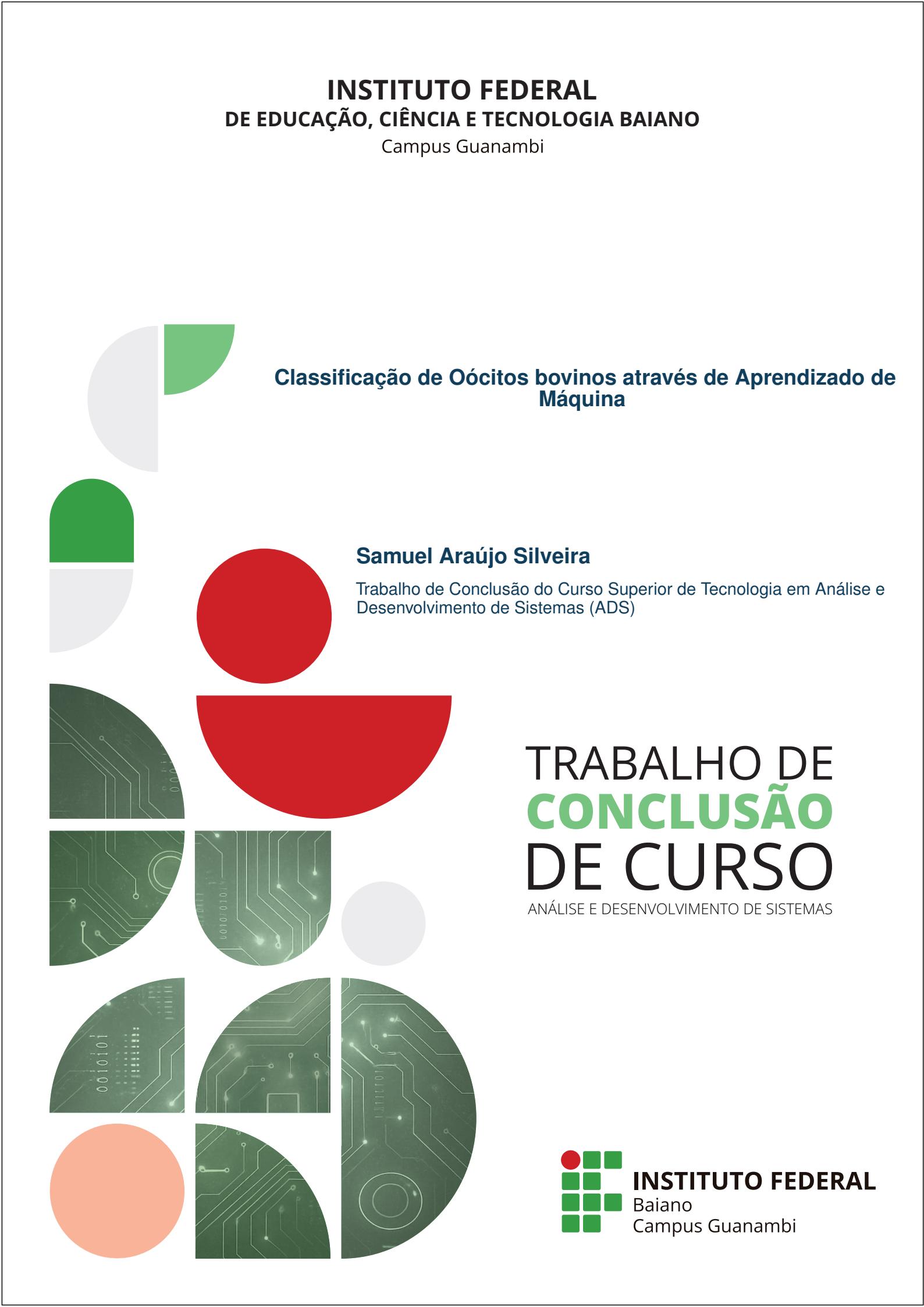


**INSTITUTO FEDERAL
DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA BAIANO**
Campus Guanambi



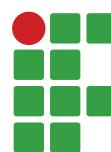
**Classificação de Oócitos bovinos através de Aprendizado de
Máquina**

Samuel Araújo Silveira

Trabalho de Conclusão do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas (ADS)

**TRABALHO DE
CONCLUSÃO
DE CURSO**

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS



INSTITUTO FEDERAL
Baiano
Campus Guanambi

BIBLIOTECA DO IF BAIANO - CAMPUS GUANAMBI

Data de Entrega: _____

Assinatura: _____

BANCA EXAMINADORA

Titulação. Nome do Professor (Orientador)
Instituto Federal Baiano - Campus Guanambi

Titulação. Nome do Professor (Membro)
Instituto Federal Baiano - Campus Guanambi

Titulação. (Membro)
Instituto Federal Baiano - Campus Guanambi

Trabalho Aprovado. Guanambi-BA, 29 de Maio de 2019

Samuel Araújo Silveira

Classificação de Oócitos bovinos através de Aprendizado de Máquina

Trabalho de Conclusão de Curso apresentado ao Instituto Federal de Educação, Ciência e Tecnologia Baiano - Campus Guanambi, ligado ao Ministério da Educação como parte dos requisitos para obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Área de Concentração: Análise e Desenvolvimento de Sistemas e Informática Acessível

Orientador: Prof. Dr. Carlos Anderson Oliveira Silva
Coorientador: Prof. Dr. José Assunção Silveira Júnior

Sobrenome, Nome

Título do Trabalho de Conclusão de Curso/ Nome Sobrenome. – Guanambi, Bahia,
2019.

1 f. : il. (alguma cor.)

Orientador: Nome do Orientador Sobrenome do Orientador

Trabalho de Conclusão de Curso (Graduação - Tecnólogo em Análise e Desenvolvimento
de Sistemas Dados) – Instituto Federal de Educação, Ciência e Tecnologia Baiano,
Campus Guanambi, 2019.

1. Palavra 1.
 2. Palavra 2.
 3. Palavra 3.
 4. Palavra 4.
- I. Sobrenome do Orientador,
 - Nome do Orientador.
 - II. Instituto Federal de Educação, Ciência e Tecnologia Baiano.
 - III. Campus Guanambi.
 - IV. Título do Trabalho de Conclusão de Curso.
-

Samuel Araújo Silveira

Classification of Bovine Oocytes through Machine Learning

Completion of course work submitted to the Federal Institute of Education, Science and Technology Baiano - Campus Guanambi, Ministry of Education as part of the requirements for obtaining the title of Technologist in Systems Analysis and Development.

Concentration Area: Systems Analysis and Development and Affordable Computing

Advisor: Prof. Dr. Carlos Anderson Oliveira Silva

Co-advisor: Prof. Dr. José Assunção Silveira Júnior

*“As invenções são, sobretudo,
o resultado de um trabalho de teimoso.”*
(Santos Dumont)

RESUMO

SILVEIRA, S. A. Classificação de Oócitos bovinos através de Aprendizado de Máquina. 2024. 48 p. Trabalho de Conclusão de Curso (Trabalho de Conclusão em Análise e Desenvolvimento de Sistemas) – Instituto Federal de Educação, Ciência e Tecnologia Baiano, Guanambi – BA, 2024.

A pecuária ocupa uma posição relevante no contexto socioeconômico brasileiro, já que o Brasil está entre os maiores exportadores de leite e carne bovina do mundo. O que torna isso possível são os métodos de reprodução animal assistida, que garantem um maior número de descendentes com boas genéticas. Um desses métodos é a produção *in vitro* cujas duas etapas iniciais são as de coleta e classificação de células reprodutoras. Atualmente, a etapa de classificação precisa ser feita de forma manual e apenas por profissionais capacitados, o que causa um prejuízo grave na eficiência do processo. Dessa forma, este trabalho desenvolveu um modelo de aprendizado de máquina capaz de classificar oócitos bovinos com o objetivo de aprimorar a eficiência de procedimentos de reprodução animal assistida. Os métodos empregados para automatizar esse processo são baseados em redes neurais convolucionais e processamento digital de imagens. Com isso, pretende-se reconhecer e classificar oócitos bovinos, aumentar a eficiência da produção *in vitro* e aprimorar a qualidade genética do gado brasileiro através de uma melhor classificação e seleção dos gametas femininos bovinos. A pesquisa demonstra uma colaboração entre computação e pecuária que pode superar práticas tradicionais através da proposta de soluções mais eficientes.

Palavras-chave: Aprendizado de Máquina, Classificação de Oócitos, Reprodução Animal Assistida, Redes Neurais Convolucionais, Agropecuária.

ABSTRACT

SILVEIRA, S. A. **Classification of Bovine Oocytes through Machine Learning.** 2024. 48 p. Trabalho de Conclusão de Curso (Trabalho de Conclusão em Análise e Desenvolvimento de Sistemas) – Instituto Federal de Educação, Ciência e Tecnologia Baiano, Guanambi – BA, 2024.

Livestock farming occupies a relevant position in the Brazilian socioeconomic context, since Brazil is among the largest exporters of milk and beef in the world. What makes this possible are the methods of assisted animal reproduction, which guarantee a greater number of descendants with good genetics. One of these methods is *in vitro* production, whose two initial stages are the collection and classification of reproductive cells. Currently, the classification stage must be done manually and only by trained professionals which causes a massive problem in the efficiency of this process. Thus, this work developed a system capable of classifying bovine oocytes with the objective of improving the efficiency of assisted animal reproduction procedures. The methods used to automate this process are based on convolutional neural networks and digital image processing. With this, the aim is to recognize and classify bovine oocytes n, increase the efficiency of *in vitro* production, and improve the genetic quality of Brazilian cattle through better classification and selection of female bovine gametes. The research demonstrates a collaboration between computing and livestock that can overcome traditional practices by proposing more efficient solutions.

Keywords: Machine Learning, Oocyte Classification, Assisted Animal Reproduction, Convolutional Neural Networks, animal husbandry.

LISTA DE FIGURAS

Figura 1 – Representação dos componentes de um oócito	18
Figura 2 – Representação de um oócito com mais de três camadas de células da granulosa	19
Figura 3 – Representação de um oócito com menos de três camadas de células da granulosa e com ooplasma heterogêneo	20
Figura 4 – Representação dos componentes de um oócito com espaço entre as células da granulosa e a zona pelúcida	21
Figura 5 – Representação dos componentes de um oócito sem camadas de células da granulosa	22
Figura 6 – Rede Neural Simples e Rede Neural Profunda	26
Figura 7 – Arquitetura da LeNet-5, uma Rede Neural Convolucional	28
Figura 8 – Etapas da Segmentação de Instância	29
Figura 9 – Esquema de funcionamento do modelo proposto	31
Figura 10 – Demonstração das componentes de um Microscópio Estereoscópio	32
Figura 11 – Arquitetura do Modelo de Aprendizado de Máquina para Identificação de Oócitos	34
Figura 12 – Imagem de Entrada e Saída no Modelo de Identificação de Oócitos	35
Figura 13 – Arquitetura do Modelo de Aprendizado de Máquina para Classificação de Oócitos	36
Figura 14 – Demonstração da classificação de um oócito bovino	37
Figura 15 – Demonstração do funcionamento do Modelo Final	38
Figura 16 – Fórmula d IoU	40

LISTA DE TABELAS

Tabela 1 – Métricas de Precisão Média (<i>Average Precision - AP</i>) do COCOEvaluator	39
Tabela 2 – Métricas Globais de Precisão Média (<i>Average Precision - AP</i>) do Modelo de Classificação.	41

SUMÁRIO

1	INTRODUÇÃO	11
2	OBJETIVOS	13
2.1	Objetivo Geral	13
2.2	Objetivos Específicos	13
3	JUSTIFICATIVA	14
4	REVISÃO BIBLIOGRÁFICA/TEÓRICA	16
4.1	Produção <i>in vitro</i>	16
4.1.1	<i>Classificação de Oócitos</i>	17
4.1.1.1	<i>Classificação I</i>	18
4.1.1.2	<i>Classificação II</i>	19
4.1.1.3	<i>Classificação III</i>	20
4.1.1.4	<i>Classificação IV</i>	21
4.2	Visão Computacional	22
4.2.1	<i>Processamento Digital de Imagens</i>	23
4.2.2	<i>OpenCV</i>	24
4.3	Inteligência Artificial	24
4.3.1	<i>Aprendizado de Máquina</i>	25
4.3.2	<i>Redes Neurais Artificiais</i>	25
4.3.2.1	<i>Redes Neurais Profundas</i>	26
4.3.2.2	<i>Redes Neurais Convolucionais</i>	26
4.3.2.3	<i>Segmentação de Instância</i>	28
4.3.2.4	<i>R-CNN</i>	29
4.3.2.5	<i>Mask R-CNN</i>	30
4.3.2.6	<i>f1-score</i>	30
5	METODOLOGIA	31
5.1	Visualização e Coleta	32
5.2	Pré-processamento	33
5.2.1	<i>OpenCV</i>	33
5.2.2	<i>LabelMe</i>	33
5.3	Construção do Modelo de Identificação	34

5.4	Construção do Modelo de Classificação	35
5.5	Modelo Final	37
6	RESULTADOS E DISCUSSÃO	39
6.1	Avaliação do Modelo de Identificação	39
6.2	Avaliação do Modelo de Classificação	41
7	CONCLUSÃO	43
	REFERÊNCIAS	44



INTRODUÇÃO

A pecuária possui grande relevância em diversos âmbitos nacionais, seja na nutrição do povo brasileiro, na geração de empregos ou em relações comerciais internacionais (Embrapa, 2024). Para sustentar essa importância, o Brasil possui 1,8 milhão de propriedades leiteiras, 3,6 milhões de empregos gerados apenas pelo segmento da pecuária leiteira (Netto; Gomes, 2021) e a exportação de carne bovina representa 6% do Produto Interno Bruto (PIB) nacional (Embrapa, 2024).

