

UNIVERSIDADE ESTADUAL DO MARANHÃO

CENTRO DE CIÊNCIAS TECNOLÓGICAS

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DA COMPUTAÇÃO E SISTEMAS

Yanna Leidy Ketley Fernandes Cruz

*Deteccão de ovócitos em imagens histológicas utilizando Redes
Neurais Convolucionais*

São Luís - MA

2022

Yanna Leidy Ketley Fernandes Cruz

*Detecção de ovócitos em imagens histológicas utilizando Redes
Neurais Convolucionais*

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia da Computação e Sistemas da Universidade Estadual do Maranhão como requisito parcial para a obtenção do grau de MESTRE em Engenharia da Computação e Sistemas.

Orientador: Prof. Dr. Ewaldo Eder Carvalho Santana

Doutor em Engenharia de Eletricidade – UFPB

Co-orientador: Prof. Dra. Raimunda Nonata Fortes Carvalho Neta

Doutora em Biotecnologia – RENORBIO/UECE

São Luís - MA

2022

Cruz, Yanna Leidy Ketley Fernandes Cruz.

Detecção de ovócitos em imagens histológicas utilizando Redes Neurais Convolucionais / Yanna Leidy Ketley Fernandes Cruz Cruz. – São Luís - MA, 2022.

71 f.

Dissertação (Mestrado) – Universidade Estadual do Maranhão, Programa de Pós-Graduação em Engenharia da Computação e Sistemas. São Luís - MA, 2022.

Orientador: Prof. Dr. Ewaldo Eder Carvalho Santana.

Coorientadora: Prof. Dra. Raimunda Nonata Fortes Carvalho Neta.

1.Deep learning. 2.Ovócitos. 3.Centropomus undecimalis. 4.YOLO. 5.Imagens histológicas.

CDU 004.932:639.3

Yanna Leidy Ketley Fernandes Cruz

*Detecção de ovócitos em imagens histológicas utilizando
Redes Neurais Convolucionais*

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Yanna Leidy Ketley Fernandes Cruz e aprovada pela comissão examinadora.

Aprovada em 1 de Julho de 2022

BANCA EXAMINADORA



Prof. Dr. Ewaldo Eder Carvalho Santana (orientador)

Doutor em Engenharia de Eletricidade - UFPE



Prof. Dra. Raimunda Nonata Fortes Carvalho Neta (coorientadora)

Doutora em Biotecnologia - RENO/BIO/UECE



Prof. Dr. Mauro Sérgio Silva Pinto

Doutor em Engenharia de Eletricidade - UFMA



Prof. Dra. Nilviane Pires Silva Sousa

Doutora em Engenharia Biomédica - UFMA

*A minha família, namorado
e amigos que ajudaram nessa
caminhada.*

Resumo

Ovócitos são células germinativas complexas que apresentam mudanças estruturais e funcionais ao longo do desenvolvimento maturacional em peixes. Deste modo, é de significativa importância o desenvolvimento de abordagens que realizem a detecção automática de ovócitos, visto que distingui-los em imagens histológicas é um procedimento demorado e suscetível a erros, pois requer a interpretação visual e subjetiva do especialista. Desta forma, este trabalho propõe um método capaz de detectá-los de forma automática. Para isso, uma base de imagens contendo 5.680 ovócitos foi utilizada, sendo esta dividida em treino, teste e validação. Técnicas para o aumento de dados foram empregadas a fim de tornar os modelos mais robustos. Para realizar a detecção as arquiteturas YOLOv3, YOLOv4 e YOLOv5 foram aplicadas à base de imagens. Os resultados obtidos foram promissores, destacando-se que o modelo YOLOv5l, na detecção de ovócitos da classe VF, atingiu os melhores valores nas métricas *precision*, *recall*, *mAP@.5* e *mAP@.95*, com 85,4%, 95,3%, 97,7% e 75,9%, respectivamente. Ressaltasse ainda, que ao levar em consideração todas classes o YOLOv5l é o modelo que, com maior frequência, atingiu os melhores resultados nas métricas analisadas.

Palavras-chaves: Deep Learning, Ovócitos, *Centropomus undecimalis*, YOLO, Imagens histológicas.

Abstract

Oocytes are complex germ cells that show structural and functional changes during maturational development in fish. Thus, the development of approaches that perform the automatic detection of oocytes is of significant importance, since distinguishing them in histological images is a time-consuming and error-prone procedure, as it requires the visual and subjective interpretation of the specialist. Thus, this work proposes a method capable of detecting them automatically. For this, an image database containing 5,680 oocytes was used, which was divided into training, testing and validation. Techniques for data augmentation were employed in order to make the models more robust. To perform the detection the YOLOv3, YOLOv4 and YOLOv5 architectures were applied to the image base. The results obtained were promising, highlighting that the YOLOv5l model, in the detection of oocytes of the VF class, reached the best values in the metrics *precision*, *recall*, *mAP@.5* and *mAP@.95*, with 85.4%, 95.3%, 97.7% and 75.9%, respectively. It should also be noted that, when taking into account all classes, YOLOv5l is the model that most frequently achieved the best results in the analyzed metrics.

Keywords: Deep Learning, Oocytes, *Centropomus undecimalis*, YOLO, Histological images.

Agradecimentos

Agradeço primeiramente a Deus, por permitir que eu atinja este sonho.

Aos meus familiares, a quem devo toda a força e motivação para conseguir meus objetivos.

Ao meu professor e orientador Ewaldo Eder Carvalho Santana, por todo o conhecimento, conselhos e paciência depositada em mim durante a orientação.

A minha coorientadora Raimunda Nonata Fortes, pela assistência e ensinamentos passados em meu mestrado.

A professora Isa Rosete Mendes pelas amostras de ovócitos disponibilizadas para realizar minha pesquisa. Ao professor José de Ribamar de Souza Torres, pela sua disponibilidade e paciência para ensinar a caracterização e identificação dos ovócitos.

Aos amigos e companheiros do mestrado que auxiliaram e motivaram nesta jornada. Em especial Marcio, Elton, Darnley, Ana Luiza, Carol e Paulo.

Ao meu namorado Antonio Phillipi, pelo companheirismo e motivação dada no mestrado.

E a todos que diretamente ou indiretamente que contribuíram para a realização deste trabalho.

*"Deus nunca disse que a jornada seria fácil, mas
Ele disse que a chegada valeria a pena."*

Max Lucado

Lista de Figuras

3.1	<i>Centropomus undecimalis</i> - Robalo Flecha. Fonte: Rodrigues [1]	30
3.2	Ovários em diferentes estádios de desenvolvimento: a) Estádio A - Imaturo, b) Estádio B - Em maturação, c) Estádio C - Maduro, d) Estádio D - Esvaziado, e) Estádio E - Repouso. Fonte: Felix et al. [2]	31
3.3	Ovócitos em Pré-vitelogênese com crescimento do núcleo contendo nucléolos proeminentes. (Aum. 200X). Fonte: A autora.	34
3.4	Ovócito em vitelogênese inicial, com vacúolos lipídicos, núcleo e nucléolos. (Aum. 200X). Fonte: A autora.	34
3.5	Ovócitos em vitelogênese final contendo grânulos proteico por todo o ovoplasma. (Aum. 200X). Fonte: A autora.	35
3.6	Rede Neural Convolutacional para processamento de imagens. Fonte: Saha [3]	36
3.7	Ilustração da camada de convolução. Fonte: Hafemann [4]	36
3.8	Ilustração da camada de subamostragem. Fonte: Hafemann [4]	37
3.9	Sequência de detecção de objetos para a arquitetura YOLO. Fonte: Redmon et al. [5].	39
3.10	Cálculo de Intersecção sobre União (IoU). Fonte: Cheng et al. [6].	42
3.11	Valores do IoU em diferentes sobreposições. Fonte: Aragão [7].	42
4.1	Etapas do método utilizado neste trabalho. Fonte: A autora.	45
4.2	Imagem capturada por microscópio ampliada em x200. Fonte: A autora.	47
4.3	Processo de rotulagem dos ovócitos na imagem. (A) Imagem rotulada pelo especialista. (B) Arquivo em formato TXT para o YOLO. Fonte: A autora.	48
5.1	Curva Precision vs Recall.	57

Lista de Tabelas

2.1	Resumo dos trabalhos relacionados referentes a identificação de objetos em imagens histológicas.	27
4.1	Detalhes do conjuntos de dados de ovócitos.	49
4.2	Composição das arquiteturas YOLOv3, YOLOv4 e YOLOv5. Fonte: [8], [9] e [10]	51
5.1	Comparação das performances dos modelos no treinamento e teste por meio dos pesos do modelo, tempo de treinamento, métrica mAP@.50 e tempo de inferência.	54
5.2	Comparação de desempenho dos modelos com base nas métricas <i>precision</i> (P), <i>recall</i> (R) e média AP para cada classe (mAP).	55
5.3	Modelos com melhores resultados	56
5.4	Resultado da detecção e localização dos ovócitos através dos modelos YOLOv3-tiny, YOLOv3-ssp e YOLOv4-tiny para cada fase ovocitária.	59
5.5	Resultado da detecção e localização dos ovócitos através dos modelos YOLOv4-esp, YOLOv5l e YOLOv5m para cada fase ovocitária.	60
5.6	Resultado da detecção e localização dos ovócitos através dos modelos YOLOv5n, YOLOv5s e YOLOv5x para cada fase ovocitária.	61

Lista de Siglas

ANN - Artificial Neural Network

AP - Average Precision

AT - Atrésico

BACH - Breast Cancer Histology Challenge

BIONORTE - Biodiversidade e Biotecnologia da Amazônia Legal

CA - Cortical Alveolar

CN-Neult - Convolutional Neural Normalized Cuts

CNN - Convolutional Neural Networks

CN-PA - Convolutional Neural Patch-wis

CN-PI - Convolutional Neural Pixel-wise

CN-SLIC - Convolutional Neural Simple Linear Iterative Clustering

CSP - Cross Stage Partial

DoG - Diferencial Gaussian

ELMOR - Extreme Learning Machine for Ordinal Regression

EPSVM - Ensemble Support Vector Machine

FN - False negative

FOS - First Order Statistics

FP - False positive

GLRLS - Grey Level Run Length Statistics

GLTD - Grey Level Dexture Descriptors

GKSVM - Gaussian Kernel Support Vector Machine

GPU - Graphics Processing Unit

HC - Haralick Coefficients

IID - Hidratado

ICTA - Integrative Colour Texture Analysis

IoU - Intersect over Union

KDLOR - Kernel Discriminant Learning Ordinal Regression

KNN - K-Nearest Neighbors

LOG - Laplacian of Gaussia

mAP - mean Average Precision

MC - *Merciccius*

MGLTF - Multiscalar Grey Level Texture Feature

MLP - MultiLayer Perceptron

NGLDS - Neighboring Grey Level Dependence Statistics

NKI - Netherlands Cancer Institute

ONN - Ordinal Neural Network

REDSVM - REDuction Support Vector Machine

RH - *Reinhardtius Hippoglossoides*

SCF - Statistical Colour Features

SGD - Stochastic Gradient Descent

SPP - Spatial Pyramid Pooling

SVM - Support Vector Machine

SVORIM - Support Vector Ordinal Regression with Implicit Constraints

SVMOD - Support Vector Machine with Ordinal Decompositions

SVR - Support Vector Regression

TL - *Trisopterus Luscus*

TN - true negative

TP - True positive

VGH - Vancouver General Hospital

VIT - Vitelovēnese

WT - Wavelet Transform

YOLO - You Only Look Once

Sumário

Lista de Figuras	ix
Lista de Tabelas	x
Lista de Siglas	xi
1 Introdução	16
1.1 Relevância e Justificativa	16
1.2 Objetivos	18
1.2.1 Objetivo Geral	18
1.2.2 Objetivos Específicos	18
1.3 Organização da Dissertação	18
2 Trabalhos Relacionados	20
2.1 Considerações sobre os trabalhos relacionados	26
3 Fundamentação Teórica	29
3.1 <i>Centropomus undecimalis</i> - Robalo Flecha	29
3.1.1 Desenvolvimento Ovariano / Morfologia dos Ovários	30
3.2 Redes Neurais Convolucionais	35
3.3 YOLO – You Only Look Once	38
3.3.1 YOLOv3	39
3.3.2 YOLOv4	40
3.3.3 YOLOv5	40
3.4 Métricas para Análise de Desempenho	41

3.4.1	<i>Intersect over Union (IoU)</i>	41
3.4.2	Matriz de confusão	42
3.4.3	Curva Precision vs Recall	43
3.4.4	Average Precision (AP)	44
4	Metodologia	45
4.1	Aquisição de Imagens	45
4.2	Divisão da base de dados	47
4.3	Data augmentation	49
4.4	Arquiteturas utilizadas	50
4.5	Ambiente para os experimentos	51
5	Resultados e Discussão	53
5.1	Performance dos modelos no treinamento e teste	53
5.2	Comparação dos Diferentes Modelos de Detecção	54
5.3	Curva Precision vs Recall	56
5.4	Resultado do Método de Detecção	57
6	Conclusão	62
6.1	Contribuições e Resultados da Pesquisa	63
6.2	Trabalhos Futuros	63
	Referências Bibliográficas	65

1 Introdução

1.1 Relevância e Justificativa

Os peixes apresentam grande importância para subsistência e geração de renda, uma vez que contribuem para o consumo alimentar de uma expressiva parcela da população e são fonte de matéria-prima de diversos produtos. Apresentam também importância ecológica, pois mantêm o ecossistema marinho equilibrado. Segundo dados da Organização das Nações Unidas para a Alimentação e a Agricultura (FAO) [11], o consumo de peixe pela população global vem aumentando cerca de 1,5% ao ano, saltando de 9 kg/habitante, em 1961, para 20,5 kg/habitante em 2020. No Brasil, este consumo obteve um crescimento de 5,93%, em relação ao ano de 2019, atingindo 802,3 mil toneladas em 2020, o segundo melhor desempenho desde 2014, de acordo com informações da Associação Brasileira da Piscicultura [12].

No que se refere à espécie *Centropomus undecimalis* [13], a ampla captura realizada através da pesca artesanal, industrial e esportiva, desenvolvida no litoral amazônico tem provocado preocupações, visto que grande parte das reservas de peixes marinhos, cerca de 70%, se encontram em áreas onde há intensa exploração. Diante desse fato, percebe-se que o risco de interrupção do ciclo de reprodução destes peixes é agravado [14]. Tal número, além de afetar a piscicultura mundial, ameaça a segurança alimentar em escala global.

Frente a esta situação, a falta de monitoramento dessa espécie de peixe impossibilita uma análise de informações referentes a aspectos biológicos e reprodutivos. Neste sentido, estudos que abordam a determinação do desenvolvimento e da maturação sexual dos peixes são fundamentais para a ciência da pesca e são pré-requisitos para a compreensão do ciclo de vida dos peixes [15]. Além disso, uma investigação das características apresentadas nos ovócitos permite compreender melhor os aspectos reprodutivos desta espécie, quantificar sua taxa de fecundidade e constatar seus níveis de estoques pesqueiros. O método mais adequado

para a determinação do ciclo reprodutivo em fêmeas é a observação de modificações no desenvolvimento gonadal ou ovocitário [16].

Sendo assim, pesquisadores requerem a utilização de imagens histológicas, que têm por objetivo fornecer informações sobre as modificações morfológicas nas células germinativas de uma espécie [17]. As células germinativas, mais conhecidas como ovócitos/oócitos, são estruturas que têm forma, geralmente esférica, sendo constituída por células foliculares, envelope vitelínico, grânulos de vitelo, citoplasma e um núcleo [18].

Várias técnicas têm sido desenvolvidas para a detecção de ovócitos em imagens histológicas, a exemplo dos trabalhos de Anta et al. [19], Cernadas et al. [20], Pérez-Ortiz et al. [21] e Pintor et al. [22], que propõem a extração de recursos na imagem, utilizando respectivamente os classificadores SVM, KNN, Adaboost e MLP, com o objetivo de reconhecer e classificar ovócitos. No entanto, destaca-se que trabalhos que fazem uso de técnicas de aprendizagem profunda, até o momento, não foram amplamente explorados para a análise de amostras de ovócitos de peixes.

Todavia, a detecção de objetos é uma importante tarefa de visão computacional que, ao utilizar imagens digitais, visa identificar e localizar objetos de uma determinada classe [23]. Para esta finalidade, técnicas que abordam redes neurais convolucionais têm sido utilizadas. Dentre estas técnicas, cita-se o YOLO (You Only Look Once), desenvolvido por Joseph Redmon em 2016 [5], que consiste em um método capaz de identificar e localizar objetos sem a necessidade de uma etapa preliminar, tendo como vantagem, a velocidade com que é capaz de fazer previsões rapidamente, permitindo o seu uso em tempo real. O algoritmo possui uma grande capacidade de generalização, permitindo ser utilizado em domínios diferentes para o qual foi treinado [8], dentre eles a identificação de ovócitos.

Deste modo, é de significativa importância o desenvolvimento de abordagens que realizem a detecção de ovócitos de forma automática, visto que distingui-los em imagens histológicas é um procedimento demorado e suscetível a erros, pois requer a interpretação visual e subjetiva do especialista.

