Google. Segundo a empresa, a arquitetura é projetada para maximizar a precisão de maneira eficaz, bem como se preocupar com os recursos restritos de um aplicativo integrado ou no dispositivo. Já a *Inception* é também uma família de arquiteturas, derivada da *GoogleLeNet*, com melhorias. A *GoogleLeNet* foi apresentada pela Google em 2014, e foi uma das ganhadoras do evento de reconhecimento de imagens *ImageNet Large-Scale Visual Recognition Challenge* de 2014. A rede tem 22 camadas, e vários módulos de inepção conectados juntos para se tornar mais profundo. Assim, a arquitetura não somente empilha as camadas, mas também aumenta a largura destas camadas utilizando os módulos *inception*. Estes módulos fazem pequenas

² Keras. <https://keras.io/>

³ Python. <https://www.python.org/>

⁴ Pytorch. <https://pytorch.org/>

convoluções, que objetivam reduzir drasticamente a quantidade de parâmetros, para que o treinamento seja facilitado (SZEGEDY, 2015).

Há também outros modelos de aprendizado profundo, como VGG (LIU e DENG, 2015), que participou do ImageNet Challenge de 2014 levando o primeiro e segundo lugares nas categorias de detecção e classificação de imagens, respectivamente, a Resnet (HE *et al*, 2016), uma rede residual de 152 camadas que venceu a competição ILSVRC 2015, a DenseNet (HUANG *et al*, 2017), que propõe uma arquitetura para trazer o máximo fluxo de informações básicas entre as camadas, entre várias outras redes.

Assim, as duas arquiteturas, *MobileNet* v1 (para potencialidade de futura aplicação embarcada) e *Inception* v3, foram utilizadas neste trabalho, tanto no processo de classificação de imagens quanto no processo de detecção de objeto na imagem.

Outra arquitetura relativamente nova é a *Faster – RCNN* (REN *et al*, 2017), uma rede classificadora de regiões associada a extratores de características, baseados em CNN.

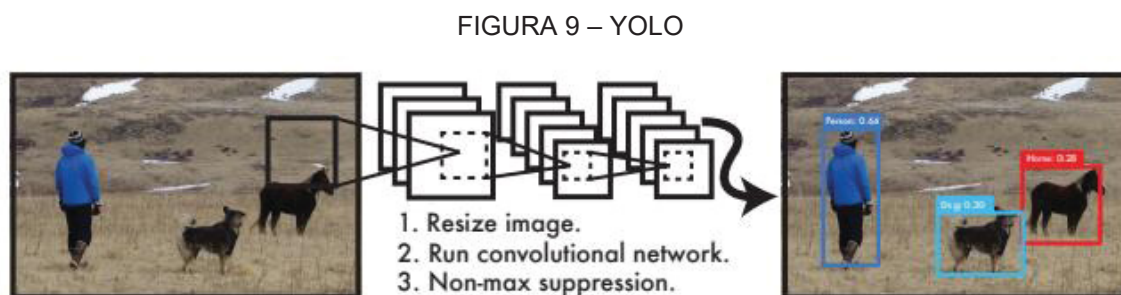
Rede convolucional baseada em região (RCNN) é um dos métodos de detecção que representa o estado-da-arte (GIRSHICK *et al*, 2014). O RCNN propôs um método que utiliza um mapa de características que é uma saída de uma rede convolucional. Na mesma época, diversas abordagens baseadas na utilização do mapa de características foram propostas (GIRSHICK, 2015; REN, 2015; ZENG *et al*, 2016; ERHAN *et al*, 2014; REDMON *et al*, 2016, LIU *et al*, 2016). Elas alcançaram alto desempenho, sendo consideradas o estado-da arte na área de detecção de objetos (ROH e LEE, 2017). A *Faster – RCNN* apresenta melhorias de suas antecessoras RCNN e Fast-RCNN. Assim, esta arquitetura, também foi escolhida e utilizada neste projeto, para a identificação de objetos.

Vale ressaltar a diferença entre as duas tarefas de identificação empregadas nesta proposta: a classificação de imagens e a identificação de objetos. A classificação de imagens consiste, na abordagem em questão, em dizer se uma imagem possui ou não *Microcystis aeruginosa*. Já o processo de identificação de objetos consiste em identificar não só a presença, mas também a localização do objeto em uma imagem.

YOLO

A YOLO (*You Only Look Once*) é uma outra abordagem na detecção de objetos, onde a mesma rede é capaz de identificar os objetos, onde estão, e a probabilidade da classe em uma única avaliação (REDMON *et al*, 2016). Segundo seus criadores, a YOLO comete mais erros de localização, mas é menos provável de prever falsos positivos no fundo.

A FIGURA 9 apresenta o funcionamento da YOLO, onde observa-se que uma única rede convolucional prevê simultaneamente os *bounding boxes* e probabilidades de classe para eles.



FONTE: REDMON *et al*, 2016.

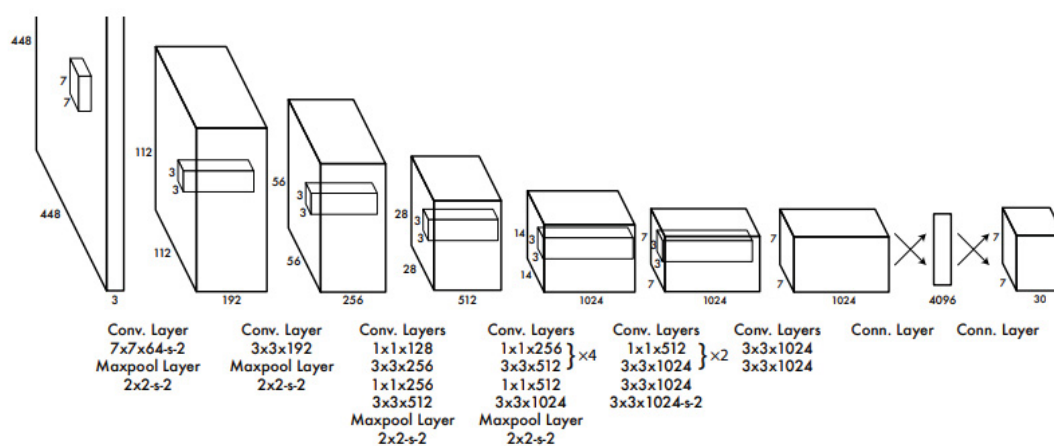
LEGENDA: (1) redimensiona a imagem de entrada para 448×448 , (2) executa uma única rede convolucional na imagem e (3) limita as detecções resultantes pela confiança do modelo.

O modelo é implementado como uma rede neural convolucional e avaliado no conjunto de dados de detecção de VOC PASCAL (EVERINGHAM, 2010). As primeiras camadas da rede extraem características, enquanto as camadas totalmente conectadas realizam a previsão das probabilidades de saída, e as coordenadas.

A arquitetura da rede é inspirada também no *GoogleLeNet*, para classificação de imagens. A rede possui 24 camadas convolucionais seguidas por 2 camadas totalmente conectadas. Em vez dos módulos iniciais usados pelo *GoogLeNet*, utiliza-se simplesmente camadas de redução 1×1 , seguidas por camadas convolucionais 3×3 , como apresenta a FIGURA 10. As camadas foram pré-treinadas no conjunto de dados da *ImageNet*, com 1000 classes. Para performar a previsão, foram adicionadas quatro camadas convolucionais e duas camadas totalmente conectadas com pesos inicializados aleatoriamente, e a resolução de entrada foi aumentada de

224x224 para 448x448. Assim, a camada final, utiliza função de ativação linear e prediz tanto a probabilidade da classe quanto as coordenadas.

FIGURA 10 – Arquitetura da YOLO



FONTE: REDMON *et al*, 2016.

LEGENDA: A rede possui 24 camadas convolucionais seguidas por 2 camadas totalmente conectadas. Camadas convolucionais 1x1 alternadas reduzem o espaço de recursos das camadas anteriores. As camadas convolucionais foram pré-treinadas na tarefa de classificação com a base *ImageNet*, com metade da resolução (imagem de entrada 224x224) e, em seguida, a resolução foi dobrada para detecção.

YOLO ainda está atrás dos sistemas de detecção de última geração em precisão. Embora possa identificar rapidamente objetos em imagens, ela não apresenta bons resultados para localizar com precisão alguns objetos, especialmente os pequenos que aparecem em grupos, como por exemplo uma imagem com um bando de pássaros juntos voando (REDMON, 2016).

Após comparação com métodos de detecção como Fast-RCNN, para os autores, YOLO é o estado da arte em detecção de objetos em sistemas de tempo-real.

A YOLOv3 (chamada de YOLO a partir deste momento no documento) traz pequenas melhorias como melhor acurácia com 0,5 de *IoU* (intercessão sobre a união, isto é, quão próximo o retângulo predito está do retângulo marcado como sendo o objeto) e com objetos pequenos, foi, por este motivo, escolhida para treinar a presente base de dados.

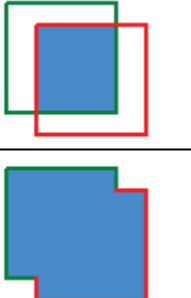
2.5 DETECÇÃO DE OBJETOS

Em detecção de objetos, a avaliação não é trivial, porque existem duas tarefas distintas para medir:

- Determinar se um objeto existe na imagem (classificação);
- Determinar a localização do objeto (localização, uma tarefa de regressão).

Em detecção de objetos, é necessário estabelecer um limite para que um objeto possa ser detectado como positivo. Para isto, é importante entender o conceito de IoU. A métrica IoU (*Intersection Over Union*) é uma medida baseada no índice Jaccard, um coeficiente de similaridade para dois conjuntos de dados (JACCARD, 1901). No escopo de detecção de objetos, o IoU mede a área de sobreposição entre o retângulo (*area of overlap*) que delimita a posição predita do objeto pelo modelo e o retângulo que delimita a posição correta do objeto, dividida pela área de união entre eles (*area of union*). A FIGURA 11 ilustra a fórmula:

FIGURA 11 – Fórmula *Intersection Over Union*

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{área de sobreposição}}{\text{área de união}}$$


FONTE: PADILLA *et al*, 2020.

O *threshold* (limite), dependendo da métrica, em geral é utilizado com valores de 50%, 75% ou 95%, mas pode-se utilizar qual se queira, dependendo do contexto de aplicação do estudo. Ele representa o limite mínimo escolhido para ser reconhecido como uma detecção positiva.

Uma IoU acima de 0,5 (50%) é considerada uma boa predição (sendo o valor mais utilizado), pois é caracterizada como sendo melhor que a probabilidade de ser positivo ou negativo (50%), portanto, tratada como positivo, e a predição perfeita é 1.

Neste trabalho, referente à métrica de IoU, os limites escolhidos foram 30%, 50%, 75% e 95%, para fins de comparação, e modo a explorar a previsibilidade dos modelos em diferentes limiares.

Também é importante destacar e entender os termos utilizados para as métricas:

- Verdadeiro Positivo (TP): Uma detecção correta. Detecção com IOU \geq limite;
- Falso positivo (FP): Uma detecção errada. Detecção com IOU $<$ limite;
- Falso negativo (FN): Um positivo não detectado;
- Verdadeiro negativo (TN): Não se aplica. Representaria uma não detecção correta. Na tarefa de detecção de objetos, há muitos retângulos delimitadores possíveis que não devem ser detectados em uma imagem. Assim, TN seriam todos os retângulos delimitadores possíveis que não foram detectados, de forma correta (tantos retângulos possíveis em uma imagem).

A métrica matriz de confusão é inteiramente constituída pelos termos descritos acima.

A matriz de confusão é uma ferramenta padrão para avaliação de modelos. Nela, é possível comparar as respostas que o modelo predisse com o valor real. Assim, a matriz separa em quatro quadrantes: primeira linha e primeira coluna: Verdadeiro positivo (TP, positivo predito como positivo); primeira linha e segunda coluna: Falso positivo (FP, negativo predito como positivo); segunda linha e primeira coluna: Falso negativo (FN, positivo predito como negativo); segunda linha e segunda coluna: Verdadeiro negativo (TN, negativo predito como negativo). Na TABELA 1 é possível observar a constituição da matriz.

TABELA 1 – Matriz de Confusão

		Valores Reais	
		Positivo	Negativo
Valores Preditos	Positivo	TP	FP
	Negativo	FN	TN

FONTE: A autora (2020).

A partir dos valores individuais da matriz de confusão, é possível calcular a também a especificidade. A especificidade se refere à porção de amostras negativas cujos resultados foram negativos, e a fórmula é dada pelos verdadeiro negativos divididos pela soma dos verdadeiros negativos mais os falso positivos, conforme a seguir:

$$\text{Especificidade} = \frac{TN}{TN + FP}$$

Deste modo, para avaliar os modelos treinados neste trabalho, foram escolhidas as seguintes métricas: *Precision* e *Recall* (que serão explicados à diante com mais detalhes) baseados em 4 limites de IoU, AP (*Average Precision*) e *loss*, além dos elementos individuais da matriz de confusão, juntamente com a especificidade.

Para calcular a AP, para uma classe específica, a curva de *Precision* x *Recall* é calculada a partir da saída de detecção do modelo, variando o limite de pontuação do modelo que determina o que é contado como uma detecção positiva prevista pelo modelo da classe (*threshold*).

Padilla *et al* (2020) desenvolveram um *survey* comparando as performances de métricas para *object detection*. Os autores ressaltam que a falta de consenso em diferentes trabalhos e implementações de AP é um problema enfrentado pela comunidade acadêmica e científica, e, portanto, eles fizeram um levantamento das métricas mais utilizadas e propuseram um padrão de implementação, que foi utilizado neste trabalho.

A avaliação dos métodos de detecção de objetos é principalmente baseada nos conceitos de *Precision* P (também conhecido como valor preditivo positivo) e *Recall* R (também conhecida como revocação ou sensibilidade). *Precision* mede a precisão das predições, ou seja, é a porcentagem de predições corretas (quantos acertos em todas as detecções feitas). Se há uma pontuação de precisão próxima a 1,0, então há uma alta probabilidade de que tudo o que o classificador prevê como uma detecção positiva seja de fato uma previsão correta. *Recall* mede capacidade de um modelo de encontrar todos os casos relevantes dentro de um conjunto de dados, ou seja, é a porcentagem de previsões positivas corretas entre todas os positivos fornecidos (quantos acertos dentre todos os positivos). Se há uma pontuação de *Recall* próxima a 1,0, quase todos os objetos que estão em seu

conjunto de dados serão detectados positivamente pelo modelo. *Precision* e *Recall* são respectivamente definidos como:

$$Precision = \frac{\text{objetos detectados corretamente}}{\text{todos os objetos detectados}}$$

$$Recall = \frac{\text{objetos detectados corretamente}}{\text{todos os objetos do ground truth}}$$

Onde *ground truth* é a classe correta.

Há uma relação inversa entre precisão e *Recall*, e essas métricas dependem do limite de pontuação do modelo definida (com *threshold* de 50% mais objetos serão detectados do que com um de 75%).

A curva *Precision* x *Recall* é uma boa maneira de avaliar detectores de objetos. Na prática, AP é a precisão média em todos os valores de *Recall* entre 0 e 1.

Um outro parâmetro de avaliação dos modelos treinados é a perda no aprendizado, conhecido também como *loss*. Quanto menor for a perda, melhor o modelo (a menos que o modelo tenha sofrido grande ajuste aos dados de treinamento, isto é, *overfitting*). A perda é calculada no treinamento e na validação e sua interpretação é o desempenho do modelo nesses dois conjuntos. Diferentemente da precisão, a perda não é uma porcentagem, mas uma soma dos erros cometidos a cada passo/época em conjuntos de treinamento e/ou validação. O objetivo do aprendizado é reduzir o valor da função de perda em relação aos parâmetros do modelo, alterando os valores do vetor de peso por meio de diferentes métodos de otimização (como a retropropagação). O valor da perda implica em quão bem ou mal um determinado modelo se comporta após cada iteração da otimização. Idealmente, espera-se redução da perda após cada uma ou várias iterações, de modo que este é um parâmetro que pode e deve ser observado durante o treinamento de uma rede, para realizar seus ajustes e encontrar melhores parâmetros de treinamento.

3 REVISÃO DE LITERATURA

Trabalhos relacionados à resolução de problemas via processamento de imagens não são recentes, assim como a utilização desta abordagem na identificação de cianobactérias.

Deste modo, a revisão de literatura se dividirá em duas partes: a primeira parte abrange trabalhos relacionados à criação e disponibilização de bases de dados para aplicações em processamento digital, e a segunda referente à utilização em pesquisas sobre cianobactérias.

3.1 CRIAÇÃO E DISPONIBILIZAÇÃO DE BASES DE DADOS

Para realizar esta etapa da revisão de literatura, foram escolhidas as seguintes bases de trabalhos científicos: Periódicos Capes⁵, Scielo⁶ e Google Acadêmico⁷. Dentre os termos pesquisados, existiram “image database”, “image dataset”, “datasets”, “microscopy datasets”, entre outros, e foram considerados trabalhos a partir de 2010, relacionados à microscopia. A seguir, são apresentadas as obras que se destacaram entre os resultados. Os trabalhos aqui relatados evidenciam a importância e necessidade de se ter bases de dados nas mais diversas áreas, de modo a proporcionar a facilidade na pesquisa aplicada utilizando técnicas computacionais de imagem.

Segundo os autores Riffle e Davis (2010), devido ao crescente desenvolvimento de métodos computacionais para analisar imagens de microscopia fluorescente automaticamente e em grande escala, algoritmos de treinamento e validação de imagens têm sido desenvolvidos na área. Eles ressaltam que geralmente para esta tarefa, grandes volumes de dados são necessários, com anotações e metadados disponíveis. Assim, projetaram uma grande base de dados, chamada o YRC PIR, que descreve a localização subcelular e a co-localização de proteínas em imagens de microscopia fluorescente. Contém 532.182 imagens TIFF de quase 85.000 experimentos separados e seus dados experimentais associados.

5 Periódicos Capes. <https://www.periodicos.capes.gov.br/>

6 Scielo. <https://scielo.org/>

7 Google Acadêmico. <https://scholar.google.com.br/>

A base está disponível gratuitamente, e para os autores, é um recurso poderoso para pesquisadores.

Diante do fato da enorme quantidade de dados retiniais obtidos, isto é, imagens de fundo de retina, permanecer armazenada localmente, os autores Zhang *et al* (2010) apresentam uma base de imagens online, com acesso aberto para que pesquisadores possam avaliar seus próprios sistemas de segmentação. Os autores ainda desenvolveram uma ferramenta própria de segmentação para a construção da base. O conjunto de imagens, batizado de ORIGA(-light), possui 650 imagens de retina, anotadas por profissionais, com foco no diagnóstico de glaucoma. Ressaltam que a base será alimentada com dados novos quando possível.

No trabalho a seguir, os autores Solli e Lenz (2010) apresentam duas grandes bases de dados, uma com 750.000 imagens e outra com 1,2 milhões. Eles geraram a base maior a partir de um mecanismo de busca que realiza a indexação das imagens, isto é, são imagens da internet categorizadas. As imagens da base menor, são provenientes de um provedor de imagens chamado Matton Images. As imagens são referentes a emoções, e contém metadados na forma de palavras-chave. Os autores realizaram experimentos de classificação de imagens para mostrar o potencial da base. A utilização de histogramas globais, isto é, descritores baseados em estatísticas globais de imagem, apresentaram resultados melhores na classificação das categorias emocionais, do que com a utilização de descritores de imagem locais.

Gupta *et al* (2010) criaram e disponibilizaram uma base de dados para utilização em reconhecimento facial e afins, chamada *Texas 3D Face Recognition Database*. A base é composta por 1.149 pares de imagens de alta resolução, provenientes de uma câmera estéreo, de 118 pessoas. São coloridas, e passaram por fases como normalização, pré-processamento e alinhamento. Há descrições nas imagens referentes a informações dos sujeitos, como gênero dos sujeitos, etnia, expressão facial e as localizações de 25 pontos fiduciais faciais antropométricos localizados manualmente. Os autores criaram também espécies de partições específicas dos dados, com o intuito de desenvolver e avaliar algoritmos de reconhecimento facial 3D.

Outra base interessante criada e disponibilizada, é a *Dresden Image Database*. Os autores Gloe e Böhme (2010) a desenvolveram especialmente para utilização com técnicas forenses aplicadas às imagens. Vale ressaltar que um dos

objetivos de técnicas forenses em imagens é responder se determinada imagem foi ou não alterada, e ainda se é possível rastrear a origem da imagem, no que se refere ao dispositivo de aquisição. A base possui mais de 14.000 imagens de várias cenas internas e externas, que foram adquiridas sob condições controladas, através de 73 câmeras digitais. Além da base original, foram coletadas imagens auxiliares para a estimativa do padrão de ruído do sensor específico do dispositivo, e um outro subconjunto de imagens para trabalhar com algoritmos de compressão JPEG. A proposta dos autores é tornar a base, que é gratuita, um recurso útil para pesquisadores e investigadores forenses.