De acordo com Oliveira, Serapião e Quintão (2014), a base da pecuária e suas ramificações, não só do Brasil, mas do mundo, é formada por técnicas de reprodução animal assistida. Isso acontece porque essas técnicas são úteis tanto para aumentar consideravelmente a quantidade de cabeças de gado, quanto para aprimorar a qualidade genética dos animais.

Nesse contexto, existe a Produção *in vitro* (PIV) como uma importante técnica de reprodução assistida. Esse método permite que haja uma seleção das melhores células reprodutoras antes de serem enviadas para os laboratórios responsáveis por finalizar o procedimento. Essa seleção precisa ser precedida de uma classificação morfológica e funcional, que ocorre logo após a coleta das células de interesse (Mello *et al.*, 2016). Para realizar a classificação, é necessário um profissional da medicina veterinária para analisar as características de célula por célula. Portanto, nota-se como esse processo está dependente de um trabalho longo, manual e condicionado à profissionais capacitados. Por conta disso, este trabalho tem por objetivo criar um modelo computacional capaz de identificar e classificar oócitos bovinos para automatizar esse processo.

Esses modelos computacionais são pertencentes à outra área de grande relevância para a atualidade, que é a Tecnologia da Informação (TI). Um de seus princípios fundamentais consiste em atuar de forma integrada com outras áreas do conhecimento, visando automatizar processos ou aprimorar a eficiência de procedimentos tradicionalmente manuais. Tal contribuição é possível porque os sistemas computacionais são capazes de processar grandes volumes de dados em curto tempo e de forma isenta do enviesamento característico das análises humanas (Henrique, 2024).

Através de tecnologias modernas de inteligência artificial, é possível desenvolver modelos de aprendizado de máquina que classifiquem automaticamente oócitos, as células reprodutoras das vacas. Com uma classificação aprimorada, é possível também melhorar a etapa de seleção dos oócitos que serão os enviados para os laboratórios de PIV. Dessa forma, o aprimoramento genético oriundo desse processo pode ocasionar uma grande melhoria na qualidade do gado brasileiro (Mueller; Eenennaam, 2022).

Por fim, a automatização do reconhecimento e classificação de oócitos bovinos é realizada através da coleta de imagens dessas células reprodutoras em laboratório, pré-processamento de cada uma delas e treinamento de dois modelos diferentes: um para identificação e outro para classificação.



OBJETIVOS

2.1 Objetivo Geral

Construir um modelo capaz de reconhecer e classificar oócitos bovinos quanto à sua morfologia e viabilidade funcional, com vistas ao processo de reprodução animal assistida.

2.2 Objetivos Específicos

Com o desdobramento do objetivo geral, pretende-se atingir os seguintes objetivos específicos:

- Aplicar técnicas de aprendizado de máquina nas tarefas e reconhecimento e classificação e oócitos;
- Estudar técnicas de processamento digital de imagens;
- Analisar a capacidade de generalização de métodos de *machine learning* aplicada a reprodução animal;
- Aplicar e avaliar métodos de Redes Neurais Convolucionais nas tarefas de reconhecimento e classificação de oócitos.



JUSTIFICATIVA

A economia brasileira é fortemente marcada por atividades econômicas que envolvem bovinos. O Brasil detém o segundo maior rebanho do mundo, é o segundo maior produtor de carne bovina (FAOSTAT, 2021), é o maior exportador e é o quinto maior produtor de leite (FAO, 2019). A criação de bovinos é muito importante para a geração de empregos, para a economia nacional e para a produção de alimentos altamente nutritivos para o Brasil e para o mundo (Teixeira; Hespanhol, 2014).

Isso se deve ao fato de que o Brasil, por estar localizado em clima tropical, não pôde copiar as tecnologias de pecuária de países de clima temperado, como os Estados Unidos da América. Essa barreira impulsionou o avanço tecnológico e fez com que os brasileiros desenvolvessem suas próprias tecnologias, que acabaram servindo de modelo para outros países de clima tropical (Camargo; Pereira, 2022).

As raças bovinas mais produtivas evoluíram e foram melhoradas geneticamente em países de clima temperado, inicialmente na Europa, posteriormente na América do Norte e em outras regiões de clima temperado. Como a maior parte do Brasil está localizada na faixa tropical, é necessário o desenvolvimento de animais adaptados à esta condição ambiental (Oosting; Udo; Viets, 2014), o que está sendo feito com enorme pioneirismo e competência por parte dos setores de produção de carne e de leite do país. Para acelerar este melhoramento genético, as técnicas de reprodução assistidas são imprescindíveis, como a inseminação artificial e a produção *in vitro* de embriões (Sales *et al.*, 2024).

A produção *in vitro* possui várias etapas no seu funcionamento e uma delas é a classificação de células reprodutoras, como os gametas femininos nomeados de oócitos (Mello *et al.*, 2016). Atualmente, essa etapa é realizada de forma manual (Shivaani; Madan, 2024), algo que ocasiona em um grande número de horas de trabalho por parte dos profissionais da medicina veterinária que realizam esse serviço.

Nesse contexto, surgem sistemas de informação que utilizam modelos inteligentes que

pode aprimorar muito a eficiência de procedimentos já existentes e realizados por humanos. Dessa forma, percebe-se como há espaço para uma parceria entre pecuária e computação com o objetivo de agilizar e aprimorar a produção *in vitro*.

Portanto, essa proposta se justifica pelo impacto social e econômico que ela pode produzir para a pecuária nacional.



REVISÃO BIBLIOGRÁFICA/TEÓRICA

Durante a evolução dos mamíferos, a necessidade de garantir a continuidade das espécies fez com que esses animais desenvolvessem e aprimorassem seus métodos reprodutivos. Essas reproduções acontecem através da união do gameta masculino, chamado de espermatozoide, com o gameta feminino, conhecido como oócito. Essa união origina um zigoto, que se divide uma série de vezes até gerar um embrião (Hafez; Hafez, 2004).

No entanto, a atualidade, junto com sua tecnologia, possui estratégias que facilitam e aprimoram todo esse processo, como métodos modernos de reprodução assistida. Dentre eles, existe a produção *in vitro* de embriões (PIV), que garante uma série de benefícios para a reprodução.

4.1 Produção *in vitro*

O processo de desenvolvimento dessa técnica originou-se entre 1877 e 1899. Nessa época, pesquisadores buscavam estabelecer estratégias que permitissem a manipulação de embriões. Isso resultou na primeira visualização da fecundação de um oócito de estrela do mar que, por sua vez, originou um zigoto. Com o passar das décadas, mais especificamente em 1959, houve o nascimento do primeiro coelho gerado a partir da técnica de PIV (Mello *et al.*, 2016).

De acordo com Mello *et al.* (2016), a PIV é uma biotécnica cujo objetivo é explorar o potencial genético de fêmeas bovinas. Ela é estratificada nas seguintes etapas: coleta, maturação, fecundação e cultivo *in vitro*. Além do mais, segundo Gouveia (2011), a PIV é importante por conta de sua importância para estudos biotecnológicos e fundamentos comerciais.

Conforme a Embrapa (2021), a PIV apresenta as seguintes vantagens:

Possibilita a utilização de bezerras pré-púberes, vacas em início de gestação, vacas com subfertilidade adquirida, vacas senis e vacas mortas accidentalmente; produção de cerca de 36 bezerros por ano a partir de uma única fêmea; avanço na multiplicação

de fêmeas bovinas de interesse para a produção animal e para a conservação de raças de animais domésticos ameaçadas de extinção; facilita o uso e aprimoramento de técnicas avançadas de reprodução animal, como: clonagem; injeção intracitoplasmática de espermatozoides; e transgenia; permite otimizar o uso de sêmen de reprodutores alto valor genético e de sêmen sexado; permite a produção de embriões com grau de sangue e sexo definidos para atender a programas específicos de produção (leite e carne), em larga escala e com menor custo.

Entre as etapas de coleta e maturação, precisa existir uma fase de classificação dos oócitos coletados (Filho *et al.*, 2011). Isso é importante porque algumas células não possuem as propriedades necessárias para seguir em frente no processo e algumas são melhores que outras.

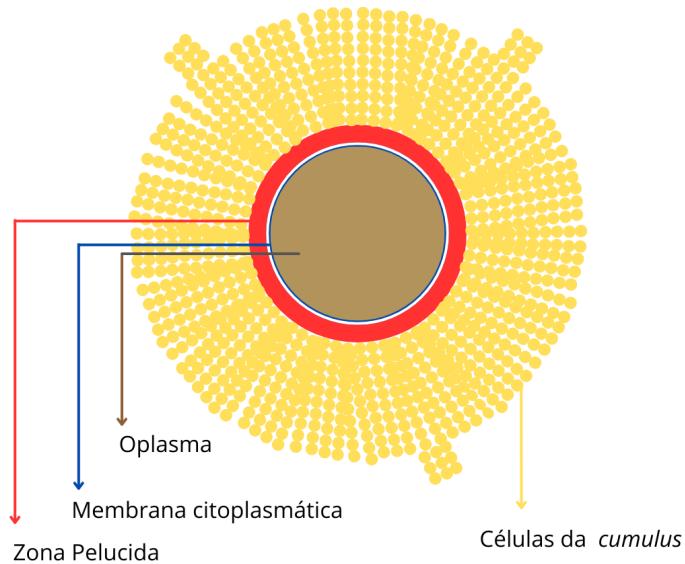
4.1.1 Classificação de Oócitos

Os oócitos coletados precisam ser divididos em quatro categorias, de acordo com as propriedades das células do cumulus e do ooplasma (Filho *et al.*, 2011). Esse método de classificação foi desenvolvido por Leibfried e First (1979).

De acordo com SPMR (2024), as células do cumulus, um dos fatores de classificação mencionados, envolvem o oócio e desempenham papéis essenciais como garantir a comunicação bidirecional e proteger o oócio. A estrutura dessas células está ilustrada na Figura 1.

Além do mais, segundo Elsevier (2024), o ooplasma, que é o segundo fator de classificação, é apenas o citoplasma do oócio. Ele é um líquido viscoso, localizado dentro da membrana citoplasmática e armazena toda a estrutura interna da célula (Elsevier, 2024). Seu posicionamento está representada na Figura 1.