Por esse motivo, este trabalho propõe uma abordagem capaz de detectar ovócitos de forma automática. Para isso, uma base de imagens foi construída, sendo esta dividida em treino, teste e validação. Além disso, técnicas para o aumento de dados

foram empregadas a fim de tornar o modelo mais robusto. Para a detecção dos ovócitos as arquiteturas YOLOv3, YOLOv4 e YOLOv5 foram aplicadas, permitindo assim, a detecção automática dos ovócitos e sua classificação em diferentes fases.

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo deste trabalho é desenvolver uma metodologia para detectar automaticamente ovócitos em imagens histológicas utilizando redes neurais convolucionais.

1.2.2 Objetivos Específicos

- Realizar uma revisão dos trabalhos presentes na literatura aplicados à detecção de ovócitos em imagens histológicas;
- Levantar conjuntos de imagens de ovócitos da espécie *Centropomus undecimalis*, sendo estes adquiridos por meios laboratoriais;
- Efetuar o treinamento de diferentes arquiteturas de redes YOLO aplicando como entrada imagens de ovócitos;
- Avaliar e comparar os resultados, de modo a selecionar as melhores arquiteturas de detecção.

1.3 Organização da Dissertação

Este trabalho, incluindo a introdução, está estruturado em seis capítulos:

- **Capítulo 2 - Trabalhos Relacionados:** apresenta de forma resumida os trabalhos mais relevantes que abordam a detecção de ovócitos em imagens histológicas;
- **Capítulo 3 - Fundamentação Teórica:** apresenta o embasamento teórico necessário para o entendimento e desenvolvimento da metodologia proposta;

- **Capítulo 4 - Metodologia Proposta:** apresenta as etapas realizadas de forma detalhada para a detecção dos ovócitos em imagens histológicas;
- **Capítulo 5 - Resultados e Discussão:** apresenta e discute os resultados obtidos pela metodologia proposta;
- **Capítulo 6 - Conclusão:** apresenta as considerações finais acerca do trabalho realizado, as contribuições do mesmo e as propostas de trabalhos futuros.

2 Trabalhos Relacionados

Neste capítulo são apresentadas descrições sobre trabalhos que abordam a detecção de ovócitos em imagens histológicas. Embora exista uma ampla quantidade de trabalhos na literatura, pesquisadores em sua maioria fazem uso de técnicas baseadas em SVM, ANN, MLP, Adaboost, KNN, etc. Por outro lado, pesquisas envolvendo a aprendizagem profunda vem sendo amplamente discutida, embora não abordem até o momento para a detecção de ovócitos de peixe. Neste caso, serão fornecidos relatos de pesquisas que possuem uma forte relação com a técnica de aprendizagem profunda para detecção de células e tecidos em imagens histológicas.

Na pesquisa de Anta et al. [19] propuseram um sistemas que reconhece automaticamente ovócitos de peixes em imagens histológicas. O conjunto de dados utilizados pela pesquisa continha 8 imagens. Para a avaliação da ferramenta, um especialista rotulou 451 ovócitos, tendo uma variação de 46 a 69 ovócito por imagem. Em seguida, um algoritmo de segmentação automático foi aplicado, sendo este, composto por quatro etapas: geração de sementes, cálculo de limite, crescimento de região e pós-processamento. No trabalho em questão, a taxa de instâncias corretas foi de 74%. Ao combinar informações de região fornecidas pelo algoritmo de crescimento de região e algoritmo baseado em borda, o método foi capaz de segmentar uma quantidade maior de células, atingindo 82% de detecção correta com um desvio padrão entre as imagens de 0.82 ± 0.13 .

O trabalho proposto por Cernadas et al. [20] tem como objetivo desenvolver um sistema, que reconheça e classifique ovócitos em imagens. O conjunto de dados continha 10 imagens histológicas de peixe da espécie *Caturilho*. Em um primeiro momento, os ovócitos de interesse são reconhecidos através de uma abordagem baseada na detecção de bordas. Para isso, o algoritmo realiza quatro etapas: captura da imagem histológica, calcula um conjunto de bordas candidatas, analisa as arestas, e mescla as informações obtidas. Uma vez que as células na imagem são reconhecidas, fez-se necessário a discriminação entre as células com núcleo e sem núcleo. Para esta etapa, o algoritmo baseado em técnicas de crescimento de região foi aplicado. Os resultados obtidos a partir de um conjunto de imagens apresentaram uma classificação

correta à uma taxa de 65%, embora, neste trabalho, utilizou-se uma base contendo poucas amostras.

Em Pérez-Ortiz et al. [21] foi proposto uma detecção dos estádios de desenvolvimento de ovócitos de peixes em imagens histológicas, através do software Govocito e de técnicas de classificação ordinais e nominais. Neste trabalho, foram obtidas 1022 e 912 células da espécie *Merluccius* (MC) e *Trisopterus Luscus* (TL), ambos com 3 estádios de desenvolvimento (cortical alveolar (CA), hidratado (HD) e vitelogênese/atrésico (VI/AT)) com 15 imagens. Além disso, foram selecionadas 7.915 células da espécie *Reinhardtius Hippoglossoides* (RH) com 6 estádios de desenvolvimento (crescimento primário, cortical alveolar (CA), vitelogênese (VI1, VI2, VI3 e VI4)) com 16 imagens. Inicialmente, o contorno das células foi desenhado manualmente por um especialista e classificado através do software Govocitos. Onze (11) abordagens ordinais foram testadas: método linear, diferentes métodos baseados em SVM: SVORIM, SVmod, REDSVM e EPSVM; um método baseado em análise discriminante: KDLOR; dois métodos baseados em Redes Neurais Artificiais: ELMOR e ONN; ORBoost, ORBoostP e SVR. Para a classificação, foram extraídas 25 características de cada ovócito, incluindo 10 de textura em escala de cinza e 15 de cores. Os experimentos demonstraram que os classificadores ordinais apresentaram maior robustez e desempenho em comparação com os métodos nominais para todas as espécies consideradas: SVMOD alcançou uma precisão de 94% e 95% para a espécie MC e TL e 80% para a espécie RH.

O trabalho proposto por Gonzalez-Rufino et al. [24] teve como objetivo avaliar a capacidade de classificadores, combinando-os com várias características de textura, a fim de determinar a classe de cada ovócito (com núcleo e sem núcleo), além dos estádios de desenvolvimento (alveolar cortical, hidratado e vitelogênese/atrésico). Nesta pesquisa, foram utilizadas 20 imagens histológicas com um total de 522 ovócitos. Em um primeiro momento foram extraídas regiões contendo os ovócitos, utilizando o algoritmo de retângulo máximo de Van-Devoorde. Em seguida foram utilizadas as características de textura: First Order Statistics (FOS), que fornece informações dos níveis de cinza distribuídos na imagem; Haralick Coefficients (HC), que fornece informações sobre as posições relativas dos níveis de cinza da imagem; Grey Level Run Length Statistics (GLRLS), que fornece um conjunto de pixels consecutivos na imagem com os mesmos níveis de cinza; Neighboring Grey Level

Dependence Statistics (NGLDS), que considera os relacionamentos entre um elemento e todos os seus elementos vizinhos em um determinado tempo, e Wavelet Transform (WT), que utiliza filtros multiescala para fornecer informações contidas na imagem no domínio da frequência. Para a classificação as classes (com núcleo e sem núcleo) e estádios dos ovócitos foram rotulados por um biólogo experiente. Em seguida, as características extraídas foram inseridas nos classificadores: K-Nearest Neighbors (KNN), Adaboost, Support Vector Machine (SVM) e MultiLayer Perceptron (MLP). Os resultados obtidos pela pesquisa mostraram que para a identificação das classes as técnicas KNN, Adaboost, SVM e MLP obtiveram respectivamente 77.3%, 73.4%, 75.6% e 78.3% de acurácia. Com relação aos estádios maturacionais as técnicas KNN, Adaboost, SVM e MLP obtiveram respectivamente 85.3%, 88.5%, 85.9% e 83.9% de acurácia.

No trabalho de Pintor et al. [22] foi aplicada a ferramenta Govocito com o objetivo de estimar o número e o tamanho dos ovócitos em diferentes estágios de desenvolvimento de peixes fêmea. Imagens das espécies *Merluccius merluccius* (pescado) e *Lutjanus synagris* (pargo) foram adquiridas. Um conjunto de 61 imagens foram analisadas, sendo 31 imagens do peixe pescado e 30 imagens do peixe pargo. Um total de 2405 ovócitos foram obtidos (1186 para pescada e 1219 para pargos). Na etapa de reconhecimento os ovócitos foram segmentados automaticamente, separando suas estruturas em ovócitos maturados do fundo da imagem. Foi aplicado um filtro Canny multi-escalar para a detecção dos contornos dos ovócitos. Em seguida, um algoritmo de detecção de bordas não supervisionado foi aplicado em duas etapas: (1) a imagem é processada por um filtro Canny multi-escalar e (2) uma análise de bordas determina quais bordas são candidatas a representar um ovócito verdadeiro ou falso. Posteriormente, o software classifica automaticamente os ovócitos de acordo com a presença ou ausência do núcleo, e de acordo com seus estágios de desenvolvimento (alvéolos corticais, hidratado, vitelino, e atrésico). Em relação ao algoritmo de detecção, o número médio de ovócitos detectados corretamente pelo técnica SVM foi de 80%. Em relação as classes (com ou sem núcleo) o classificador obteve uma acurácia de 84% e para a classificação dos estágios de desenvolvimento o classificador obteve 87,1%.

Gonzalez-Rufino et al. [25] propuseram a classificação de ovócitos de peixe de acordo com a presença ou ausência de núcleo, e seus estádios de desenvolvimento

através da análise de textura e cor em imagem histológica. Para a validação da pesquisa, 47 imagens histológicas de 12 peixes da espécie *Merluccius merluccius* foram utilizadas. No total 1022 células foram extraídas para a classificação. Sua distribuição por classe foi: 337 células com núcleo e 685 sem núcleo. Para os estádios de desenvolvimento: 259 estão no estágio cortical alveolar (AC), 61 estão no estágio hidratado (HD) e 702 estão no estágio vitelogênese/atrésico (VII/AT). A base foi dividida em um conjunto de 80% para treino, 10% para validação e 10% para teste, selecionados aleatoriamente. As características extraídas foram compostas por textura e cor, sendo: Statistical Colour Features (SCF), Grey Level Dexture Descriptors (GLTD), Multiscalar Grey Level Texture Feature (MGLTF) e Integrative Colour Texture Analysis (ICTA). Além disso, um total de 17 classificadores foi utilizado, dentre esses, a técnica (SVM) alcançou os melhores resultados para identificação dos núcleos com 99,0% de precisão e para estágios de desenvolvimento com 99,6%.

Pintor et al. [26] apresentam um ferramenta (chamada Govocito) para detectar e classificar automaticamente ovócitos em diferentes estágios de desenvolvimento (cortical alveolar, hidratado, vitelogênese, e atrésico). A arquitetura do software é composta por três camadas: interface do usuário (GUI), lógica da aplicação e armazenamento e comunicação de dados. Para a segmentação foram aplicadas as técnicas multi-scalar Canny, Sorbel, Prewitt, Roberts, Laplaciano, Laplacian of Gaussian (LOG) e Diferencial of Gaussian (DoG). Para a classificação, o Govocito utilizou o algoritmo Gaussian Kernel Support Vector Machine (GKSVM), recebendo com entrada: estatísticas de primeira ordem da imagem RGB colorida (12 características) e padrões binários locais (10 características). As espécies de peixes utilizada para a pesquisa foram o pescada (*Merluccius merluccius*) e o pargo (*Lutjanus synagris*). Um total de 2405 ovócitos foram utilizados, sendo 1186 de pescada (31 imagens) e 1219 de pargo (30 imagens). Em relação à classificação dos ovócitos, a acurácia para as classes (com/sem núcleo) foi de 83,8 % em média para ambas as espécies e 87,1 % para a classificação do estágio de desenvolvimento.

Em Inik et al. [27] propõem um método que envolve a segmentação e classificação aplicada a imagens histológicas de folículos de tecido ovariano, com base em uma rede neural convolucional. O método de contagem de folículos consiste em duas etapas. Na primeira etapa, a segmentação e a localização dos folículos são determinadas. Na segunda, os folículos determinados são classificados, utilizando

a rede Faster R - CNN. Neste, a segmentação e a detecção de folículos ovarianos são realizadas em quatro etapas: obtenção das subimagens de tamanho 32x32 pixels, segmentação, limpeza dos ruídos e detecção. Para a segmentação, um total de 1972 imagens foram obtidas com ampliação de 10, 20 e 40 vezes. Para a classificação 1750 imagens de folículos primordiais, primários, pré-antrais, secundários e terciários foram utilizadas para treinamento. O método proposto atingiu níveis de acurácia na contagem de folículos do tipo primordial de 86,67%, primário: 95,35%, pré-antral: 97,06%, secundário: 97,69% e terciário: 100%. De modo geral, a precisão média do método proposto foi de 95,35%.

No trabalho de Sirinukunwattana et al. [28] uma Rede Neural Convolutacional Espacialmente Restrita foi utilizada para realizar a detecção de núcleos em imagens histológicas. Foram avaliados um conjunto de 100 imagens histológicas de adenocarcinoma colorretal com tamanho comum de 500 × 500 pixels. Um total de 29.756 núcleos foram marcados por um especialista para fins de detecção. Destes, 7.722 são núcleos epiteliais, 5.712 são núcleos fibroblastos, 6.971 são núcleos inflamatórios e 2.039 são núcleos diversos. Na detecção dos núcleos a rede obteve uma recall de 82.7%, F1-score de 80.2% e uma precisão de 78.3%.

Em Xu et al. [29] foi aplicado quatro arquiteturas de redes neurais convolucionais: Convolutional Neural Pixel-wise (CN-PI), Convolutional Neural Patch-wise (CN-PA), Convolutional Neural Simple Linear Iterative Clustering (CN-SLIC) e Convolutional Neural Normalized Cuts (CN-Ncult), com objetivo de segmentar e classificar imagens provenientes das instituições Netherlands Cancer Institute (NKI) e Vancouver General Hospital (VGH). Para a identificação das atípia nuclear, as imagens foram divididas em três datasets com ampliações diferentes: 10x, 20x, 40x, cujos os tamanhos utilizados foram respectivamente, 769x688, 1539x1376 e 3078x2752 pixels. Enquanto que, para a segmentação dos tecidos mamários epiteliais e estromais as redes foram treinadas com um conjunto de imagens com ampliação de 20x e tamanho 1128x720 pixels. Dentre as arquiteturas utilizadas, o modelo que apresentou melhor desempenho foi a CN-PI que obteve 94.30% de acurácia e uma área da curva ROC de 0.9694, para imagens da instituição VGHL.

Em Chang et al. [30] foi utilizada uma abordagem para classificação de núcleos baseada em aprendizagem profunda. Em um primeiro momento é realizada a segmentação com o objetivo de extrair características morfológicas na imagens,

através de um conjunto de filtros de Gabor, com diferentes frequências e orientações. Em seguida, cada pixel é agrupado conforme sua vizinhança. Por fim, para obter os núcleos de forma individual aplicou-se operações morfológicas matemáticas. As imagens resultantes são geradas com dimensão de 52 x 52 pixels centradas no núcleo. A fim de melhorar a generalização da rede, foi utilizado técnicas de aumento de dados, ou seja, as imagens geradas foram submetidas a diferentes operações aleatórias: rotação, inversão e adição de ruído. Com isso, 99.676 células cancerosas e 15.566 células normais selecionadas para o treinamento e 36.000 células para o teste. A arquitetura utilizada da rede conta com 3 camadas convolucionais, duas camadas de max-pooling sem sobreposição e quatro camadas totalmente conectadas com 31 neurônios. Esta arquitetura atingiu 91,3% de acurácia, 89,9% de sensibilidade e 92,8% de especificidade.

Roy et al. [31] desenvolveram um método baseado em uma rede neural convolucional para a classificação em imagens histológicas de mama, divididas em quatro classes: normal, benigno, in situ e carcinoma invasivo; e duas classe: maligno e benigno. O conjunto de imagens utilizados foram obtidos a partir da Conferência Internacional de Análise e Reconhecimento de Imagens - ICIAR 2018. A base contém 400 imagens para fins de treinamento e 100 imagens para teste. As imagens de entrada possuem tamanho de 2048x1536 pixels. Para a extração, patches individuais de tamanho quadrado não sobreposto de 512x512 pixels foram extraídos em cada imagem, resultando em um total de 12 patches em uma única imagem. O aumento do conjunto de dados foi utilizado visando expandir o conjunto de treinamento, sendo este, através da translação, rotação e inversão dos patches. Na pesquisa uma Rede Neural Convolucional hierárquica foi implementada contendo as seguintes configurações: seis camadas de convolução, cinco camadas de max-pooling e três camadas completamente conectadas. Resultados experimentais mostram que a técnica proposto alcançou uma precisão de 90% para 4 classes e 92,5% para 2 classes.