Praus e Praks (2010) utilizaram técnicas matemáticas para a recuperação de imagens semelhantes relacionadas à água, de um banco de dados de teste de 126 imagens RGB. A base era composta por fotografias digitais dos vários níveis de água, incluindo também imagens semelhantes de áreas de superfície e vegetação. Os autores utilizaram análise de componentes principais (PCA) com base na decomposição de valores singulares (SVD), com 6 componentes principais. As imagens foram agrupadas e divididas em 2 grupos principais em termos de cores semelhantes dos objetos exibidos. Verificaram então que a abordagem MIA-LSI complementada com um método de agrupamento adequado é capaz de reconhecer as imagens semelhantes de águas superficiais de acordo com a cor e a forma dos objetos flutuantes. Segundo os autores, esta descoberta pode ser utilizada no emprego de monitoramento visual automático via imagens digitais da qualidade da água de superfície.

Barbosa *et al* (2016) desenvolveram um estudo do tipo *survey* para demonstrar a utilização de imagens digitais na análise visual de plantas na área agrícola. Eles ressaltam que a análise via imagens é um método não destrutivo, rápido e reprodutível, e que permite o armazenamento das imagens para posterior avaliação. Citam ainda que as análises através de programas e algoritmos computacionais aumentam a possibilidade de extração de informações, facilitando o processo de aquisição e análise dos dados. Os autores descrevem diferentes meios de aquisição das imagens no que refere às câmeras e à captação, suas aplicações na análise de plantas, e concluem que além das possibilidades de avaliação, a utilização da análise digital permite processar um grande número de amostras, tendo seu potencial expandido pela utilização em dispositivos móveis e drones.

O diagnóstico médico com auxílio de computador é uma área amplamente explorada. O trabalho desenvolvido por Yuan *et al* (2011) apresenta uma abordagem para recuperação de imagens médicas com base em conteúdo. A proposta realiza a aquisição das imagens junto com seus relatórios de diagnóstico, em seguida são limpas, validadas e aprimoradas. São então utilizados algoritmos de processamento de imagens, para que apenas regiões anatômicas sejam consideradas para a correspondência das imagens. Extrai-se também um vetor de recursos, essencial na recuperação de imagem por conteúdo, segundo os autores. Para eles, o método de construção de banco de dados para sistemas de diagnóstico auxiliado por computador é comprovadamente eficiente. Assim, criaram um banco de imagens de tomografia computadorizada do cérebro. Eles utilizaram técnicas de extração de características, com o intuito de fornecer uma base de imagens previamente diagnosticadas à radiologistas, e obter apoio no diagnóstico. A base apresenta 3 *datasets*, com imagens de diagnóstico normal, com tumor e com derrame, totalizando 285 imagens.

Ziehl-Neelsen é um teste bacteriológico de microscopia de coloração para a detecção da tuberculose. Segundo Shan *et al* (2017), sua sensibilidade é baixa, mas diagnóstico por meio da captura automatizada da imagem em foco (autofoco), costura de campos de visão para formar mosaicos (*autostitching*) e segmentação automática de bacilos (gradação) pode melhorar significativamente a sensibilidade. Assim, os autores Shah *et al* (2017) construíram um banco de imagens que permeiam estes três domínios. A base é oriunda de 3 microscópios diferentes, e as imagens estão em 7 categorias de diversos conjuntos de dados, totalizando mais de 4200 imagens. A base foi validada com diferentes algoritmos automatizados, e está disponibilizada gratuitamente.

Hakro e Tallib (2016) afirmam que o reconhecimento e o manuseio de caracteres por imagem são campos ativos de pesquisa que envolvem compreensão, interpretação, manuseio eficiente e encaminhamento de documentos, bem como sua recuperação. Segundo eles, ainda há muito a se fazer na pesquisa sobre os *scripts* cursivos (conectados), e muitos pesquisadores estão trabalhando na área. A escrita Sindi tem a maior extensão do alfabeto árabe original entre as línguas que adotam a escrita árabe; ele contém 52 caracteres, em comparação com 28 caracteres do alfabeto árabe original, a fim de acomodar mais sons para o idioma. Assim, os autores desenvolveram uma base com mais de 4 bilhões de palavras e 15 bilhões

de caracteres, em 150 fontes diversas, com 4 estilos e 4 espessuras diferentes. Como fontes da base, utilizaram sites, livros e teses. Desenvolveram ainda um aplicativo para criar uma imagem de texto a partir de um documento de texto compatível. O banco de dados está disponível gratuitamente de forma parcial ou completa.

Para Öhlschläger e Vö (2016), embora houvesse o crescente interesse da comunidade acadêmica em regras semânticas e sintáticas referentes à uma cena, e em como elas permitem uma percepção e orientação de objeto eficazes, ainda nenhum banco de dados de imagens contendo modificações de cena de objeto controladas estava disponível publicamente à época de seus trabalhos. Assim, propuseram o primeiro banco de dados desse tipo, o SCEGRAM. Fotografaram 62 cenas internas do mundo real, com variações: seis condições de consistência que contêm violação semântica e sintática, bem como suas combinações. A base foi validada e, também apresenta classificações de consistência para cada condição de cena de objeto, bem como cenas de ausência de objeto e imagens somente de objeto. Para os autores, o banco de dados de imagens SCEGRAM é o primeiro a conter inconsistências semânticas e sintáticas bem controladas de objeto-cena que podem ser usadas em uma ampla gama de paradigmas cognitivos.

Semelhante à proposta do presente trabalho, Sen *et al* (2015) propuseram uma base de dados com 627 imagens de algas de água doce da Turquia, primeira base deste tipo do país, segundo os autores. As imagens foram adquiridas de estudos prévios, e posteriormente padronizadas de acordo com o banco de dados central de herbários turcos e um banco de dados de algas de água doce turcas. A base conta ainda com informações sobre a taxonomia das algas. Os dados que a base apresenta são referentes às *Bacillariophyta* (202 dados), *Chlorophyta* (255 dados), *Chrysophyta* (3 dados), *Cryptophyta* (2 dados), *Cyanophyta* (87 dados), *Dinophyta*, (12 dados), *Euglenophyta* (61 dados), *Prasinophyta* (1 dado), *Rhodophyta* (1 dado) e *Xanthophyta* (3 dados).

A grande variedade de áreas onde pesquisadores criam e disponibilizam bases de dados salienta a popularidade e eficiência de se utilizar técnicas de processamento de imagens para solucionar problemas reais. Vale ressaltar ainda, que não foram encontradas bases similares à proposta no presente trabalho.

3.2 PROCESSAMENTO DE IMAGENS APLICADO À CIANOBACTÉRIAS

Nesta segunda etapa da revisão de literatura, foram escolhidas as mesmas bases, e os termos pesquisados foram “image processing cyanobacteria” e “cyanobacteria classification”, “cyanobacteria automated recognition”, “cyanobacteria neural networks”, entre outros, bem como uma combinação deles. Aqui buscou-se relatar os trabalhos desenvolvidos utilizando computação para solucionar problemas relacionados à cianobactérias, sem restrição de período.

O trabalho mais recente publicado na literatura até a confecção do presente trabalho data de agosto de 2020. Este é o trabalho que mais se assemelha ao trabalho proposto. Baek *et al* (2020) desenvolveram um modelo de deep learning utilizando o modelo baseado em regiões Fast R-CNN para classificar cinco espécies de cianobactérias (*Microcystis aeruginosa*, *Microcystis wesenbergii*, *Oscillatoria*, *Aphanizomenon*, and *Dolichospermum*). O modelo baseado em regiões permite a detecção de múltiplos objetos ao mesmo tempo quando comparado ao modelo de CNN que detecta apenas um por vez. Posteriormente, aplicaram um modelo de CNN para a contagem de células de *Microcystis*. As 1250 imagens do estudo, que não estão disponibilizadas publicamente, foram capturadas via microscópio, e foram advindas de amostras de água de rios locais (Coréia do Sul), preservadas com lugol. As identificações foram realizadas por especialistas. As precisões obtidas para a classificação foram de 0,929, 0,973, 0,829, 0,890 e 0,890 para *Microcystis aeruginosa*, *Microcystis wesenbergii*, *Dolichospermum*, *Oscillatoria* e *Aphanizomenon*, respectivamente. Quanto à contagem, a avaliação foi feita pelas métricas coeficiente de determinação R^2 (indica quanto o modelo foi capaz de explicar os dados coletados) e RMSE (*root mean squared error*), medida que calcula a raiz quadrática média dos erros entre valores observados (reais) e previsões (hipóteses). Os valores obtidos foram 0,78 para R^2 e 0,85 para RMSE. A contagem apresentou alguns problemas de estimação para mais e para menos em alguns casos, mas segundo os autores, estes são valores aceitáveis. Para os autores, é necessário realizar um pré-processamento na base de imagens para alcançar uma alta resolução, o que pode melhorar os resultados da contagem.

No trabalho de Pyo *et al* (2020), do mesmo grupo de pesquisa do trabalho anterior, os autores utilizaram redes neurais convolucionais para predição de concentrações de cianobactérias em rio. Utilizaram um modelo de EFDC

(Environmental Fluid Dynamics Code), que é um modelo hidrodinâmico tridimensional de qualidade da água, que possui diversas características, como transporte de sedimentos, qualidade da água e transporte de sedimentos tóxicos. A saída deste modelo pode gerar imagens e desta forma, o estudo utilizou imagens temporais de alta frequência, com dados sintéticos, como um modelo 3D. O objetivo era investigar a capacidade de uma rede neural convolucional prever a proliferação de cianobactérias prejudiciais. A CNN obteve uma NSE (*Nash-Sutcliffe Efficiency*, métrica de performance) de 0,84. Adicionando ruído aos dados, a acurácia diminuiu para 0,76. Segundo os autores, o estudo demonstrou a capacidade da CNN para predição das concentrações de cianobactérias usando imagens temporais de alta frequência.

Thiel e Wiltshire (1995) desenvolveram uma metodologia para identificar cianobactérias por meio de imagens digitais. Eles utilizaram uma base com imagens de lâminas de microscópio, com amostras de água contendo cianobactérias, com um total de 31 amostras. As amostras eram provenientes de lagos e reservatórios da Grã-Bretanha. As imagens passaram por várias técnicas de processamento de imagens, analisando forma e textura. Por fim, eles obtiveram uma precisão superior à 90% para *Anabaena* e *Oscillatoria*. Os autores propunham à época, acrescentar mais cianobactérias para identificação.

Walsby e Avery (1996) propuseram um método de análise de imagem para medir o comprimento total dos filamentos de cianobactérias coletadas em filtros de membrana de amostras de populações naturais ou culturas. As imagens eram visualizadas por microscopia de epifluorescência e foram digitalizadas. Em seguida, foram analisadas com um programa de análise de imagens. O comprimento total dos filamentos, considerando correção de erros decorrentes de orientação, cruzamento e sobreposição, são então determinados. Para os autores, o sistema fornece um método preciso e rápido para determinar o comprimento total do filamento em um campo de visão.

Gandola *et al* (2016) desenvolveram um método automático para a identificação de cinco gêneros de cianobactérias. Eles realizaram a aquisição de 100 imagens com uma câmera acoplada a um microscópio, provenientes de um lago. Então fizeram o pré-processamento para reduzir ruídos e eliminar outros organismos que não eram interessantes e, com informações morfológicas desenvolveram um algoritmo para reconhecer filamentos cruzados ou “quebrados”, e para extrair o

padrão das informações. Usaram então 17 indicadores de padrões específicos como entrada de um algoritmo de aprendizado de máquina, específicos para a identificação das cinco espécies de cianobactérias tóxicas, com estruturas filamentosas, isto é, “compridas” (*Aphanizomenon*, *Cylindrospermopsis*, *Dolichospermum*, *Limnothrix* e *Planktothrix*). O método foi validado comparando as leituras do sistema com manuais, o que mostrou que o sistema, chamado ACQUA, é eficaz.

Zeder *et al* (2009) desenvolveram uma técnica baseada em microscopia de fluorescência para quantificar e medir filamentos não ramificados de *Planktothrix* e outras cianobactérias, desde que sejam morfológicamente semelhantes. O método usou simulação de Monte Carlo para estimar o tamanho e a densidade do filamento fora do campo de visão e, em seguida, criaram o *software* para a estimativa em um microscópio automatizado. É capaz, depois da detecção, de contar os filamentos e medi-los individualmente ou quando conectados. Os autores usaram uma abordagem de análise de imagem orientada a objeto baseada em modelo. O método foi validado comparando-o com procedimentos manuais, ou seja, centenas de filamentos de sete imagens foram medidos manualmente para mensuração posterior. A precisão da validação para eventos de sobreposição de filamentos foi de 86% e 55% para filamentos paralelos. Apesar de algumas discrepâncias, como quando aplicado em imagens de filamentos ramificados, o que gera resultados errôneos, o sistema se mostrou eficaz para os autores, que também destacam a economia de tempo.

Os autores Mosleh *et al* (2012) realizaram um estudo preliminar para desenvolver um método para reconhecer cinco gêneros de algas e cianobactérias (Bacillariophyta (*Navicula*), Chlorophyta (*Scenedesmus*) e Cyanobacteria (*Chroococcus*, *Microcystis* e *Oscillatoria*)). Eles usaram técnicas de processamento de imagens, incluindo pré-processamento para remoção de ruído, segmentação com detecção de borda, extração de características de forma e textura de algas selecionadas (como forma, área, perímetro, eixos menores e maiores e o espectro de Fourier com componentes principais de análise (PCA) e, finalmente, reconhecimento com redes neurais artificiais, que obtiveram 93% de acerto para 100 imagens utilizadas nos testes preliminares.

Apesar dos trabalhos citados apresentarem alguns resultados efetivos para um número muito limitado de casos e um domínio de aplicação específico, nenhum

deles apresenta uma abordagem robusta para identificar e contar células de cianobactérias. Neste trabalho apresenta-se uma base rotulada de cianobactérias da espécie *Microcystis aeruginosa*, e um exemplo de utilização que consiste na identificação e contagem de suas células em imagens de lâminas de microscópio, utilizando as recentes Redes Neurais Convolucionais (CNN). A riqueza do conteúdo revisado é de grande importância para o desenvolvimento do trabalho, auxiliando nos diversos aspectos que abrangem a factibilidade da proposta.

4 MATERIAIS E MÉTODOS

Esta seção está dividida em duas partes distintas. Na primeira, são explanados os passos realizados para a formulação da base de imagens, sendo esta dividida ainda em aquisição e rotulação. Na segunda, são detalhados os processos de aplicação de *deep learning*, dividindo a etapa em classificação e em contagem de células de *Microcystis aeruginosa*.

4.1 FORMULAÇÃO DA BASE

Nesta seção são explanadas as etapas que levaram a formulação da base de dados.

4.1.1 Aquisição das Imagens

As amostras desta pesquisa foram provenientes de reservatórios de Curitiba e região metropolitana, no estado do Paraná. Algumas delas, foram cedidas pela companhia de abastecimento de água do estado, e outras foram cedidas pelo departamento de Botânica da Universidade Federal do Paraná.

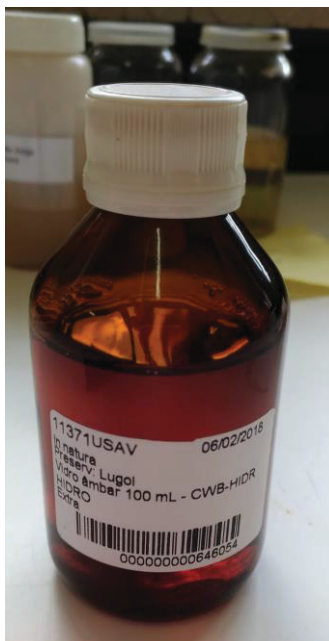
O primeiro passo foi o recolhimento da amostra de água bruta nos reservatórios, nos pontos de captação onde havia ou não floração de cianobactérias de modo a obter duas classes para comparação. Várias coletas foram feitas em diferentes locais. Esta ação foi executada por pessoas responsáveis da companhia de saneamento, bem como da UFPR. Em algumas coletas, a autora participou do processo para entender melhor a dinâmica da captação das amostras.

Para a diversificação e o enriquecimento da base, as amostras são provenientes de mananciais e reservatórios situados nas áreas de Piraquara, Iraí, Foz do Areia, Iguaçu e São José dos Pinhais.

Posteriormente, as amostras foram manipuladas, reduzidas em volume e colocadas em vidros de cor âmbar, com lugol (solução para fixação e conservação da amostra), que foi então rotulado para identificação. A FIGURA 12 mostra um dos

vidros identificados. Este processo é baseado nas normas da Cetesb, de manipulação de amostras de fitoplâncton.

FIGURA 12 – Vidro de cor âmbar com amostra de água bruta manipulada e identificada.



FONTE: A autora (2020).

Em seguida, a amostra foi levada ao laboratório de Ficologia da Universidade Federal do Paraná, para dar prosseguimento ao processo de aquisição das imagens.

A partir desta amostra, foram preparadas várias lâminas para sua posterior análise no microscópio. A FIGURA 13 mostra o início da preparação de duas lâminas.

FIGURA 13 – Preparação de lâminas de microscópio



FONTE: A autora (2018).

As lâminas preparadas foram então analisadas no microscópio, para averiguação da presença da cianobactéria, bem como para realizar a aquisição das imagens padronizadas.

O microscópio utilizado é da marca Olympus, com lente modelo UPlanFL N, de aumento de 100x, chamada de lente de imersão e foi utilizado um óleo de imersão na lâmina. Em microscopia ótica, quando o aumento da lente é superior à 40x, é necessário utilizar uma interface líquida entre a lâmina e a lente, para que o índice de refração seja igual, tanto na lâmina quanto na lente.

As imagens foram capturadas com a câmera do microscópio, modelo Olympus DP-71, com auxílio de programa proprietário da própria Olympus (DP-BSW Ver.3.1), no tamanho de 1360x1024 pixels, resolução de 12,5 megapixels e 144 DPI, velocidade de ISO de 800. Com estes valores, obtém-se o tamanho do pixel de aproximadamente 80 micrômetros.

Além de capturar imagens das amostras com florações de cianobactérias, foram também capturadas imagens do fitoplâncton em geral, contendo outras algas e organismos que ocorrem comumente na água de reservatórios e mananciais. Os mesmos parâmetros foram estabelecidos para ambas as amostras.

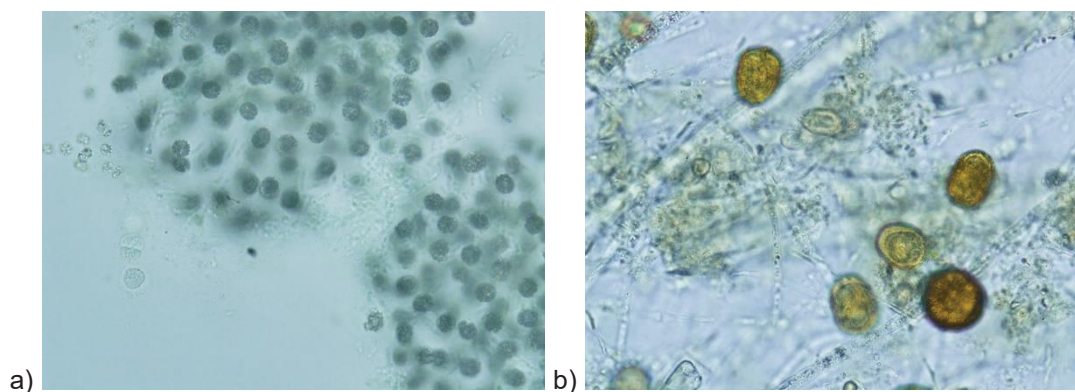
Cada lâmina gerava algumas dezenas de imagens, por processo de varredura, linha a linha. Uma imagem era capturada, e a lâmina era direcionada para o próximo campo de visão, horizontalmente. Ao final da lâmina, pulava-se uma linha, observando para que as células de uma imagem não estivessem na imagem seguinte. O foco também era ajustado, de modo a deixar as imagens das células, se se sobrepunham umas às outras, o mais nítido possível, conforme orientação do especialista.

O processo de aquisição resultou em 2.265 imagens de floração de *Microcystis aeruginosa*, e 1.657 imagens de organismos diversos no fitoplâncton, totalizando 3.922 imagens.

4.1.2 Classificação e Rotulação da Base

A classificação primária foi separar as imagens em duas classes: *Microcystis aeruginosa* e Outros, onde a classe Outros se refere às imagens do fitoplâncton onde não há *Microcystis aeruginosa*, como apresentado na FIGURA 14.