Figura 1 – Representação dos componentes de um oócito



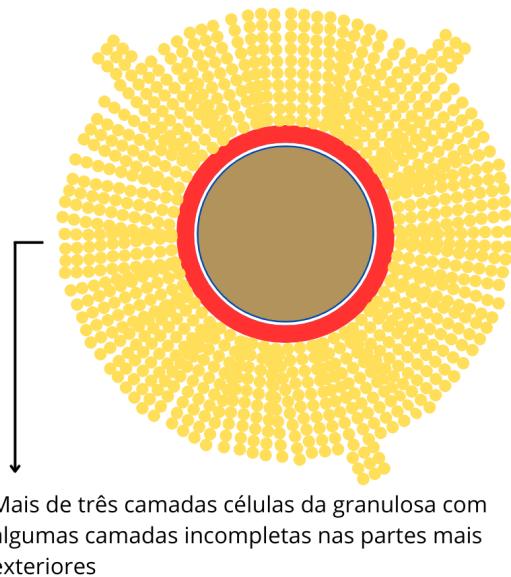
Fonte: Elaboração própria.

Por fim, de acordo com Filho *et al.* (2011), os oócitos viáveis são os de classificação I a III, enquanto os de classificação IV são desconsiderados.

4.1.1.1 Classificação I

Nesta categoria, as células do cumulus ou células da granulosa precisam formar, no mínimo, mais de três camadas (Filho *et al.*, 2011). Enquanto isso, o ooplasm precisa possuir coloração marrom com granulações finas e homogêneas (Filho *et al.*, 2011), conforme ilustrado na Figura 2

Figura 2 – Representação de um oócito com mais de três camadas de células da granulosa



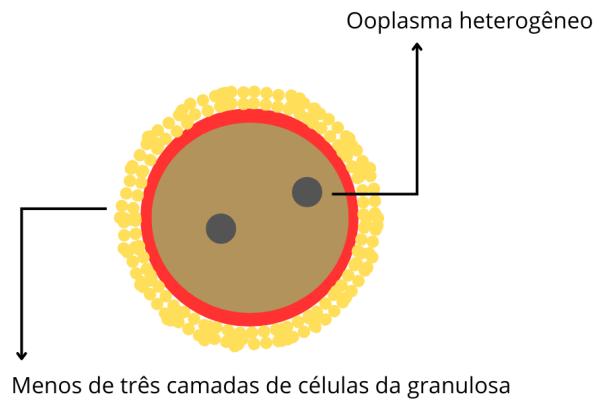
Mais de três camadas células da granulosa com algumas camadas incompletas nas partes mais exteriores

Fonte: Elaboração própria.

4.1.1.2 Classificação II

Nesta classificação, os oóцитos possuem menos de 3 camadas de células da granulosa, enquanto o ooplasma possui granulações distribuídas heterogeneamente (Filho *et al.*, 2011), conforme ilustrado na Figura 3.

Figura 3 – Representação de um oócito com menos de três camadas de células da granulosa e com ooplasma heterogêneo

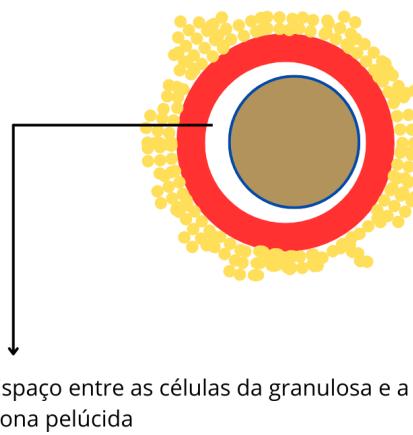


Fonte: Elaboração própria.

4.1.1.3 Classificação III

Nesta classificação, os oócitos possuem camadas de células da granulosa. No entanto, há um espaço entre a membrana celular e a zona pelúcida (Filho *et al.*, 2011), conforme ilustrado na figura 4.

Figura 4 – Representação dos componentes de um oócito com espaço entre as células da granulosa e a zona pelúcida

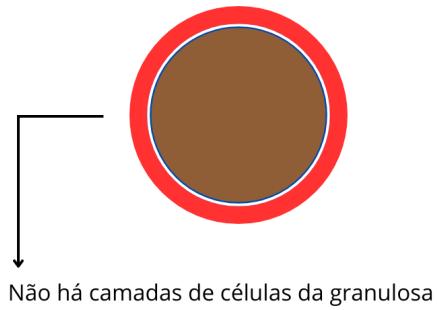


Fonte: Elaboração própria.

4.1.1.4 Classificação IV

Nesta classificação, os oócitos não possuem células da granulosa ou possuem citoplasma com cor e granulação fora do normal (Filho *et al.*, 2011), conforme a Figura 5.

Figura 5 – Representação dos componentes de um oócito sem camadas de células da granulosa



Fonte: Elaboração própria.

4.2 Visão Computacional

As ciências da computação possuem um ramo de estudo chamado de visão computacional (Ballard; Brown, 1982). Essa área visa aprimorar as capacidades computacionais de compreensão e interpretação de conteúdos visuais (Ballard; Brown, 1982). Isso permite a criação de sistemas robustos e inteligentes que podem ser utilizados em diversos campos do conhecimento humano.

O processamento de imagens é o cerne da visão computacional. Ele tem como objetivo coletar das imagens os tipos de dados definidos como parâmetros, ou seja, os que serão utilizados pelo sistema. Finalizada essa etapa, acontece o procedimento de classificação de entidades presentes na imagem (Jähne, 2002). Segundo (Jähne, 2002), esse processamento pode ser estratificado em uma série de passos até o resultado final ser alcançado. Alguns desses passos são dispensáveis, mas outros não.

De acordo com Jähne (2002), o primeiro passo deve ser o ato de aquisição de imagem, algo que pode ser feito através de câmeras digitais ou analógicas. No entanto, essa última forma torna obrigatório o processo de digitalização da imagem para que se torne adequada para o processo. Logo em seguida, efetua-se o pré-processamento cujo objetivo é eliminar eventuais imperfeições das imagens coletadas. Com essas etapas finalizadas, o reconhecimento de objetos pode ser realizado através da extração de características.

4.2.1 Processamento Digital de Imagens

A explicação sobre o que é uma imagem acontece através de uma função bidimensional, cuja representação é $f(x,y)$. Nessa função, x e y são coordenadas em um plano e a amplitude de f em qualquer ponto desse plano é conhecida como intensidade. Caso as quantidades de x , y e os valores de intensidade de f sejam finitas e discretas, a imagem será definida como digital. Além disso, a composição de uma imagem digital é um número finito de elementos cujas localizações e valores são particulares para um pixel (Gonzalez; Woods, 2010).

Segundo Gonzalez e Woods (2010), existem passos fundamentais que devem ser seguidos durante o processamento digital de imagens, sendo elas:

- **Aquisição de Imagem:** é a primeira etapa do processo e envolve o recebimento de imagens em formato digital. Também pode haver a necessidade de um pré-processamento para adequar a imagem ao sistema que a analisará;
- **Filtragem e Realce de Imagens:** esta etapa consiste em manipular uma imagem para que ela se torne mais apropriada para o processamento de um sistema específico. É importante frisar que esta fase varia de acordo com as necessidades de cada contexto;
- **Restauração de Imagens:** esta área também trata do aprimoramento visual de imagens. No entanto, há uma objetividade maior aqui, tendo em vista que seus métodos de aplicação são oriundos da matemática ou estatística;
- **Processamento de Imagens Coloridas:** a cor pode ser um importante fator descritivo de determinado objeto, ou seja, o sistema que puder fazer uso desse artifício pode vir a ser mais eficiente no processo de identificação;
- **Extração de Características:** também conhecida como processamento morfológico, essa etapa é responsável pela extração de componentes cuja importância é vital para a representação dos atributos da imagem;
- **Reconhecimento de Objetos:** da mesma forma que um humano é capaz de identificar objetos com o sentido da visão, uma máquina também pode ser capaz de realizar essa mesma identificação. No entanto, isso é apenas possível por meio da codificação de todas as etapas do processamento de imagens (Jain; Kasturi; Schunck, 1995);

Dito isso, é importante elucidar sobre um aspecto crucial para o desenvolvimento da acurácia do reconhecimento, que é o *dataset*. Essa tecnologia consiste em um grande aglomerado de dados que servirão para a realização de análises. Nessa coleção de dados, existem alguns que serão utilizados pelo sistema como uma referência positiva, para que seja possível deliberar se determinado objeto de análise poderá ser classificado como objeto de interesse. Enquanto

isso, as referências negativas servirão para que o sistema possa saber quais aspectos ignorar no processo de análise (Jain; Kasturi; Schunck, 1995).

As características de uma imagem deverão ser avaliadas individualmente assim que a fase de reconhecimento for iniciada. Primeiramente, o sistema irá verificar se a imagem possui certo grau de características candidatas a serem parte do objeto de interesse. Se esse resultado for positivo, a imagem será comparada com as do *dataset*. Dessa forma, se as comparações forem válidas, a imagem será classificada como um objeto de interesse (Jain; Kasturi; Schunck, 1995).

4.2.2 OpenCV

De acordo com a documentação OpenCV (2025), a OpenCV é uma sigla cujo significado é biblioteca de visão computacional e de código aberto. Essa ferramenta é gratuita, algo que a torna ideal para o desenvolvimento de projetos acadêmicos. Além do mais, ela é amplamente utilizada em *softwares* que aplicam o aprendizado de máquina.

A biblioteca OpenCV consiste em cinco aspectos fundamentais (Bradski; Kaehler, 2008):

- **CV:** esse componente possui o objetivo de realizar o processamento de imagens através da visão computacional;
- **MLL:** esse componente é responsável pelo aprendizado de máquina, ou seja, ele irá executar modelos treinados e manipular as saídas por meio dos parâmetros definidos;
- **HighGUI:** esse componente trata das questões relacionadas a imagens e vídeos. Além do mais, ele possui uma funcionalidade de geração de janelas de visualização das mídias que serão adicionadas ao código;
- **CXCore:** esse componente possui a responsabilidade de englobar todos os aspectos citados em uma apresentação simples e sucinta para que o desenvolvedor possa ter uma compreensão mais adequada;
- **CvAux:** permite que o programador utilize funções de reconhecimento facial e reconhecimento de gestos.

4.3 Inteligência Artificial

A Inteligência Artificial (IA) é um ramo das ciências da computação cujo objetivo é desenvolver sistemas de computadores capazes de pensar e agir como seres humanos. Essa tecnologia é capaz de desenvolver uma série de ferramentas capazes de aprimorar a produção nas mais diversas áreas do trabalho humano, como por exemplo: extração e processamento de dados, automação de processos, etc (Carvalho, 2021).

Dessa forma, percebe-se a evolução que a IA é capaz de realizar diante de algoritmos convencionais tendo em vista que eles apenas leem e executam as linhas sequenciais de um código, ou seja, não são capazes de aprender de nenhuma forma. Portanto, sua aplicabilidade é extremamente limitada (Sichman, 2021).

Anteriormente à IA, o ser humano era o único responsável pelo aprendizado e desenvolvimento de soluções para problemas complexos. No entanto, a inteligência artificial surge como uma alternativa para essa realidade (Sichman, 2021).

4.3.1 Aprendizado de Máquina

A Inteligência Artificial (IA) possui uma ramificação nomeada de Aprendizado de Máquina (AM) cujo foco é utilizar grandes quantidades de dados para criar e treinar modelos. Esses modelos, por sua vez, simulam o comportamento humano de resolução de problemas complexos que necessitam de interpretações para serem solucionados, ou seja, um algoritmo convencional não tem como resolvê-los (Smola; Vishwanathan, 2010).

De acordo com Deisenroth, Faisal e Ong (2020), o cerne do aprendizado de máquina gira em torno de dados, modelos e processamento. Os dados possuem uma característica basilar para o AM, já que são eles os responsáveis pela etapa crucial de treinamento. Os modelos, por sua vez, são responsáveis por lidar com situações não apresentadas no *dataset*, ou seja, aqui há a etapa de aprendizado.

Além do mais, segundo Agrawal (2020), o AM possui dois tipos de variáveis vitais para o aprendizado do modelo: os hiperparâmetros e os parâmetros. O primeiro é de caráter manual, pois é necessário determiná-los previamente em relação ao treinamento. O segundo é volátil, uma vez que o próprio modelo é incumbido da responsabilidade de alterá-los à medida que novas informações vão sendo alimentadas.