Liu et al. [32] propuseram um método de classificação de imagens histológicas de câncer de mama baseado em Redes Neurais Convolucionais, nomeado DeepBC. Foi utilizada 7.909 imagens histológicas de mama em diferentes ampliações (40x, 100x, 200x e 400x) da base BreakHist. Cada imagem possui dimensão de 700x460 pixels com 3 canais RGB. As imagens foram divididas em duas classes: maligno e benigno. Na pesquisa, 70% das imagens foram empregadas para treinamento,

enquanto que os 30% restantes foram utilizadas para teste. As imagens histológicas de entrada possuem um tamanho de 227x227 pixels. Os resultados na avaliação do método mostraram que o DeepBC alcançou taxas de acurácia de 92% para a classe maligna e 96,43% para a classe benigna.

No trabalho de Golatkar et al. [33] um método baseado em aprendizagem profunda foi aplicado para classificação de imagens de tecido mamário por meio de uma rede neural convolucional Inception-v3. As imagens foram obtidas pela Breast Cancer Histology Challenge (BACH) de 2018. Esta base, consiste em imagens histológicas de alta resolução em cores RGB e tamanho 2048x1536 pixels e são rotuladas em quatro (normal, benigna, in situ e invasivo) e duas (carcinoma e não carcinoma) categorias. O conjunto de dados contém 100 imagens. Nas imagens são extraídos patches de tamanho 299x299 pixels centradas no núcleo. Além disso, com o objetivo do aumento destes dados, os patches são invertidos horizontalmente, verticalmente e deslocados em 180 graus. Neste trabalho, foram obtidos uma acurácia de 85% para as quatro classes e 93% para as duas classes.

2.1 Considerações sobre os trabalhos relacionados

Os trabalhos apresentados neste capítulo, embora utilizem técnicas e abordagens distintas, apresentaram resultados significativos na identificação de objetos em imagens histológicas. Dessa forma, a seguir são ressaltadas algumas de suas particularidades, além disso, é apresentado na Tabela 2.1 um resumo dos trabalhos relacionados:

- Na pesquisa realizada por Pérez-Ortiz et al. [21] a identificação dos ovócitos foi preliminarmente realizada de forma manual, sendo esta, uma avaliação custosa e suscetível a erros.
- Algumas abordagens como a de Anta et al. [19], Cernadas et al. [20], Pérez-Ortiz et al. [21], Gonzalez-Rufino et al. [24], Pintor et al. [22], Gonzalez-Rufino et al. [25], Pintor et al. [26], Sirinukunwattana et al. [28] e Golatkar et al. [33] demonstraram a utilização das respectivas técnicas por meio de uma quantidade limitada de amostra de imagens histológicas. Dessa forma, a validação torna-se restrita.

Tabela 2.1: Resumo dos trabalhos relacionados referentes a identificação de objetos em imagens histológicas.

Trabalhos	Imagens	Técnicas	Características
Anta et al. [19]	8	Geração de sementes Crescimento de região	—
Cernadas et al. [20]	10	Deteção de bordas Crescimento de região	—
Pérez-Ortiz et al. [21]	31	SVMOD	Textura em escala de cinza e cor
Gonzalez-Rufino et al. [24]	20	KNN, Adaboost, SVM e MLP	Textura em escala de cinza, cor e domínio da frequência
Pinfior et al. [22]	61	SVM	Filtro de borda
Gonzalez-Rufino et al. [25]	41	SVM	Textura e cor
Pinfior et al. [26]	61	GKSVM	Textura, cor, filtros de borda e domínio da frequência
Inik et al. [27]	3,722	Faster R-CNN	—
Sirinukanwattana et al. [28]	100	CNN	—
Xu et al. [29]	51	CN-PI	—
Chang et al. [30]	150,000	CNN	—
Roy et al. [31]	100	CNN	—
Liu et al. [32]	7,909	DeepFC	—
Golatkar et al. [33]	100	Inception-v3	—

- As Técnicas de aumento de dados são mecanismos necessários para a generalização de um modelo baseado em aprendizagem profunda, visto que, estas criam novas imagens a serem utilizadas na etapa de treinamento [34]. Entretanto, somente os trabalhos de Roy et al. [31], Chang et al. [30] e Golatkar et al. [33] aplicam técnicas de redes neurais convolucionais em conjunto com mecanismos de aumento de dados.

- Para realizar a identificação dos ovócitos, nos trabalhos de Gonzalez-Rufino et al. [25], Pérez-Ortiz et al. [21] e Gonzalez-Rufino et al. [24], foram empregadas uma grande quantidade de técnicas. Essa diversidade de operações torna estas técnicas mais complexas e, conseqüentemente, mais custosas computacionalmente.
- Ao utilizar inúmeras técnicas de extração de características, a exemplo dos trabalhos de Gonzalez-Rufino et al. [25], Pérez-Ortiz [21], Gonzalez-Rufino et al. [24] e Pintor et al. [26], surge a necessidade da seleção de quais destas características melhor definem o objeto no processo de identificação. Diferentemente das técnicas citadas, o uso de redes neurais convolucionais é capaz de extrair implicitamente características com o objetivo de detectar e classificar objetos, independentemente de pose, escala e iluminação [35].
- Uma peculiaridade encontrada em todos os trabalhos foi o fato de não realizarem a localização dos objetos na imagem. Uma alternativa capaz de, ao mesmo tempo, detectar e localizar objetos é o emprego da abordagem YOLO.
- As redes neurais convolucionais possuem uma ampla capacidade de generalização, possibilitando assim classificar objetos em imagens nos mais variados domínios. Dessa forma, esse tipo de abordagem tem se tornado cada vez mais frequente em pesquisas que buscam identificar objetos.

Diante disso, é proposta uma abordagem para detectar automaticamente ovócitos da espécie *C. undecimalis* utilizando diferentes arquiteturas da rede YOLO.

3 Fundamentação Teórica

Neste capítulo é apresentado o referencial teórico utilizado para a elaboração desta pesquisa. As seções seguintes abordam o peixe *Centropomus undecimalis* - Robalo Flecha, Redes Neurais Convolucionais, a técnica YOLO e as métricas de desempenho utilizadas para avaliação da metodologia proposta.

3.1 *Centropomus undecimalis* - Robalo Flecha

A espécie *Centropomus undecimalis* Bloch 1792 é conhecida por robalo-flecha ou camorim e, nos países de língua inglesa, por "common snook" [36], pertencem à família Centropomidae, da ordem dos Perciformes. Apresentam distribuição tipicamente tropical e subtropical, exclusivamente no litoral do continente americano e são encontradas desde o sul da Flórida ao o sul do Brasil [37]. São organismos eurihalinas, podem ser encontradas tanto no mar como nas águas salobras estuarinas, lagunas, desembocadura de rios e lagoas de água doce, com acesso temporário ou permanente [17].

A abundância deste peixe é modificada pelo período reprodutivo, ou seja, os indivíduos maduros que se encontram em rios e lagos da água doce regressam ao mar para desovar (catádromos), depois da desova permanecem um temporada no mar para regressar logo depois ao ambiente estuarino [38]. Sua adaptação a diferentes habitats e salinidade os caracterizam como peixe rústicos e gregários que aceitam bem o cativeiro [17]. São carnívoros, alimentam-se principalmente de peixes de pequeno porte, crustáceos e moluscos.

Possui corpo alongado e comprido, o que o torna um nadador extremamente ágil. Tem boca oblíqua com vários dentes finos e maxilar inferior que ultrapassa nitidamente o superior. O dorso apresenta coloração cinzento-prateado, com reflexos esverdeados, e ventre esbranquiçado. Apresenta uma linha lateral formada por um listra longitudinal escura que se estende ao longo do corpo até o final da nadadeira caudal. As nadadeiras são amareladas, as vezes com pontos pretos,

podendo apresentar uma faixa escura na linha lateral [37] [38]. Na fase adulta o robalo-flecha pode alcançar 25kg e 1,4m de comprimento total [39]. Estas características podem ser observadas na Figura 3.1.

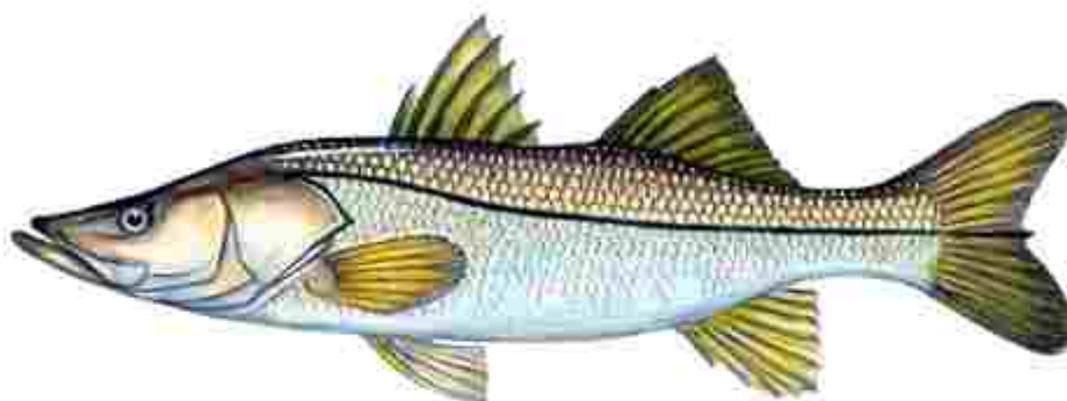


Figura 3.1: *Centropomus undecimalis* - Robalo Flecha. Fonte: Rodrigues [1]

Segundo Sousa [40] uma das características sexuais mais notáveis dos *C. undecimalis* é o fato de consistirem em espécies hermafroditas protândricas, que maturam primeiramente como machos e posteriormente, mudam de sexo e permanecem, em sua maioria, como fêmeas pelo resto de suas vidas. Esta característica segundo Taylor [41] foi identificada neste peixe por meio da presença simultânea na gônada de lamelas ovígeras e tecido espermático degenerativo. Além disso, as fêmeas tornam-se sexualmente maduras com três anos de idade. Sua reprodução é externa, as larvas eclodem em 36 horas e em três dias já consumiram toda a reserva vitelínica, iniciando-se a alimentação exógena [42].

3.1.1 Desenvolvimento Ovariano / Morfologia dos Ovários

As gônadas são estruturas pares, alongadas, apresentando lóbulos semelhantes em tamanho e estão situadas dorsalmente na cavidade corporal, relacionadas látero ventralmente com a bexiga natatória, sendo ligados entre si através do peritônio, com a extremidade anterior (cefálica) livre, ligando-se na porção posterior final (anal) para formar o oviduto, que se abre para o exterior através do poro urogenital. Apresentam coloração esbranquiçada e irrigação reduzida na fase imatura. Quando em início de maturação, assumem forma cilíndrica, apresentando uma coloração amarelo-clara, com vascularização mais evidente; e, em final de maturação, os ovócitos são vistos a olho nu, apesar do seu diminuto tamanho. O estroma de

sustentação no interior do órgão comporta lamelas ovulíferas, onde estão localizadas as células germinativas ou ovócitos em diferentes estádios de maturação, conforme pode ser observado na Figura 3.2 [43] [2].



Figura 3.2: Ovários em diferentes estádios de desenvolvimento: a) Estádio A - Imaturo, b) Estádio B - Em maturação, c) Estádio C - Maduro, d) Estádio D - Esvaziado, e) Estádio E - Repouso. Fonte: Félix et al. [2]

Segundo Mendonça [17], em estudos de reprodução de peixes normalmente faz-se necessário o conhecimento das fases de desenvolvimento gonadal da espécie, sendo este baseado em inspeção visual da aparência externa da gônada. Uma análise mais detalhada e precisa requer a utilização de métodos histológicos, que fornecem informações sobre as modificações celulares de cada espécie.

Classificação Macroscópica dos Estádios Maturacionais

Uma vez aberta a cavidade celomática e identificado o sexo, as gônadas devem ser observadas, com o objetivo de se avaliar seu grau de desenvolvimento, ou seja, seu estágio de maturidade [44]. Sendo assim, nos estudos iniciais da reprodução desta espécie, utiliza-se uma escala de aplicação geral composta, basicamente, de 5 estádios: Estádio A - Imaturo, Estádio B - Maturação inicial, Estádio C - Maturação avançada, Estádio D - Maduro e Estádio E - Repouso [40] [17] [2], conforme observado na Figura 3.2.

Estádio A - Imaturo: os ovários são muito pequenos, ocupando menos de 1/3 da cavidade celomática; são filamentosos, translúcidos, sem sinais de vascularização; não se observam ovócitos a olho nu [44]. Possuem ovogônias¹ e

¹Pequenas células localizadas na periferia das lamelas ovíferas, formando grupos celulares ou isoladamente, possui citoplasma escasso, núcleo grande e nucléolo em posição central.

ovócitos pré-vitelogênicos (cromatina-nucleolar², perinuclear inicial³ e final⁴); as ovogônias e cromatina-nucleolar apresentam-se isoladas ou em forma de ninhos [17].

Estádio B - Em Maturação: os ovários são maiores, ocupando de 1/3 a 2/3 da cavidade celomática, intensamente vascularizados, aproximando-se mais do poro genital, sendo que o oviduto apresenta-se como uma lâmina delgada, em forma de tubo, transparente e vazia. A olho nu observam-se ovócitos opacos, pequenos e médios [44]. Possui uma grande quantidade de ovócito perinuclear inicial e final e ovócitos cortical-alveolares⁵ [17].

Estádio C - Maduro: os ovários apresentam-se túrgidos, ocupando de 2/3 a, praticamente, toda a cavidade celomática, sendo visível um grande número de ovócitos grandes opacos e/ou translúcidos que podem ocupar, inclusive, os ovidutos; sua vascularização, inicialmente, é reduzida e, no final, torna-se imperceptível [44]. Observa-se também a predominância de ovócitos em vitelogênese final e inicial, os ovócitos encontram-se grandes, esféricos, repletos de grânulos de vitelo e ovócitos em hialinização⁶, prontos para a desova. O tecido espermatogênico torna-se apenas vestigial [17].

Estádio D - Esvaziado: os ovários apresentam-se flácidos, com membranas distendidas, de tamanho relativamente grande mas não volumosos, ocupando menos da metade da cavidade celomática; observam-se poucos ovócitos, em estado de absorção, muitas vezes formando grumos esbranquiçados; a característica mais marcante é a presença de zonas hemorrágicas [44]. Apresenta aspecto desorganizado com as lamelas ovigeras rompidas, distendidas, causando espaços vazios entre si. Ocorrem restos foliculares, ovócitos

²São ovócitos que se separam do grupo celular, o citoplasma é bem definido, apresenta 1 a 2 nucléolos que migram para a periferia do núcleo.

³Perinucleolar inicial - Os ovócitos apresentam-se poliédricos, com muitos núcleos dispostos aleatoriamente na periferia da membrana nuclear, o citoplasma encontra-se fortemente basófilo.

⁴Perinucleolar final - Os ovócitos apresentam-se com tamanhos variados, com núcleo grande e nucléolos dispersos em seu interior. Nas células maiores, o citoplasma apresenta um aspecto granular.

⁵Os ovócitos apresentam o surgimento de alvéolos corticais no citoplasma. O núcleo encontra-se com todos os nucléolos dispersos periféricamente; os alvéolos aumentam de tamanho e número, até formar várias filas de citoplasma.

⁶Apresentam aumento em seu volume devido à incorporação de água que ocorre em seu interior, os grânulos de vitelo se tornam maiores e amorfos e as células foliculares tornam-se achatadas.

vitelogênicos, ovócitos atrésicos⁷ e uma discreta quantidade de ovócitos pré-vitelogênicos [17].

Estádio E - Repouso: os ovários apresentam tamanho reduzido, ocupando cerca de 1/3 da cavidade celomática, sendo claramente maiores que os imaturos (A); são translúcidos, com fraca vascularização, não se observando ovócitos a olho nu. Histologicamente, são bastante similares àqueles imaturos (Estádio A); ocorrem apenas células germinativas jovens. As lamelas ovígeras são mais longas devido ao aumento de volume sofrido no decorrer do ciclo anterior [44].

Fases de Desenvolvimento Ovocitário

O ovário consiste de lamelas ovulíferas, contendo inúmeros folículos que cobrem os ovócitos em seus diferentes estádios. Segundo Mendonça [17] e Félix et al. [2], podem ser identificadas seis fases de desenvolvimento ovocitário: cromatina-nuclear, perinuclear inicial, perinuclear final, cortical-alveolar, vitelogênese inicial, vitelogênese final e hialinizado. Entretanto, devido a fatores físicos, químicos, ambientais e precipitação pluviométrica influenciarem diretamente no ciclo reprodutivo da espécie [17], as proporções de ovócitos não foram completamente abundantes para contemplar todas as fases dos ovócitos. Neste caso, são descritas neste trabalho apenas as que apresentam predominância nas amostras obtidas para a pesquisa.

Pré-vitelogênese: (células nas fases cromatina-nuclear, perinuclear inicial, perinuclear final e cortical-alveolar) separam-se do agrupamento em "ninho", provavelmente, devido ao aumento de seu volume. O citoplasma é bem definido e mais basófilo que na fase anterior; o núcleo apresenta, inicialmente, nucléolos esféricos e intensamente basófilos, que se vão tornando mais numerosos e volumosos, migrando para a periferia nuclear. Estão presentes em gônadas em todos os estádios de maturidade, sendo inicialmente arredondados e, posteriormente, devido à pressão de uns contra os outros aparecem triangulares,

⁷A atresia folicular é um processo degenerativo onde ovócito de vários estágios de desenvolvimento são absorvidos pelo ovário. Assume contornos irregulares; os grânulos de vitelo, quando presentes, formam uma massa amorfa.

retangulares ou ovais. Destes ovócitos do "estoque de reserva" irão originar-se as populações celulares que vão iniciar a vitelogênese (Figura 3.3).