FIGURA 14 – a) Classe *Microcystis aeruginosa* b) Classe Outros



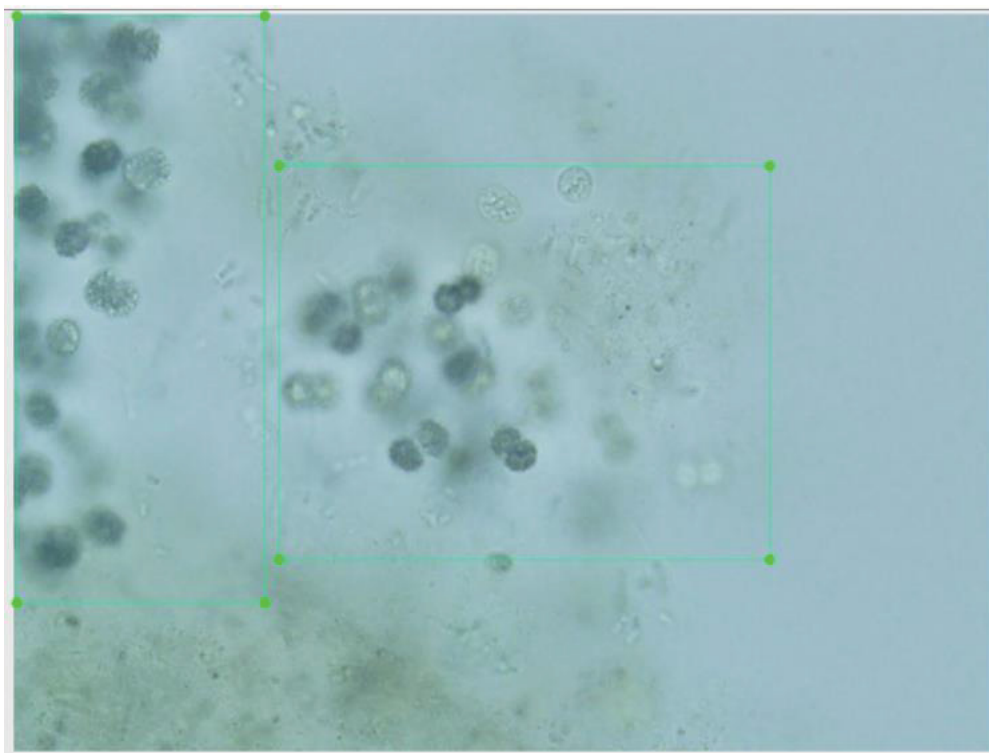
Fonte: A autora (2020)

Para realizar a contagem por meio de detecção de objetos, é necessário que a base esteja anotada/rotulada, isto é, que esteja marcada e identificada. Assim, a rotulação da base se deu com o auxílio da ferramenta *Labellmg*⁸. Este *software* permite que se criem caixas retangulares nos locais desejados, e então se rotule a

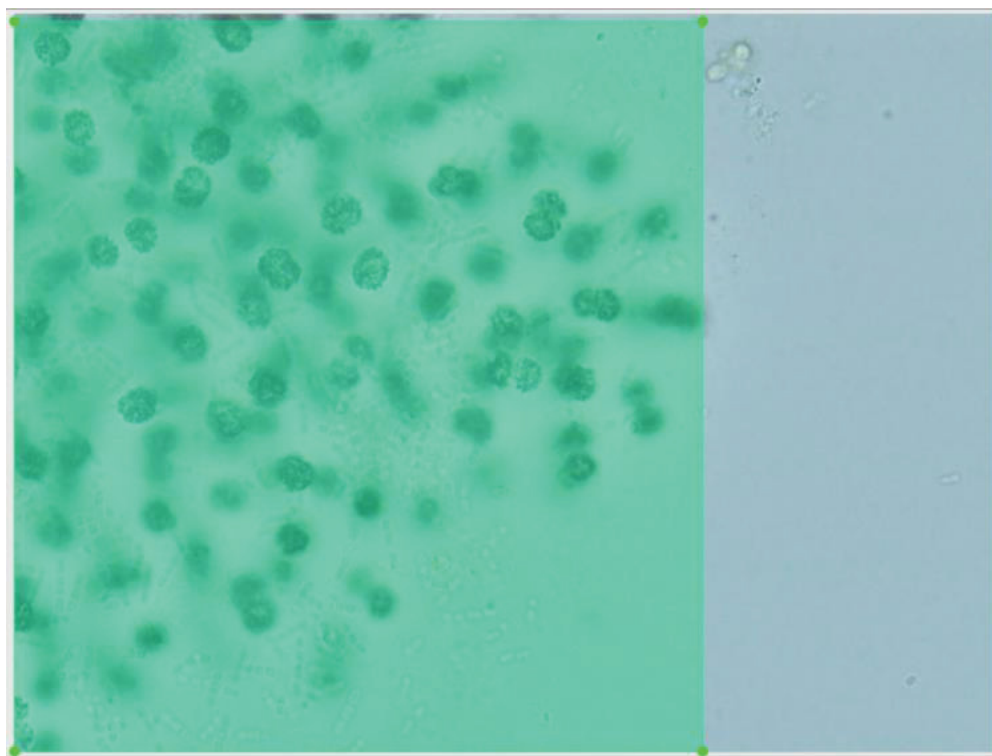
8 *Labellmg*. <https://tzutalin.github.io/labellmg/>

região selecionada. Posteriormente, ele cria um arquivo de extensão XML com as coordenadas e a classificação do objeto na imagem. Inicialmente foi executado um processo de rotulação de colônia de *Microcystis*, criando uma demarcação sobre toda a colônia, como um quadrilátero de linha cor verde, delimitando a região da colônia, como pode ser observado nas FIGURAS 15 e 16.

FIGURA 15 – Rotulação de Colônia de *Microcystis* 1



Fonte: A autora (2020).

FIGURA 16 – Rotulação de Colônia de *Microcystis* 2

Fonte: A autora (2020).

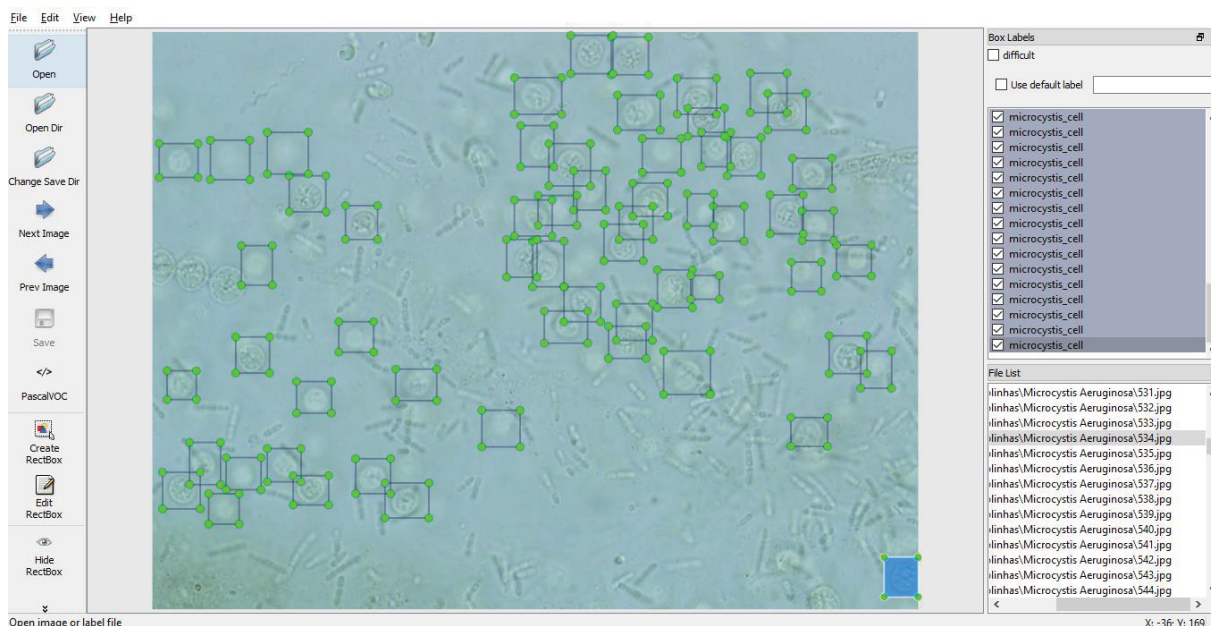
Esta marcação da colônia inteira posteriormente foi substituída pela rotulação das células de *Microcystis* de forma individual, em todas as imagens de floração.

A rotulação individual deu-se criando o retângulo de marcação de modo a abranger toda a área da célula, isto é, sendo possível identificar sua borda, de forma que o retângulo ficasse o mais próximo possível do tamanho da célula. Para isto, utilizou-se o recurso de *zoom* durante o processo de rotulação. Também evitou-se rotular apenas partes de células que apareciam nas bordas das imagens, cuja identificação não estava clara.

A FIGURA 17 exhibe a ferramenta *LabelImg* e o processo de rotulação de uma das imagens. Na janela da imagem, podem-se observar os recursos que a ferramenta oferece, como abrir e alterar a pasta onde as imagens são salvas, criar e editar os retângulos, entre outros. Basicamente, escolhe-se a pasta onde estão armazenadas as imagens, escolhe-se um diretório para que os arquivos XML gerados sejam guardados, e então inicia-se o processo de rotulação de cada célula. Busca-se abrangê-la totalmente, de modo a evitar que o retângulo se exceda muito

além de seu tamanho, bem como evitando pedaços de células nas bordas onde não era possível a identificação.

FIGURA 17 – Processo de Rotulação



FONTE: A autora (2020).

Este processo de rotulação se deu de forma inteiramente manual, célula a célula, com a orientação de especialistas, cujas indicações foram para marcar todas as células visivelmente distinguíveis, considerando o foco e os cortes sofridos devido ao campo de visão que a imagem abrange. Assim, todas as células onde é possível identificar sua estrutura, foram anotadas.

4.1.3 Processo de *Data Augmentation*

Quanto mais dados um algoritmo de aprendizado de máquina e aprendizado profundo tiver na etapa de treinamento, mais eficaz ele pode ser. Mesmo quando os dados são de qualidade inferior, os algoritmos podem realmente ter um desempenho melhor, desde que dados úteis possam ser extraídos pelo modelo do conjunto de dados original. Utilizar a técnica de *data augmentation* é uma maneira de reduzir o *overfitting* nos modelos, de modo que se aumenta a quantidade de dados de

treinamento utilizados pelos algoritmos. Uma prática corrente e bem aceita para aumentar os dados da imagem é realizar alterações geométricas e de cores, como espelhar a imagem, cortar e rotacionar e alterar a paleta de cores da imagem (PEREZ e WANG, 2017).

As alterações escolhidas para aumentar a base de imagens foram o espelhamento vertical e o espelhamento horizontal, e isto totalizou 6.795 imagens de *Microcystis aeruginosa*. A técnica foi também aplicada à base das imagens correspondentes ao fitoplâncton, totalizando 4.971 imagens, entretanto, na etapa de *object detection*, esta base não foi utilizada, pois não haviam células de *Microcystis* propositalmente. Vale ressaltar, que o processo de *data augmentation* para a etapa de *object detection* foi realizado nas imagens juntamente com seus rótulos.

4.2 APLICAÇÃO DE DEEP LEARNING

Nesta seção, as atividades relacionadas à aplicação de *deep learning* do trabalho, são apresentadas.

4.2.1 O processo de Classificação

Este processo consiste em classificar se as imagens contêm ou não *Microcystis aeruginosa*. Vale lembrar que para esta etapa, a base utilizada ainda não estava com as células individualmente rotuladas, isto é, foram utilizadas as imagens originais com a cianobactéria, e as imagens sem a cianobactéria.

Para detectar imagens com potencial presença de *Microcystis aeruginosa*, foi utilizado o *framework* de aprendizado profundo *TensorFlow* (desenvolvido e mantido pelo Google). Duas redes neurais foram escolhidas, *Mobilenet_v1* (HOWARD *et al*, 2017) e *Inception_v3* (SZEGEDY *et al*, 2016) (aqui também chamadas simplesmente de *Mobilenet* e *Inception*), ambas utilizadas para determinar a existência de objetos na cena.

Em suma, *Mobilenet* é uma rede otimizada para dispositivos móveis, o que pode reduzir ligeiramente a precisão nas previsões. A rede é uma rede otimizada com recursos que ajudam a evitar *overfitting*, o que compromete a capacidade de generalização das previsões. Com o objetivo de reduzir o tempo de treinamento e

melhorar os resultados, foi utilizada a técnica *transfer learning* (ZHUANG *et al*, 2020), que seleciona uma parte de um modelo pré-treinado em uma tarefa relacionada e a reutiliza em um novo modelo. As redes utilizadas foram previamente treinadas com *ImageNet*, uma base de imagens onde existem mais de 14 milhões de imagens rotuladas em mais de 20 mil categorias.

Para o treinamento, foi utilizada a base de imagens aumentada descrita na seção 4.1.3. A base de imagens foi dividida e rotulada em duas categorias distintas: aquelas que continham *Microcystis* e as que não continham, respectivamente rotuladas como “Microcystis” e “Outros”. A resolução das imagens utilizadas no treinamento foi ajustada em função da solicitação de entrada de cada modelo, resultando em 224x224px para *Mobilenet* e 229x229px para *Inception*.

Os parâmetros de treinamento foram: 400 períodos de treinamento; taxa de aprendizagem de 0,01 e fração de 0,50 do tamanho do modelo. Vale ressaltar que o processo de validação é automático e aleatório, feito pela rede, que escolhe as imagens a cada etapa para validação.

Todos os modelos foram treinados sem GPU (*Graphics Processing Unit*) em um computador com Intel Core i7, 2,9Ghz, 8GB de RAM e Windows 10 de 64bits como sistema operacional.

Foram realizados testes de treinamento com 200, 300, 400 e 500 épocas para ambas as redes. A curva da taxa de aprendizagem foi utilizada como parâmetro para analisar a configuração da rede. Segundo James *et al* (2013), ela é uma métrica que pode ser utilizada para evitar eventualmente o problema de *overfitting* durante o treinamento e validação dos modelos. Deste modo, ao final, 400 épocas foram selecionadas, uma vez que o *Mobilenet* apresentou maior instabilidade, tendendo a uma possível estabilização a partir de 300 épocas, enquanto a estabilidade foi observada a partir da curva de taxa de aprendizagem em *Inception_v3* já a partir de 200 épocas (FIGURA 18) durante o treinamento simples (sem *cross validation*).

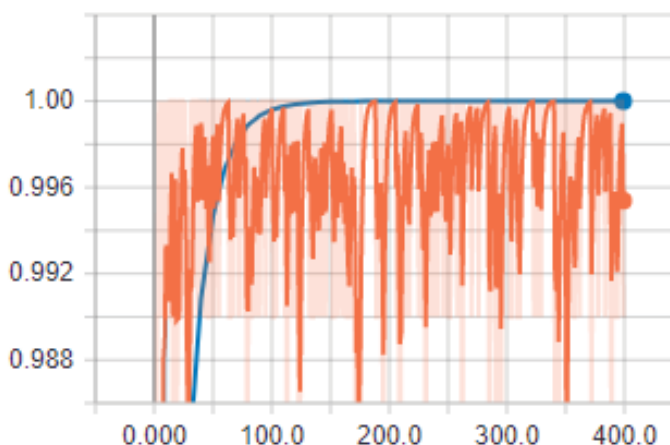
Em média, cada modelo levou 2 dias para treinar. Dois tipos de treinamento foram realizados com as duas redes: treinamento com todas as imagens, exceto 20% separadas para teste (aqui referido como treinamento simples ou modelo simples), e treinamento com validação cruzada (aqui referido como treinamento com validação cruzada ou modelo com validação cruzada).

O treinamento de validação cruzada foi realizado para entender o aprendizado do modelo e verificar se não há *overfitting*. Ele consistiu em separar o conjunto total de imagens em 5 partes aleatoriamente, chamados de *folds*, sendo que uma das partes (20%) é destinada ao futuro teste. Assim, foram realizados 5 treinamentos, de modo que cada treinamento é realizado com uma parcela diferente das imagens nos 4 *folds* e treinado em no *fold* 5, gerando uma acurácia diferente para cada um deles. Assim, a acurácia é a média das 5 acurácias obtidas. A validação cruzada é amplamente utilizada com bancos de dados menores, a fim de também evitar *overfitting*. A ideia de utilizá-lo nesta proposta, além de evitar *overfitting*, é verificar como o modelo se comportaria em cenários diversos. Vale ressaltar que a validação pós treino do modelo é realizada de forma automática.

Como mencionado, para verificar a eficácia da predição das redes treinadas em ambas as formas, 20% da base foi reservada para verificação posterior nos testes, ou seja, essa porção das imagens não foi incluída no conjunto de treinamento.

FIGURA 18 – Curva de aprendizado de *Inception_v3*

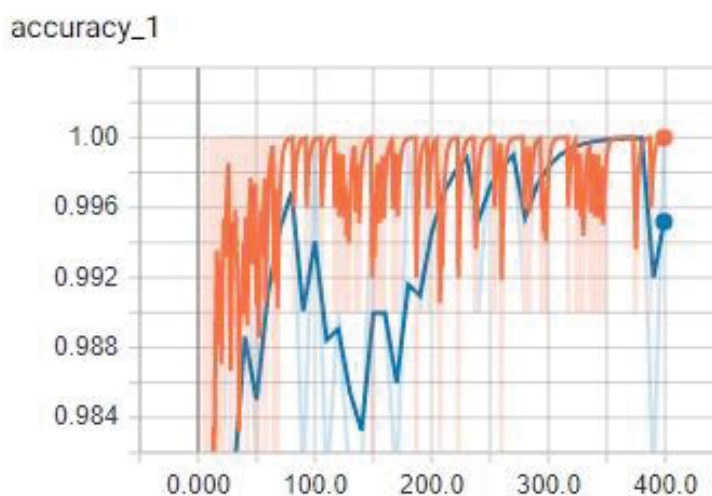
accuracy_1



FONTE: A autora (2020).

LEGENDA: A linha azul se refere a validação, e a laranja se refere ao treinamento. O eixo x se refere às épocas.

A curva de aprendizado para treinamento com a rede *Mobilenet*, em treinamento simples é mostrada na FIGURA 19.

FIGURA 19 – Curva de aprendizado de *Mobilenet*

FONTE: A autora (2020).

LEGENDA: A linha azul se refere a validação, e a laranja se refere ao treinamento.

4.2.2 O Processo de Detecção de Objeto e Contagem de Células

A detecção de objetos em imagens se difere do processo de classificação de imagens. No processo de classificação, apresentado no item anterior, o objetivo é classificar uma imagem no que se refere a ela conter ou não a cianobactéria. No processo de detecção de objetos, a cianobactéria é localizada e também classificada.

De acordo com Liu *et al* (2020), um dos problemas mais fundamentais e desafiadores da visão computacional, a detecção de objetos, tem por objetivo determinar se existem quaisquer instâncias de objetos de determinadas categorias (como humanos, carros, bicicletas, cães ou gatos) em uma imagem e, se houver, retornar a localização espacial e a extensão de cada instância de objeto.

No *Tensorflow*, há uma ferramenta de código aberto chamada *Object Detecion API*⁹ (*Aplication Programming Language*) específica para esta tarefa. Ela suporta uma variedade de arquiteturas e modelos para tarefas de treinamento e reconhecimento de imagens.

9 TensorflowObject Detecion API. https://github.com/tensorflow/models/tree/master/research/object_detection

Os modelos escolhidos foram os mesmos utilizados para a classificação, isto é, *Mobilenet* e *Inception*. Adicionalmente, com o objetivo de mostrar a potencialidade de utilização para pesquisas da base de imagens proposta, é relevante aplicar a rede neural convolucional que representa o estado-da-arte na área de detecção de objetos, a rede *Faster-RCNN* (*Fast Region-based Convolutional Network*) (REN *et al*, 2017), bem como a rede convolucional YOLO, estado-da-arte em detecção de objetos em sistemas de tempo real.

Para a rede *Mobilenet*, foram empregados os seguintes parâmetros de treinamento: resolução das imagens 224 x 224px, 500 épocas de treinamento, taxa de aprendizado inicial de 0,004, um *batch* de tamanho 24 e taxa de *momentum* de 0,9. Já para a *Inception*, foram os seguintes: resolução das imagens 300 x 300 px, 500 épocas de treinamento, taxa de aprendizado inicial de 0,004, um *batch* de tamanho 24 e taxa de *momentum* também de 0,9. Vale ressaltar que o *batch size* ou tamanho do lote se refere ao número de exemplos de treinamento utilizados em uma iteração, e termo *momentum* quantifica o grau de importância da variação de peso do ciclo anterior ao ciclo atual, assim, o cálculo para o ajuste dos pesos torna o aprendizado mais estável e acelera a convergência (RUSSELL e NORVIG, 2004).

Os parâmetros de treinamento da rede *Faster – RCNN* foram: resolução das imagens: 600 x 1024 px; 500 épocas de treinamento, taxa de aprendizado inicial de 0,0003, tamanho de lote 1, e *momentum* de 0,9. Este treinamento também se deu com a base aumentada.

Um experimento adicional foi realizado, executando o treinamento com os mesmos parâmetros para as três redes, entretanto, alterando a quantidade de épocas para 3000. O objetivo é explorar diferentes cenários de treinamento.