4.3.2 Redes Neurais Artificiais

De acordo com Haykin (2001), o Redes Neurais Artificiais (RNA) são estudadas e aplicadas pelo fato de que o cérebro humano é capaz de processar informações de forma muito mais eficiente do que um computador por causa das propriedades de complexidade, não-linearidade e paralelismo presentes no cérebro humano. Portanto, um sistema de aprendizado possui muito a ganhar caso consiga reproduzir tal estrutura.

Além do mais, as habilidade de generalização é outro fator que aumenta consideravelmente o poder computacional de uma RNA. A generalização é a capacidade de fornecer saídas adequadas para entradas que não estavam presente no *dataset* utilizado para treinar o sistema (Haykin, 2001).

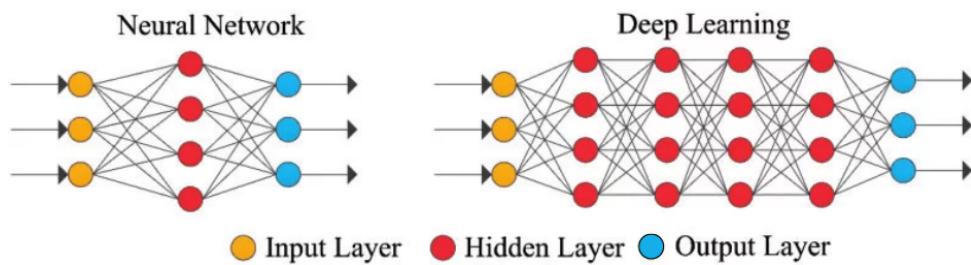
No entanto, as redes neurais não são capazes de fornecer resultados trabalhando isoladamente. Elas precisam ser integradas em um sistema maior. Isso acontece da seguinte forma: um

problema complexo e dividido em várias tarefas menores e simples, e então grupos de subtarefas serão entregues para redes neurais cuja especialização mais se assemelha às necessidades da tarefa (Haykin, 2001).

4.3.2.1 Redes Neurais Profundas

Embora Redes Neurais Profundas (RNP) sejam semelhantes à Redes Neurais Artificiais, há um acréscimo de camadas empilhadas que as diferencia, como ilustrado na figura 6.

Figura 6 – Rede Neural Simples e Rede Neural Profunda



Fonte: Adaptado de Bohani, Suliman, Saripuddin, Syarmila e Salleh (2021).

Em relação à essas camadas adicionais, quanto mais delas existirem na RNP, mais complexo o sistema será, mais recursos computacionais precisarão ser utilizados e mais tempo será necessário para treinar o modelo (Subasi, 2020).

De acordo com Mohanasundaram, Malhotra, Arun e Periasamy (2019), cada camada irá trabalhar com características diferentes baseadas na saída da camada anterior. Isso significa que, quanto mais se avançar na rede, mais complexas serão as características que as camadas reconhecerão tendo em vista que elas se rearranjam e aprendem com as características das camadas anteriores. A vantagem desse processo é que as redes conseguem modelar relações não lineares e complexas, algo que permite o tratamento de dados não rotulados e não estruturados.

Portanto, nota-se como as redes neurais profundas são capazes de coletar dados brutos, não rotulados, desestruturados e ainda conseguir agrupá-los e realizar o processamento deles (Mohanasundaram; Malhotra; Arun; Periasamy, 2019).

4.3.2.2 Redes Neurais Convolucionais

De acordo com Vargas, Paes e Vasconcelos (2016), uma Rede Neural Convolucional (RNC) nacional é uma variação das Redes Neurais Profundas cujo funcionamento é inspirado no processamento biológico de dados visuais. Ademais, semelhantemente a outros métodos da visão computacional, uma RNC também aplica filtros em dados visuais ao mesmo tempo que armazena os relacionamentos de vizinhança entre os *pixels* da imagem com o decorrer do processamento. Segundo Zhang, Yang, Chen e Li (2018), esse é o modelo mais utilizado de RNP

em aprendizado de características para classificação e reconhecimento de imagens em grande escala.

O funcionamento de uma RNC é dividido em três etapas principais, que são: a convolução, subamostragem e classificação. (Zhang; Yang; Chen; Li, 2018). A convolução, por sua vez, necessita de componentes básicos para funcionar adequadamente, que são: dados de entrada, filtros e mapas de características.

Segundo Alves (2018), as entradas da convolução são consideradas como matrizes tridimensionais cujas duas primeiras dimensões são a altura e a largura da imagem, enquanto a profundidade da terceira dimensão é definida por quantos canais de cores a imagem tem. Por exemplo: se determinada imagem seguir o padrão RGB, a profundidade da terceira dimensão será de três canais.

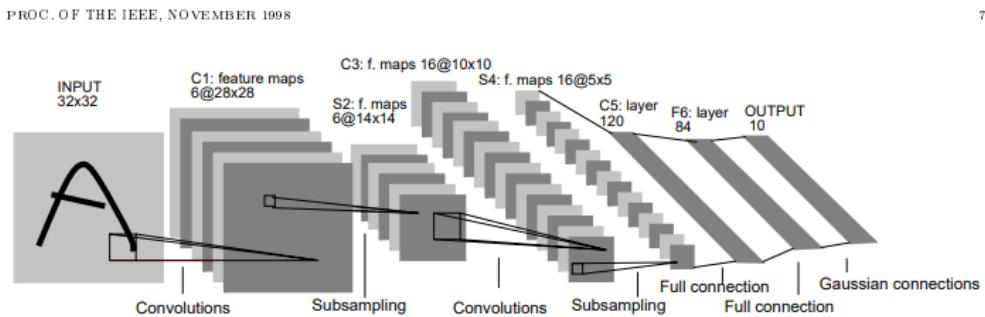
Já o filtro, é uma matriz bidimensional com altura e largura menores que as da entrada cuja função é percorrer a imagem para assimilar as principais características. Ela percorre a imagem de acordo com uma *stride*, que são os saltos feitos pelo filtro para se locomover. Quando todo o trajeto tiver sido percorrido, é gerado um mapa de características, que será a primeira *hidden layer* (Alves, 2018). Dessa forma, à cada convolução, a RNC amplia sua complexidade e identifica trechos maiores da mídia visual que foi fornecida como entrada. Isso significa que as primeiras camadas se concentram em aspectos simples, enquanto as camadas convolucionais finais irão reconhecer características mais complexas (IBM, 2024).

Logo em seguida, ocorre a subamostragem (*subsampling*). Ela irá simplificar a saída da camada anterior através de uma matriz bidimensional que irá percorrer toda a saída. À medida que essa matriz se locomove, ela escolhe o maior valor dos trechos percorridos e vai gerando outra matriz formada apenas pelos maiores valores. Isso resulta em uma generalização que irá consumir menos recursos computacionais nas etapas seguintes e impede o *overfitting*, que acontece quando um modelo de aprendizado de máquina se adapta excelentemente a um conjunto de dados, mas não funciona adequadamente com dados novos.

Por fim, a classificação ocorre na camada totalmente conectada através da combinação das características assimiladas pelas camadas anteriores (IBM, 2024).

Para ilustrar melhor esse procedimento, a Figura 7 abaixo mostra a estrutura de uma das primeiras Redes Neurais Convolucionais:

Figura 7 – Arquitetura da LeNet-5, uma Rede Neural Convolucional



Fonte: LeCun, Bottou, Bengio e Haffner (1998).

Como pode ser visto na imagem, uma matriz com vários valores percorre a imagem de entrada e quais serão esses valores depende exclusivamente do tipo de resultado desejado. A partir disso, os valores da matriz serão multiplicados por cada número que representa o *pixel*, os resultados das multiplicações serão somados e a somatória de todo esse processo ocupará um pixel da imagem de saída. Isso é repetido até a imagem inteira ser percorrida e um *feature map* ser gerado (Kumar, 2025).

Com isso, o procedimento de *subsampling* começa para reduzir as dimensões do *feature map* ao mesmo tempo que valores relevantes são mantidos com o objetivo de generalizar as informações importantes para o aprendizado do modelo, algo que impede o problema de *overfitting*, que ocorre quando um modelo se adapta demais a uma imagem específica e não funciona bem com outros objetos de interesse da mesma classificação (Kumar, 2025).

Logo em seguida, as etapas de *full connection* realizam a classificação com base nas características relevantes extraídas. Esse procedimento é feito da seguinte forma: as matrizes de cada *feature map* são achatadas para se tornarem unidimensionais e são transformadas em um único grande vetor (Kumar, 2025).

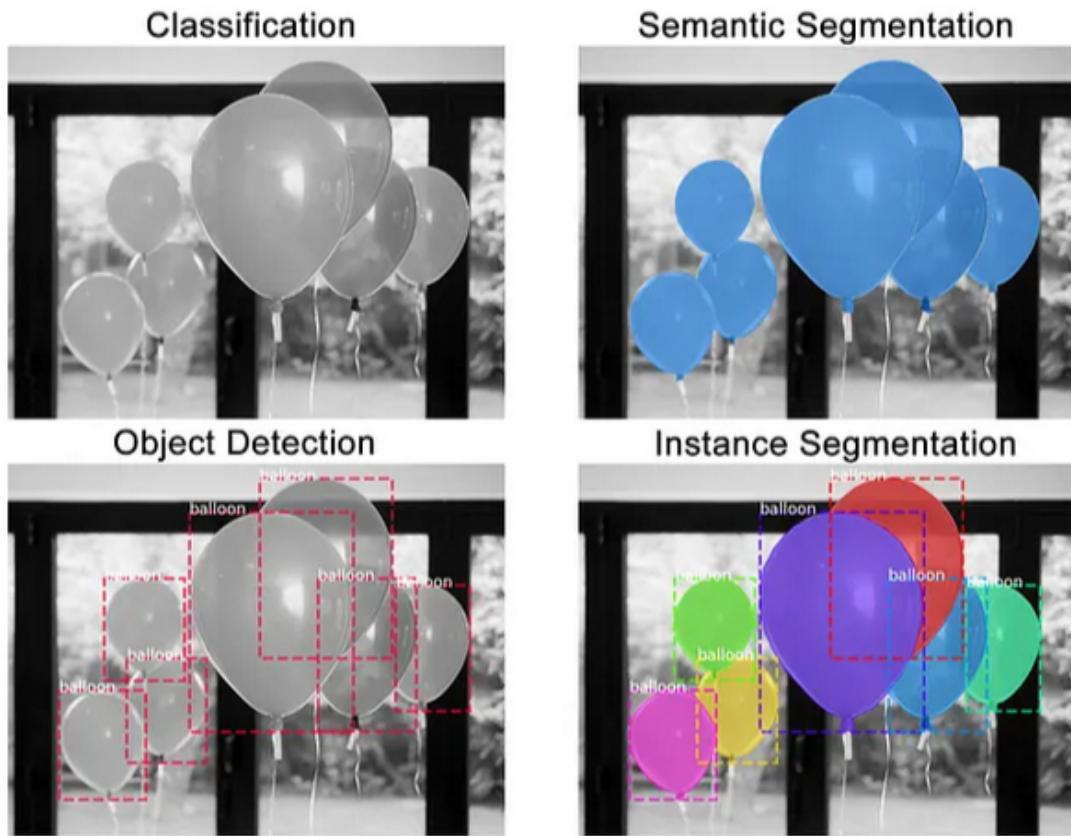
Os valores desse vetor são processados para que as características sejam aprendidas e associadas a probabilidades para cada classe de interesse (Kumar, 2025).

4.3.2.3 Segmentação de Instância

De acordo com Abdulla (2018), segmentação de instância é o processo de identificar o contorno de um objeto no nível dos *pixels*.

Primeiramente, ocorre a classificação, ou seja, percebe-se que objetos de determinada classe estão presentes na imagem. Logo após, ocorre a segmentação de semântica cujo objetivo é identificar todos os *pixels* que pertencem à classe de interesse. Depois, a detecção de objetos nota a presença de sete instâncias, mas ainda não consegue diferenciar os *pixels* que se sobreponem. Por fim, a segmentação de instância é capaz de discernir os contornos de cada objeto, ou seja, o problema de sobreposição foi solucionado. Esse procedimento está descrito na Figura 8.

Figura 8 – Etapas da Segmentação de Instância



Fonte: Abdulla (2018).