Figura 3.3: Ovócitos em Pré-vitelogênese com crescimento do núcleo contendo nucléolos proeminentes.(Aum. 200X). Fonte: A autora.

Vitelogênese inicial: a vitelogênese, processo celular pelo qual o citoplasma acumula substâncias de reserva para posterior utilização na alimentação do embrião, caracteriza-se por um acelerado crescimento citoplasmático. O núcleo cresce, mas não na mesma proporção do citoplasma, contendo vários nucléolos alongados e achatados (evidência de intensa síntese), localizados na periferia do núcleo, nem sempre bem identificados. Observa-se o surgimento de uma película acidófila contínua, que é a membrana vitelina, acoplada à membrana citoplasmática do ovócito. Esta aparece rodeada por uma camada de células foliculares, com núcleo de posição central, um nucléolo e citoplasma basófilo. O conjunto ovócito + membrana vitelina + camada de células foliculares é chamado de folículo ovocitário, que é a unidade fundamental do ovário. Estes ovócitos ocorrem em ovários em estágio de maturação inicial (Figura 3.4).

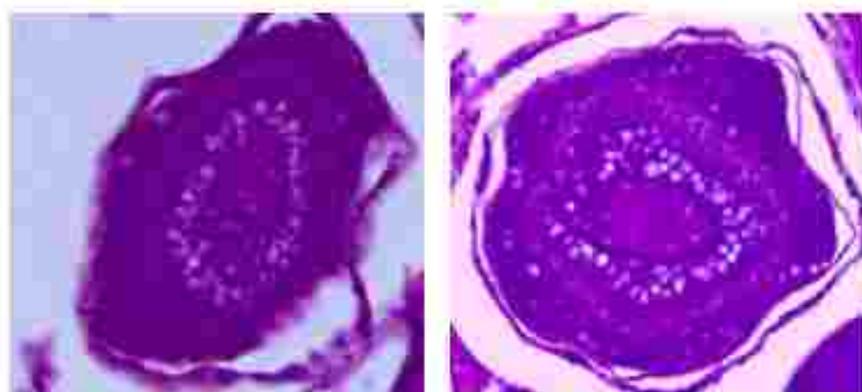


Figura 3.4: Ovócito em vitelogênese inicial, com vacúolos lipídicos, núcleo e nucléolos.(Aum. 200X). Fonte: A autora.

Vitelogênese final: o ovócito aumenta rapidamente de tamanho, em função do aumento acelerado do número de grânulos de vitelo acidófilo. Já não mais se percebem as vesículas lipídicas; os grânulos protéicos aumentam de tamanho, mudando totalmente o aspecto do ovoplasma. A basofilia desaparece quase totalmente. O núcleo mantém suas características, com nucléolos bem pequenos, perde sua forma esférica e contrai-se. A membrana vitelina pode tornar-se mais espessa e apresentar estrias de disposição radial, visíveis sob grandes aumentos, quando pode ser denominada de "zona radiata". Aparentemente, são as células mais maduras observadas na linhagem germinativa, ocorrendo em ovários em maturação avançada. Quando os ovócitos atingem esta fase inicia-se a desintegração e migração do núcleo para onde será o polo animal da célula (Figura 3.5).

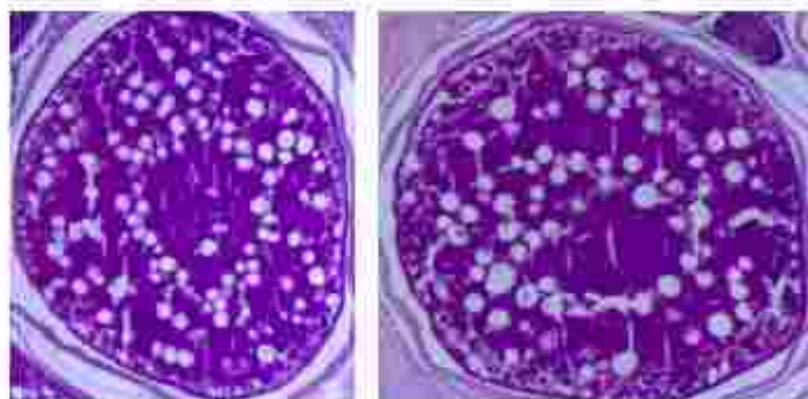


Figura 3.5: Ovócitos em vitelogênese final contendo grânulos protéico por todo o ovoplasma. (Aum. 200X). Fonte: A autora.

3.2 Redes Neurais Convolucionais

Rede Neural Convolucional, do inglês *Convolutional Neural Networks* (CNN) é uma rede biologicamente treinável que aprende características invariantes através de hierarquias multiníveis de características. A rede extrai implicitamente características de padrões visuais apresentados como entrada e os classifica em categorias a partir dos recursos extraídos, independentemente de pose, escala e iluminação [35].

Normalmente, as Redes Convolucionais são compostas por arquiteturas de treinamento em múltiplos estágios. Cada estágio é composto por três camadas: Uma camada convolucional, uma camada de subamostragem e uma camada

completamente conectada [45]. As camadas que realizam a convolução e a subamostragem, englobando o processo implícito de extração de características, localizam-se na base da rede. Enquanto que, no topo da CNN encontra-se a camada completamente conectada que é responsável pela classificação dos padrões de entrada. Uma estrutura típica de uma CNN é mostrada na Figura 3.6.

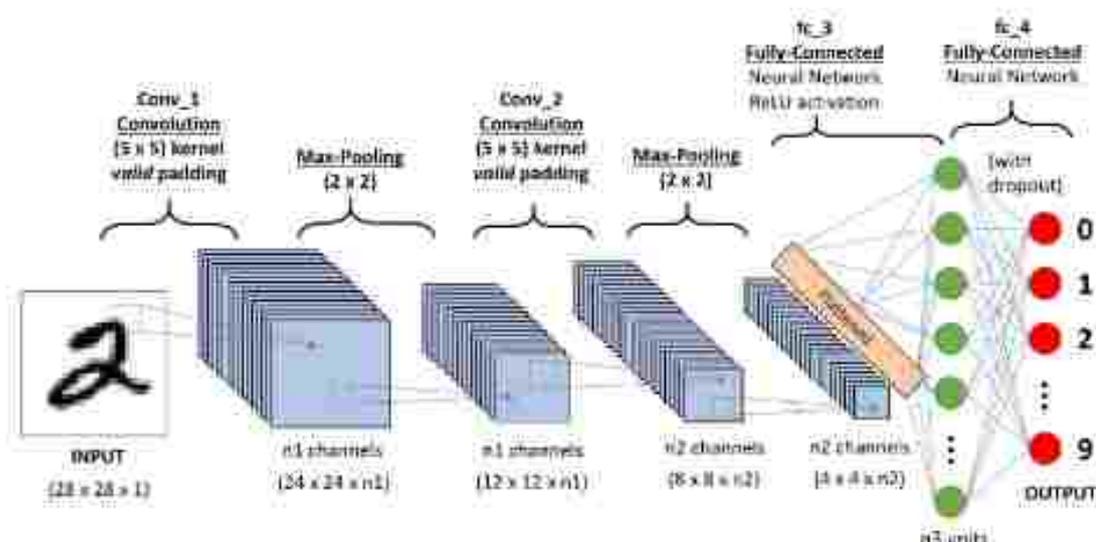


Figura 3.6: Rede Neural Convolutional para processamento de imagens. Fonte: Saha [3]

A camada Convolutional é responsável por extrair características através da imagem de entrada, como mostra a Figura 3.7. Uma operação de convolução é dada pelo deslocamento de uma matriz de tamanho pré-definido chamada 'filtro' ou 'kernel' sobre toda a imagem. Cada filtro detecta uma característica particular que ocorre em qualquer parte da entrada [46]. Esta camada é responsável por preservar a relação espacial entre os pixels aprendendo características individuais da imagem [47].

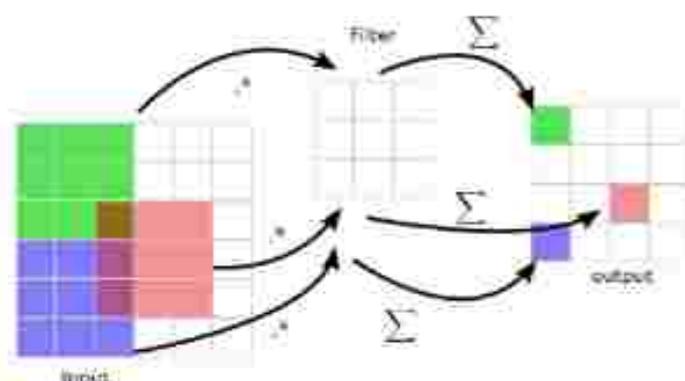


Figura 3.7: Ilustração da camada de convolução. Fonte: Hafemann [4]

A camada de subamostragem é responsável por reduzir progressivamente o tamanho do mapa de características usando alguma função para resumir sub-regiões, como o valor máximo ou médio, retendo as informações mais importantes. A subamostragem funciona percorrendo uma janela através da entrada e alimentando o conteúdo da janela para uma função de *pooling*. A Figura 3.8, fornece um exemplo para a camada de subamostragem [4].

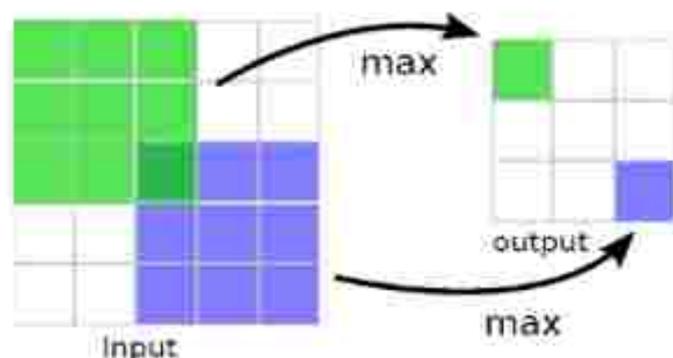


Figura 3.8: Ilustração da camada de subamostragem. Fonte: Hafemann [4]

Ao término das sequências de camadas de convolução e subamostragem que têm a tarefa de extrair características da imagem, os pixels de todos os mapas de características da camada anterior são dados como entrada às camadas completamente conectadas.

A camada totalmente conectada é um *MultiLayer Perceptron* (MLP) tradicional que usa uma função de ativação *softmax* na camada de saída [48]. A camada totalmente conectada é usada para classificar a imagem de entrada da CNN em várias classes com base no conjunto de dados de treinamento. O treinamento supervisionado é realizado usando uma forma de descida gradiente estocástica (SGD) para minimizar a discrepância entre a saída desejada e a saída atual da rede, com base em alguma função de perda [35]. Todos os pesos de todos os filtros em todas as camadas são atualizados simultaneamente com o algoritmo *backpropagation*.

Essas redes neurais treinadas com *backpropagation* admitem uma grande variedade de arquiteturas específicas aplicáveis a uma ampla gama de aplicações em imagens. Unidas, as camadas de uma CNN controlam as informações úteis das imagens, inserindo a extração de características, a não linearidade e a redução da dimensionalidade das imagens de forma a tornar essas características invariantes a

rotação, translação e escala [35]. Atualmente, diversas arquiteturas de CNN são desenvolvidas e publicadas para aplicações em diferentes problemas, sendo a detecção de objetos uma de suas principais tarefas no âmbito de visão computacional [49].

3.3 YOLO – You Only Look Once

O YOLO (You Only Look Once), desenvolvido por Joseph Redmon em 2016, consiste em um dos primeiros métodos para detecção de objetos em um estágio. O algoritmo baseia-se na arquitetura das Redes Neurais Convolucionais, onde através desta é feita simultaneamente a predição das caixas delimitadoras e suas respectivas classes, permitindo uma única avaliação da imagem. Esse tipo de abordagem possui grande capacidade de generalização, permitindo ser utilizado em domínios diferentes para o qual foi treinado [5].

A motivação de Redmon foi construir um modelo unificado de todas as fases em uma rede neural. Com a imagem de entrada contendo (ou não) os objetos, após passar por uma única rede neural de múltiplas camadas convolucionais, o sistema produz vetores preditivos correspondentes a cada objeto que aparece na imagem. Em vez de iterar o processo de classificação de diferentes regiões na imagem, o sistema YOLO calcula todos os recursos da imagem e faz previsões para todos os objetos ao mesmo tempo.

O YOLO funciona dividindo a imagem fornecida em uma grade de células com tamanho $S \times S$. Em seguida, para obter a detecção de objetos, cada célula da grade deverá prevê B caixas delimitadoras [5]. A caixa delimitadora é definida como duas coordenadas relativas à matriz da imagem correspondendo à posição central do objeto (x e y), e as duas dimensões de largura (w) e altura (h). Células e âncoras, em tais regiões, irão prever objetos em áreas específicas da imagem. Cada caixa delimitadora será acompanhada por uma “pontuação de objetividade” que definirá o grau de confiança do modelo em relação à caixa delimitadora que contém um objeto [8]. A Figura 3.9 ilustra o funcionamento do YOLO.

Houve cinco versões do YOLO publicadas até o momento. Cada versão foi atualizada e integrada com as ideias mais avançadas da comunidade de pesquisa em

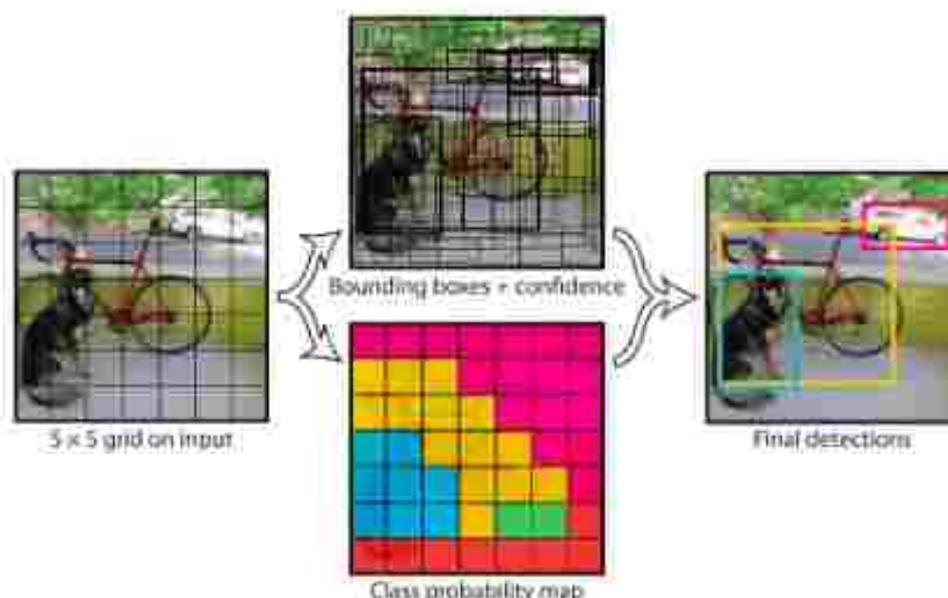


Figura 3.9: Sequência de detecção de objetos para a arquitetura YOLO. Fonte: Redmon et al. [5].

visão computacional. Além disso, algumas ideias também foram removidas devido à incapacidade de atingir o desempenho e a precisão necessários do algoritmo [10].

Neste trabalho, modelos de detecção baseados no YOLO foram aplicados com o objetivo de classificar e localizar os ovócitos em imagens histológicas. Foram selecionados para os experimentos o YOLO em suas versões 3, 4 e 5 devido ao fato de que esses modelos possuem a capacidade de identificar variações de objetos de mesma classe em diferentes dimensões, além de serem as versões mais recentes na literatura [50], [5] e [8].

3.3.1 YOLOv3

O YOLOv3 lançado em 2018 por Redmon e Farhadi [8] [51] consiste em um modelo dividido em dois componentes principais: um extrator e um detector de recursos. Para extrair recursos de uma imagem de entrada o YOLOv3 faz uso de um backbone com Darknet53. O backbone é uma rede neural profunda composta por uma camada de convolução cuja função é extrair características essenciais da imagem de entrada. Para realizar a detecção o YOLOv3 anexa as camadas de previsão ao lado da rede em vez de empilhá-la nas últimas camadas. A característica mais notável do modelo é que ele faz detecções em 3 escalas diferentes, sendo estas em dimensões de

52x52, 26x26 e 13x13 pixels. Por fim, como resultado o YOLOv3 prevê a classe dos objetos para cada caixa delimitadora usando regressão logística.