Já para o treinamento da rede YOLO, as características foram diferentes, devido às particularidades da rede. O treinamento foi realizado clonando-se o repositório com as bibliotecas *python* necessárias, e não na ferramenta *TensorFlow*. Foram utilizadas as mesmas imagens anotadas, entretanto, foi necessário rodar um script específico para deixar as anotações das imagens de treinamento no formato da YOLO. Após isto, foi realizado o *download* da rede pré-treinada *darknet*, e então foi possível realizar o treinamento da rede com as imagens da base. Após o treino, foram realizados os testes, sendo considerada a mesma quantidade das demais redes (20%). Os critérios do treinamento foram padrão, sendo 51 épocas de

treinamento, com um *batch size* (número de elementos treinados após uma iteração) de 32.

5 RESULTADOS

O capítulo 5 traz todos os resultados obtidos após o desenvolvimento de cada etapa envolvida no trabalho.

5.1 RESULTADOS DA MONTAGEM DA BASE

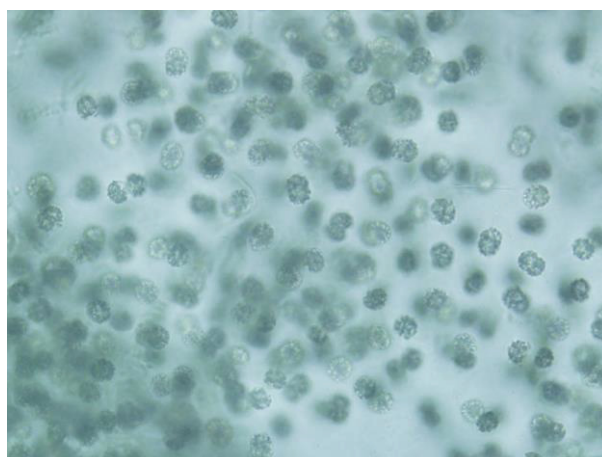
Nesta seção é possível observar as imagens resultantes dos processos de aquisição e processamento das imagens, para a montagem da base.

5.1.1 Imagens Originais

Como resultado do processo de aquisição de imagens, as figuras aqui apresentadas mostram a diversidade das amostras coletadas. Conforme mencionado na seção 4.1.1, foram obtidas 2.265 imagens de floração de *Microcystis aeruginosa*, e 1.657 imagens de organismos diversos no fitoplâncton, um total de 3.922 imagens.

A FIGURA 20 exibe um exemplo de variedade de *Microcystis aeruginosa*, assim como as FIGURAS 21 até a 25. A partir da FIGURA 26 até a 30, são exibidas variedades de organismos presentes no fitoplâncton, que não contém *Microcystis aeruginosa*.

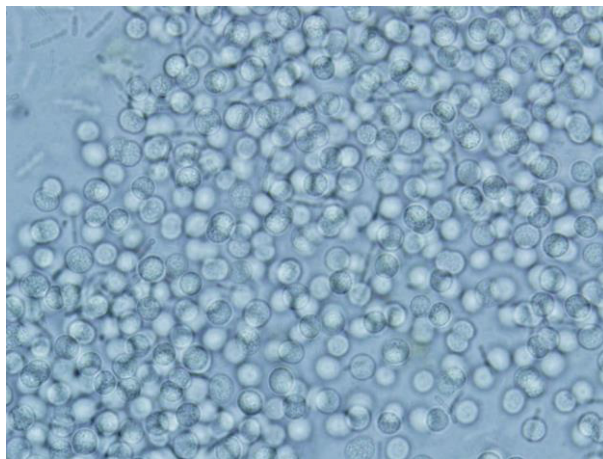
FIGURA 20 – Variedade de *Microcystis aeruginosa* 1



FONTE: A autora (2020).

Na FIGURA 21, pode-se observar as pseudo-anabaenas, que são os pequenos filamentos na imagem. Estes organismos são presença característica da mucilagem das *Microcystis*.

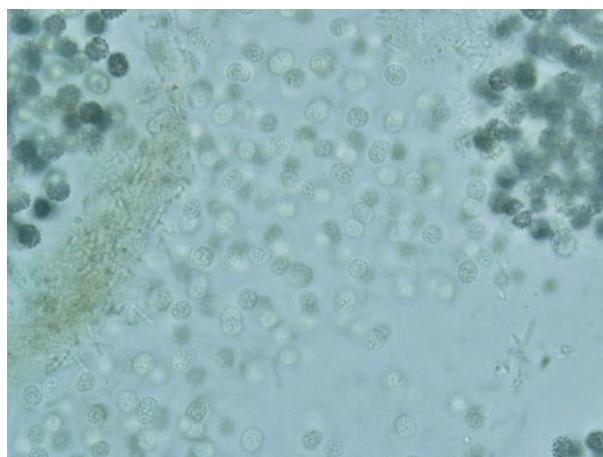
FIGURA 21 – Variedade de *Microcystis aeruginosa* 2



FONTE: A autora (2020).

A FIGURA 22 apresenta *Microcystis* de duas cores, entretanto, segundo os especialistas, ambas são da espécie *aeruginosa*, podendo a mudança de cor ser ocasionada pelas substâncias aplicadas nas amostras, a fim de preservá-las.

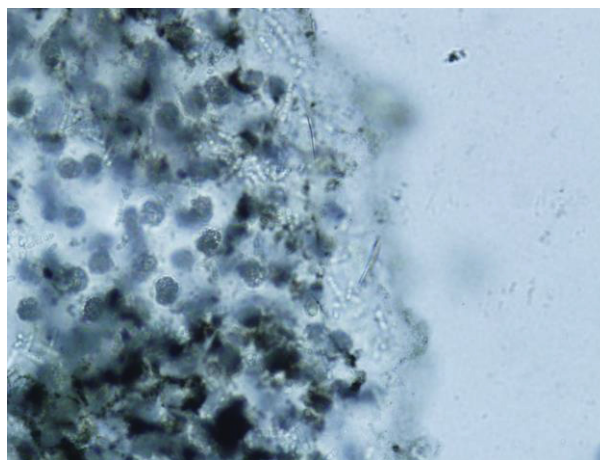
FIGURA 22 – Variedade de *Microcystis aeruginosa* 3



FONTE: A autora (2020).

Na FIGURA 23, aparecem células mais escuras. Isto se deve à utilização de um corante chamado Nankin, aplicado à algumas das amostras cedidas ao trabalho. Este corante que serve para que os organismos fiquem mais evidentes, prática comum no manuseio de amostras de fitoplâncton. Também é possível observar nesta imagem a borda repleta de pseudo-anabaenas, bem como alta densidade de células da cianobactéria.

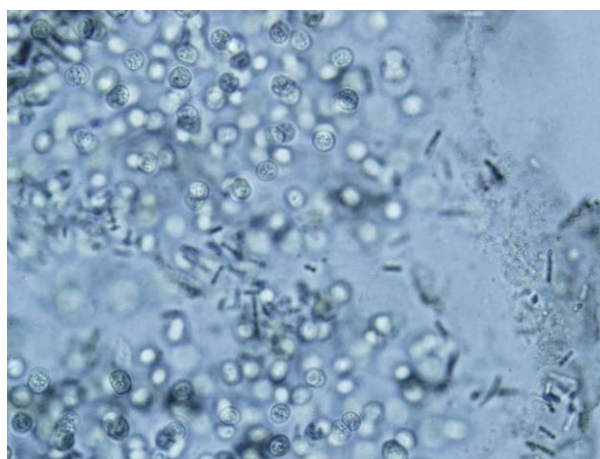
FIGURA 23 – Variedade de *Microcystis aeruginosa* 4



FONTE: A autora (2020).

Na FIGURA 24, é possível ver células nítidas e não tão nítidas. Isto se deve ao fato da mudança de calibração do foco, para os níveis onde se encontram as células. Devido a esta sobreposição das células, é necessário regular o foco para visualizá-las.

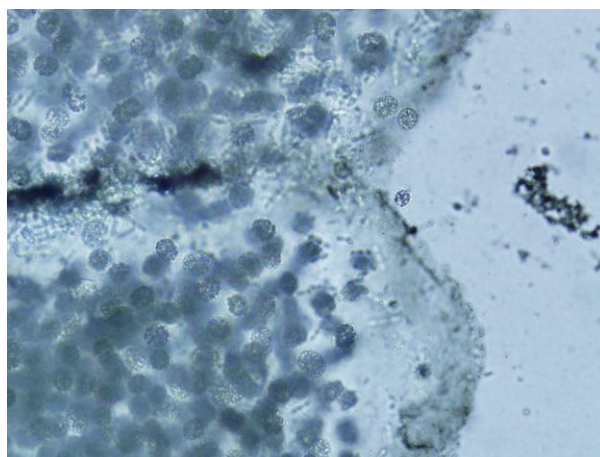
FIGURA 24 – Variedade de *Microcystis aeruginosa* 5



FONTE: A autora (2020).

Na FIGURA 25 fica evidente a alta densidade da amostra, pois se vê muitas células aglomeradas, o que torna difícil sua visualização e análise.

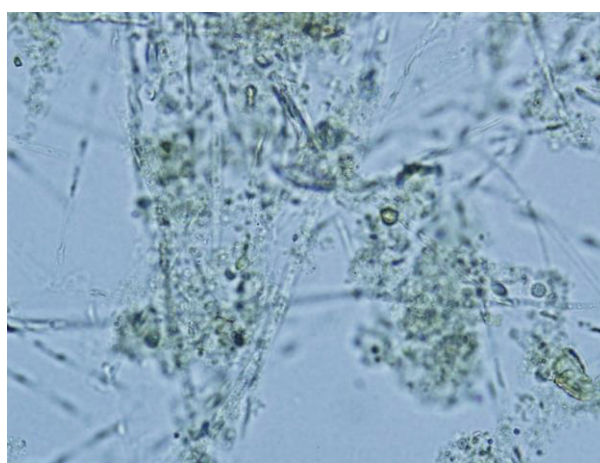
FIGURA 25 – Variedade de *Microcystis aeruginosa* 6



FONTE: A autora (2020).

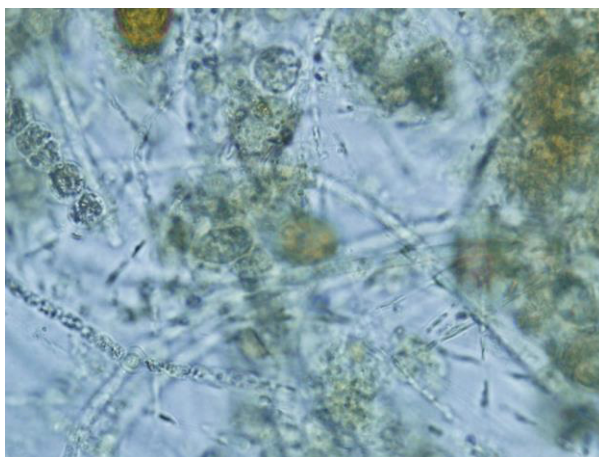
As FIGURAS 26 até a 29 exibem diferentes organismos em uma lâmina, incluindo diatomáceas, algas, sedimentos, outras cianobactérias e outros organismos.

FIGURA 26 – Variedade de Fitoplâncton 1



FONTE: A autora (2020).

FIGURA 27 – Variedade de Fitoplâncton 2



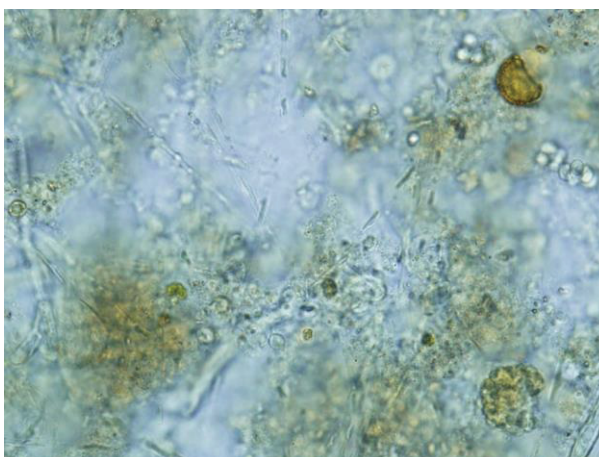
FONTE: A autora (2020).

FIGURA 28 – Variedade de Fitoplâncton 3



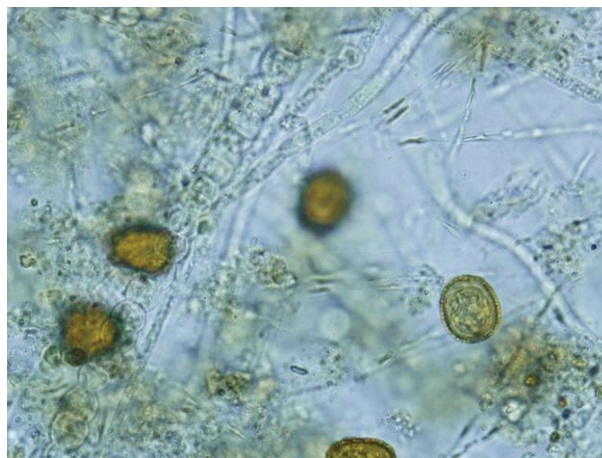
FONTE: A autora (2020).

FIGURA 29 – Variedade de Fitoplâncton 4



FONTE: A autora (2020).

FIGURA 30 – Variedade de Fitoplâncton 5



FONTE: A autora (2020).

5.1.2 Imagens Rotuladas

Os rótulos criados nas imagens são visíveis por meio da ferramenta *Labelimg*, e o resultado da rotulação para fins práticos, é um arquivo XML, contendo todas as coordenadas de cada retângulo criado para a marcação de cada célula.

Importante ressaltar que o processo de rotulação teve acompanhamento e validação de especialista.

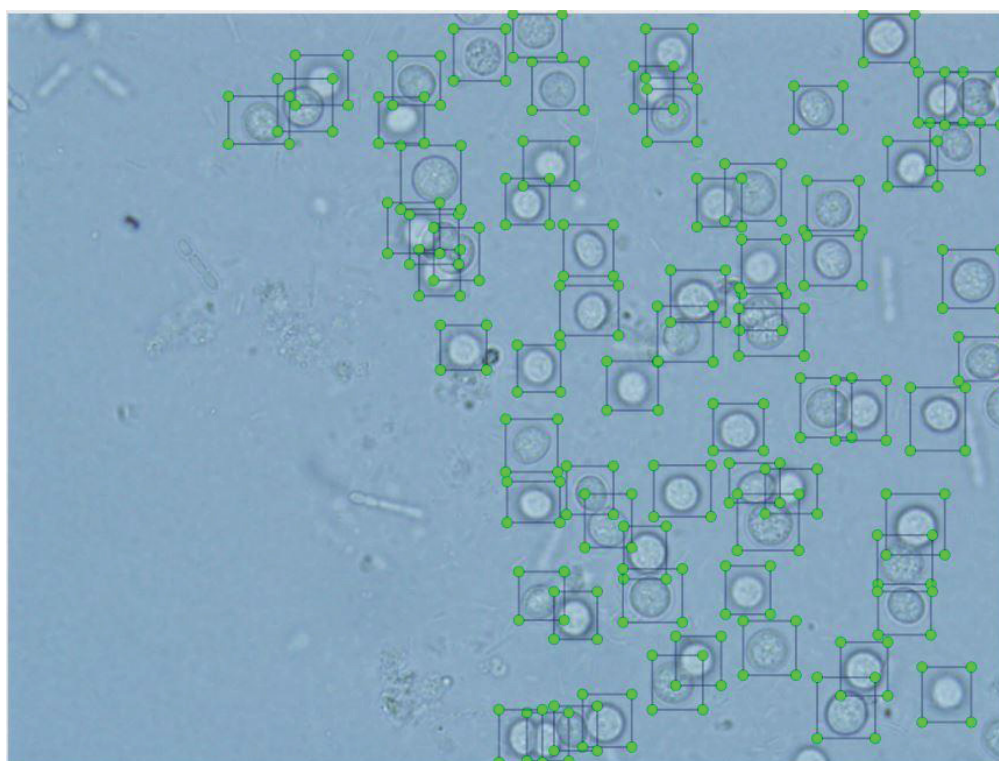
Para a aplicação prática, os métodos computacionais se utilizam da imagem original juntamente com seu arquivo XML correspondente, com as marcações feitas, para realizar a localização das células e então executar a tarefa para o qual está sendo programado.

Para fins ilustrativos as FIGURAS 31 a 35 mostram exemplos do resultado da rotulação das imagens.

É possível observar nas imagens que em cada célula há um quadrilátero envolvendo-a completamente, e de tamanho justo, isto é, próximo do tamanho da célula. Também se nota que células que aparecem apenas parcialmente nas bordas das imagens não possuem a marcação com o quadrilátero. A ideia aqui é não rotular células onde não seja possível identificar características da *Microcystis aeruginosa*, para evitar que seja confundida com outros organismos. Isto pode acontecer, por exemplo, se uma célula que se encontra na borda tenha uma parte muito pequena

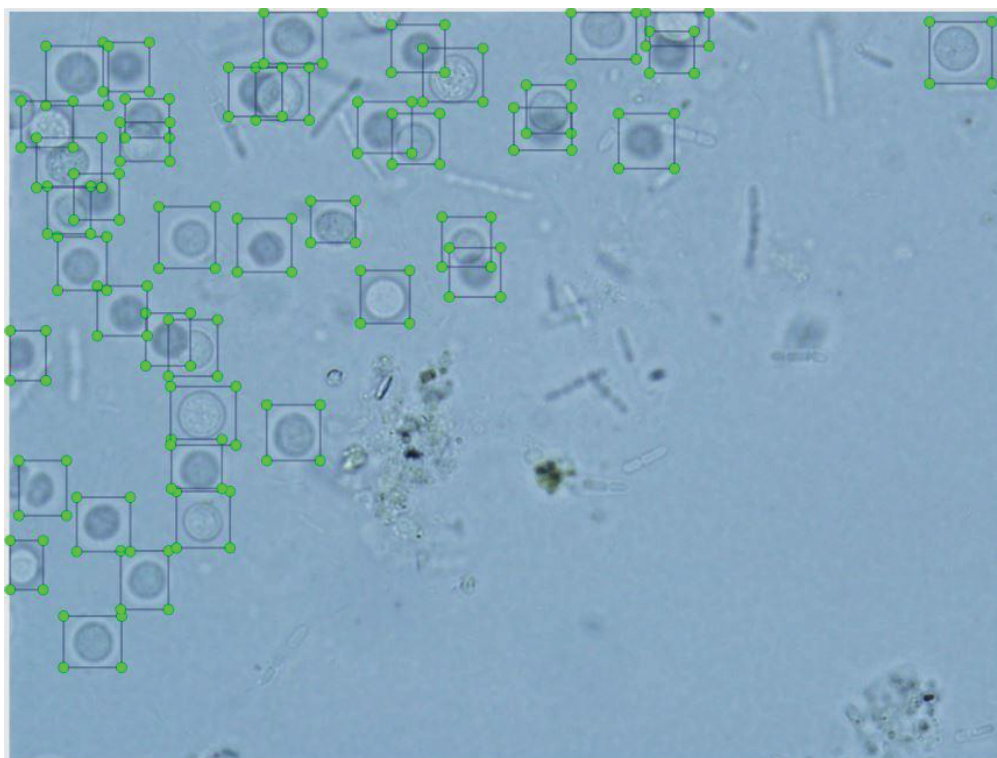
visível. Assim, cada quadrilátero corresponde a uma célula de *Microcystis aeruginosa* e foi rotulada desta forma.

FIGURA 31 – Exemplo de Imagem Rotulada 1



FONTE: A autora (2020).

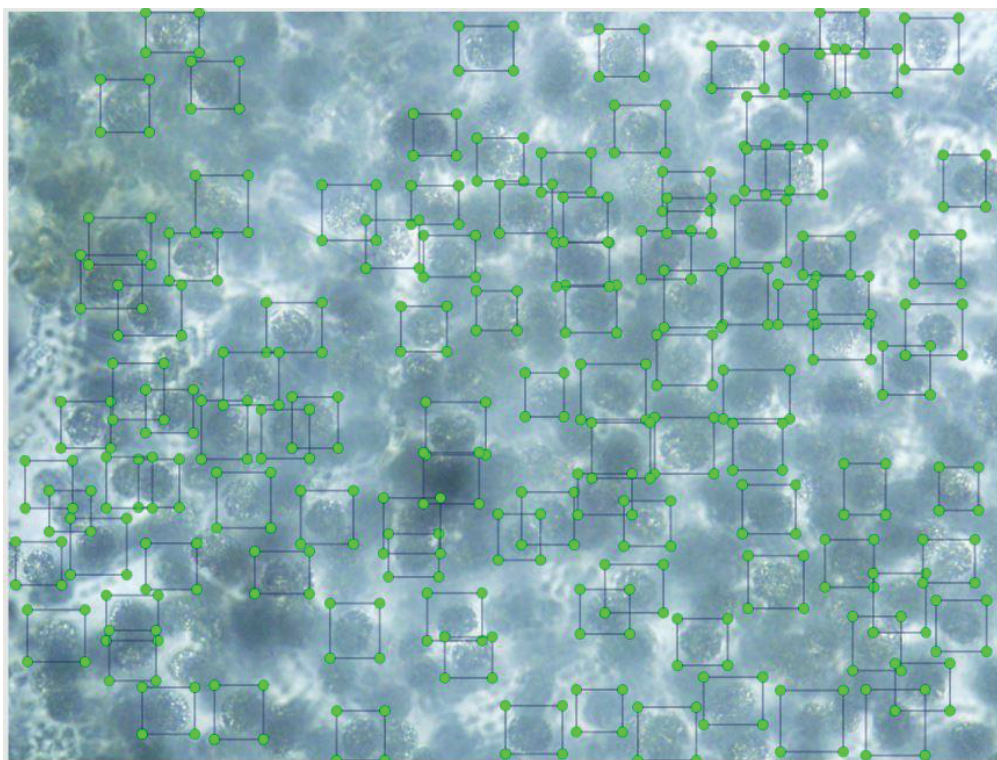
FIGURA 32 – Exemplo de Imagem Rotulada 2



FONTE: A autora (2020).