4.3.2.4 R-CNN

De acordo com Yenigün (2025), a CNN utiliza uma matriz que percorre uma imagem inteira até encontrar um objeto de interesse. Tendo em vista que esse processo é pouco eficiente e exige muitos recursos computacionais, a *Region-based Convolutional Neural Networks* (R-CNN) aparece como uma solução para otimizar esse busca a partir de determinações das regiões mais prováveis que o objeto irá aparecer.

Dessa forma, a R-CNN possui 4 principais etapas:

- **Proposta de Região:** um conjunto de propostas de regiões é gerado cada uma delas funciona como uma matriz que pode ou não conter objetos de interesse ;
- **Extração de Características:** cada região é proposta é redimensionada para um tamanho fixo porque um classificador de rede neural necessita de dimensões padronizadas;
- **Classificação:** verificar se há objetos de interesse dentro de cada região e classificá-los de acordo com suas classes caso existam;

- **Régressão de caixa delimitadora:** as matrizes que possuírem objetos de interesse dentro delas serão redimensionadas para que se enquadrem adequadamente aos limites dos objetos.

4.3.2.5 Mask R-CNN

O Mask R-CNN, por sua vez, é uma evolução do R-CNN porque possui duas características adicionais: uma rede treinável de proposta de região e segmentação de instância (Yenigün, 2025).

Dessa forma, a rede aprenderá quais são as regiões mais prováveis de conter objetos de interesse, enquanto a segmentação de instância para definir os contornos dos objetos ao nível do pixel, algo que facilita a compreensão de objetos que se sobrepõem (Yenigün, 2025).

4.3.2.6 *f1-score*

Após o desenvolvimento do modelo de aprendizado de máquina, torna-se vital averiguar sua eficiência. Para isso, há três importantes passos (KASHYAP, 2024):

- **Precisão:** Mede a acurácia das previsões positivas. Ou seja, cada região que o modelo classificar como contendo um objeto de interesse será analisada para verificar se há predição condiz com a verdade ou não;
- **Recall (Sensitividade):** Mede a capacidade do modelo de identificar todos os objetos de interesse definidos no pré-processamento;
- ***f1-score*:** Balanceia as duas métricas em um único valor.

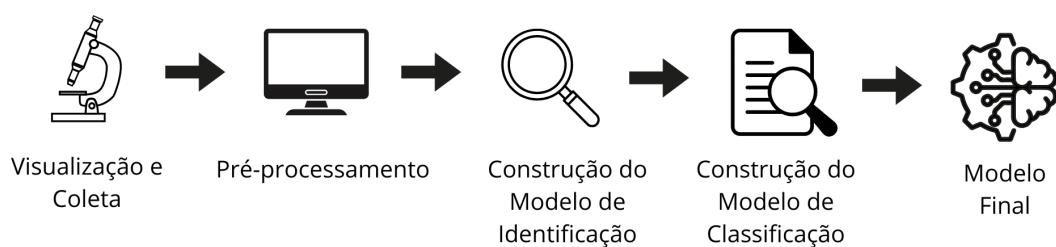
CAPÍTULO
5

METODOLOGIA

Neste capítulo, serão apresentadas todas as ferramentas e os procedimentos utilizados para a construção do modelo computacional de reconhecimento e classificação oócitos bovinos.

No modelo proposto, os oócitos são visualizados através de instrumentos de aproximação por conta de seu tamanho microscópico e as suas imagens são capturadas e armazenadas. Em seguida, as imagens captadas são pré-processadas para a retirada de ruídos e padronização de camadas; processadas em uma rede neural para reconhecer as regiões onde se encontram os oócitos; construção do modelo para classificá-los; e, finalmente, um modelo final que une essas duas funcionalidades. Esse processo pode ser visualizado na figura 9 e será descrito com mais detalhes nas seções seguintes.

Figura 9 – Esquema de funcionamento do modelo proposto



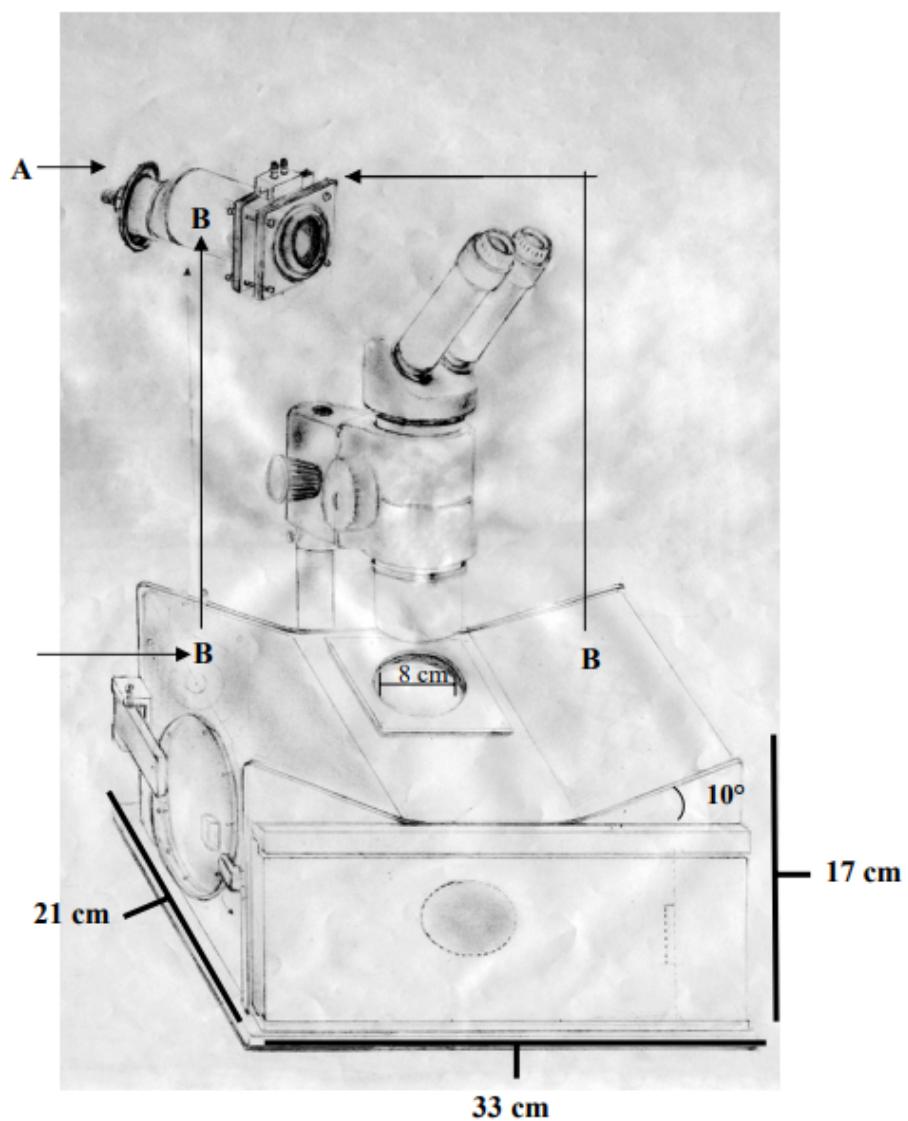
Fonte: Elaboração própria

5.1 Visualização e Coleta

Os oócitos utilizadas neste trabalho foram fornecidas pela empresa parceira BioInova (Bioinnova, 2024). Ela possui experiência com produções *in vitro* desde 2019 e, por conta disso, possui um vasto *dataset* de oócitos que será necessário para realizar o processo de aprendizado do sistema.

A visualização e captura das imagens são feitas utilizando um instrumento de aproximação microscópico. Ele é demonstrado na Figura 10

Figura 10 – Demonstração das componentes de um Microscópio Estereoscópico



Fonte: Gomes, Cora e Dusi (2005).

Ele é equipado com uma câmera, o que permite a coletada das imagens de maneira apropriada. Essa é composta por um conjunto óptico de três câmeras, disponíveis num celular

Samsung Galaxy s20 fe, que segundo o fabricante, é composta 12, 64 e 12 megapixels (MP) e aberturas de f1.8, f2.0 e f2.2, respectivamente (SAMSUNG, 2024).

5.2 Pré-processamento

Nessa seção, todas as etapas necessárias para preparar as condições mímicas necessárias para o desenvolvimento do modelo de identificação e classificação de oócitos serão abordadas.

5.2.1 OpenCV

O pré-processamento consiste em corrigir aspectos ruidosos nas imagens, tais como: sujeiras, áreas ofuscadas, dentre outros. Essas correções podem ser radiométricas ou geométricas (Chaki; Dey, 2019) e são realizadas através da biblioteca OpenCV e seus métodos.

Algumas imagens passaram por uma correção radiométrica (Chaki; Dey, 2019). Esse tipo de problema surge quando há um mal posicionamento do objeto capturado diante da luz ou falta de calibração adequada nos sensores. Isso faz com que alguns *pixels* da imagem não sejam constituídos, ou seja, será necessário reconstituir os de forma artificial através dos *pixels* mais próximos como referência (Chaki; Dey, 2019). Para tratar esse tipo de problema, os fundamentos utilizados são equalização de histograma e conversão de espaço de cores (OpenCV, 2025).

Houve também casos em que movimentos durante o processo de captura e/ou lentes distorcidas no dispositivo de captação ocasionam um posicionamento distorcido de determinados *pixels* da imagem (Chaki; Dey, 2019). Para resolver isso, foi necessário reposicionar os *pixels* que foram desvirtuados através das técnicas de transformações geométrica da OpenCV, que são: correção de distorção de lente, transformação de perspectiva, e transformação afim (OpenCV, 2025).

Para ajustes na iluminação, utilizou-se filtros de gradiente e equalizações de histograma, que servem para regularizar a intensidade de *pixels* (Krig, 2014).

Por fim, houve a necessidade de realizar ajustes relacionados ao foco de determinadas regiões da imagem captada. Por exemplo: se alguma região de interesse estiver desfocada ou se alguma região irrelevante estiver com muito foco, ambas essas situações precisarão ser readjustadas para que o processamento ocorra adequadamente (Krig, 2014).

5.2.2 LabelMe

O *LabelMe* foi utilizado para rotulação (Russell; Torralba; Murphy; Freeman, 2008). Um círculo foi atribuído em cada oóbito para fazer com que o aprendizado identificasse qual é o objeto de aprendizado. Logo depois, um JSON é gerado com todas as informações pertinentes da rotulação, no entanto, ainda foi necessário usar o comando *labelme2coco*, que traduz o JSON

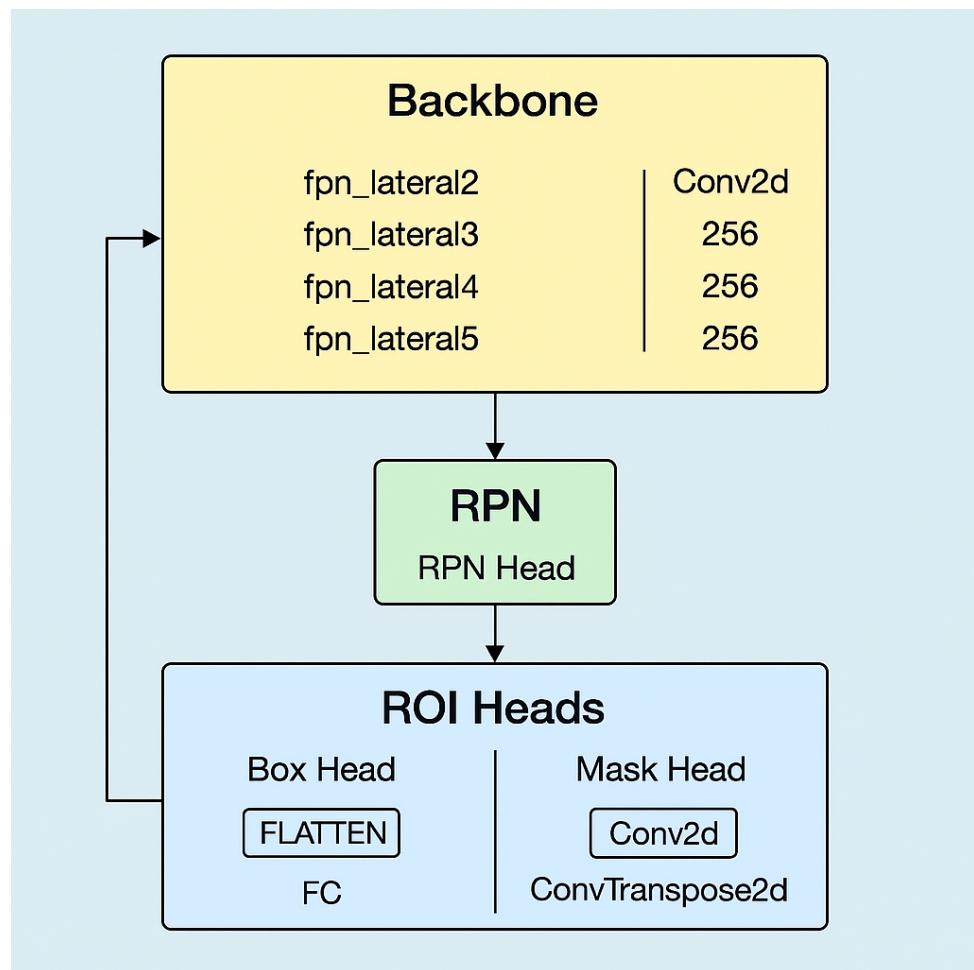
gerado para o formato esperado pelo Detectron2, a CNN utilizada na tarefa de identificação das células e que falaremos mais na seção 5.3.