3.3.2 YOLOv4

Em abril de 2020, Alexey Bochkovskiy, pesquisador e engenheiro russo que construiu o framework Darknet e 3 arquiteturas YOLO anteriores em C com base nas ideias teóricas de Joseph Redmon, cooperou com Chien Yao e Hori-Yuan e publicou YOLOv4 [9] [10]. No YOLOv4 um ponto notável em sua estrutura é a divisão de seus componentes em três partes para a obtenção da detecção, o backbone, o neck e o head. O backbone do modelo YOLOv4 foi considerado pelo autor um CSPDarknet53 que consiste em um extrator de recursos. A camada CSP - Cross Stage Partial divide os recursos de entrada em dois grupos: um grupo é processado pela camada convolucional, enquanto o segundo evita as camadas convolucionais e é incluído na entrada para a camada seguinte [52]. O neck é composto por um SPP - Spatial Pyramid Pooling e uma PANet que têm por objetivo aumentar o campo receptivo e separar as características importantes do backbone [50]. Por último, o head realiza as previsões composta por um vetor contendo as coordenadas da caixa delimitadora, a pontuação de confiança e a probabilidade das classes.

3.3.3 YOLOv5

YOLOv5 foi publicado pela primeira vez em maio de 2020 por Glenn Jocher pesquisador e CEO da Ultralytics LLC. O autor nomeou a rede YOLOv5 para evitar conflitos de nomenclatura devido ao lançamento anterior do YOLOv4 [9]. No entanto, YOLOv5 não deve ser confundido com um descendente da série YOLO original baseada em DarkNet. Novos recursos e melhorias no YOLOv5 estão focados principalmente na incorporação do estado da arte para redes de aprendizagem profunda, como funções de ativação e aumento de dados. Estes foram parcialmente adotados do YOLOv4, como o backbone CSPNet [52], com outros elementos originados de contribuições anteriores do YOLOv4 pelo mantenedor do YOLOv5.

Outro fato importante mencionado por Jocher é a capacidade de aprendizagem automática, utilizando o K-Means, para selecionar as caixas de âncora que melhor se adaptem aos conjuntos de dados [10]. Em termos de métricas de desempenho, o YOLOv5 permite uma detecção rápida com a taxas de precisão similares ao do YOLOv4. Da mesma forma, que suas versões anteriores, o YOLOv5 foi lançado em diferentes tamanhos (s, m, n, l e x) com precisão e velocidade de detecção distintas [53].

3.4 Métricas para Análise de Desempenho

Esta seção descreve as métricas frequentemente utilizadas, que calculam o desempenho em relação aos resultados dos treinamentos e predições das redes. As subseções seguintes abordam as métricas *Intersect over Union*, Matriz de confusão, Curva Precision vs Recall e Average Precision.

3.4.1 *Intersect over Union (IoU)*

É uma métrica utilizada para definir o quanto uma predição da localização de um objeto na imagem coincide com a sua verdadeira localização. Esta medida é baseada no índice Jaccard ⁸, que avalia a sobreposição entre duas caixas delimitadoras. Requer uma caixa delimitadora com valores verdadeiros anotados e uma caixa delimitadora prevista pelo modelo, processo ilustrado na Figura 3.10. Ao aplicar o IoU, podemos dizer se uma detecção é válida (Verdadeiro Positivo) ou não (Falso Positivo) [49].

O valor do IoU pode corresponder ao intervalo entre 0 até 1, quanto maior for o valor do IoU, mais preciso será a seleção de um objeto. Já um valor baixo, faz com que o algoritmo tenha uma margem de aceitação baixa. A Figura 3.11 mostra a relação dos valores do IoU.

⁸Estadística usada para medir a similaridade e diversidade de conjuntos de amostras [49].

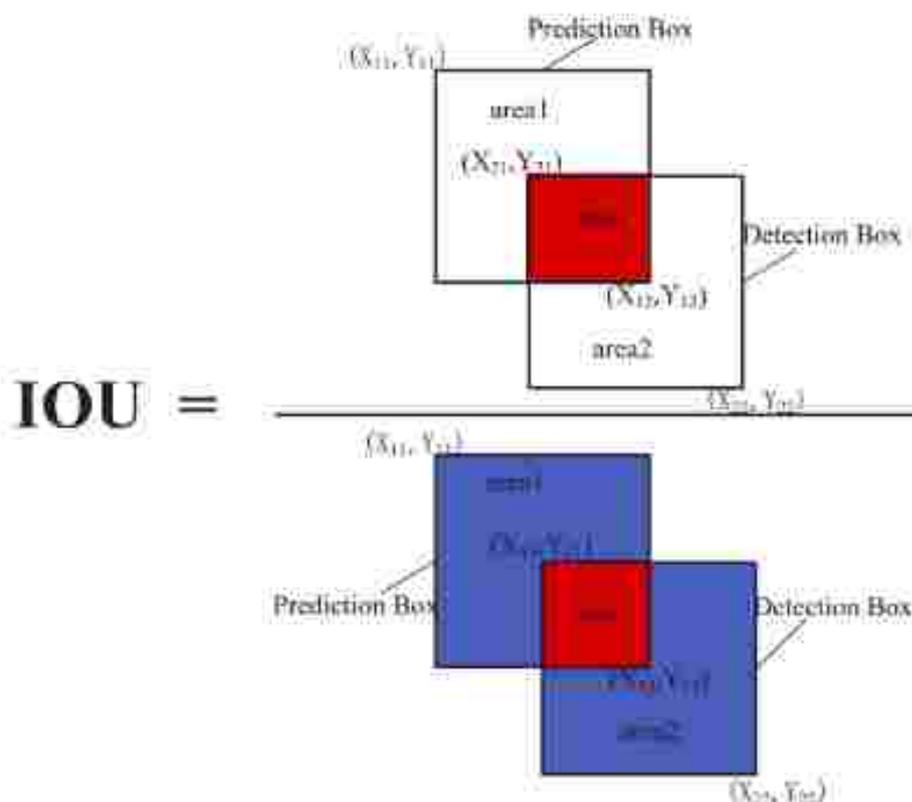


Figura 3.10: Cálculo de Intersecção sobre União (IoU). Fonte: Cheng et al. [6].

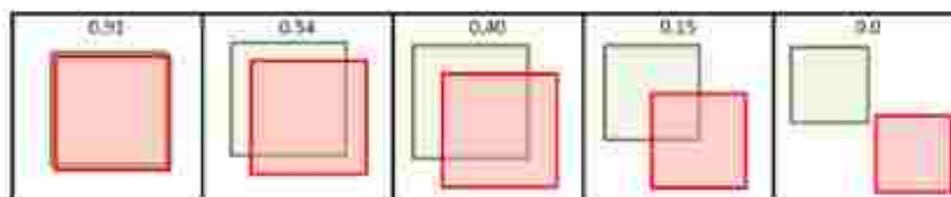


Figura 3.11: Valores do IoU em diferentes sobreposições. Fonte: Aragão [7].

3.4.2 Matriz de confusão

Geralmente, as métricas utilizadas para validação de um modelo derivam do resultado de uma matriz de confusão. É importante destacar que as combinações utilizando a matriz de confusão realizadas neste trabalho são no conceito de detecção de objetos [49]:

- True Positive (TP): Ocorre quando o retângulo envolvente da previsão e o do real estão com sobreposição do IoU maior que um threshold pré-estabelecido para aquela classe de objeto, ou seja, é uma detecção correta do objeto. São valores para o threshold: 0.5, 0.75 e 0.95. Sendo o mais utilizado o primeiro, que representa 50%.

- False Positive (FP): Ocorre quando o retângulo envolvente da previsão e o do real estão com o valor de IoU menor que o threshold, ou seja, é uma detecção incorreta do objeto da mesma classe.
- False Negative (FN): Ocorre quando o modelo não detectou o objeto.
- True Negative (TN): Não se aplica. Ocorre quando uma previsão correta é feita de onde não existe o real. Sendo assim, todos os TN serão classificados como FP. Este valor não é usado nas métricas.

3.4.3 Curva Precision vs Recall

A *precision* pode ser definida como a fração de elementos relevantes selecionados dentro de um grupo de elementos selecionados. Ou seja, em uma classificação binária, a *precision* é a razão entre o número de verdadeiros positivos selecionados por um modelo e o número de elementos que foram classificados como positivos pelo modelo (Equação 3.1).

$$\frac{TP}{TP + FP} \quad (3.1)$$

O *recall* é definido como a razão entre os elementos relevantes selecionados e todos os itens relevantes dentro do universo de elementos disponíveis. Em outras palavras, o *recall* é definido como a fração de verdadeiros positivos selecionados dentro de todos os positivos disponíveis no grupo de elementos (Equação 3.2).

$$\frac{TP}{TP + FN} \quad (3.2)$$

A curva *precision vs recall* é uma boa maneira de avaliar o desempenho de um modelo de detector de objetos à medida que a confiabilidade é alterada de uma forma decrescente traçando uma curva para cada classe de objeto. Um modelo utilizado por uma determinada classe é considerado bom se seu *precision* se mantiver alto (o mais próximo de um) enquanto o *recall* também estiver alto, o que significa que se houver uma variação no limite da confiabilidade, os valores do *precision* e *recall* ainda permanecerão altos [49].

3.4.4 Average Precision (AP)

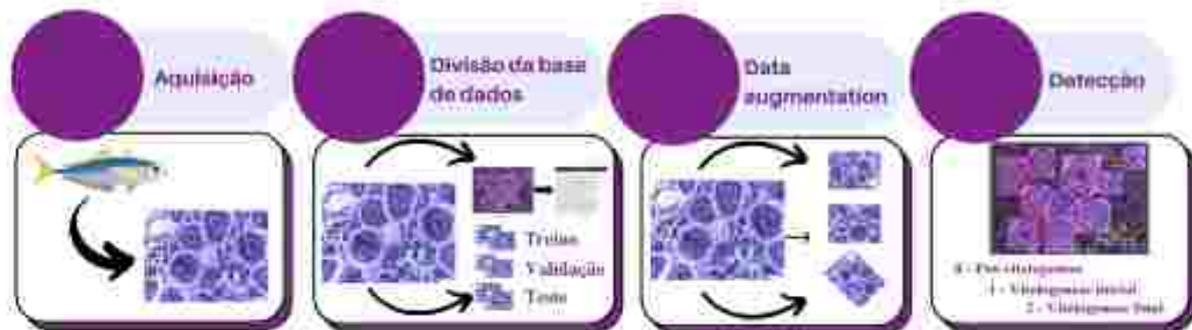
Average Precision (AP) é o resultado da razão entre a *precision* e *recall* coletados durante o treinamento de um algoritmo. Ao final do treinamento, depois de obter-se um conjunto de APs, o *mean Average Precision (mAP)* é calculado (Equação 3.3) através da média de todos os resultados do AP para cada classe do modelo de detecção de objetos, de forma que, na escala de porcentagem, quanto maior for o valor de mAP maior é a precisão do modelo de detecção [54] [7].

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3.3)$$

4 Metodologia

Nesta seção descrevemos os detalhes da metodologia aplicada neste estudo. Em um primeiro momento buscou-se, através de uma pesquisa exploratória, fundamentar este trabalho, desenvolvendo o seu referencial teórico, conforme observado na seção anterior. Quanto à abordagem científica adotada ao problema apontado, desenvolveu-se uma análise através de métodos quanti-qualitativos. Para isso, foi aplicado um método para detecção automática de ovócitos em imagens histológicas. Este método proposto é composto por quatro etapas: 1) Aquisição; 2) Divisão da base de dados; 3) Data augmentation e 4) Detecção. Por fim, é realizada uma avaliação da capacidade das arquiteturas YOLO em identificar estruturas correspondentes às fases de desenvolvimento ovocitária (pré-vitelogênese, vitelogênese inicial e vitelogênese final). A Figura 4.1 ilustra as etapas do método proposto para detecção automática de ovócitos.

Figura 4.1: Etapas do método utilizado neste trabalho. Fonte: A autora.



4.1 Aquisição de Imagens

As amostras das imagens histológicas das gônadas dos peixes foram obtidas em parceria entre o Programa de Pós-Graduação em Biodiversidade e Biotecnologia da Amazônia Legal - BIONORTE e o Programa em Engenharia da Computação e Sistemas na Universidade Estadual do Maranhão - UEMA. A aquisição de espécimes

de *C. undecimalis* foi autorizada pelo Sistema de Autorização e Informação em Biodiversidade - SISBIO (n° 34.683/2017) e todos os procedimentos metodológicos para manejo de peixes foram aprovados pelo Comitê de Ética da Universidade Estadual do Maranhão - UEMA (n° 4.476.902/2020).

A coleta dos peixes da espécie *C. undecimalis* foi realizada por meio de compra direta no Mercantil do Peixe, localizado na cidade de Tutóia (2°45'44" S; 42°16'28" W), Maranhão. Todas as amostras foram adquiridas mensalmente durante o período de Janeiro de 2019 a Fevereiro de 2020.

Os peixes foram armazenados em isopor com gelo, vedados com fita adesiva e encaminhados ao Laboratório de Pesca e Ecologia Aquática da Universidade Estadual do Maranhão, onde foram inicialmente identificados. Em seguida, os espécimes foram contados, medidos, pesados e processados para retirada de gônadas e outras vísceras. Algumas gônadas foram armazenadas em frascos com solução de Bouin por 24 horas e posteriormente transferidas para uma solução de álcool 70% para posterior análise histológica.

No processo de aquisição das imagens, utilizou-se procedimentos histológicos para seccionar os ovários. As secções foram coradas com Hamatoxilina e Eosina (H&E), em seguida as imagens foram capturadas por um microscópio LEICA DM 500 com uma ampliação de x200. Uma câmera digital LEICA EC4 foi usada diretamente conectada ao microscópio. Além do Software de aquisição de imagens LEICA LAS EZ compatível com o equipamento e um monitor Sony para visualização.

Cada imagem adquirida é um arquivo no formato JPEG, com composição colorida RGB apresentando 16 bits por pixel e resolução de 1.024 x 768 pixels. A Figura 4.2 pertence ao conjunto de imagens adquiridas, nela pode-se identificar facilmente os ovócitos que serão detectados em uma etapa posterior. Para melhorar o desempenho da rede e reduzir os custos computacionais, todas as imagens foram redimensionadas para 640 x 640 pixels.

Foram adquiridas para a construção da base 305 imagens histológicas. Esta base, poderá ser disponibilizada posteriormente, permitindo que mais pesquisas sejam conduzidas neste domínio, ampliando o monitoramento desta espécie, visto

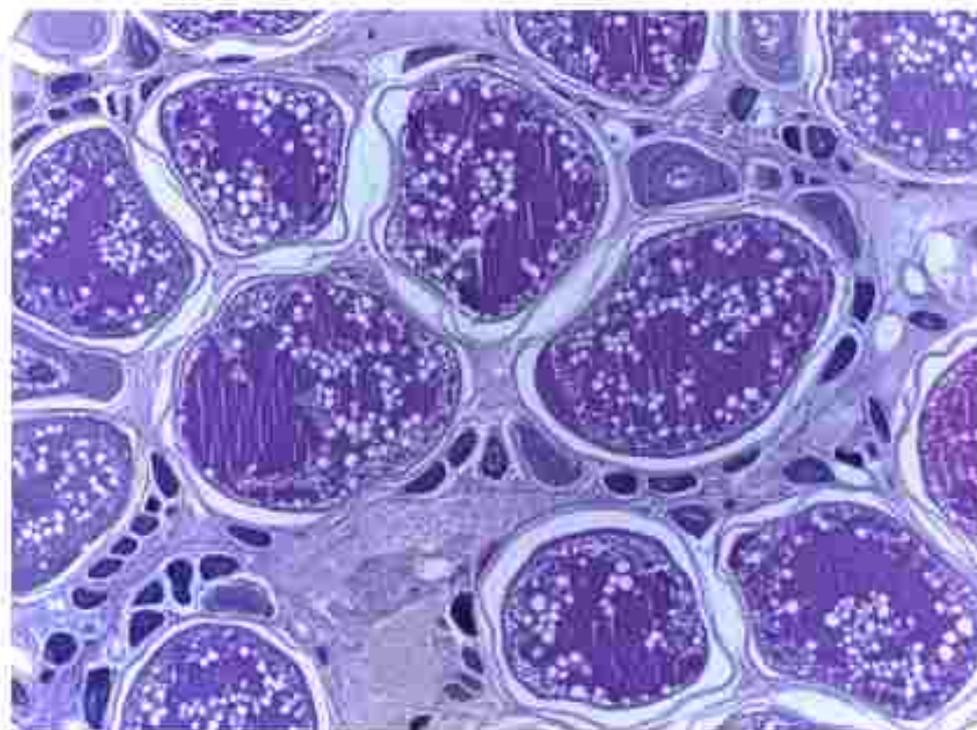


Figura 4.2: Imagem capturada por microscópio ampliada em x200. Fonte: A autora.

que, o peixe em estudo apresenta uma característica de sexualidade peculiar, o hermafroditismo protândrico¹

4.2 Divisão da base de dados

Os conjuntos de dados têm desempenhado um papel fundamental ao longo da história da pesquisa de reconhecimento de objetos, não apenas como um terreno comum para medir e comparar o desempenho de algoritmos, mas também impulsionar o campo em direção a problemas cada vez mais complexos e desafiadores [55]. Tendo isso em vista, foram empregadas para a base de dados, imagens com recortes histológicos do peixe *C. undecimalis*. A fim de organizar, rotular e dividir as imagens, neste trabalho, foi utilizado o software Roboflow [56].

Após a etapa de aquisição, o especialista, fazendo uso da caixa delimitadora, rotulou em formato TXT, 305 imagens, seguindo padrões específicos para o YOLOv3, YOLOv4 e YOLOv5. O arquivo gerado consiste em um conjunto de coordenadas que identificam os objetos em uma imagem com anotações de forma padronizada [57].