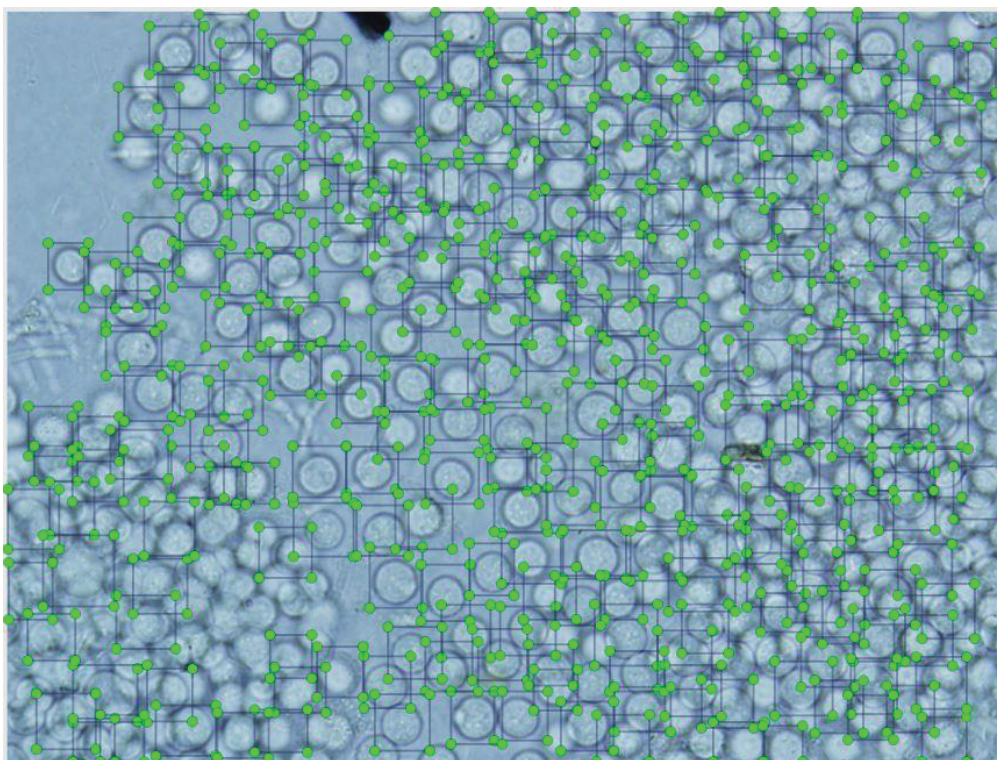
Na FIGURA 33 observa-se sua densidade, e conforme orientações de especialistas, apenas as células com seus contornos visíveis são contadas. Quanto ao restante onde não se pode identificar células individuais, são feitas estimativas de sua quantidade.

FIGURA 33 – Exemplo de Imagem Rotulada 3



FONTE: A autora (2020).

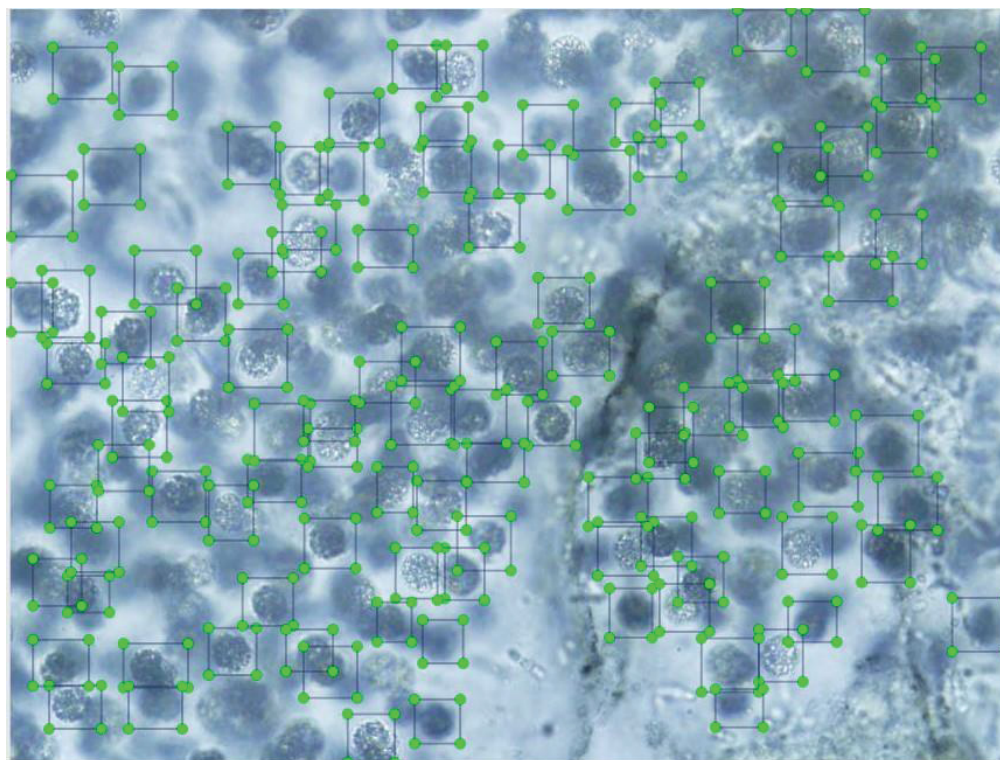
FIGURA 34 – Exemplo de Imagem Rotulada 4



FONTE: A autora (2020).

Também na FIGURA 35 é possível observar a densidade da amostra, o que torna o trabalho do especialista uma tarefa difícil para identificação e especialmente contagem destas células. Do mesmo modo, é possível observar o grande desafio de rotular todas as células de todas as imagens da maneira mais próxima da realidade exercida pelos especialistas.

FIGURA 35 – Exemplo de Imagem Rotulada 5



FONTE: A autora (2020).

Já a FIGURA 36 apresenta o a rotulação em um arquivo XML.

Neste arquivo, para cada retângulo que foi gerado de maneira visual nas imagens originais, foi criado um conjunto de coordenadas de localização, sendo a altura, a largura e a profundidade que neste caso representa a espessura da linha. Se na imagem houver 80 retângulos contornando as células, o XML terá 80 conjuntos de coordenadas, refletindo suas determinadas localizações.

Vale lembrar que a base disponibilizada é composta das imagens originais mais os arquivos XML de cada imagem, referentes às anotações.

FIGURA 36 – Arquivo XML

```

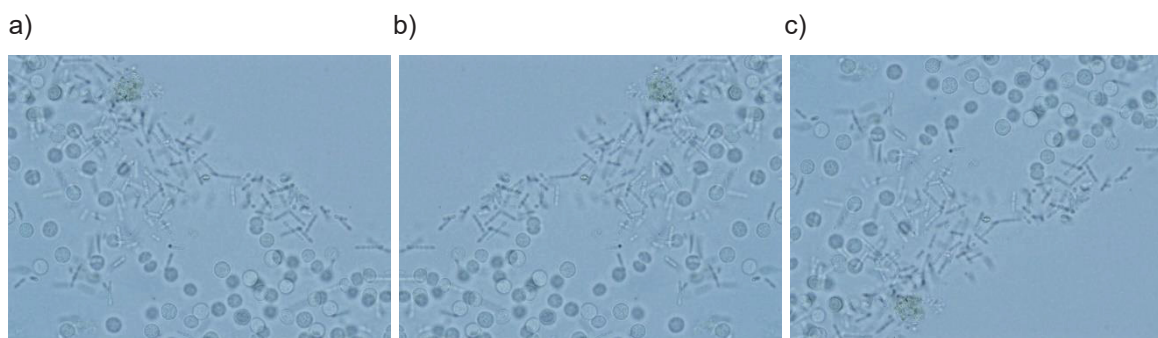
<annotation>
  <folder>Microcystis Aeruginosa</folder>
  <filename>149.jpg</filename>
  <path>C:\Doutorado\Microcystis Aeruginosa\149.jpg</path>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>1360</width>
    <height>1024</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>microcystis_cell</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>135</xmin>
      <ymin>120</ymin>
      <xmax>212</xmax>
      <ymax>189</ymax>
    </bndbox>
  </object>
  <object>
    <name>microcystis_cell</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>497</xmin>
      <ymin>512</ymin>
      <xmax>569</xmax>
      <ymax>585</ymax>
    </bndbox>
  </object>
</annotation>

```

FONTE: A autora (2020).

A FIGURA 37 apresenta uma imagem conforme foi capturada, isto é, sendo em a) a original, e suas duas variações de *data augmentation*, horizontal em b) e vertical em c). Conforme citado na seção 4.1.3, este processo totalizou 6.795 imagens de *Microcystis aeruginosa*, e em 4.971 de fitoplâncton.

FIGURA 37 – Imagem em a) Original, em b) Espelhamento Horizontal e em c) Espelhamento Vertical



FONTE: A autora (2020).

5.2 RESULTADOS DA APLICAÇÃO DE DEEP LEARNING

5.2.1 Classificação de Imagens

Nesta seção são apresentados os resultados do processo de classificação de imagens, onde o objetivo é saber se uma imagem contém ou não cianobactérias.

Ao final do treinamento dos modelos, as acurácias obtidas foram as seguintes: para o modelo *Mobilenet*, 100% de acurácia para o modelo treinado de forma simples (sem validação cruzada), e de 99,78% de acurácia para o modelo treinado com validação cruzada; para o modelo *Inception*, no treinamento simples 99,8% de acurácia, e no treinamento com validação cruzada, o modelo apresentou uma média de acurácia de 99,73%. Estas informações são apresentadas na TABELA 2.

TABELA 2 – Acurácias dos Modelos de Classificação

	<i>Mobilenet</i>	<i>Inception</i>
Acurácia Modelo Simples	100%	99,8%
Acurácia Modelo Validação Cruzada	99,78%±0,12	99,73%±0,13

FONTE: A autora (2020).

Para testar os modelos de classificação treinados, foram utilizadas como parâmetros para avaliação de desempenho a Matriz de Confusão dos modelos, a sensibilidade e a especificidade, métricas características para classificadores.

A seguir, apresentam-se as matrizes de confusão resultantes dos testes com os 20% previamente separados, isto é, dados não incluídos durante o treinamento.

Na TABELA 3 é possível observar o número de verdadeiros positivos e verdadeiros negativos, para o modelo simples (sem *cross-validation*) para as redes *Mobilenet* e *Inception*. Já na TABELA 4, a matriz de confusão apresenta a média dos testes entre todos os *folds*, para ambas as redes também, utilizando validação cruzada (*cross-validation*).

TABELA 3 – Matriz de Confusão dos Testes com Modelo Simples

Testes com Modelo Simples				
	<i>Mobilenet</i>		<i>Inception</i>	
	Microcystis	Outros	Microcystis	Outros
Positivo	1356	3	1347	2
Negativo	3	992	12	993
TOTAL	1359	995	1359	995
Sensibilidade	99,78		99,11	
Especificidade	99,70		99,80	

FONTE: A autora (2020).

TABELA 4 – Matriz de Confusão dos Testes com Validação Cruzada

Testes com Modelo Validação Cruzada				
	<i>Mobilenet</i>		<i>Inception</i>	
	Microcystis	Outros	Microcystis	Outros
Positivo	1355,8±2	1,8±1,78	1351,8±4,65	1,4±0,54
Negativo	3,2±2,2	992,4±2,0	7,2±4,65	992,8±0,7
TOTAL	1359	994,2	1359	994,2
Sensibilidade	99,76 ±0,16		99,47 ±0,34	
Especificidade	99,81 ±0,17		99,85 ±0,05	

FONTE: A autora (2020).

Nos testes do modelo treinado com validação cruzada, o modelo *Mobilenet* apresentou média de 99,76 imagens e desvio padrão de $\pm 0,16$ na sensibilidade e média de 99,81 com $\pm 0,17$ de desvio padrão na especificidade nos testes. O método *Inception* obteve média de 99,47 com $\pm 0,34$ de desvio padrão para sensibilidade, e 99,85 de média com $\pm 0,05$ de desvio padrão para especificidade.

Desta forma, pode-se resumir a acurácia dos testes nos modelos com treinamento simples, conforme a TABELA 5. A rede *Mobilenet* apresentou 99,77% de acurácia para a classe *Microcystis*, e 99,69% de acurácia para a classe *Outros*.

Já a rede *Inception* obteve 99,11% para a classe *Mycrocystis*, e 99,79% para a classe Outros.

TABELA 5 – Acurácia nos Testes Simples

Acurácia dos Testes do Modelo Simples		
	<i>Mobilenet</i>	<i>Inception</i>
Microcystis	99,77%	99,11%
Outros	99,69%	99,79%

FONTE: A autora (2020).

Também é possível verificar a acurácia dos testes nos modelos com treinamento de validação cruzada, como exibido na TABELA 6. Para a classe *Mycrocystis*, a rede *Mobilenet* obteve acurácia de 99,78% com desvio padrão de 0,16, e a rede *Inception* obteve 99,46% e desvio padrão de 0,34. Já para a classe Outros, a *Mobilenet* alcançou 99,81% com desvio padrão de 0,18, e a rede *Inception* 99,85% e desvio padrão de 0,05.

TABELA 6 – Acurácia nos Testes com Validação Cruzada

Acurácia dos Testes do Modelo com Validação Cruzada				
	<i>Mobilenet</i>		<i>Inception</i>	
Fold Nº	Microcystis	Outros	Microcystis	Outros
1	99,77%	99,59%	99,04%	99,89%
2	100%	99,69%	99,26%	99,79%
3	99,85%	100%	99,85%	99,89%
4	99,77%	99,79%	99,41%	99,89%
5	99,55%	100%	99,77%	99,79%
Média/DP	99,78 ±0,16%	99,81 ±0,18%	99,46 ±0,34%	99,85 ±0,05%

FONTE: A autora (2020).

Comparando os resultados, é possível notar que ambas as redes obtiveram excelente acurácia nos testes, com ligeira diferença entre ambas.

Considerando que as provas estatísticas permitem concluir os dados da pesquisa e verificar a validade das hipóteses, com o intuito de confirmar que,

estatisticamente não há diferença na escolha das redes *Mobilenet* e *Inception* para a classificação das imagens, foi efetuado o teste de distribuição T de Student sob a amostra nos testes de validação cruzada. O teste é indicado para pequenos conjuntos de amostras. A hipótese que se adota para a execução do teste, é a de que não há diferença significativa entre as médias apresentadas, isto é, a hipótese de nulidade (diferença = 0), indicando que, portanto, não há diferença estatística na escolha das redes. Assim, a hipótese alternativa é a de que há sim diferença entre as médias (diferença \neq 0). Ao executar o teste, o resultado fornece um valor que é usado para rejeitar ou não a hipótese de nulidade. Rejeita-se a hipótese de nulidade quando o valor obtido com o teste é menor que o nível de significância (em geral 5%, valor assumido no presente estudo) (ALVES, 2017).

A fórmula do teste t para dados pareados se dá por:

$$T = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$$

Onde:

- \bar{d} = média das diferenças pareadas.
- μ_d = média populacional de todas as diferenças pareadas. Ao testar amostras pareadas, a hipótese nula é que μ_d é igual a 0, e a hipótese alternativa é que $\mu_d < 0$, $\mu_d > 0$, ou $\mu_d \neq 0$.
- s_d = desvio padrão das diferenças pareadas.

Assim, o teste T foi realizado via ferramenta Microsoft Office Excel, através da fórmula:

$$TESTET(\text{matriz1}; \text{matriz2}; \text{caudas}; \text{tipo})$$

onde para o primeiro teste, a matriz1 corresponde às médias de *Microcystis* obtidas com o modelo *Mobilenet*, a matriz 2 corresponde às médias de *Microcystis* obtidas com o modelo *Inception*. Já no segundo teste, matriz1 corresponde às médias da classe Outros obtidas com *Mobilenet*, e as médias da classe Outros com o *Inception* são inseridas na matriz2. O termo caudas recebe o valor 2, que corresponde à análise de duas caudas da distribuição, utilizado quando a hipótese de nulidade é considerada, como é o caso. Por fim, tipo recebe o valor 1 que corresponde a uma amostra pareada, isto é, a segunda amostra (matriz2) está pareada com a primeira (matriz1).

Os valores obtidos comparando a identificação de *Microcystis* para os dois modelos foi de 0,169. Já para a identificação da classe *Outros*, para os dois modelos foi de 0,707. Ambos estão acima do valor de 0,05 do nível de significância. Portanto, não se pode rejeitar a hipótese de nulidade, isto é, é possível dizer que não há diferença entre as escolhas dos modelos para esta tarefa.

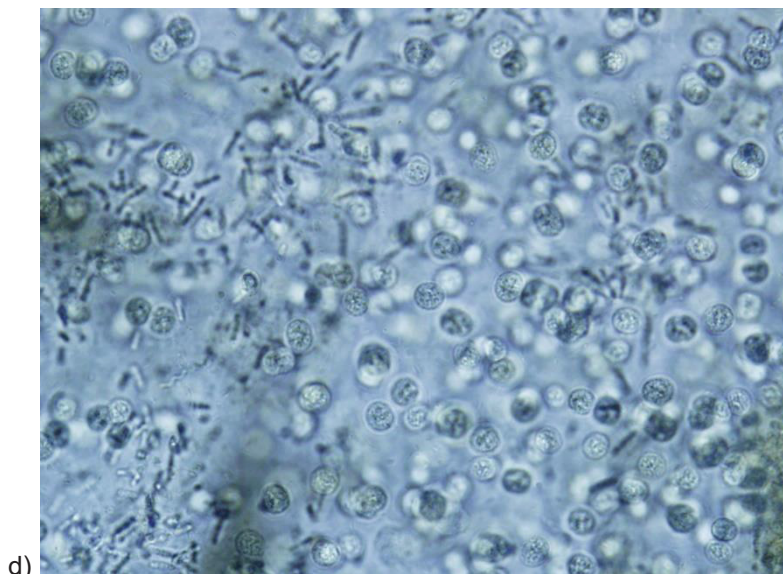
5.2.2 Identificação de Objeto e Contagem

Nesta seção, são apresentados os resultados referentes à identificação de objeto, o que pode ser utilizado para a contagem de células de *Microcystis aeruginosa*.

A avaliação de desempenho de modelos de *object detection*, isto é, treinados para a identificação de objetos, possui uma série de métricas que podem ser utilizadas. Algumas delas foram apresentadas na seção de Fundamentação Teórica.

A seguir, a mesma figura testada nos quatro modelos treinados (*Mobilenet*, *Inception*, *Faster-RCNN* e *YOLO*) é apresentada, para cada um deles, no intuito de demonstrar visualmente o resultado de detecção.