5.3 Construção do Modelo de Identificação

Com o conjunto de imagens completamente pré-processado e rotulado, foi possível utilizar o Detectron2, que de acordo com (Lad, 2024), é um *framework* para diversos algoritmos de detecção de objetos e, entre eles, encontra-se o Mask R-CNN.

Logo após, um modelo de aprendizado de máquina para identificar oócitos bovinos foi gerado com o Detectron2. A arquitetura desse modelo está representada na Figura 11

Figura 11 – Arquitetura do Modelo de Aprendizado de Máquina para Identificação de Oócitos



Fonte: Elaboração Própria.

Percebe-se uma arquitetura baseada no *Mask R-CNN* utilizado pelo *Detectron2* e adaptada especificamente para detecção e segmentação de oócitos bovinos.

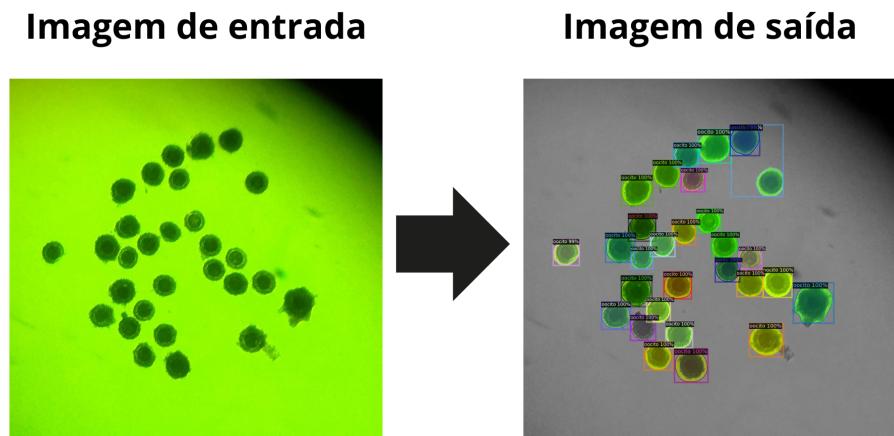
Essa rede é composta por três módulos principais:

- **Backbone (ResNet + FPN)**: as imagens dos oócitos foram processadas e suas características, extraídas. A *Feature Pyramid Network* (FPN) foi responsável por combinar informações em múltiplas escalas (fpnlateral2 – 5) com convoluções de 256 canais, algo que permitiu detectar oócitos de diferentes tamanhos e contrastes;
- **Region Proposal Network (RPN)**: aqui, regiões de interesse foram geradas através de probabilidade de cada região conter um objeto de interesse;
- **ROI Heads (Box Head e Mask Head)**: as propostas geradas na etapa anterior foram refinadas através de classificações e ajustes das caixas delimitadoras dos oócitos.

Logo após, houve a construção do modelo de identificação com os conjuntos de nome *oocitosTrain* e *oocitosVal* configurados para 1 classe (oóцитos) e com 1000 iterações.

Com o modelo para identificar oócitos treinado, tornou-se possível desenvolver a etapa de classificação. Pode-se averiguar o funcionamento de modelo de identificação na Figura 12

Figura 12 – Imagem de Entrada e Saída no Modelo de Identificação de Oócitos



Fonte: Elaboração Própria.

Para que esse modelo pudesse ser treinamento corretamente, foi utilizada uma proporção de 80% da base de dados para treinamento e 20% para validação, que é uma divisão padrão.

Após a finalização dessa etapa, tornou-se necessário avaliar a eficiência do modelo de identificação. Com esse intuito, o Detectron2 utiliza o método *COCOEvaluator*.

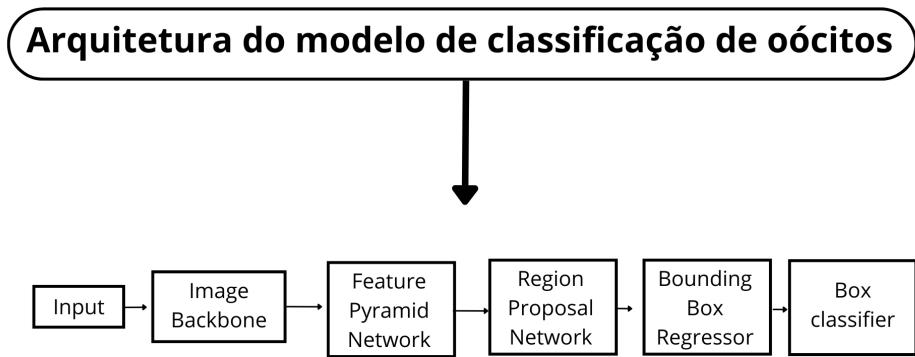
5.4 Construção do Modelo de Classificação

Para a construção do modelo de classificação foi necessário isolar através de recortes os oócitos contidos nas imagens. Isso foi feito para facilitar o aprendizado das características

específicas de cada tipo. Em seguida, a partir do VGG Image Annotator (VIA), 4 rótulos foram criados e todos os objetos de interesse isolados foram atribuídos à sua respectiva divisão. Nesse sentido, foi possível gerar um JSON no formato adequado para o Detectron2, algo que o VIA gera automaticamente. Entretanto, o algoritmo *cocosplit.py* ainda foi importante para dividir o JSON gerado em outros dois JSONS: um para treinamento e outro para teste a partir da proporção de 80% para 20%. Para garantir que cada classe fosse bem representada nas duas divisões, o *cocossplit.py* precisou ser modificado para realizar proporções inteligentes para cada classificação tendo em vista que seu código original não faz essa diferenciação.

Logo após, o modelo de aprendizado de máquina para classificação de oócitos bovinos foi construído. Sua arquitetura pode ser visualizada na Figura 13

Figura 13 – Arquitetura do Modelo de Aprendizado de Máquina para Classificação de Oócitos



Fonte: Elaboração Própria.

Como está representado na imagem, o oóbito isolado foi processado por uma série de camadas convolucionais na etapa de *backbone*. Esse processo extraí mapas de características da classificação.

Em seguida, as informações pertinentes são extraídas pelo *backbone* e processados pela *Feature Pyramid Network* (FPN). Essa etapa combina mapas de características de diferentes resoluções. Isso serve para melhorar a capacidade do modelo em detectar objetos de tamanhos variados.

Após a extração e combinação das características, os mapas gerados pela FPN são enviados à *Region Proposal Network* (RPN). Ela serve para gerar regiões com probabilidade de obter objetos de interesse, ou seja, oóbitos bovinos de determinada classificação. Além do mais, ela analisa as áreas de interesse e ajusta as coordenadas das (*bounding boxes*).

Antes das regiões propostas serem enviadas para as etapas de *bounding box regressor*

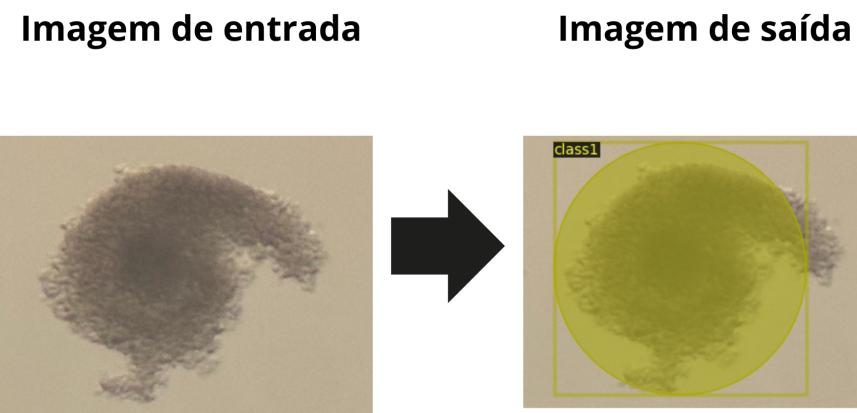
e *box classifier*, há um procedimento intermediário executado pelo *ROI Align*, que serve para regularizar as coordenadas que classificam os oócitos.

Logo após, a *bounding box regressor* realiza a atribuição de rótulos de classificação aos oócitos com determinadas características e refina ainda mais as coordenadas.

A partir da finalização desse processo, o modelo é capaz de produzir três tipos de resultados: a classificação de cada objeto de interesse, as caixas delimitadoras e as máscaras de segmentação, que contornam cada pixel da instância.

Com o modelo treinado, cada oóbito isolado passou a ser classificado. Pode-se visualizar essa etapa na Figura 14

Figura 14 – Demonstração da classificação de um oóbito bovino



Fonte: Elaboração Própria.

5.5 Modelo Final

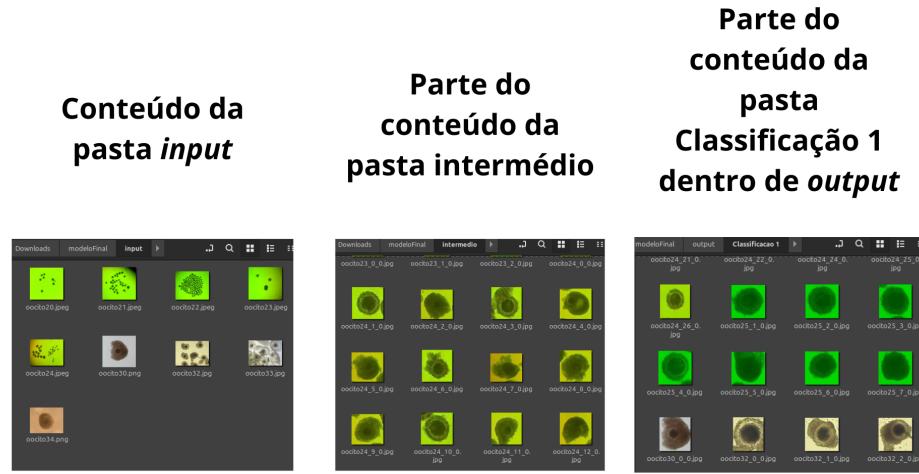
Após o treinamento dos dois modelos, tornou-se necessário desenvolver um algoritmo através da linguagem de programação *Python* que une as funcionalidades de identificação e classificação. Primeiramente, ele carrega as funções pertinentes dos dois modelos que estão nos seguintes arquivos: *model_final_identificacao.pth*, *config_identificacao.yaml*, *model_final_classificacao.pth* e *config_classificacao.yaml*.

A partir desses dados, foi desenvolvida uma função que analisa todas as imagens dentro um diretório nominado de *input* e utiliza o modelo de identificação para recortar os oócitos de imagens que possuem vários objetos de interesse. Logo após, as imagens de cada oóbito isolado são armazenadas em uma pasta chamada de intermédio. Com isso, outra função é chamada para classificar cada instância e, no fim do processo, diretórios referentes a cada classificação são criados dentro de uma pasta chamada de *output*. Em seguida, cada oóbito classificado é

armazenado de acordo com seu diretório.