¹Maturam primeiramente como machos e posteriormente, mudam de sexo e permanecem, em sua maioria, como fêmeas pelo resto de suas vidas [40].

Na estrutura interna do arquivo TXT é apresentado um objeto por linha, em cada linha são fornecidas informações como número da classe, centro em x, centro em y, largura e altura do objeto. As classes são indexadas por uma numeração, na qual é iniciada pelo índice 0 (zero), como exemplo; a classe 0 (PV - pré-vitelogênese), classe 1 (VI - vitelogênese inicial) e classe 2 (VF - vitelogênese final). Segundo Do Thuan [10], as coordenadas da caixa delimitadora do objeto devem estar organizadas em xywh normalizadas entre [0,1]. A Figura 4.3 apresenta os ovócitos rotulados na imagem com suas respectivas classes e coordenadas.

Um dos atributos presentes nos rótulos representa as fases de desenvolvimento maturacional ovocitário. As fases utilizadas nesta pesquisa foram escolhidas devido a frequência destes ovócitos nas imagens, pois segundo Claro et al. [58] técnicas baseadas em aprendizagem profunda necessitam de uma grande quantidade de objetos para uma detecção mais eficiente.



Figura 4.3: Processo de rotulagem dos ovócitos na imagem. (A) Imagem rotulada pelo especialista. (B) Arquivo em formato TXT para o YOLO. Fonte: A autora.

Para validar o modelo de detecção foi subdividida aleatoriamente em três conjuntos, conforme a pesquisa de Nie et al. [59]: treino, validação e teste. O conjunto de treino contém 244 imagens (2.393 PV, 633 VI, 1.500 VF), o conjunto de validação contém 31 imagens (299 PV, 51 VI, 139 VF) e o conjunto de teste contém 30 imagens (374 PV, 98 VI, 193 VF). Dessa forma, um conjunto de 5.680 ovócitos foram rotulados por um especialista. Ao todo, a classe 0 (PV) possui 3.066 ovócitos, a classe 1 (VI) possui 782 ovócitos e a classe 2 (VF) possui 1.832 ovócitos. A Tabela 4.1 demonstra de forma detalhada a divisão da base:

Tabela 4.1: Detalhes do conjuntos de dados de ovócitos.

Dataset	Número de Imagens	Número de ovócitos por imagem			Total de ovócitos
		PV	VI	VF	
Treino	244	2.393	633	1.500	4.526
Validação	31	299	51	139	489
Teste	30	374	98	193	665
Total	305	3.066	782	1.832	5.680

4.3 Data augmentation

O aumento de dados é uma conjunto de técnicas utilizadas para criar novos exemplos de treinamento a partir de dados existentes [34]. Estas técnicas têm a finalidade de auxiliar na generalização do modelo de aprendizagem de máquina. Para esta pesquisa, a base de imagens construída até o momento contém apenas 305 amostras, sendo esta, considerada pequena para o treinamento de um modelo de aprendizagem profunda. Portanto, através do software Roboflow [56], empregamos as seguintes técnicas de aumento de dados visando uma previsão robusta:

- Flip: consiste em inverter a imagem em duas orientações (horizontal e vertical).
- Rotação em 90°: consiste em adicionar rotações de 90° graus na imagem. As rotações são aplicadas em três sentidos (horário, anti-horário e cabeça para baixo).
- Crop: consiste em recortes aleatórios na imagem original.
- Rotação aleatória: consiste em girar aleatoriamente a imagem original no sentido horário ou anti-horário em alguns graus, mudando a posição do objeto no quadro.

4.4 Arquiteturas utilizadas

As arquiteturas de detecção foram utilizadas com o objetivo de classificar e localizar os ovócitos em imagens de acordo com suas fases. As versões 3, 4 e 5 do YOLO foram selecionadas para os experimentos devido ao fato dessas arquiteturas terem a capacidade de identificar variações de objetos de uma mesma classe em diferentes dimensões [50], [5] e [8].

Na versão YOLOv3, foram aplicados os modelos YOLOv3-tiny e YOLOv3-ssp. Na versão YOLOv4, foram aplicados os modelos YOLOv4-tiny e YOLOv4-csp. Na versão YOLOv5, foram aplicados os modelos YOLOv5s, YOLOv5m, YOLOv5n, YOLOv5l e YOLOv5x. Para a escolha desses modelos foram consideradas as características de alto desempenho em aplicações de tempo real, que segundo Silva, Narciso e Gonçalves [60] se destacam pelo uso de classificadores lógicos independentes para classificação múltipla com sobreposição de classes. Luo et al. [61] destaca que essas versões apresentam maior desempenho ao lidar com objetos de tamanhos pequenos.

Um ponto comum de todas as arquiteturas de detecção de objetos YOLO é que os recursos da imagem de entrada são compactados através de um extrator de recursos chamado backbone e então encaminhados para um detector de objetos (Detection Neck e Detection Head). O Neck funciona como um agregador de recursos que tem a função de misturar e combinar os recursos formados no backbone. Em seguida, o Detection Head realiza a detecção, incluindo a localização e classificação dos objetos [10].

Para extrair informações essenciais da imagem de entrada, os modelos YOLOv3-tiny e YOLOv3-ssp usam o backbone Tiny e Darknet53, respectivamente. Nos modelos YOLOv4-tiny e YOLOv4-csp, os backbones Tiny e CSPDarknet53 são usados, respectivamente. Os modelos YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s e YOLOv5x usam um backbone CSPDarknet53 mais uma camada Spatial Pyramid Pooling (SPPF).

Para aumentar o campo receptivo e separar os recursos mais importantes, os modelos YOLOv3-tiny, YOLOv3-ssp e YOLOv4-tiny usam a Neck Filter Pyramid Network (FPN). Enquanto os modelos YOLOv4-csp, YOLOv5l, YOLOv5m, YOLOv5n,

YOLOv5s e YOLOv5x usam a Neck Path Aggregation Network (PANet). Para realizar a previsão final, os modelos YOLOv3-tiny, YOLOv3-ssp, YOLOv4-tiny, YOLOv4-csp, YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s e YOLOv5x usam uma camada YOLO para gerar um vetor contendo as coordenadas da caixa delimitadora: largura, altura, rótulo de classe e probabilidade de classe. A Tabela 2 resume a composição das arquiteturas YOLOv3, YOLOv4 e YOLOv5.

Tabela 4.2: Composição das arquiteturas YOLOv3, YOLOv4 e YOLOv5. Fonte: [8], [9] e [10]

	Backbone	Neck	Head
YOLOv3			
YOLOv3-Tiny	Tiny	FPN	YOLO
YOLOv3-ssp	Darknet53	FPN	
YOLOv4			
YOLOv4-Tiny	Tiny	FPN	YOLO
YOLOv4-csp	CSPDarknet53	PANet	
YOLOv5			
YOLOv5l			
YOLOv5m			
YOLOv5n	CSPDarknet53 com uma camada SPPF	PANet	YOLO
YOLOv5s			
YOLOv5x			

4.5 Ambiente para os experimentos

Os experimentos foram conduzidos utilizando a plataforma gratuita Google Colaboratory com execução em uma Unidade de Processamento Gráfico (GPU). As configurações de hardware compreendem em uma unidade de processamento gráfico de núcleo tensor NVIDIA Tesla P100 (GPU), 12GB de RAM e 68GB de disco. Além disso, o Google Colaboratory fornece um ambiente configurado com Python 3. Foi necessário instalar manualmente as bibliotecas *Keras*, *OpenCV*, *TensorFlow*, *matplotlib*, *numpy* e *scipy*.

Os modelos foram treinados no conjunto de dados de imagens para 300 épocas, conforme a documentação do YOLO e trabalhos presentes na literatura [62] [52] [63] [59], com um tamanho de lote igual a 16 imagens. Nestes modelos foram utilizados os seguintes hiperparâmetros:

otimizador = SGD, lr0 = 0.01, lrf = 0.2, momentum = 0.937, weight_decay = 0.0005, warmup_epochs = 3.0, warmup_momentum = 0.8, warmup_bias_lr = 0.1, box = 0.05, cls = 0.5, cls_pw = 1.0, obj = 1.0, obj_pw = 1.0, iou_t = 0.2, anchor_t = 4.0, fl_gamma = 0.0, hsv_h = 0.015, hsv_s = 0.7, hsv_v = 0.4, degrees = 0.0, translate = 0.1, scale = 0.5, shear = 0.0, perspective = 0.0, flipud = 0.0, fliplr = 0.5, mosaic = 1.0, mixup = 0.0.

5 Resultados e Discussão

Neste capítulo são apresentados os resultados obtidos a partir da detecção dos ovócitos, realizada pelos modelos YOLO, aplicados no conjunto de dados de imagens histológicas, seguindo o método proposto neste trabalho.

A seção 5.1 deste capítulo apresenta a performance dos modelos utilizados durante a etapa de treinamento e teste. A seção 5.2 apresenta uma comparação de desempenho entre os diferentes modelos de detecção, levando em consideração as métricas *precision*, *recall* e média AP (mAP) para cada uma das classes. Por fim, a seção 5.3 apresenta uma comparação entre os modelos através da curva *precision vs recall*.

5.1 Performance dos modelos no treinamento e teste

Para cada treinamento foram gerados pesos (parâmetros treináveis e não treináveis do modelo) expressos em milhões, que foram, respectivamente de 61, 62, 68, 72, 46, 20, 17, 7 e 86 para o YOLOv3-tiny, YOLOv3-ssp, YOLOv4-tiny, YOLOv4-csp, YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s e YOLOv5x.

O tempo médio de treinamento para os modelos YOLOv3-tiny e YOLOv3-ssp foi estimado em 0h 46min e 4h 19min, respectivamente. Para os modelos YOLOv4-tiny e YOLOv4-csp foi de 4h 25min e 3h 16min, respectivamente. Enquanto que, para os modelos YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s e YOLOv5x o tempo de treinamento foi de 3h 23min, 2h 16min, 0h 37min, 0h 54min e 5h 35min, respectivamente.

Em seguida, a precisão média para 50% de confiança (mAP@.5) na detecção dos ovócitos, de cada modelo, é calculada em relação à todas as fases de desenvolvimento ovocitária (PV, VI e VF). As melhores mAP@.5 alcançadas foram para os modelos YOLOv4-csp com 74.8%, seguido pelo YOLOv5s com 74.1%, YOLOv3-tiny com 72.6% e YOLOv4-tiny com 72.3%. O modelo YOLOv3-ssp obteve 71.8% de precisão média, enquanto que, os modelos YOLOv5l, YOLOv5m, YOLOv5n e YOLOv5x alcançaram respectivamente 65.0%, 71.4%, 68.9% e 71.6%.

Após os modelos serem treinados, o tempo de execução das inferências (expresso em milissegundos, ms) nas imagens de teste foram de 2.1, 2.2, 1.2, 1.6, 0.9, 0.8, 0.9, 0.9, 0.7 ms para os respectivos modelos YOLOv3-tiny, YOLOv3-ssp, YOLOv4-tiny, YOLOv4-csp, YOLOv5l, YOLOv5m, YOLOv5n, YOLOv5s e YOLOv5x. A Tabela 5.1 demonstra a performance dos modelos.

Tabela 5.1: Comparação das performances dos modelos no treinamento e teste por meio dos pesos do modelo, tempo de treinamento, métrica mAP@.50 e tempo de inferência.

Model	Model Weight (M)	Training Time (h min)	mAP@.5 (%)	Inference Time (ms)
YOLOv3				
YOLOv3-tiny	61	0h 46min	72.6	2.1
YOLOv3-ssp	62	4h 19min	71.8	2.2
YOLOv4				
YOLOv4-tiny	68	4h 25min	72.3	1.2
YOLOv4-csp	72	3h 16min	74.8	1.6
YOLOv5				
YOLOv5l	46	3h 23min	65.0	0.9
YOLOv5m	20	2h 16min	71.4	0.8
YOLOv5n	17	0h 37min	68.9	0.9
YOLOv5s	7	0h 54min	74.1	0.9
YOLOv5x	86	5h 35min	71.6	0.7

5.2 Comparação dos Diferentes Modelos de Detecção

A fim de verificar a eficácia das arquiteturas propostas neste trabalho, os modelos YOLOv3, YOLOv4 e YOLOv5 foram comparados, levando em consideração para este fim as métricas *precision* (P), *recall* (R) e média AP (mAP) para cada uma das classes. Para isso, todos os modelos foram treinados e testados utilizando o mesmo conjunto de dados e o mesmo ambiente nos experimentos. A seguir na Tabela 5.2 são apresentados os resultados de detecção dos modelos.

Tabela 5.2: Comparação de desempenho dos modelos com base nas métricas *precision* (P), *recall* (R) e média AP para cada classe (mAP).

Modelos	P(%)			R(%)			mAP@.5(%)			mAP@.95(%)		
	PV	VF	VI	PV	VF	VI	PV	VF	VI	PV	VF	VI
YOLOv3												
YOLOv3-tiny	67.0	68.1	71.1	73.5	96.4	65.3	72.6	89.2	70.0	36.5	52.6	35.3
YOLOv3-ssp	85.0	78.8	77.2	74.7	89.1	77.6	80.5	90.8	82.8	48.4	59.3	57.2
YOLOv4												
YOLOv4-tiny	69.9	70.7	74.9	51.0	96.9	62.2	56.6	89.7	70.6	16.3	55.8	40.1
YOLOv4-csp	79.7	76.0	79.5	55.1	92.2	75.5	65.2	90.7	81.9	24.8	62.3	56.4
YOLOv5												
YOLOv5l	83.0	85.4	79.3	64.2	95.3	85.7	78.0	95.7	90.2	61.7	75.9	67.9
YOLOv5m	82.9	70.3	89.4	79.3	97.4	77.6	85.9	94.5	86.2	59.2	73.6	64.1
YOLOv5n	79.0	56.2	50.7	41.8	97.1	68.6	61.4	88.3	61.7	27.9	58.5	31.8
YOLOv5s	68.6	58.5	52.3	48.7	95.3	70.6	57.5	84.6	65.1	24.7	57.5	36.7
YOLOv5x	78.5	79.6	82.2	66.8	94.4	84.7	75.4	95.4	86.2	36.0	73.4	58.7

Para a identificação da classe PV, o modelo YOLOv3-ssp atingiu a maior taxa de precisão, com 85%. Por outro lado, o modelo YOLOv5m atingiu a maior taxa de *recall*, com 79.3%. O modelo YOLOv5m também atingiu a maior taxa para a média mAP@.5 com 85.9% e umas das mais altas para a média mAP@.95, com 59.2%, ficando atrás, nesse quesito, apenas do modelo YOLOv5l que obteve uma média mAP@.95 de 61.7%.

No que se refere à identificação da classe VF, o modelo YOLOv5l atingiu a maior taxa de precisão, com 85.4%. Por outro lado, o modelo YOLOv5m atingiu a maior taxa de *recall*, com 97.4%. O modelo YOLOv5m também atingiu umas das mais altas taxas para a média mAP@.5, com 94.5%, ficando atrás, nesse quesito, apenas dos modelos YOLOv5l e YOLOv5x que obtiveram uma média mAP@.5 de 95.7% e 95.4%, respectivamente. O modelo YOLOv5l também atingiu a maior taxa para a média mAP@.95, com 75.9%.

Por sua vez, para a identificação da classe VI, o modelo YOLOv5m atingiu a maior taxa de precisão, com 89.4%. Enquanto que o modelo YOLOv5l obteve a maior taxa de *recall*, com 85.7%. O modelo YOLOv5l também atingiu a maior taxa para as médias mAP@.5 e mAP@.95, com 90.2% e 67.9%, respectivamente.

É possível observar na Tabela 5.2, destacado em cinza, que a classe VF, no geral, foi a classe em que os modelos atingiram as maiores taxas de recall, mAP@.5 e mAP@.95, isso ocorreu devido aos classificadores lidarem melhor com ovócitos que apresentam, em média, diâmetros de $388.3 \mu m$ [18], essa característica vai de encontro com os resultados obtidos nas pesquisas de [64], [65] e [66], que destacam que objetos com maiores dimensões apresentam resultados mais satisfatórios na tarefa de detecção.

Além do mais, os ovócitos na fase PV apresentam a forma arredondada e a coloração uniforme. Já na fase VI, sua estrutura externa perde a forma esférica, apresentam a coloração heterogênea e sua estrutura interna contém nucléolos bem pequenos e esbranquiçados. Essas características podem ser um indicativo que, mesmo a classe VI, contendo ovócitos de pequenas dimensões, permite aos modelos atingirem resultados mais expressivos do que na classe PV.

Na Tabela 5.3 é possível observar os modelos que obtiveram os melhores resultados para cada uma das métricas utilizadas nas respectivas classes. Destaca-se aqui os modelos YOLOv5m e YOLOv5l que atingiram os melhores desempenhos.