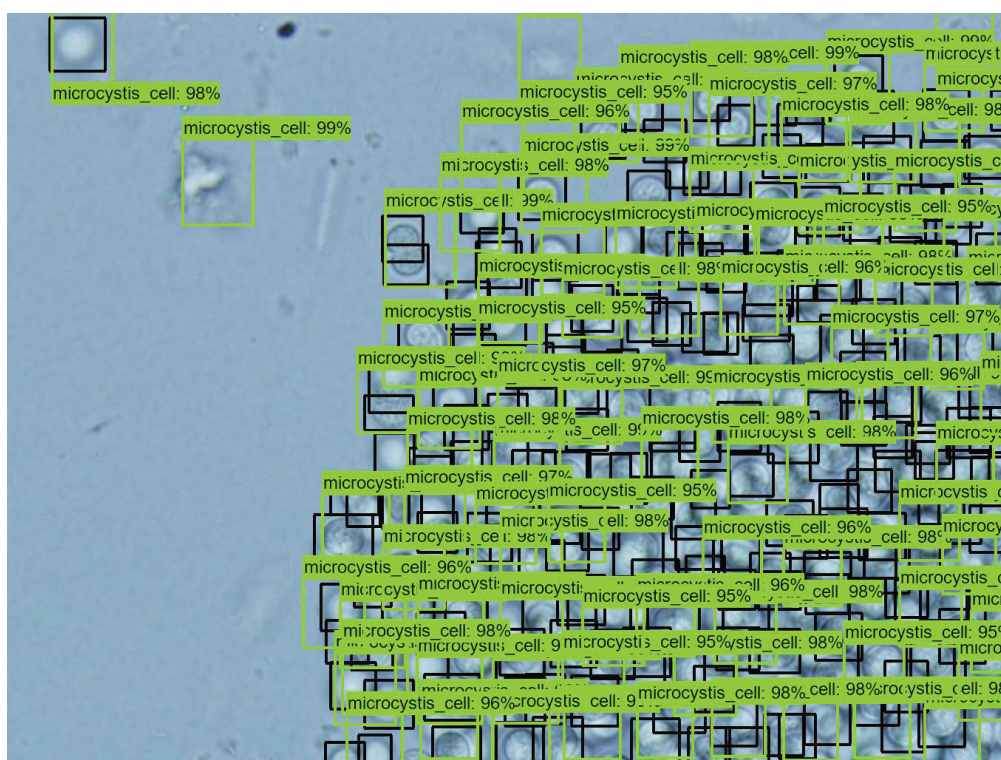
Assim, a FIGURA 38 apresenta uma imagem testada no modelo de *object detection Mobilenet* em a). Em b) é exibido um exemplo de detecção da mesma imagem do modelo *Inception*, em c) apresenta-se a detecção do modelo *Faster – RCNN*, por fim, da *YOLO* em d). A caixa em volta das células possui o percentual de certeza da predição do objeto.



FONTE: A autora (2020).

A seguir, a FIGURA 39 exibe uma imagem representando os *ground-truth boxes* com os *predicted boxes* do modelo *Mobilenet*. Para exemplificar os termos para as avaliações métricas, os *boxes* verdes são as predições do modelo, enquanto os *boxes* pretos são os *boxes* verdadeiros.

FIGURA 39 – *Ground-truth boxes x Predicted Boxes*



FONTE: A autora (2020).

As perdas dos modelos durante seu treinamento (*loss(es)*), para 500 e para 3000 épocas, são apresentadas na TABELA 7 a seguir:

TABELA 7 – Perda dos Modelos

Perda Total dos Modelos		
	500 épocas	3000 épocas
<i>MobileNet</i>	4,251	4,291
<i>Inception</i>	4,522	3,853
<i>Faster-RCNN</i>	1,294	1,384

FONTE: A autora (2020).

Adicionalmente, a perda do modelo YOLO, em suas 51 épocas, finalizou em 128.1333, e considerando que esta perda precisa estar o mais próximo de zero, este valor extremamente elevado é ruim, significa que o modelo está muito longe do esperado.

Não se pode considerar apenas a perda dos modelos. Espera-se se uma perda baixa, contudo, a precisão também é importante. A perda deve ser utilizada para o ajuste fino dos treinamentos, já a precisão que os modelos já treinados apresentam com a realização dos testes, refletem a realidade aplicada, isto é, quão precisos eles são para a tarefa de predição.

Portanto, para avaliar a precisão dos modelos de detecção de objetos treinados foram realizados testes com 20% das imagens previamente separadas via algoritmo. O total de imagens utilizadas nos testes foi de 1.359, e o total de células anotadas para estas imagens foi de 166.342. Desta forma, para a contagem, foi verificado quanto cada modelo encontrou do total de células de *Microcystis aeruginosa*.

Quanto às métricas de desempenho escolhidas para a avaliação dos modelos treinados, a seguir na TABELA 8, são apresentados resultados de testes dos três modelos (*MobileNet*, *Inception* e *Faster-RCNN*) com as IoU de 30%, 50%, 75% e 95%, para cada um deles, com o modelo treinado em 500 épocas. Resultados da YOLO serão apresentados em seguida. Os resultados incluem os valores da matriz de confusão, onde TP são os verdadeiros positivos detectados no teste, FN correspondem aos falso negativos e os FP representam os falso positivos

encontrados. Além disso, valores de *Precision* e a *Recall* também compõem a tabela.

TABELA 8 – Resultados 500 épocas

Resultados com 500 épocas de treinamento				
		MOBILENET	INCEPTION	FASTER
IoU 30%	TP	126367	123278	124509
	FN	39975	43064	41833
	FP	98871	30789	10872
	PRECISION	0,5610	0,8001	0,9196
	RECALL	0,7596	0,7411	0,7485
IoU 50%	TP	124980	103600	89825
	FN	41362	62742	76517
	FP	83216	15239	2756
	PRECISION	0,6003	0,8717	0,9702
	RECALL	0,7513	0,6228	0,5400
IoU 75%	TP	121104	71295	31775
	FN	45238	95047	134567
	FP	60740	5198	190
	PRECISION	0,6659	0,9320	0,9940
	RECALL	0,7280	0,4286	0,1910
IoU 95%	TP	46344	7992	0
	FN	119998	158350	0
	FP	68799	16308	0
	PRECISION	0,4024	0,3288	1
	RECALL	0,2786	0,0480	0

FONTE: A autora (2020).

A TABELA 9 apresenta os resultados dos testes com o experimento adicional treinado com 3000 épocas, também para cada um dos modelos, e cada uma das IoUs propostas.

TABELA 9 – Resultados 3000 épocas

Resultados com 3000 épocas de treinamento				
		MOBILENET	INCEPTION	FASTER
IoU 30%	TP	126223	123335	130056
	FN	40119	43007	36286
	FP	94815	21900	11798
	PRECISION	0,5710	0,8492	0,9168
	RECALL	0,7588	0,7414	0,7818
IoU 50%	TP	123668	109088	112729
	FN	42674	57254	53613
	FP	69100	11699	6125
	PRECISION	0,6415	0,9031	0,9484
	RECALL	0,7434	0,6558	0,6776
IoU 75%	TP	116969	83604	81585
	FN	49373	82738	84757
	FP	40502	3750	2037
	PRECISION	0,7427	0,9570	0,9756
	RECALL	0,7031	0,5026	0,4904
IoU 95%	TP	59607	35200	20944
	FN	106735	131142	145398
	FP	34563	2471	316
	PRECISION	0,6329	0,9344	0,9851
	RECALL	0,3583	0,2116	0,1259

FONTE: A autora (2020).

Já a TABELA 10 exibe os resultados dos testes com a rede YOLOv3, sendo também a matriz de confusão, *Precision* e *Recall*.

TABELA 10 – Resultados YOLOv3

Resultados YOLO 51 épocas				
	IoU 30%	IoU 50%	IoU 75%	IoU 95%
TP	1284	1254	698	3
FN	164447	164447	164447	164447
FP	611	641	1197	1892
PRECISION	0,6775	0,6617	0,368338	0,001583
RECALL	0,0077	0,0075	0,004227	0,000018

FONTE: A autora (2020).

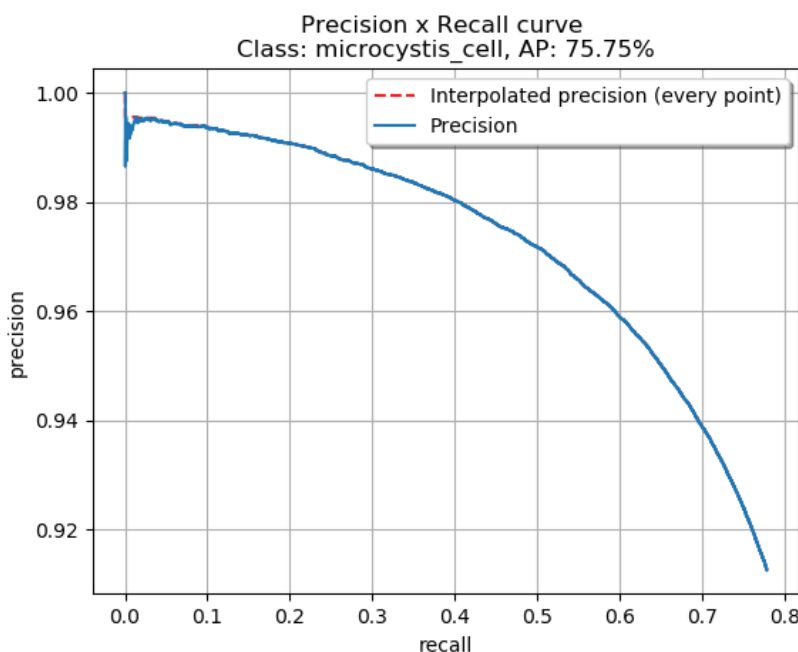
Para interpretar os resultados trazidos é importante lembrar que uma precisão ideal é próxima de 1, ao passo que um *Recall* também próximo de 1 é o tipo de resultado que se espera de um bom modelo de detecção de objetos.

Sabendo disso, não se pode olhar apenas para a *Precision* ou apenas para a *Recall*. A curva PR produzida a partir destes dois valores gera a AP (*average precision*), isto é, a precisão média dos modelos, uma métrica utilizada na avaliação de modelos de aprendizado de máquina. Assim, analisando os resultados obtidos com os testes, observam-se como sendo as melhores precisões as dos modelos *Inception* e *Faster-RCNN*, nas seguintes condições:

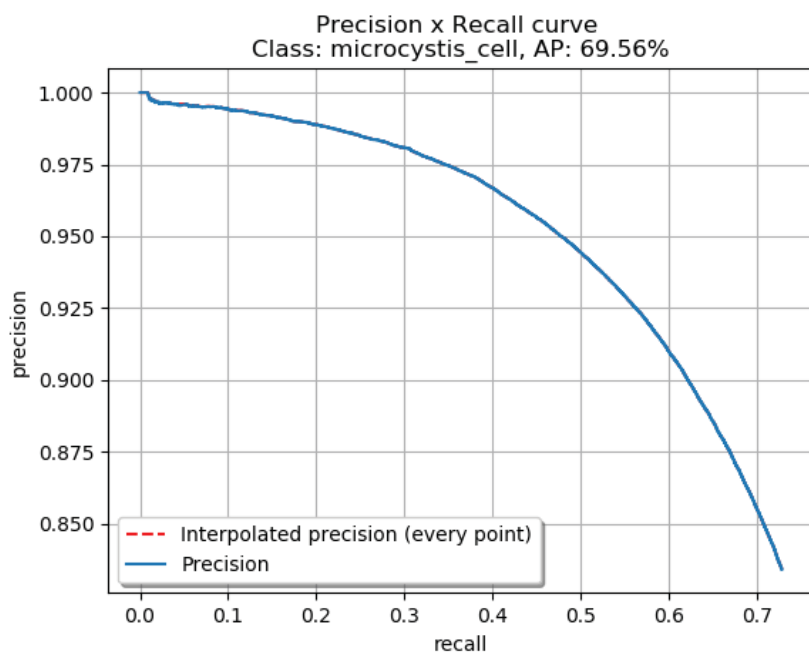
- *Faster-RCNN*: 3000 épocas de treinamento, IoU de 30%, apresentando AP de 75,75% (FIGURA 40);
- *Inception*: 3000 épocas de treinamento, IoU de 30%, apresentando AP de 69,56% (FIGURA 41);
- *Faster-RCNN*: 3000 épocas de treinamento, IoU de 50%, apresentando AP de 66,15% (FIGURA 42).

Quanto ao melhor resultado apresentado pelo *Faster-RCNN*, significa que o modelo tem precisão de 75,75% a um *threshold* de 50%, isto é, uma detecção é considerada positiva se há 50% ou mais de probabilidade que seja positiva.

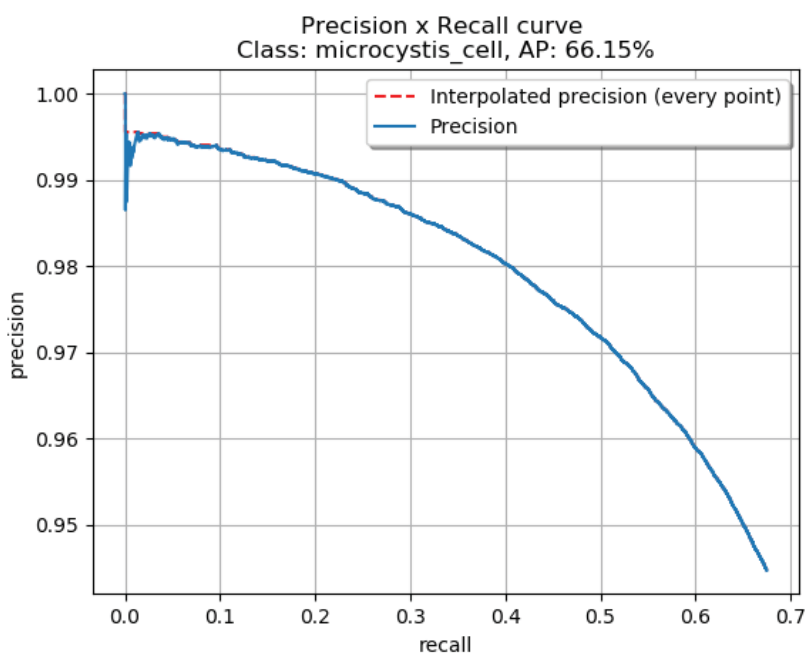
FIGURA 40 – Curva PR *Faster-RCNN* 3000 épocas (IoU 30%)



FONTE: A autora (2020).

FIGURA 41 – Curva PR *Inception* 3000 épocas (IoU 30%)

FONTE: A autora (2020).

FIGURA 42 – Curva PR *Faster-RCNN* 3000 épocas (IoU 50%)

FONTE: A autora (2020).

Todos os demais valores de AP para as IoU's de 30%, 50%, 75% e 95% para todos os modelos de rede estão na seção Apêndice 1.

É possível observar que os resultados são satisfatórios e promissores, uma vez que uma AP acima de 50% é considerada boa, como dito anteriormente, por ser melhor que a probabilidade de 50% de ser ou não ser, entretanto quando mais próximo de 100% estiver, melhor. O melhor resultado alcançou uma AP superior à 75%. Outros resultados promissores na faixa dos 60% também foram observados.

Adicionalmente, as precisões médias obtidas com a rede YOLO, foram de 0,23% para a IoU de 30%, de 0,02% para a IoU de 50%, e de 0% para 75% e 95% de IoU.

Quanto aos resultados obtidos com a rede YOLO, pode-se dizer que não foram satisfatórios, e não são considerados promissores para a base em questão. Os resultados da YOLO corroboram com o que os pesquisadores citados anteriormente na seção de fundamentação teórica relatam, quanto ao fato de que ela não traz bons resultados na identificação de objetos pequenos que aparecem em grupo. Como não é claro de quão pequeno se tratam os objetos, e a versão 3 do modelo trazia melhorias nesta questão, foi pertinente desenvolver os testes com a rede. Entretanto, não se recomenda a utilização desta rede para a identificação de colônias de células de *Microcystis aeruginosa*.

Sendo assim, demonstrou-se a aplicabilidade da base Microaer_ufpr, e considerando os resultados obtidos após os testes com as várias redes aqui citadas, pode-se recomendar aos pesquisadores interessados a dar continuidade nas pesquisas com as redes que alcançaram os melhores resultados nos testes, citadas anteriormente. Os pesquisadores podem produzir mais testes a partir dos aqui apresentados, e aplicar o refinamento personalizado dos parâmetros da rede, uma vez que aqui foi exposta uma comparação entre todos os modelos treinados sob as mesmas condições. O ajuste fino pode ser uma tarefa exaustiva, e varia de acordo com cada aplicação e até mesmo com os modelos escolhidos. Vale ressaltar, que até mesmo o corte da IoU deve ser decidido em acordo com os especialistas interessados na aplicação. Contudo, estes ajustes realizados de forma empírica e personalizada possibilitará alcançar resultados ainda mais satisfatórios.

5.2.3 Disponibilização da base

A *Microaer_ufpr*, sendo a base com as imagens originais e seus rótulos, está disponibilizada gratuitamente, e para ter acesso, basta enviar um e-mail para a autora geisla_albuquerque@yahoo.com.br, para ter acesso ao *link* de *download*. A *Microaer_ufpr* estará disponível também na página do programa de pós-graduação da Universidade Federal do Paraná.

6 CONSIDERAÇÕES FINAIS

As cianobactérias ocorrem nos corpos d'água devido à eutrofização, processo que pode ser desencadeado por ação natural ou humana, e que gera um aumento nos nutrientes da água, favorecendo seu aparecimento. Elas representam uma preocupação mundial, devido à sua potencial toxicidade. Suas toxinas podem causar prejuízo para o indivíduo que tenha contato com elas, podendo este contato ser inalando, ingerindo, tocando ou até mesmo via hemodiálise. As cianobactérias foram relacionadas à casos de mortes de vários tipos de animais e de pessoas ao redor do mundo. A *Microcystis aeruginosa* por exemplo, possui potencial tóxico, e há relatos de mortes humanas devido à exposição a ela. Devido à sua importância, no Brasil, a lei prevê que as companhias de saneamento básico façam o controle e o monitoramento de cianobactérias, e isto inclui analisar amostras no microscópio para sua identificação e contagem, processo inteiramente manual.

As redes neurais artificiais e as técnicas de processamento de imagens, vêm ao longo dos anos ajudando a solucionar problemas nas mais diversas áreas, sendo já uma realidade no dia a dia do país e do mundo. As redes neurais profundas (CNN) vêm mostrando suas potencialidades em diversas áreas, como na identificação de objetos em cenas. Estudos vem sendo desenvolvidos para auxiliar na identificação e no monitoramento de cianobactérias, considerando sua importância para a saúde pública.

Embora seja um tema relevante, não há uma base de imagens de cianobactérias livre que possa auxiliar a comunidade acadêmica e científica no desenvolvimento de estudos com processamento de imagens.

Atendendo aos objetivos geral e específicos da pesquisa, este trabalho apresentou a *Microaer_ufpr*, uma base composta por 2.265 imagens de floração de *Microcystis aeruginosa*, rotuladas célula a célula, e também por 1.657 imagens de organismos diversos no fitoplâncton, sendo um total de 3.922 imagens. As amostras de água são provenientes de mananciais e reservatórios situados nas áreas de Piraquara, Iraí, Foz do Areia, Iguaçu e São José dos Pinhais.

Também se comprovou, através de experimentos com redes neurais convolucionais, que a base tem potencial abrangente na área. Apresentou excelentes resultados quanto à classificação de objetos, alcançando acurácia acima de 99% em todos os testes, para os modelos de CNN *Mobilenet* e *Inception*. Gerou

resultados igualmente promissores quanto à detecção de objetos, chegando a uma precisão média de 75,75% com a rede *Faster-RCNN* com *threshold* de 30% (células são consideradas positivas com probabilidade superior a 30%) e de 66,15% com *threshold* de 50%. A rede *Inception* trouxe uma precisão média de 69,56% (*threshold* de 30%).

A criação e disponibilização de bases de dados para estudos computacionais é uma área que tem notável crescimento de contribuições, devido sua importância para a comunidade acadêmica. Em parte, isto se deve à conhecida necessidade de grande volume de dados que as redes neurais convolucionais demandam quando se trata de processamento de imagens.

Outra contribuição pode ser observada na variedade de testes com redes neurais convolucionais aqui apresentada, a fim de nortear pesquisadores que utilizarão a *Microaer_ufpr*. Desta forma, é possível auxiliar na decisão da rede e focar no ajuste fino necessário para melhorar a acurácia de predição, ou ainda definir questões sobre extração de características, entre outras possibilidades, uma vez que resultados prévios são exibidos.

Microaer_ufpr busca ser um bom recurso para a realização de pesquisas que possam contribuir na área de identificação das cianobactérias.

O trabalho teve o artigo intitulado “*Image classification of cyanobacteria Microcystis aeruginosa in raw water samples in Curitiba’s region*”, com resultados preliminares sobre os estudos com a base de dados, publicado no *IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, em 2020. O evento tem reconhecimento internacional, sendo referência na área de sistemas médicos e relacionados à saúde em geral. Outro artigo será desenvolvido para a publicação oficial da base de dados.

6.1 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Como trabalhos futuros, há uma vasta gama de possibilidades. A princípio, estuda-se a possibilidade de rotular a base existente com imagens do fitoplâncton, identificando os organismos presentes.

Outra possibilidade é a coleta de amostras para criar outra base, com outras espécies de cianobactérias, incluindo filamentosas. Cianobactérias de morfologia

filamentosa apresentam rompimentos em seus filamentos, o que dificulta a contagem.

Em se tratando de utilização prática da Microaer_ufpr, estuda-se a possibilidade de investir fortemente no treinamento dos modelos para se chegar aos melhores resultados, e a partir daí, criar uma interface, onde os usuários poderão carregar as imagens para que o software identifique e conte as células, e ao final gere o total de cada amostra. Alternativamente, pode ser desenvolvido de modo a realizar a varredura em tempo real da amostra, criando uma interface capaz de contar as células sem repetição, quando o especialista varre a amostra pelo microscópio.