Todo esse processo pode ser visualizado na Figura 15

Figura 15 – Demonstração do funcionamento do Modelo Final



Fonte: Elaboração Própria.

CAPÍTULO

6

RESULTADOS E DISCUSSÃO

Este tópico abordará os testes realizados nos modelos de aprendizado de máquina para identificação e classificação de oócitos bovinos.

6.1 Avaliação do Modelo de Identificação

A Tabela 1 demonstra as informações geradas pelo método de avaliação *COCOEvaluator*, o avaliador nativo do *Detectron2*.

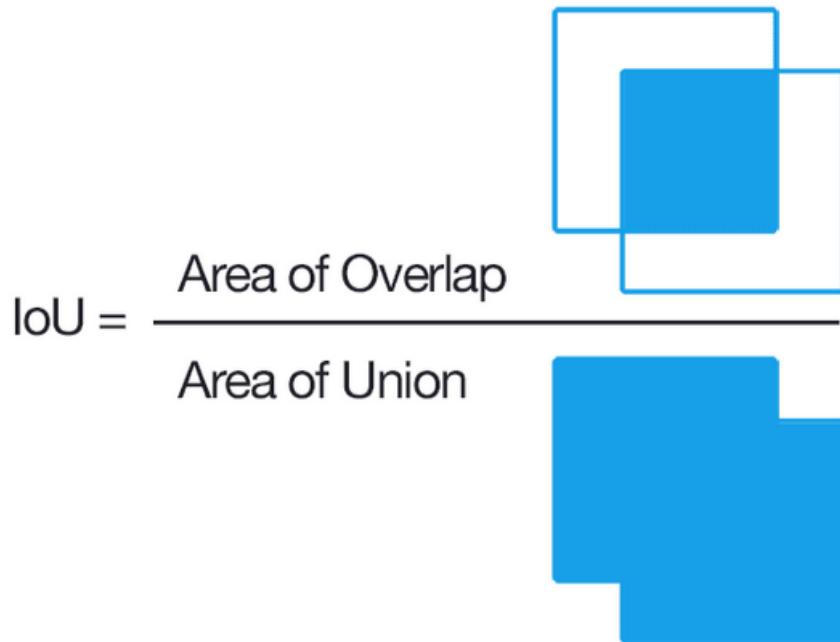
Tabela 1 – Métricas de Precisão Média (*Average Precision - AP*) do COCOEvaluator.

Métrica	Descrição	Caixas Delimitadoras	Segmentação
AP	AP Principal (IoU = 0.50 a 0.95, média)	80.974%	78.405%
AP ⁵⁰	AP com IoU ≥ 0.50 (Detecção "fácil")	97.783%	97.783%
AP ⁷⁵	AP com IoU ≥ 0.75 (Detecção "rigorosa")	96.033%	94.945%
AP _M	AP para objetos médios	81.938%	79.493%
AP _L	AP para objetos grandes	81.022%	78.314%

Fonte: Elaboração Própria.

Primeiramente, é preciso explicar os termos da tabela para compreende-la adequadamente. Dessa forma, *Intersection over Union* (IoU), que é a base para todas as medidas, serve para calcular a proporção entre a área de intersecção e a área da união entre o espaço previsto pelo modelo e a anotação real feita no pré-processamento. Portanto, o resultado ideal para essa proporção é 1.0 e quanto mais perto desse valor, melhor. Isso acontece porque um resultado de 1.0 indica que a área prevista está exatamente onde a anotação foi feita. Essa fórmula pode ser visualizada na Figura 16.

Figura 16 – Fórmula d IoU



Fonte: Rosebrock (2016).

Além do mais, as caixas delimitadoras definem o quanto bem o modelo consegue desenhar um retângulo ao redor do oócito, enquanto a segmentação verifica o quanto bem o modelo desenha uma máscara nos contornos exatos do oócito, ou seja, *pixel por pixel*. E, por fim, *Average Precision* (AP), é simplesmente o resultado médio realizado a partir de vários valores diferentes de IoU.

Dito isso, pode-se destrinchar os elementos da tabela:

- **AP (IoU=0.50 a 0.95, média):** Demonstra uma alta precisão média do modelo (>75%). A Detecção (80.974%) é ligeiramente maior que a Segmentação (78.405%). Isso significa que o modelo possui um pouco mais de facilidade em delimitar os retângulos do que os contornos dos oóцитos;
- **AP⁵⁰ (IoU≥0.50):** Há uma excelente Localização porque o modelo consegue identificar praticamente todos os objetos corretamente nos dois processos com um IoU moderado;
- **AP⁷⁵ (IoU≥0.75):** Mesmo com um limite consideravelmente mais rigoroso, a delimitação das caixas e dos contornos continua excelente e precisa;
- **AP_M (Objetos Médios):** Há uma precisão adequada para ambos os processos quando se trata de oóцитos de tamanho médio, no entanto, a delimitação da caixa é ligeiramente melhor;

- **AP_L (Objetos Grandes):** Apesar da precisão continuar boa, há uma pequena queda em comparação à oócitos médios;

Dessa forma, conclui-se que o modelo de identificação de oócitos bovinos está funcionando de maneira ideal.

6.2 Avaliação do Modelo de Classificação

A Tabela 2 demonstra as informações geradas pelo método de avaliação *COCOEvaluator*.

Tabela 2 – Métricas Globais de Precisão Média (*Average Precision - AP*) do Modelo de Classificação.

Métrica	Descrição	Detecção (bbox AP)	Segmentação (segm AP)
AP	AP Principal (IoU = 0.50 a 0.95)	88.128%	94.243%
AP ⁵⁰	AP com IoU \geq 0.50 (Detecção "Fácil")	100.000%	100.000%
AP ⁷⁵	AP com IoU \geq 0.75 (Detecção "Rigorosa")	100.000%	100.000%
AP _M	AP para objetos médios ($32^2 < \text{Area} < 96^2$)	88.762%	92.673%
AP _L	AP para objetos grandes (Area $> 96^2$)	88.960%	95.384%

Fonte: Elaboração Própria (Baseado nos resultados do COCOEvaluator).

Dessa forma, é possível desctrinchar os elementos dessa tabela:

- **AP (IoU=0.50 a 0.95, média):** O modelo atinge uma alta precisão no geral, mas possui mais facilidade com a segmentação;
- **AP⁵⁰ (IoU \geq 0.50):** Há uma localização perfeita, o que demonstra que houve uma memorização do conjunto de validação;
- **AP⁷⁵ (IoU \geq 0.75):** Apesar do limiar rigoroso, a localização perfeita permanece, algo que corrobora com a memorização do conjunto de validação;
- **AP_M (Objetos Médios):** Grande perfomance, especialmente na segmentação;
- **AP_L (Objetos Grandes):** O modelo lida melhor com objetos de interesse grandes;

A principal conclusão acerca desses dados é que o modelo possui um problema de *overfitting*, ou seja, o modelo não conseguiu generalizar seu aprendizado. Ele apenas memorizou as imagens que foram usadas para validá-lo. Por conta disso, uma série de estratégias foram utilizadas para lidar com essa óbice. Elas são:

- **Aumento do Dataset:** A primeira versão do modelo foi treinada com, em média, 16 oócitos para classificação. Esse valor foi escolhido porque a classificação 3 de oócitos, que foi a mais difícil de achar, continha apenas 16 oócitos em todo o *dataset*, enquanto as outras classificações possuíam consideravelmente mais imagens. Como o *overfitting* aconteceu, foi decidido alterar o balanceamento e utilizar o total de imagens que cada classe possuía para si;
- **Regularização L2:** A variável *Weight Decay* presente na configuração do treinamento foi diminuída com o intuito de aumentar a generalização do aprendizado, tendo em vista que um grande valor torna o treinamento muito sensível para variações entre diferentes imagens de oócitos;
- **Regularização de Lote:** Diminuir a variável *BATCH_SIZE_PER_IMAGE* força o treinamento a se adaptar para características mais amplas e combater a memorização das imagens de validação;
- **Otimização de Treinamento:** Reduzir a variável *BASE_LR* serve pra estabilizar o aprendizado e diminuir a intensidade que o modelo se corrige durante o processo de treinamento;
- **Parada/Controle:** Essa estratégia permite manipular a duração do treinamento. Ele verifica a partir de qual ponto o *overfitting* começa através do monitoramento da variável AP e interrompe o treinamento assim que o problema de memorização começa.

Portanto, nota-se como nenhuma das estratégias foi suficiente para resolver o *overfitting*. Conclui-se que, para resolver esse problema, é necessário que haja um *dataset* consideravelmente maior e que contenha imagens de alta qualidade dos oócitos isolados. A empresa parceira, BioInova, não conseguiu proporcionar essas imagens em tempo hábil e não havia como realizar essa coleta por conta própria tendo em vista que não há estrutura para isso na região.



CONCLUSÃO

A construção de um modelo adequado de aprendizado de máquina para identificação de oócitos bovinos e a análise dos problemas enfrentados para a construção do modelo de classificação são as principais contribuições desse trabalho. Isso representa um grande avanço para automatizar a etapa de classificação da PIV, algo de grande importância para a economia brasileira.

O desenvolvimento desses modelos foi dependente de tecnologias específicas, como visão computacional, processamento de imagens, aprendizado de máquina e programação. Elas permitiram pré-processar as imagens, rotulá-las, processá-las e, por fim, a união das funcionalidades através de um algoritmo. Os resultados obtidos representam uma contribuição significativa para futuros trabalhos que queiram unir a tecnologia da informação com a pecuária.

Para a contribuição na construção de modelos futuros, é importante considerar a importância de uma alta qualidade das imagens, algo que permite uma visão adequada dos fatores que compõem as métricas da qualidade de um oóbito.

Enfim, percebe-se como a união de diferentes áreas da computação, como visão computacional, processamento de imagens e aprendizado de máquina, possuem um grande potencial para desenvolver trabalhos em parceria com outras áreas do conhecimento humano. Demonstrou-se com sucesso o potencial e capacidade que a tecnologia da informação possui para automatizar e, como consequência, aumentar a eficiência de procedimentos que ainda precisam ser feitas de maneira manual.

REFERÊNCIAS

- ABDULLA, W. **Splash of Color: Instance Segmentation with Mask R-CNN and TensorFlow.** 2018. Disponível em: <https://engineering.matterport.com/splash-of-color-instance-segmentation-with-mask-r-cnn-and-tensorflow-7c761e238b46>. Citado nas páginas 28 e 29.
- AGRAWAL, T. **Hyperparameter Optimization in Machine Learning: Make your machine learning and deep learning models more efficient.** 1. ed.. ed. Berkeley, CA: Apress, 2020. XIX, 166 p. ISBN 978-1-4842-6578-9. Citado na página 25.
- ALVES, G. **Entendendo Redes Convolucionais (CNNs).** 2018. <https://medium.com/neuronio-br/entendendo-redes-convolucionais-cnns-d10359f21184>. Publicado em Neurônio BR, Acesso em: 23 ago. 2024. Citado na página 27.
- BALLARD, D. H.; BROWN, C. M. **Computer Vision.** Englewood Cliffs, NJ: Prentice Hall, 1982. 544 p. ISBN 0131653164. Citado na página 22.
- Bioinnova. **Bioinnova - Biotecnologia e Inovação.** 2024. <https://www.bioinnova.com.br/>. Acessado em: 15 ago. 2024. Citado na página 32.
- BOHANI, F. A.; SULIMAN, A.; SARIPUDDIN, M.; SYARMILA, S.; SALLEH, N. S. M. A comprehensive analysis of supervised learning techniques for electricity theft detection. **Journal of Electrical and Computer Engineering**, v. 2021, 07 2021. Citado na página 26.
- BRADSKI, G.; KAEHLER, A. **Learning.** [S.l.]: O'Reilly Media, Inc., 2008. ISBN 9780596554040. Citado na página 24.
- CAMARGO, L. S. de A.; PEREIRA, J. F. Genome-editing opportunities to enhance cattle productivity in the tropics. **CABI Agriculture and Bioscience**, v. 3, p. 8, 2022. Disponível em: <https://doi.org/10.1186/s43170-022-00075-w>. Citado na página 14.
- CARVALHO, A. C. P. d. L. F. Inteligência artificial: riscos, benefícios e uso responsável. **Estudos Avançados**, v. 35, n. 101, p. 21–36, apr 2021. Disponível em: <https://www.revistas.usp.br/eav/article/view/185020>. Citado na página 24.
- CHAKI, J.; DEY, N. **A Beginner's Guide to Image Preprocessing Techniques.** Boca Raton: Taylor & Francis, 2019. (Intelligent Signal Processing and Data Analysis). ISBN 9781138339316. Citado na página 33.
- DEISENROTH, M. P.; FAISAL, A. A.; ONG, C. S. **Mathematics for Machine Learning.** Cambridge, United Kingdom: Cambridge University Press, 2020. Citado na página 25.
- ELSEVIER. **Citoplasma inferior do óocito secundário.** 2024. Acesso em: 6 ago. 2024. Disponível em: <https://www.elsevier.com/resources/anatomy/secondary-oocyte/cytoplasm/cytoplasm-inferior/15466>. Citado na página 17.