Tabela 5.3: Modelos com melhores resultados

	Precision	Recall	mAP@.5%	mAP@.95%
PV	YOLOv3-ssp	YOLOv5m	YOLOv5m	YOLOv5l
VF	YOLOv5l	YOLOv5m	YOLOv5l	YOLOv5l
VI	YOLOv5m	YOLOv5l	YOLOv5l	YOLOv5l

5.3 Curva Precision vs Recall

A curva de *Precision vs Recall* mostra a relação entre precisão (*precisão*) e a revocação (*recall*) para diferentes limites gerados a partir dos dados de teste. Uma área alta sob a curva representa alta precisão e alta revocação, ou seja, a alta precisão está relacionada a uma baixa taxa de falsos positivos e a alta revocação está relacionada a uma baixa taxa de falsos negativos [67]. Sendo assim, a figura 5.1 ilustra as curvas de *Precision vs Recall* geradas para cada modelo YOLO. A área sob a curva foi estimada com a mAP@.5. Nesta, constata-se que os modelos YOLOv5l, YOLOv5m e YOLOv5x obtiveram respectivamente 0.880, 0.859 e 0.857 de área sob a curva, constituindo assim

os modelos com os melhores resultados. Destaca-se ainda que estes resultados vão de encontro com o observado nas Tabelas 5.2 e 5.3.

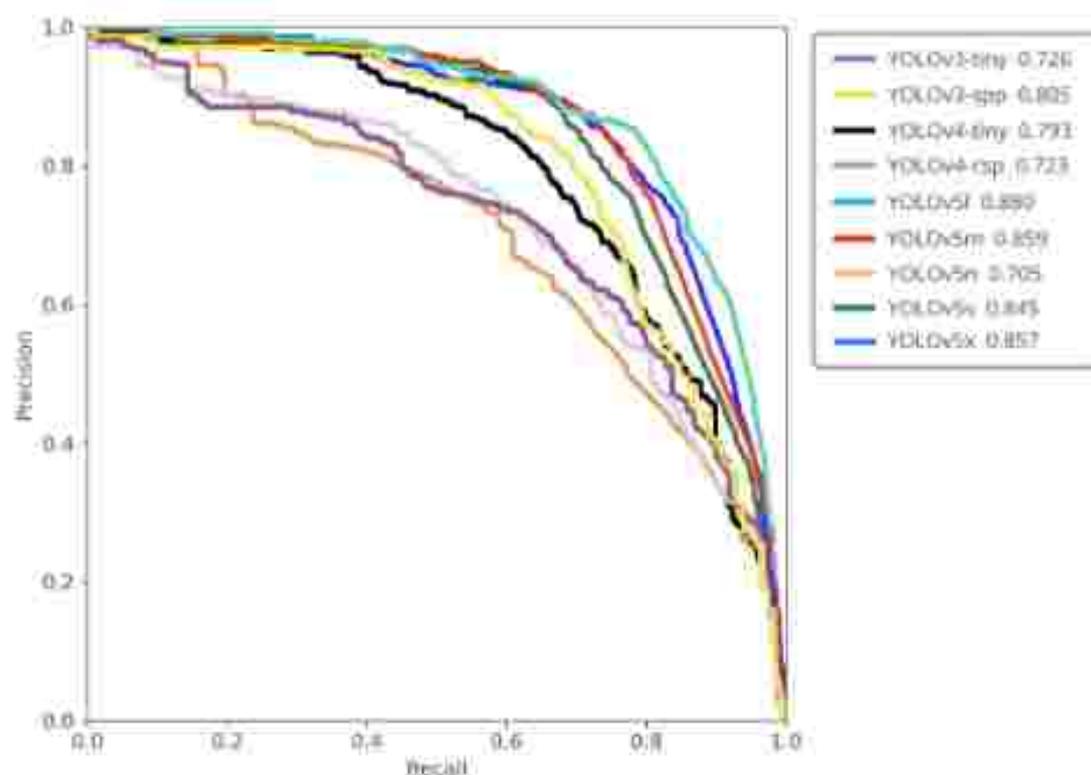


Figura 5.1: Curva Precision vs Recall.

5.4 Resultado do Método de Detecção

Quando os modelos são treinados e, posteriormente, testados, cada modelo é aplicado nos mesmos conjuntos de dados, resultando em imagens contendo a detecção. Um exemplo deste tipo de resultado é ilustrado nas Tabelas 5.4, 5.5 e 5.6. Na imagem a caixa delimitadora em cor vermelha representa a fase PV, a caixa delimitadora em cor laranja representa a fase VI e a caixa delimitadora em cor rosa clara representa a fase VF. Para cada respectiva caixa é gerada sua percentagem referente ao IoU, ou seja, o quanto essa predição da localização do ovócito coincide com a sua verdadeira localização.

Como pode ser observado os modelos YOLOv3-ssp, YOLOv4-tiny, YOLOv5l e YOLOv5m aplicados nas imagens (Imagem 1 e Imagem 2), a fase VF obteve uma percentagem IoU superior a 90%. O YOLOv3-ssp, em sua maioria, apresentou percentagens superior a 80% para a fase PV. O YOLOv3-ssp, YOLOv5l, YOLOv5m

e YOLOv5x obtiveram em sua maior parte ovócitos nas fases VI uma porcentagem elevada, atingindo valores de 80%. De modo geral, observa-se que os resultados obtidos com a detecção destes ovócitos e ilustrados aqui, estão alinhados com o que foi apresentado no tópico 5.2.

Tabela 5.4: Resultado da detecção e localização dos ovócitos através dos modelos YOLOv3-tiny, YOLOv3-ssp e YOLOv4-tiny para cada fase ovocitária.

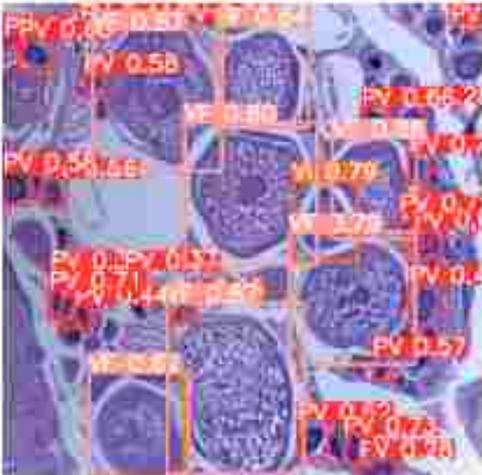
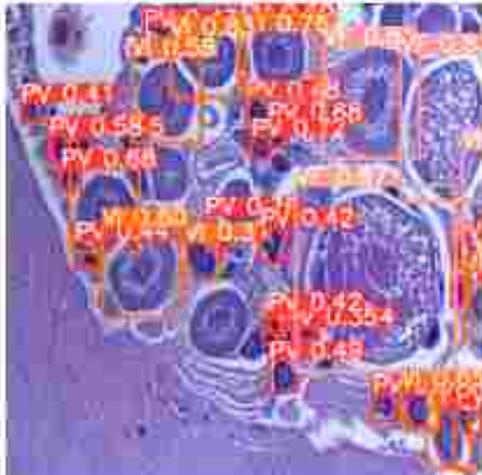
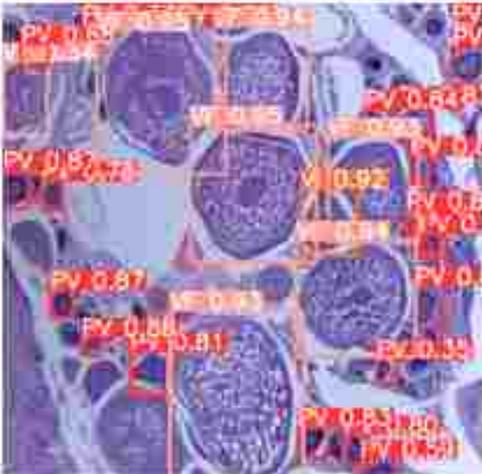
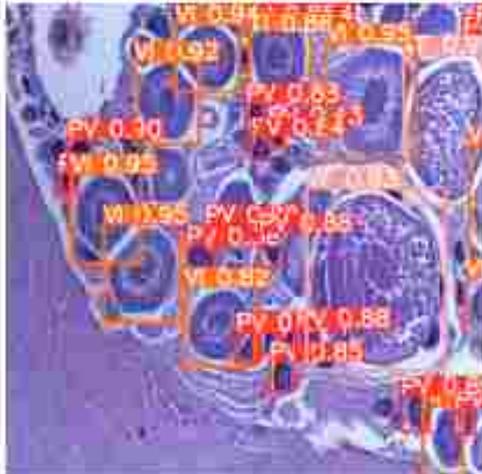
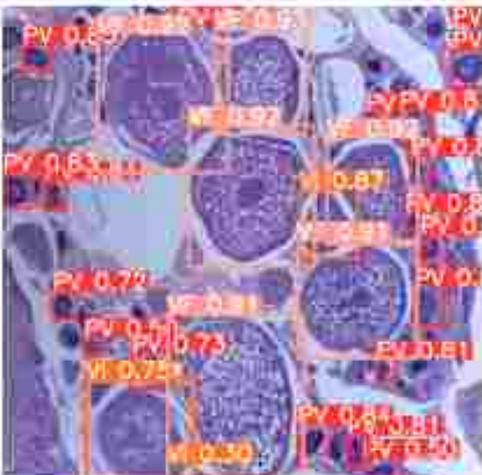
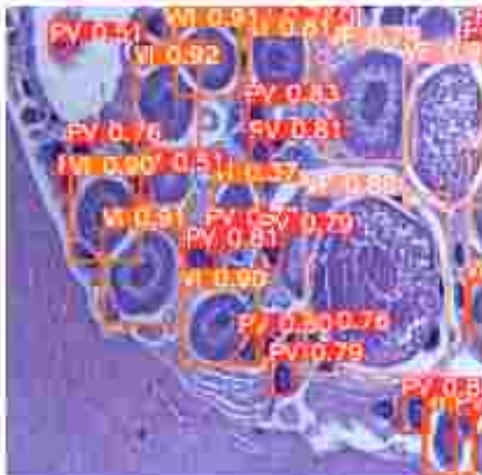
Model	Imagem 1	Imagem 2
YOLOv3-tiny		
YOLOv3-ssp		
YOLOv4-tiny		

Tabela 5.5: Resultado da detecção e localização dos ovócitos através dos modelos YOLOv4-csp, YOLOv5l e YOLOv5m para cada fase ovocitária.

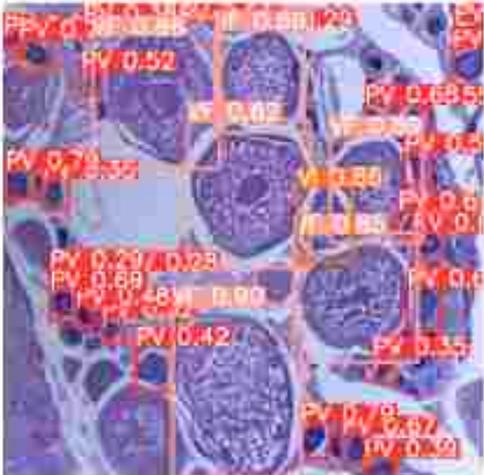
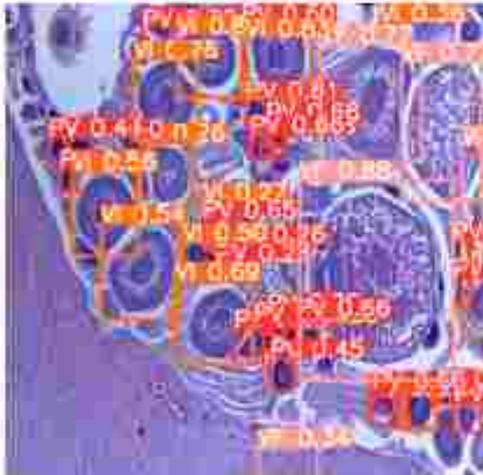
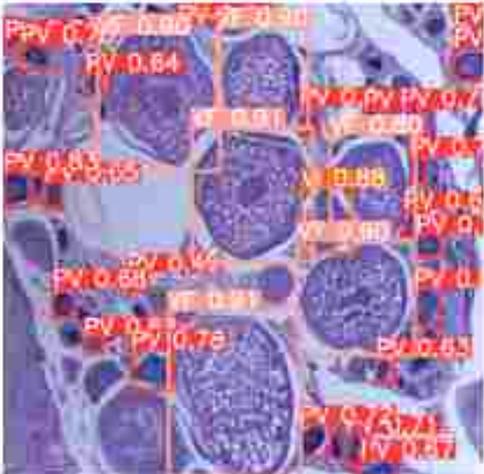
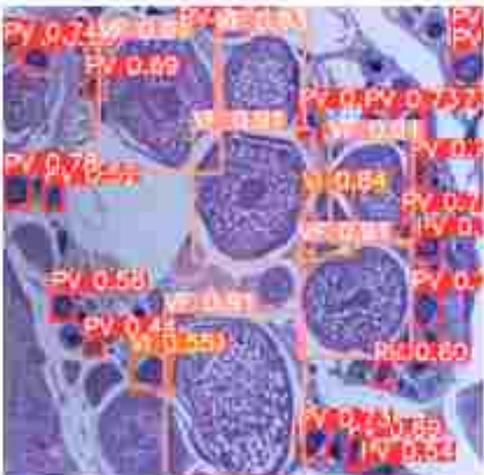
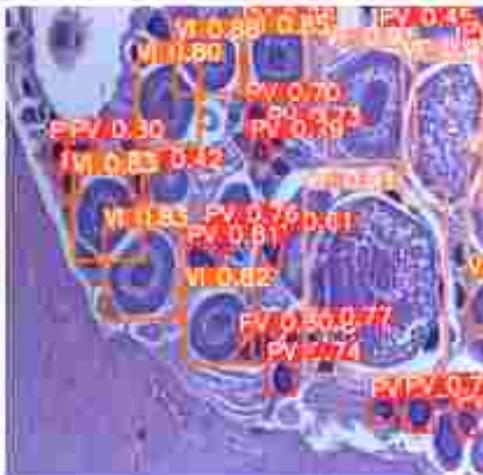
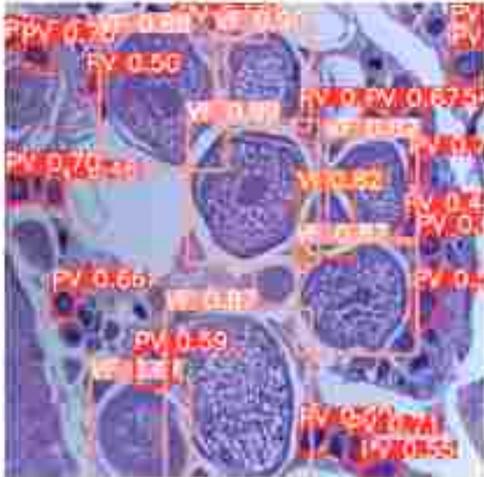
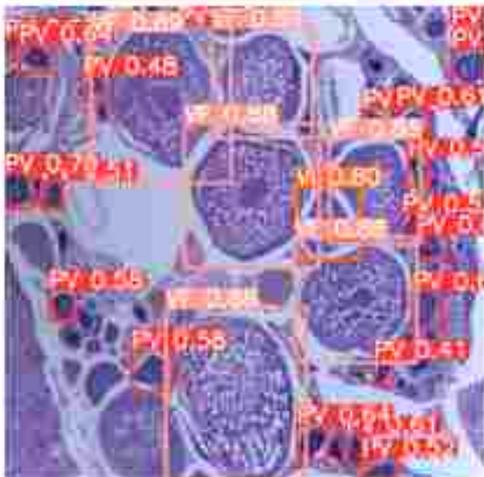
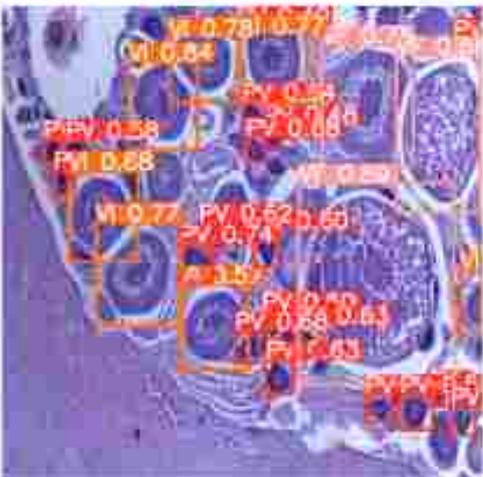
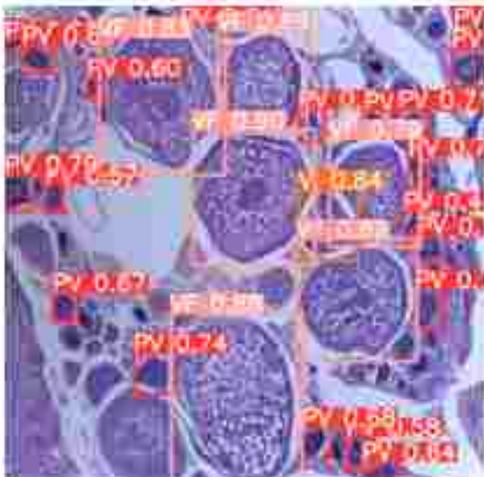
Model	Imagem 1	Imagem 2
YOLOv4-csp		
YOLOv5l		
YOLOv5m		

Tabela 5.6: Resultado da detecção e localização dos ovócitos através dos modelos YOLOv5n, YOLOv5s e YOLOv5x para cada fase ovocitária.