Com uma base de imagens, as possibilidades de estudos são amplas, e não estão apenas limitadas à utilização de redes neurais convolucionais.

REFERÊNCIAS

- ALBUQUERQUE, V. *et al.* Sistema de segmentação de imagens para quantificação de microestruturas em metais utilizando redes neurais artificiais. *Matéria* (Rio J.), Rio de Janeiro, v. 12, n. 2, 2007.
- ALVES, M C. Teste t de Student. Piracicaba-SP, 2017. Disponível em: <http://cmq.esalq.usp.br/wiki/lib/exe/fetch.php?media=publico:syllabvs:lcf5759a:teste_t.pdf>. Acesso em: 28 fev. 2021.
- ANDREOLI, C. V. Mananciais de Abastecimento: planejamento e gestão. Estudo de caso do Altíssimo Iguaçu. Curitiba: Sanepar Finep, 2003.
- ANDREOLI, C. V.; CARNEIRO, C. Gestão Integrada de Mananciais de abastecimento Eutrofizados. Curitiba: Sanepar Finep. 2005.
- BAEK, S.; PYO, J.; PACHEPSKY, Y.; PARK, Y.; LIGARAY, M.; AHN, C.; KIM, Y.; CHUN, J. A.; CHO, K. H. Identification and enumeration of cyanobacteria species using a deep neural network, *Ecological Indicators*, Volume 115, 2020.
- BARBOSA, J. Z. *et al.* Uso de imagens digitais obtidas com câmeras para analisar plantas. *Rev. de Ciências Agrárias*, Lisboa, v. 39, n. 1, p. 15-24. 2016.
- BOCHKOVSKIY, Alexey; WANG, Chien-Yao; LIAO, Hong-Yuan Mark. Yolov4: Optimal speed and accuracy of object detection, 2020.
- BOTELHO, M. F; CENTENO, J. A. S. Integração de dados do laser scanner com a banda pan-cromática do sensor QuickBird II para a identificação de edificações através das redes neurais numa abordagem orientação a regiões. *Acta Scientiarum*, Maringá, v. 7, n. 1. 2005.
- BRAGA, A. P., CARVALHO, A. C. P. L. F., LUDEMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. LTC, 2007.
- BRASIL. Ministério da Saúde. Gabinete do Ministro. Portaria Nº 2.914, de 12 de dezembro de 2011. Brasília, 2011.
- BROWN, B. Deep Learning Helps Pathologists Detect Cancer. *Health Tech Insider* by convolutional neural networks. In: *Proceedings of the 8th International Automotive*. 2017.
- CHORUS I, *et al.* Health risk caused by freshwater cyanobacteria in recreational waters. *J Toxicol Environ*, 3:323-347. 2000.
- CHORUS, I., BARTRAM, J. *Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management*. London: E&FN Spon. 1999.
- CODD, G. A., BELL, S. G., and BROOKS, W. P. Cyanobacterial toxins in water. *Water Sci. Technol.* 21:1-13. Congress. p. 162-170. 2011. 1989.

COPASA. Copasa Orienta – Cianobactérias: o perigo das algas azuis. Saneamento, tratamento e abastecimento de água. Programa Chuá: Educação Sanitária e Ambiental da Copasa. Disponível em: <http://www.copasa.com.br/wps/wcm/connect/771d71f8-d24f-4cdb-abf8e3461660d5f6/COPASA_Agua.pdf?MOD=AJPERES>. Acesso em: 17 mai. 2019.

DENG, J. *et al.* ImageNet: a Large-Scale Hierarchical Image Database. IEEE Conference on Computer Vision and Pattern Recognition. 2009.

DUY, T. N., *et al.* "Toxicology and risk assessment of freshwater cyanobacterial (blue-green algal) toxins in water." In: Review of Environmental and Contamination and Toxicology, 163: 113-186. 2000.

ERHAN, D. *et al.*: Scalable Object Detection Using Deep Neural Networks. In: Proc. CVPR. 2014.

ESTEVEZ, F. A. Fundamentos de Limnologia. Editora Interciência, 2ª ed., Rio de Janeiro, 602 p. 1998.

EVERINGHAM, M, *et al.* The Pascal Visual Object Classes (VOC) challenge. International Journal of Computer Vision. 88. 303-338. 2010.

FACELI, K. *et al.* Inteligência artificial: uma abordagem de aprendizado de máquina. Rio de Janeiro: LTC, 2011.

FAPESP. Pesquisadores caracterizam substâncias produzidas por cianobactérias aquáticas. São Paulo. Disponível em: <http://agencia.fapesp.br/pesquisadores_caracterizam_substancias_produzidas_por_cianobacterias_aquaticas/24620/>. Acesso em: 18 mai. 2018.

FRANCIS, G. Poisonous Australian lakes. Nature 18:11–12, 1878.

FURTADO, M. I. Redes neurais artificiais [recurso eletrônico]: uma abordagem para sala de aula. Ponta Grossa (PR): Atena Editora, 2019.

GANDOLA E. *et al.* ACQUA: Automated Cyanobacterial Quantification Algorithm for toxic filamentous genera using spline curves, pattern recognition and machine learning. J. Microbiol. Methods. 124:48–56. 2016.

GIRSHICK, R *et.al.*: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proc. CVPR. 2014.

GIRSHICK, R: Fast R-CNN. In: Proc. ICCV. 2015.

GLOE, T., BÖHME, R. The Dresden Image Database for Benchmarking Digital Image Forensics. Journal of Digital Forensic Practice. 3. 1584-1590. 2010.

GUPTA, Shalini *et al.* Texas 3D Face Recognition Database. 97 - 100. 2010.

HAKRO, D., TALIB, A. Printed Text Image Database for Sindhi OCR. ACM Transactions on Asian and Low-Resource Language Information Processing. 2016.

HAYKIN, S. Redes Neurais Princípios e Prática. Porto Alegre, Bookman, 900p. 2001.

HE, K., ZHANG, X., REN S. and SUN, J. “Deep Residual Learning for Image Recognition”. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.

HOWARD, Andrew G. *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications. 2017.

HUANG, G., LIU, Z., VAN DER MAATEN, L and WEINBERGER, K. Q. “Densely Connected Convolutional Networks”. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261–2269, 2017.

JACCARD, P. “Étude comparative de la distribution florale dans une portion des alpes et des jura,” Bulletin de la Societe Vaudoise des Sciences Naturelles, vol. 37, pp. 547–579. 1901.

JAMES G, *et al.*, “An Introduction to Statistical Learning: with Applications in R”. Springer Texts in Statistics 103. 2013.

JOCHIMSEN, E. M., *et al.* Liver failure and death following exposure to microcystin toxins at a dialysis center in Brazil. N. Engl. J. Med. 338:873–878. 1998.

KERAS. Keras. Disponível em: < <https://keras.io/> >. Acesso em: 15 out. 2021.

KRIENITZ, L., BALLOT, A., KOTUT, K., WIEGAND, C., PÜTZ, S., METCALF, J. S., CODD, G. A., STEPHAN, P. Contribution of hot spring cyanobacteria to the mysterious deaths of Lesser Flamingos at Lake Bogoria, Kenya, FEMS Microbiology Ecology, Volume 43, Issue 2, Pages 141–148, 2003.

LECUN, Y. *et al.* “Gradient-based learning applied to document recognition,” in Proceedings of the IEEE, pp. 2278–2324. 1998.

LECUN, Y., BENGIO Y., HINTON. G. “Deep Learning.” Nature, v521, p436-444. 2015.

LIU, L., *et al.* Deep Learning for Generic Object Detection: A Survey. Int J Comput Vis 128, 261–318. 2020.

LIU, S and DENG, W. “Very deep convolutional neural network based image classification using small training sample size”. In 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 730–734, 2015.

LIU, W. *et al.*: SSD: Single Shot MultiBox Detector. In: Proc. ECCV. 2016.

MACEDO, M. M. G. Uso da transformada de hough na vetorização de moldes e outras aplicações. Dissertação de Mestrado. Niterói, Rio de Janeiro. 2005. Manager. Version 4.1. Oxford: The Cochrane Collaboration. 2000.

MATSUNAGA, H., HARADA, K.-I., SENMA, M., ITO, Y., YASUDA, N., USHIDA, S., KIMURA, Y., Possible cause of unnatural mass death of wild birds in a pond in Nishinomiya, Japan: Sudden appearance of toxic cyanobacteria. *Nat. Toxins*, 7: 81-84. 1999.

MOSLEH, M. A. A., *et al.* A preliminary study on automated freshwater algae recognition and classification system. *BMC Informatics* 13 (Suppl 17), S25. 2012.

ÖHLSCHLÄGER, S., VÕ, M. SCEGRAM: An image database for semantic and syntactic inconsistencies in scenes. *Behavior Research Methods*. 2016.

OLIVEIRA, E. C. Introdução à biologia vegetal. São Paulo: EDUSP, 1996.

PADILLA, R., NETTO, S., DA SILVA, E. A Survey on Performance Metrics for Object-Detection Algorithms. 2020.

PEEMEN, M.; MESMAN, B.; CORPORAL, H. Speed sign detection and recognition. 2011.

PEREZ, L., WANG, J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. 2017.

PRAUS, P., PRAKS, P. Hierarchical clustering of RGB surface water images based on MIA-LSI approach. *Water S.A.* 36. 143-149. 2010.

PRÓXIMO NÍVEL EMBRATEL. 5 aplicações cotidianas que utilizam Deep Learning. Disponível em: <<https://proximonivel.embratel.com.br/5-aplicacoes-cotidianas-que-utilizam-deep-learning/>>. Acesso em: 10 dez. 2020.

PYO, J.; PARK, L. J.; PACHEPSKY, Y.; BAEK, S.; KIM, K.; CHO, K. H. Using convolutional neural network for predicting cyanobacteria concentrations in river water, *Water Research*, Vol. 186, 2020.

PYTHON. Python. Disponível em: < <https://www.python.org/> >. Acesso em: 15 out. 2021.

PYTORCH. Pytorch. Disponível em: < <https://pytorch.org/> >. Acesso em: 15 out. 2021.

REDMON, J. *et al.*: You Only Look Once: Unified, Real-Time Object Detection. In: *Proc. CVPR*. 2016.

REN, S. *et al.*: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *NIPS*. 2015.

- REN, S., HE, K., GIRSHICK, R and SUN, J. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149. 2017.
- RIFFLE, M., DAVIS, T. N. The yeast resource center public image repository: a large database of fluorescence microscopy images. BMC Bioinformatics 11, 263. 2010.
- ROH, M., LEE, J. Refining faster-RCNN for accurate object detection. 514-517. 2017.
- RUSSELL, S.; NORVIG, P. Inteligência Artificial. Elsevier. 2004
- SANT'ANNA, C. L. *et al.* Atlas de cianobactérias e microalgas de águas continentais brasileiras. São Paulo, Instituto de Botânica, 175p. 2012.
- SEN, B. *et al.* A computerized image database for freshwater algae recorded in Turkey. TURKISH JOURNAL OF BOTANY. 39. 198-204. 2015.
- SHAH, M. *et al.* Ziehl-Neelsen sputum smear microscopy image database: A resource to facilitate automated bacilli detection for tuberculosis diagnosis. 2017.
- SILVA, G. H. S., *et al.* Um anjo para ELA: arquitetura baseada em visão computacional aplicada ao monitoramento on-line de parâmetros respiratórios de pacientes com Esclerose Lateral Amiotrófica (ELA) em ambiente hospitalar. Rev. Bras. Eng. Bioméd., Rio de Janeiro, v. 29, n. 2, p. 206-210. 2013.
- SIQUEIRA, D. B., OLIVEIRA-FILHO, E. C. Cianobactérias de água doce e saúde pública: uma revisão. Universitas: Ciências da Saúde. Vol. 3, n.1, p109-127. Uniceub. 2005.
- SOLLI, M., LENZ, R. Emotion related structures in large image databases. CIVR 2010 - 2010 ACM International Conference on Image and Video Retrieval. 2010.
- SZEGEDY, C. *et al.* "Going deeper with convolutions," in CVPR 2015, 2015.
- SZEGEDY, C. *et al.* Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 2818-2826. 2016.
- TENSORFLOW. Uma plataforma completa de código aberto para machine learning. Disponível em: <<https://www.tensorflow.org/?hl=pt-br>>. Acesso em: 15 out. 2021.
- THIEL, S., WILTSHIRE, R. J. The automated detection of cyanobacteria using digital image processing techniques, Environment International, Volume 21, Issue 2, Pages 233-236. 1995.
- UDACITY. Conheça 8 aplicações de deep learning no mercado. Disponível em: <<https://br.udacity.com/blog/post/aplicacoes-deep-learning-mercado>>. Acesso em: 22 mai. 2018.

VERONEZI, C. C. D. *et al.* *Análise computacional para auxílio ao diagnóstico de osteoartrite de coluna lombar baseado em redes neurais artificiais*. Revista Brasileira de Ortopedia, São Paulo, v. 46, n. 2. 2011.

WALSBY, AE., & AVERY, A. Measurement of filamentous cyanobacteria by image analysis. *Journal of Microbiological Methods*, 26, 11 - 20. 1996.

YAN, J. *et al.* IoU-Adaptive Deformable R-CNN: Make Full Use of IoU for Multi-Class Object Detection in Remote Sensing Imagery. *Remote Sensing*. 2019.

YUAN, K. *et al.* Brain CT image database building for computer-aided diagnosis using content-based image retrieval. *Information Processing & Management*. 47. 176-185. 2011.

ZEDER, M., *et al.* A small population of planktonic Flavobacteria with disproportionately high growth during the spring phytoplankton bloom in a prealpine lake. *Environ. Microbiol.* 2009.

ZENG, X. *et al.*: Gated Bi-directional CNN for Object Detection. In: *Proc. ECCV*. 2016.

ZHANG, Z. *et al.* ORIGA(-light): An Online Retinal Fundus Image Database for Glaucoma Analysis and Research. *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*. 2010.

ZHUANG, F. *et al.* A comprehensive survey on transfer learning. *Proceedings of the IEEE*, v. 109, n. 1, p. 43-76, 2020.

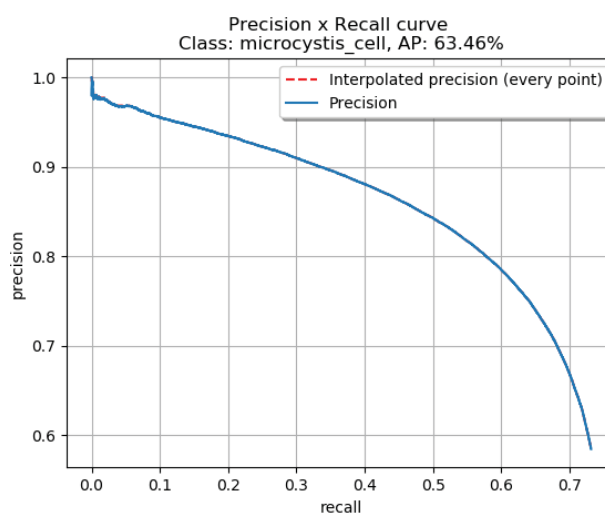
APÊNDICE 1 – GRÁFICOS CURVA PR

A seguir, são apresentados os gráficos gerados com as curvas PR para as IoU's de 30%, 50%, 75% e 95%, para todos os modelos, isto é, *Mobilenet*, *Inception* e *Faster – RCNN*.

Modelos treinados com 500 épocas – IoU 30%

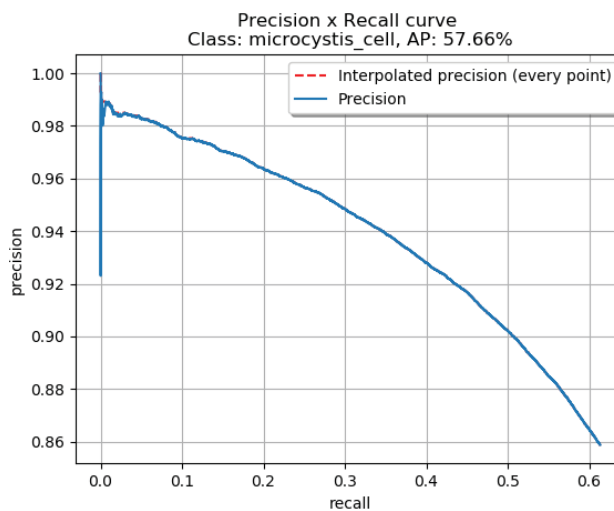
Na FIGURA A1 tem-se a curva PR para o modelo *Mobilenet* (IoU 30%), e uma AP de 63,46%.

FIGURA A1 - Curva PR *Mobilenet* (IoU 30%)



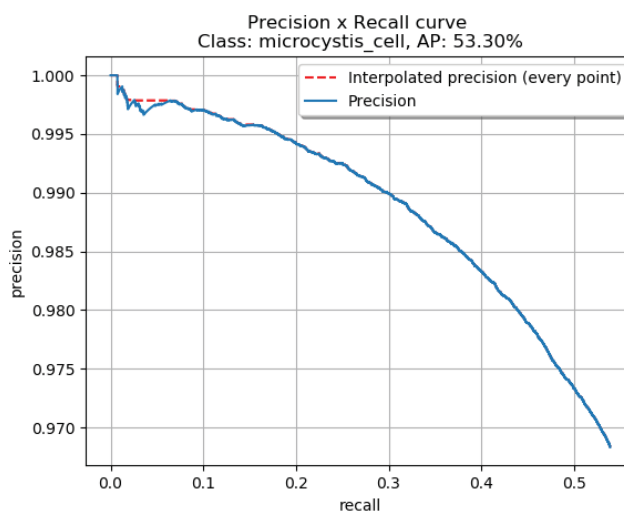
FONTE: A autora (2020).

Com a IoU de 30%, a Figura A2, mostra uma AP de 57,66%, é exibida a curva PR do modelo *Inception*.

FIGURA A2 - Curva PR *Inception* (IoU 30%)

FONTE: A autora (2020).

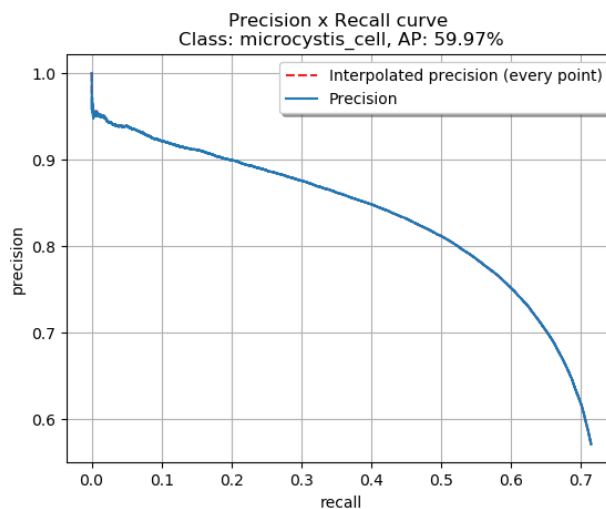
Finalizando com a IoU de 30%, a FIGURA A3 mostra a curva PR gerada pelo modelo *Faster - RCNN*, e uma AP de 53,30%.

FIGURA A3 - Curva PR *Faster - RCNN* (IoU 30%)

FONTE: A autora (2020).

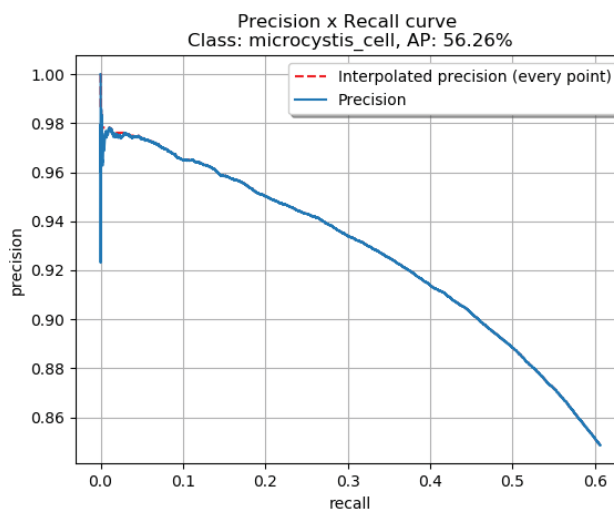
Modelos treinados com 500 épocas – IoU 50%

Na FIGURA A4, pode-se observar a curva PR para o modelo *Mobilenet*, com um *threshold* de 50%, e uma AP de 59,97%.