- Embrapa. **PIVE. Conteúdo migrado na íntegra em: 22/12/2021. Embrapa Recursos Genéticos e Biotecnologia, Laboratório de Reprodução Animal.** 2021. Disponível em: <https://www.embrapa.br/agencia-de-informacao-tecnologica/tematicas/reproducao-animal/tecnologias/pive>. Acesso em: 04 ago. 2024. Citado na página 16.
- EMBRAPA. **Qualidade da carne bovina.** 2024. <https://www.embrapa.br/qualidade-da-carne/carne-bovina>. Acesso em: 23 ago. 2024. Citado na página 11.
- FAO. **FAO Statistical Database.** 2019. <http://www.fao.org/faostat>. Acesso em: 23 de agosto de 2024. Citado na página 14.
- FAOSTAT. **Carne Bovina - Produção (milhões de toneladas).** 2021. <http://fenix.fao.org/faostat/internal/en/#data>. Acesso em: 19 jul. 2022. Citado na página 14.
- FILHO, J. M. P. *et al.* **Produção de embriões bovinos in vivo e in vitro.** [S.l.: s.n.]: Universidade Federal de Viçosa-Departamento de Zootecnia, 2011. Citado nas páginas 17, 18, 19, 20 e 21.
- GOMES, A. C. M. M.; CORA, J.; DUSI, D. M. de A. **Câmara de Exaustão para utilização em microscópio estereoscópico.** 2005. Citado na página 32.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento digital de imagem.** [S.l.: s.n.]: Pearson, 2010. v. 10. ISBN 8576054019. Citado na página 23.
- GOUVEIA, F. F. **A produção in vitro de embriões bovinos.** [S.l.: s.n.], 2011. Citado na página 16.
- HAFEZ, E. S. E.; HAFEZ, B. **Reprodução animal.** [S.l.: s.n.], 2004. Citado na página 16.
- HAYKIN, S. **Redes Neurais: Princípios e Prática.** [S.l.: s.n.]: Bookman Editora, 2001. Citado nas páginas 25 e 26.
- HENRIQUE, G. **Processamento de Dados: como funciona e para quê serve?** [S. l.], 2024. Acesso em: 23 out. 2025. Disponível em: <https://faculdade.grancursosonline.com.br/blog/processamento-de-dados/>. Citado na página 11.
- IBM. **O que são redes neurais convolucionais?** 2024. Acesso em: 15 ago. 2024. Disponível em: <https://www.ibm.com/br-pt/topics/convolutional-neural-networks>. Citado na página 27.
- JAIN, R.; KASTURI, R.; SCHUNCK, B. G. **Machine Vision.** [S.l.: s.n.]: McGraw-Hill, Inc., 1995. ISBN 0-07-032018-7. Citado nas páginas 23 e 24.
- JÄHNE, B. **Digital Image Processing.** Berlin; Heidelberg; New York; Barcelona; Hong Kong; London; Milan; Paris; Tokyo: Springer, 2002. (Engineering Online Library). ISBN 3-540-67754-2. Citado na página 22.
- KASHYAP, P. **Understanding Precision, Recall, and F1 Score Metrics.** 2024. [S. l.], 2 dez. 2024. Disponível em: <https://medium.com/@piyushkashyap045/understanding-precision-recall-and-f1-score-metrics-ea219b908093>. Citado na página 30.
- KRIG, S. **Computer Vision Metrics: Survey, Taxonomy, and Analysis.** 1st. ed. Berkeley, CA: Apress, 2014. 508 p. ISBN 978-1-4302-5929-9. Citado na página 33.

- KUMAR, G. M. K. **Basic CNN Architecture: A Detailed Explanation of the 5 Layers in Convolutional Neural Networks.** [S. l.], 2025. Acesso em: 16 out. 2025. Disponível em: <https://www.upgrad.com/blog/basic-cnn-architecture/>. Citado na página 28.
- LAD, R. **An introduction to Detectron2.** 2024. Medium. Publicado em 9 de agosto de 2024. Disponível em: <https://medium.com/@ravina.lad01/an-introduction-to-detectron2-1b15e2f39b19>. Citado na página 34.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-Based Learning Applied to Document Recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 28.
- LEIBFRIED, L.; FIRST, N. L. Characterization of bovine follicular oocytes and their ability to mature in vitro. **Journal of Animal Science**, v. 48, p. 76–86, 1979. Citado na página 17.
- MELLO, R. R. C. *et al.* Produção in vitro (piv) de embriões em bovinos. **Rev. bras. reprod. anim.**, p. 6458–64, 2016. Citado nas páginas 11, 14 e 16.
- MOHANASUNDARAM, R.; MALHOTRA, A. S.; ARUN, R.; PERIASAMY, P. Chapter 8 - deep learning and semi-supervised and transfer learning algorithms for medical imaging. In: SANGAIAH, A. K. (ed.). **Deep Learning and Parallel Computing Environment for Bioengineering Systems**. Academic Press, 2019. p. 139–151. ISBN 978-0-12-816718-2. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780128167182000154>. Citado na página 26.
- MUELLER, M. L.; EENENNAAM, A. L. V. Synergistic power of genomic selection, assisted reproductive technologies, and gene editing to drive genetic improvement of cattle. **CABI Agriculture and Bioscience**, v. 3, p. 13, 2022. Disponível em: <https://doi.org/10.1186/s43170-022-00080-z>. Citado na página 12.
- NETTO, V. N.; GOMES, A. T. **Importância Econômica e Social.** 2021. <https://www.embrapa.br/qualidade-da-carne/carne-bovina>. Conteúdo migrado na íntegra em: 08 dez. 2021. Citado na página 11.
- OLIVEIRA, C. S.; SERAPIÃO, R. V.; QUINTÃO, C. C. R. **Biotécnicas da reprodução em bovinos: minicursos ministrados durante o 3º Simpósio ‘Biotécnicas da Reprodução em Bovinos’ no Laboratório de Reprodução Animal do Campo Experimental Santa Mônica.** Juiz de Fora, MG: Embrapa Gado de Leite, 2014. (Documentos, 175). Outubro/2014. ISBN ISSN 1516-7453. Citado na página 11.
- OOSTING, S.; UDO, H.; VIETS, T. Development of livestock production in the tropics: farm and farmers' perspectives. **Animal**, Cambridge University Press, v. 8, n. 8, p. 1238–1248, 2014. Citado na página 14.
- OPENCV. **OpenCV modules.** 2025. Citado nas páginas 24 e 33.
- ROSEBROCK, A. **Intersection over Union (IoU) for object detection.** 2016. Acesso em: 20 out. 2025. Disponível em: <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>. Citado na página 40.
- RUSSELL, B. C.; TORRALBA, A.; MURPHY, K. P.; FREEMAN, W. T. Labelme: a database and web-based tool for image annotation. **International journal of computer vision**, Springer, v. 77, n. 1, p. 157–173, 2008. Citado na página 33.

- SALES, J. N. d. S.; PUGLIESI, G.; CARVALHO, L. R.; SIMÕES, L. M. S.; LEMOS, L. A.; VICENTE, M. P.; SILVA, R. R. R.; BARUSELLI, P. S. Evolution over the last 40 years of the assisted reproduction technologies in cattle-the brazilian perspective i-timed artificial insemination. **Animal Reproduction**, SciELO Brasil, v. 21, n. 3, p. e20240034, 2024. Citado na página 14.
- SHIVAANI, M.; MADAN, P. Application of imaging and spectroscopy techniques for grading of bovine embryos - a review. **Frontiers in Veterinary Science**, v. 11, p. 1364570, 2024. Disponível em: <https://doi.org/10.3389/fvets.2024.1364570>. Citado na página 14.
- SICHMAN, J. S. Inteligência artificial e sociedade: avanços e riscos. **Estudos Avançados**, v. 35, n. 101, p. 37–50, apr 2021. Disponível em: <https://www.revistas.usp.br/eav/article/view/185024>. Citado na página 25.
- SMOLA, A.; VISHWANATHAN, S. **Introduction to Machine Learning**. The Pitt Building, Trumpington Street, Cambridge, United Kingdom: Cambridge University Press, 2010. ISBN 0521825830. Citado na página 25.
- SPMR. **Células do Cumulus**. 2024. Acesso em: 5 ago. 2024. Disponível em: <https://www.spmr.pt/index.php/18-repropedia/c/97-celulas-do-cumulus#:~:text=As%20c%C3%A9lulas%20do%20cumulus%20s%C3%A3o,prote%C3%A7%C3%A3o%20e%20matura%C3%A7%C3%A3o%20do%20o%C3%B3cito>. Citado na página 17.
- SUBASI, A. Chapter 3 - machine learning techniques. In: SUBASI, A. (ed.). **Practical Machine Learning for Data Analysis Using Python**. Academic Press, 2020. p. 91–202. ISBN 978-0-12-821379-7. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780128213797000035>. Citado na página 26.
- TEIXEIRA, J. C.; HESPAÑOL, A. N. A trajetória da pecuária bovina brasileira. **Caderno Prudentino de Geografia**, v. 2, n. 36, p. 26–38, 2014. Citado na página 14.
- VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: **Proceedings of the XXIX Conference on Graphics, Patterns and Images**. [S.l.: s.n.]: s.n., 2016. Citado na página 26.
- YENİGÜN, O. **R-CNN, Fast R-CNN, Faster R-CNN, and Mask R-CNN**. Medium, 2025. Disponível em: <https://towardsdev.com/r-cnn-fast-r-cnn-faster-r-cnn-and-mask-r-cnn-e7cd2e6f0a82>. Citado nas páginas 29 e 30.
- ZHANG, Q.; YANG, L. T.; CHEN, Z.; LI, P. A survey on deep learning for big data. **Information Fusion**, v. 42, p. 146–157, 2018. ISSN 1566-2535. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1566253517305328>. Citado nas páginas 26 e 27.

**INSTITUTO FEDERAL
DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA BAIANO**
Campus Guanambi

TRABALHO DE **CONCLUSÃO** DE CURSO

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS



INSTITUTO FEDERAL
Baiano
Campus Guanambi

"A missão do IF Baiano é oferecer educação profissional, científica e tecnológica pública, gratuita e de excelência em diferentes níveis e modalidades, voltada ao desenvolvimento humano, social, econômico, cultural, tecnológico e científico de todos e de todas, em diferentes regiões da Bahia e do Brasil."

LEGISLAÇÃO

- A Lei de Criação: **Lei nº 8.670/93**, publicada no D.O.U em dia 01 de julho de 1993.
- Lei de Autorização: **Lei nº 8.731/93**, publicada o D.O.U no dia 17 de novembro de

**Instituto Federal Baiano- Campus
Guanambi**

Zona Rural – Distrito de Ceraíma. Guanambi
BA. CEP: 46430-000.
Telefone: (77) 3493 - 2100
Site:

Instituto Federal Baiano- Campus Guanambi
<https://www.ifbaiano.edu.br/unidades/guanambi>



Coordenação do Curso

Telefone: (77) 3493-2100 (Ramal 357)

E-mail:

superior.ads@guanambi.ifbaiano.edu.br

Secretaria de Registros Acadêmicos

Alencastre Honório Moura

E-mail: alencastre.moura@ifbaiano.edu.br

Assessora Pedagógica

Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).