Model	Imagem 1	Imagem 2
YOLOv5n		
YOLOv5s		
YOLOv5x		

6 Conclusão

Este trabalho descreveu um método para a detecção de ovócitos da espécie *C. undecimalis* a partir de imagens histológicas. A abordagem também descreveu detalhadamente as etapas necessárias para realizar o processo de detecção por meio da técnica YOLO. Este recurso não é coberto por nenhuma outra abordagem do estado da arte apresentada nesta pesquisa.

Para isso foi construído uma base de imagens histológicas contendo três fases de desenvolvimento ovocitário, fases essas rotuladas por um especialista. A fim de tornar o modelo mais robusto, técnicas de aumento de dados foram aplicadas durante o treinamento. Em seguida, as arquiteturas de detecção de um único estágio, a YOLO, em suas versões 3, 4 e 5, foram aplicadas.

Os resultados obtidos foram promissores, destacando-se que os modelos YOLOv5m e YOLOv5l atingiram os melhores desempenhos na detecção de ovócitos, ao serem avaliados através das métricas *precision*, *recall* e *mAP@.5* e *mAP@.95*. Por meio destas métricas foi possível observar que o modelo YOLOv5m atingiu 89,4% de precisão para a classe VI. O modelo YOLOv5m também atingiu a taxa de 79,3% de *recall* para a classe PV e 97,4% para a classe VF. Além disso, modelo YOLOv5m atingiu 85,9% de *mAP@.5* na classe PV. Por sua vez, o modelo YOLOv5l alcançou 85,4% de precisão para a classe VF, 95,7% de média *mAP@.5* e 75,9% de *mAP@.95* para esta mesma classe. Além disso, o modelo YOLOv5l atingiu 85,7%, 90,2% e 67,9%, respectivamente de *recall*, *mAP@.5* e *mAP@.95* para a classe VI e 61,7% de média *mAP@.95* para a classe PV.

Observou-se também, que a quantidade de ovócitos não é um fator significativamente relevante para o aumento dos valores das métricas estabelecidas, mais sim o tamanho em que os ovócitos se encontram nas imagens, expondo que o uso de uma arquitetura que melhor se adapte à objetos de pequenas dimensões faz-se necessário.

Dessa forma, compreende-se que o método proposto é uma solução adequada para a identificação de ovócitos, permitindo, assim, uma melhor

investigação dos aspectos reprodutivos desta espécie, além de possibilitar o controle a nível nacional do combate à degradação dos estoques pesqueiros, nos locais onde há intensa exploração deste peixe.

6.1 Contribuições e Resultados da Pesquisa

As principais contribuições desta pesquisa foram:

- Uma análise comparativa das técnicas estudadas, ressaltando suas principais características bem como suas limitações.
- Construção de uma base de imagens histológicas contendo três fases de desenvolvimento ovocitário, da espécie *Centropomus undecimalis*, rotuladas por um especialista.
- Desenvolvimento de um método capaz de detectar ovócitos de forma automática.
- Comparação entre os modelos YOLO em suas versões 3, 4 e 5.
- Avaliação e análise de quais modelos apresentam os melhores desempenhos na tarefa de detecção dos ovócitos.

6.2 Trabalhos Futuros

Embora os resultados obtidos sejam satisfatórios para a detecção e localização de ovócitos em imagens, melhorias podem ser realizadas na metodologia proposta em alguns pontos específicos. Algumas sugestões são apresentadas a seguir:

- Balancear a base, visto que, as classes apresentam proporções distintas em suas respectivas amostras.
- Completar a base existente, contemplando-a com todas as fases de desenvolvimento ovocitário, permitindo uma ampla verificação dos modelos na tarefa de detecção.

- Avaliar o método proposto em comparação à técnicas baseadas em redes neurais convolucionais com dois estágios, visto que, a arquitetura YOLO consiste em único estágio
- Incluir e avaliar métodos que estimem os hiperparâmetros a serem utilizados pela rede YOLO.
- Realizar recortes de diferentes dimensões na imagem original, ampliando a quantidade de amostras, aumentando assim a capacidade de generalização do modelo.

Referências Bibliográficas

- [1] RODRIGUES, S. *Mundo do pescador*. [S.l.]: Blogspot, 2022. <http://mundopescadorweb.blogspot.com/2014/11/peixe-robalo-flexa.html>. Access on 06 de Abril de 2022.
- [2] FÉLIX, R. T. d. S. et al. Desenvolvimento ovariano de *plagioscion squamosissimus* (heckel, 1840)(actinopterygii, perciformes), no reservatório de pedra, rio de contas, bahia. *Biota Neotropica*, SciELO Brasil, v. 9, n. 3, p. 131–136, 2009.
- [3] SAHA, S. A comprehensive guide to convolutional neural networks. *LURL*<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>), 2018.
- [4] HAFEMANN, L. G. An analysis of deep neural networks for texture classification. 2014.
- [5] REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 779–788.
- [6] CHENG, S.; ZHAO, K.; ZHANG, D. Abnormal water quality monitoring based on visual sensing of three-dimensional motion behavior of fish. *Symmetry*, Multidisciplinary Digital Publishing Institute, v. 11, n. 9, p. 1179, 2019.
- [7] ARAGÃO, M. V. S. P. de. *Seleção de retângulos envolventes em arquiteturas para detecção de objetos*. Tese (Doutorado) — PUC-Rio, 2021.
- [8] REDMON, J.; FARHADI, A. *YOLOv3: An Incremental Improvement*. 2018.
- [9] BOCHIKOVSKIY, A.; WANG, C.-Y.; LIAO, H.-Y. M. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.
- [10] THUAN, D. Evolution of yolo algorithm and yolov5: the state-of-the-art object detection algorithm. 2021.

- [11] PAULY, D.; ZELLER, D. Comments on faos state of world fisheries and aquaculture (sofia 2016). *Marine Policy*, Elsevier, v. 77, p. 176–181, 2017.
- [12] PEIXE, B. *Anuário 2020 Peixe BR da Piscicultura*. [S.L.]: São Paulo: Peixe Br, 2020.
- [13] PEREIRA, M. E. G. de S. et al. Contribuição dos atores da pesca para a bioecologia do camurim *centropomus undecimalis* (centropomidae–perciformes) capturado no litoral amazônico brasileiro. *Research, Society and Development*, v. 9, n. 10, p. e4119108691–e4119108691, 2020.
- [14] ROSA, R. S.; LIMA, F. C. Os peixes brasileiros ameaçados de extinção. *Livro vermelho da fauna brasileira ameaçada de extinção*, MMA Brasília, v. 2, p. 9–275, 2008.
- [15] SIVAKUMARAN, K. P. et al. Maturation and reproductive biology of female wild carp, *cyprinus carpio*, in victoria, australia. *Environmental Biology of Fishes*, Springer, v. 68, n. 3, p. 321–332, 2003. ISSN 1573-5133.
- [16] KARLOU-RIGA, C.; ECONOMIDIS, P. S. Ovarian atretic rates and sexual maturity of european horse mackerel, *trachurus trachurus* (l.), in the saronikos gulf (greece). *Fishery. Bulletin.*, v. 94, p. 66–76, 1996.
- [17] MENDONÇA, M. C. F. B. d. et al. Autoecologia do camorim, *centropomus undecimalis* (bloch, 1792),(perciformes: Centropomidae) em ambiente hipersalino em galinhos, rn, brasil. Universidade Federal de São Carlos, 2004.
- [18] RHODY, N. et al. Assessing reproductive condition in captive and wild common snook stocks: a comparison between the wet mount technique and histological preparations. *Transactions of the American Fisheries Society*, Taylor & Francis, v. 142, n. 4, p. 979–988, 2013.
- [19] ANTA, P. et al. Combining region and edge information to extract fish oocytes in histological images. In: *7th IASTED International Conference on Visualization, Imaging and Image Processing*. [S.L.: s.n.], 2007. p. 82–87.
- [20] CERNADAS, E. et al. Recognize and classify fish oocytes in histological images. In: *8th IASTED International Conference on Visualization, Imaging and Image Processing*. [S.L.: s.n.], 2008. p. 180–186.

- [21] PÉREZ-ORTIZ, M. et al. On the use of nominal and ordinal classifiers for the discrimination of states of development in fish oocytes. *Neural Processing Letters*, Springer Science and Business Media LLC, v. 44, n. 2, p. 555–570, oct 2015.
- [22] PINTOR, J. et al. Govocitos: A software tool for estimating fish fecundity based on digital analysis of histological images. *Computers and Electronics in Agriculture*, Elsevier BV, v. 125, p. 89–98, jul 2016.
- [23] ZOU, Z. et al. Object detection in 20 years: A survey. *ArXiv*, abs/1905.05055, 2019.
- [24] GONZÁLEZ-RUFINO, E. et al. Exhaustive comparison of colour texture features and classification methods to discriminate cells categories in histological images of fish ovary. *Pattern Recognition*, Elsevier BV, v. 46, n. 9, p. 2391–2407, sep 2013.
- [25] GONZÁLEZ-RUFINO, E. et al. Statistical and wavelet based texture features for fish oocytes classification. In: *Pattern Recognition and Image Analysis*. [S.L.]: Springer Berlin Heidelberg, 2011. p. 403–410.
- [26] PINTOR, J. M. et al. A multi-platform graphical software for determining reproductive parameters in fishes using histological image analysis. In: *Pattern Recognition and Image Analysis*. [S.L.]: Springer International Publishing, 2015. p. 743–750.
- [27] İNİK Özkan et al. A new method for automatic counting of ovarian follicles on whole slide histological images based on convolutional neural network. *Computers in Biology and Medicine*, Elsevier BV, v. 112, p. 103350, sep 2019.
- [28] SIRINUKUNWATTANA, K. et al. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Transactions on Medical Imaging*, Institute of Electrical and Electronics Engineers (IEEE), v. 35, n. 5, p. 1196–1206, may 2016.
- [29] XU, J. et al. Deep learning for histopathological image analysis: Towards computerized diagnosis on cancers. In: *Deep Learning and Convolutional Neural Networks for Medical Image Computing*. [S.L.]: Springer International Publishing, 2017. p. 73–95.

- [30] CHANG, Y. H. et al. Deep learning based nucleus classification in pancreas histological images. In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.]: IEEE, 2017.
- [31] ROY, K. et al. Patch-based system for classification of breast histology images using deep learning. *Computerized Medical Imaging and Graphics*, Elsevier BV, v. 71, p. 90–103, jan 2019.
- [32] WENZHONG, L. et al. Classifications of breast cancer images by deep learning. Cold Spring Harbor Laboratory, jun 2020.
- [33] GOLATKAR, A.; ANAND, D.; SETHI, A. Classification of breast cancer histology using deep learning. In: *Lecture Notes in Computer Science*. [S.l.]: Springer International Publishing, 2018. p. 837–844.
- [34] HUSSAIN, Z. et al. Differential data augmentation techniques for medical imaging classification tasks. v. 2017, p. 979, 2017.
- [35] LECUN, Y.; KAVUKCUOGLU, K.; FARABEL, C. Convolutional networks and applications in vision. In: *IEEE. Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*. [S.l.], 2010. p. 253–256.
- [36] NUNES, Z. M. P. et al. Bioecologia do robalo-flexa, *centropomus undecimalis*, em lagoa costeira tropical no norte do brasil. *Boletim do Instituto de Pesca*, v. 41, n. 3, p. 457–469, 2018.
- [37] MENEZES, N. A.; FIGUEIREDO, J. L. Manual de peixes marinhos do sudeste do brasil. v. 4, p. 23–26.
- [38] RIVAS, L. R. Systematic review of the perciform fishes of the genus *centropomus*. *Copeia*, JSTOR, p. 579–611, 1986.
- [39] FISHBASE. *Centropomus Undecimalis*. 2019. Disponível em: <<http://www.fishbase.org/summary/Centropomus-undecimalis.html>>.
- [40] SOUSA, M. L. N. M. d. *Caracterização morfológica de gônadas de machos do robalo, centropomus undecimalis (BLOCK, 1970) oriundos de cultivo e de ambiente natural*. Dissertação (Mestrado) — Dissertação de Mestrado em Ciências Veterinárias - Universidade Estadual do Ceará, 2011.

- [41] TAYLOR, R. G. et al. Age, growth, maturation, and protandric sex reversal in common snook, *centropomus undecimalis*, from the east and west coasts of south florida. *Fishery Bulletin*, National Marine Fisheries Service, v. 98, n. 3, p. 612–612, 2000.
- [42] BARROSO, M. V. et al. Estratégias de conservação das populações de robalos *centropomus spp.* na foz do rio doce, linhares, espírito santo. *Cadernos de Agroecologia*, v. 2, n. 2, 2007.
- [43] VIEIRA, A. V.; TOIGO, F. T. Classificação bi-rads: categorização de 4.968 mamografias. *Radiol Bras*, SciELO Brasil, v. 35, n. 4, p. 205–8, 2002.
- [44] VAZZOLER, A. E. A. d. M. [S.l.: s.n.].
- [45] KANG, K.; WANG, X. Fully convolutional neural networks for crowd segmentation. *arXiv preprint arXiv:1411.4464*, 2014.
- [46] DUMOULLIN, V.; VISIN, F. A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*, 2016.
- [47] NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*. [S.l.: s.n.], 2010. p. 807–814.
- [48] BRIDLE, J. S. Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters. In: *Advances in neural information processing systems*. [S.l.: s.n.], 1990. p. 211–217.
- [49] SILVA, J. G. J. d. et al. Redes neurais profundas para detecção e contagem automática de moscas brancas no contexto agrícola. Universidade Federal de Goiás, 2021.
- [50] HE, K. et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 37, n. 9, p. 1904–1916, 2015.
- [51] YIN, Y.; LI, H.; FU, W. Faster-yolox: An accurate and faster object detection method. *Digital Signal Processing*, Elsevier, v. 102, p. 102756, 2020.

- [52] WANG, C.-Y. et al. Cspnet: A new backbone that can enhance learning capability of cnn. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. [S.l.: s.n.], 2020. p. 390–391.
- [53] YANG, G. et al. Face mask recognition system with yolov5 based on image recognition. In: *IEEE. 2020 IEEE 6th International Conference on Computer and Communications (ICCC)*. [S.l.], 2020. p. 1398–1404.
- [54] MARQUES, B. H. P. *Avaliação de algoritmos baseados em Deep Learning para Localizar placas veiculares brasileiras em ambientes complexos*. Dissertação (B.S. thesis) — Brasil, 2019.
- [55] LIU, L. et al. Deep learning for generic object detection: A survey. *International journal of computer vision*, Springer, v. 128, n. 2, p. 261–318, 2020.
- [56] NELSON, J. *Roboflow [Computer software]*. [S.l.]: Roboflow, 2022. <https://roboflow.com>. Access on 26 de Janeiro de 2022.
- [57] EVERINGHAM, M. et al. The pascal visual object classes (voc) challenge. *International journal of computer vision*, Springer, v. 88, n. 2, p. 303–338, 2010.
- [58] CLARO, M. et al. Utilização de técnicas de data augmentation em imagens: Teoria e prática. *Sociedade Brasileira da Computação-SBC*. Available at: <https://sol.sbc.org.br/livros/index.php/sbc/catalog/download/48/217/455-1>, 2020.
- [59] NIE, Y. et al. Automatic detection of melanoma with yolo deep convolutional neural networks. In: *IEEE. 2019 E-Health and Bioengineering Conference (EHIB)*. [S.l.], 2019. p. 1–4.
- [60] SILVA, F. *Um Modelo de Recuperação de Informação para a Web Semântica*. Dissertação (Mestrado) — Dissertação de Mestrado em Engenharia de Eletricidade - Universidade Federal do Maranhão, 2009.
- [61] LUO, H.-W. et al. Contextual-yolov3: Implement better small object detection based deep learning. In: *IEEE. 2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*. [S.l.], 2019. p. 134–141.
- [62] ULTRALYTICS. *Tips for Best Training Results*. [S.l.]: YOLOv5 Documentation, 2022. <https://docs.ultralytics.com/tutorials/training-tips-best-results/>. Access on 26 de Janeiro de 2022.

- [63] ABAS, S. M.; ABDULAZEEZ, A. M.; ZEEBAREE, D. Q. A yolo and convolutional neural network for the detection and classification of leukocytes in leukemia. *Indonesian Journal of Electrical Engineering and Computer Science*, v. 25, n. 1, p. 200–213, 2022.
- [64] SOZZI, M. et al. Automatic bunch detection in white grape varieties using yolov3, yolov4, and yolov5 deep learning algorithms. *Agronomy*, Multidisciplinary Digital Publishing Institute, v. 12, n. 2, p. 319, 2022.
- [65] KUBERA, E. et al. Detection and recognition of pollen grains in multilabel microscopic images. *Sensors*, MDPI, v. 22, n. 7, p. 2690, 2022.
- [66] WANG, D. et al. Using an improved yolov4 deep learning network for accurate detection of whitefly and thrips on sticky trap images. *Transactions of the ASABE*, American Society of Agricultural and Biological Engineers, v. 64, n. 3, p. 919–927, 2021.
- [67] BUCKLAND, M.; GEY, F. The relationship between recall and precision. *Journal of the American society for information science*, Wiley Online Library, v. 45, n. 1, p. 12–19, 1994.