FIGURA A4 - Curva PR *Mobilenet* (IoU 50%)

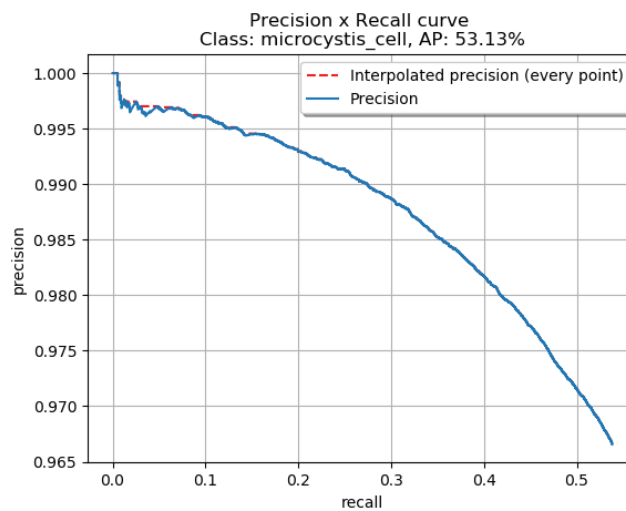
FONTE: A autora (2020).

Já na FIGURA A5, com um limite mínimo de IoU de 50%, é apresentada a curva PR para o modelo *Inception*, com uma AP de 56,26%.

FIGURA A5 - Curva PR *Inception* (IoU 50%)

FONTE: A autora (2020).

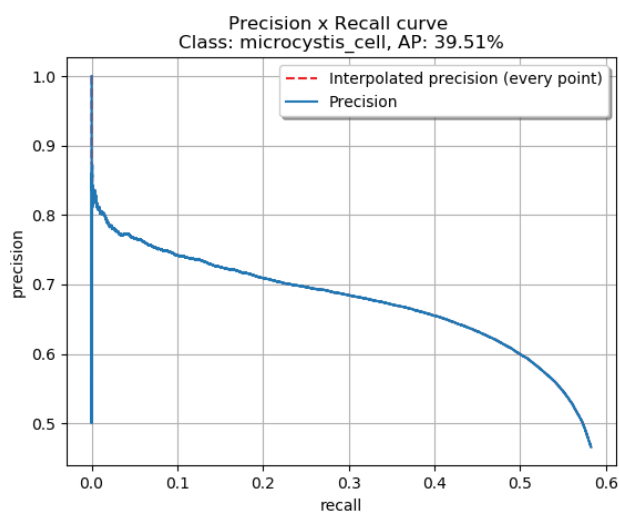
A FIGURA A6 exibe a curva PR do modelo *Faster – RCNN*, e sua AP de 53,13% para uma IoU de 50%.

FIGURA A6 - Curva PR *Faster - RCNN* (IoU 50%)

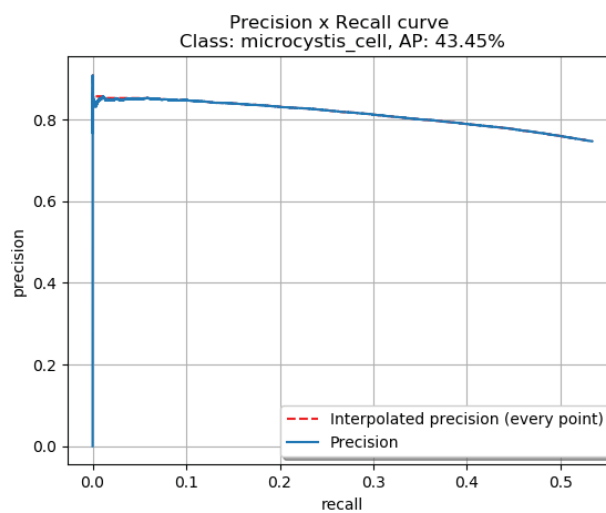
FONTE: A autora (2020).

Modelos treinados com 500 épocas – IoU 75%

Aqui apresentam-se as curvas PR com *threshold* de 75% para os modelos *Mobilenet* e *Inception* e suas AP's, respectivamente nas FIGURAS A7 e A8.

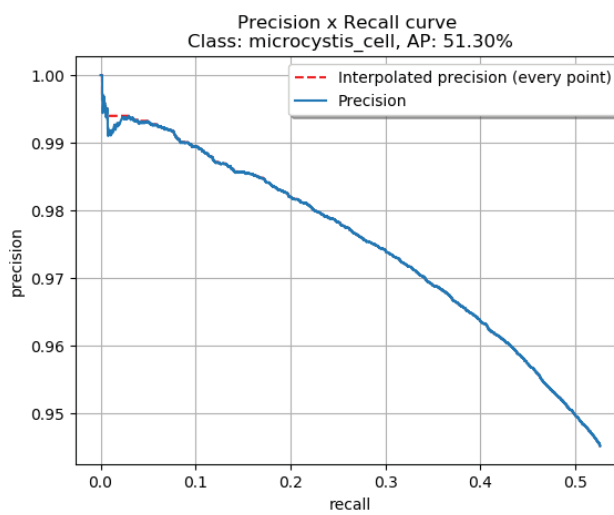
FIGURA A7 - Curva PR *Mobilenet* (IoU 75%)

FONTE: A autora (2020).

FIGURA A8 - Curva PR *Inception* (IoU 75%)

FONTE: A autora (2020).

Na FIGURA A9 é possível observar uma AP de 51,30% para um *threshold* de 75% para o *Faster-RCNN*.

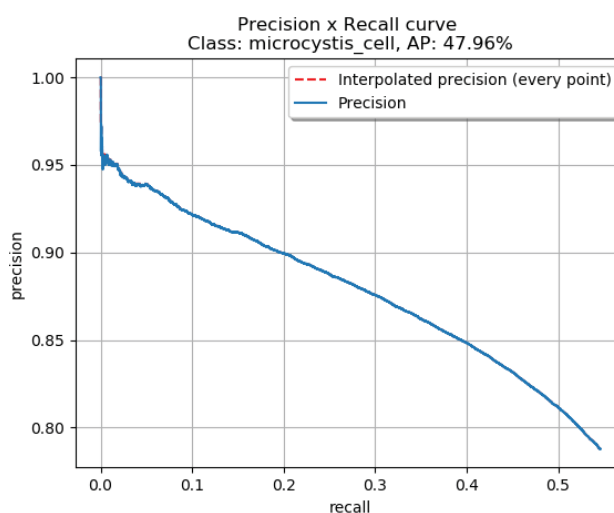
FIGURA A9 - Curva PR *Faster-RCNN* (IoU 75%)

FONTE: A autora (2020).

Modelos treinados com 500 épocas – IoU 95%

Observa-se na FIGURA A10 uma AP de 47,96% para o modelo *Mobilenet*, a uma IoU de 95%.

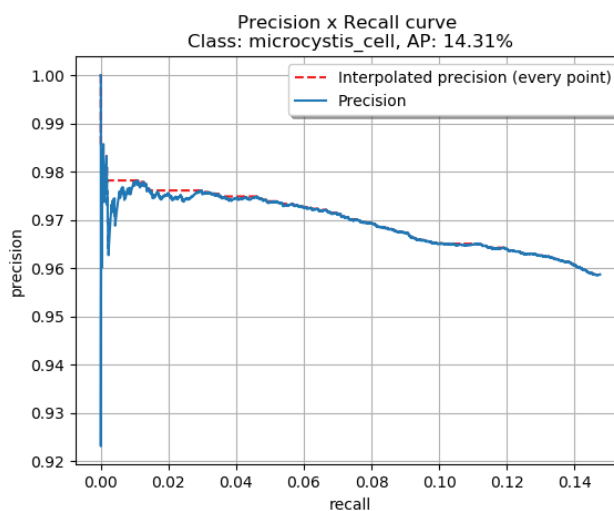
FIGURA A10 - Curva PR *Mobilenet* (IoU 95%)



FONTE: A autora (2020).

A FIGURA A11 exibe a curva PR para o modelo *Inception*, que corresponde a 14,31% de AP.

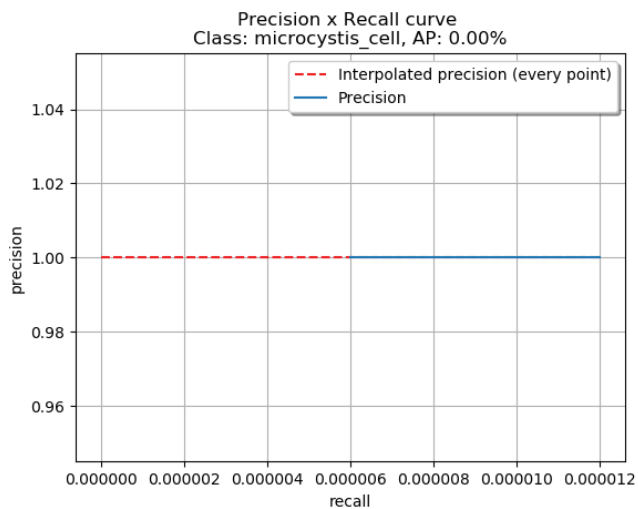
FIGURA A11 - Curva PR *Inception* (IoU 95%)



FONTE: A autora (2020).

Já na FIGURA A12, tem-se a curva PR do modelo *Faster-RCNN* para uma IoU de 95%, onde a AP fica em 0%.

FIGURA A12 - Curva PR *Faster - RCNN* (IoU 95%)

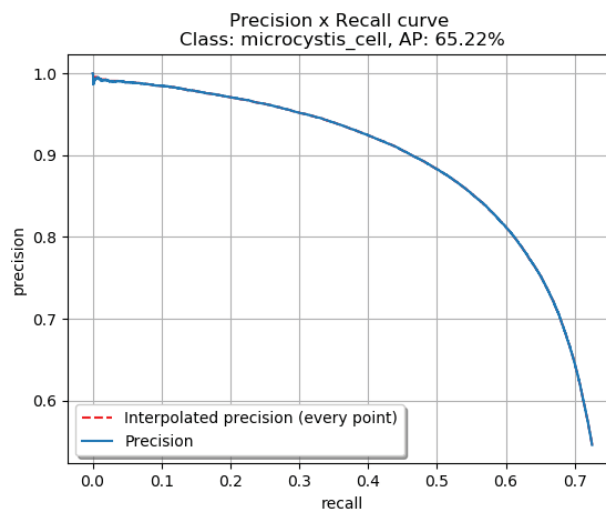


FONTE: A autora (2020).

Modelos treinados com 3000 épocas – IoU 30%

A seguir, são apresentadas as curvas PR para os três modelos treinados com 3000 épocas.

A FIGURA A13 exhibe uma AP de 65,22% para o modelo *Mobilenet*, treinado com 3000 épocas, a uma IoU de 30%.

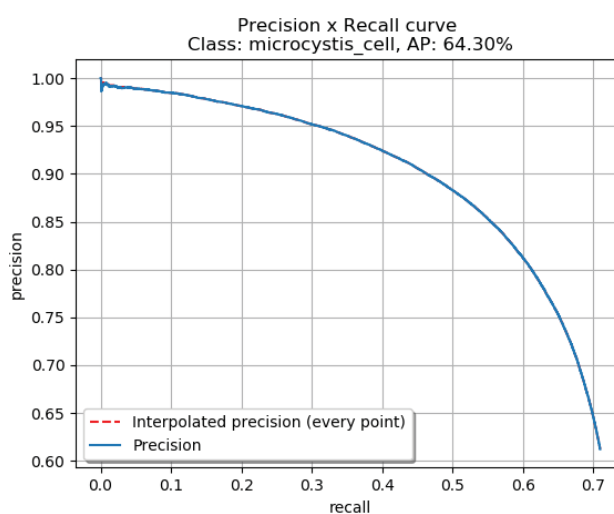
FIGURA A13 - Curva PR *Mobilenet* 3000 épocas (IoU 30%)

FONTE: A autora (2020).

Modelos treinados com 3000 épocas – IoU 50%

Aqui são exibidas as curvas para os modelos treinados com 3000 épocas e uma IoU de 50%.

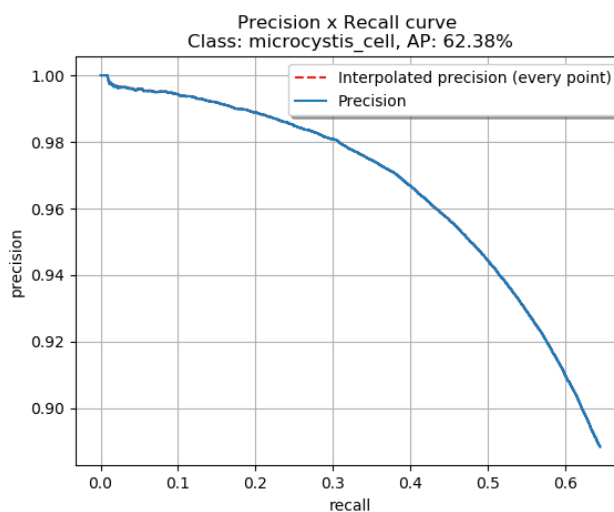
Na FIGURA A14 tem-se a curva PR do *Mobilenet*, que obteve uma AP de 64,30%.

FIGURA A14 - Curva PR *Mobilenet* 3000 épocas (IoU 50%)

FONTE: A autora (2020).

Na FIGURA A15 observa-se na curva PR uma AP de 62,38% para o modelo *Inception*.

FIGURA A15 - Curva PR *Inception* 3000 épocas (IoU 50%)



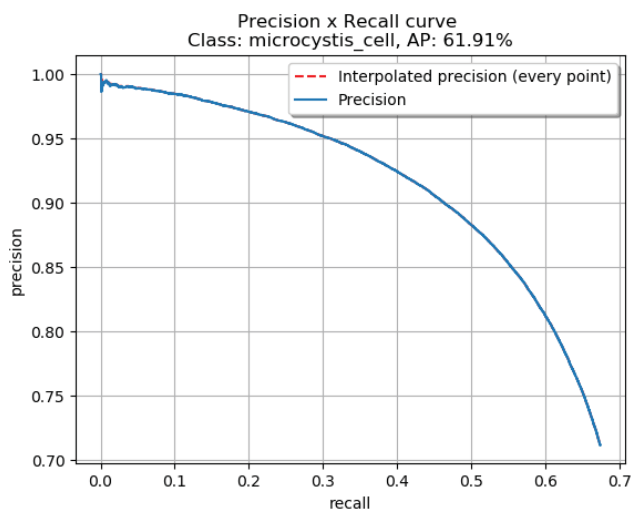
FONTE: A autora (2020).

Modelos treinados com 3000 épocas – IoU 75%

As curvas PR dos modelos treinados com 3000 épocas e uma IoU de 75% são mostrados a seguir.

Na FIGURA A16, observa-se que o modelo *Mobilenet* obteve uma AP de 61,91%.

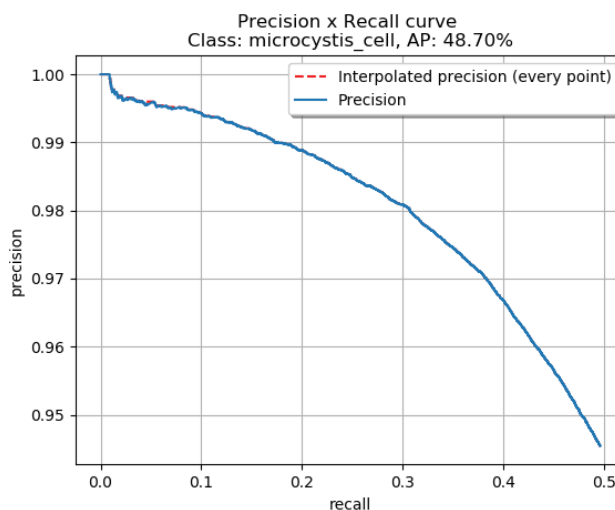
FIGURA A16 - Curva PR *Mobilenet* 3000 épocas (IoU 75%)



FONTE: A autora (2020).

Já com uma AP de 48,70%, a curva PR do modelo *Inception* é apresentada na FIGURA A17.

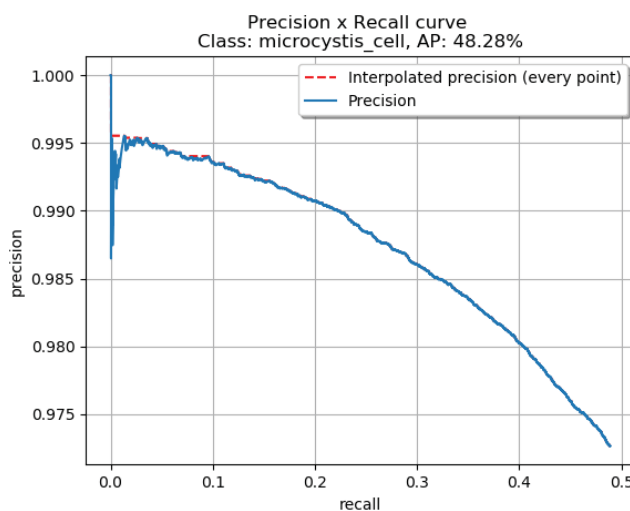
FIGURA A17 - Curva PR *Inception* 3000 épocas (IoU 75%)



FONTE: A autora (2020).

O modelo *Faster-RCNN* alcançou uma AP de 48,28%, como se pode observar na FIGURA A19.

FIGURA A19 - Curva PR *Faster R-CNN* 3000 épocas (IoU 75%)



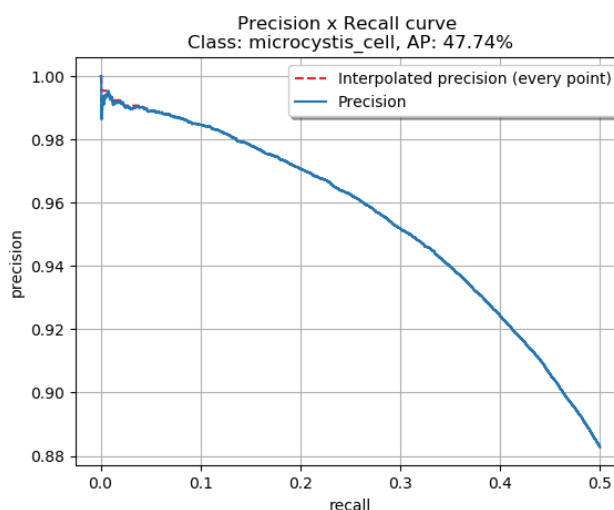
FONTE: A autora (2020).

Modelos treinados com 3000 épocas – IoU 95%

A seguir as curvas PR bem com as precisões médias dos modelos treinados com 3000 épocas e IoU de 95% são apresentadas.

Na FIGURA A19, o modelo *Mobilenet* exibe uma AP de 47,74% com uma IoU de 95%.

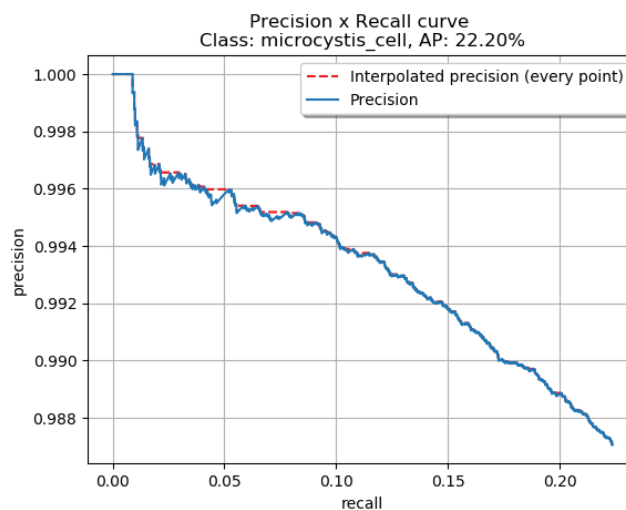
FIGURA A19 - Curva PR *Mobilenet* 3000 épocas (IoU 95%)



FONTE: A autora (2020).

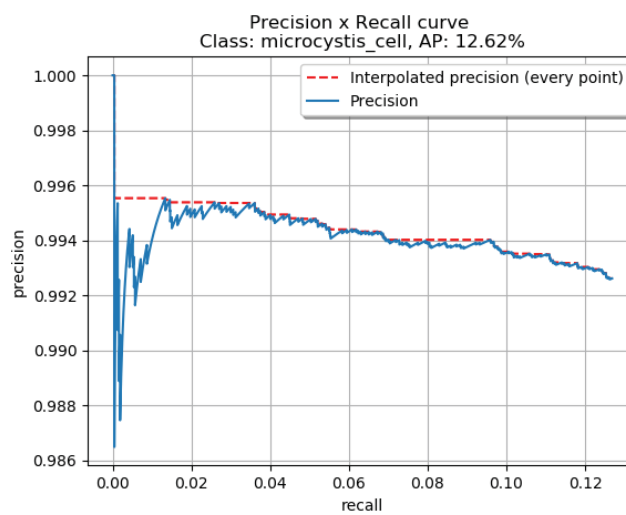
A FIGURA A20 mostra a curva PR do modelo *Inception*, que alcançou uma AP de 22,20% com o *threshold* de 95%.

FIGURA A20 - Curva PR *Inception* 3000 épocas (IoU 95%)



FONTE: A autora (2020).

Por fim, observa-se na FIGURA A21 uma AP de 12,62% para o modelo *Faster-RCNN*, a uma IoU de 95% com 3000 épocas de treinamento.

FIGURA A21 - Curva PR *Faster R-CNN* 3000 épocas (IoU 95%)

FONTE: A autora (2020).