UNIVERSIDADE FEDERAL DO PAMPA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

DIEGO SILVA DOS SANTOS

CLASSIFICAÇÃO AUTOMÁTICA DE ESCORE DE PIGMENTAÇÃO OCULAR BOVINA USANDO DEEP LEARNING

Bagé
2026

DIEGO SILVA DOS SANTOS

CLASSIFICAÇÃO AUTOMÁTICA DE ESCORE DE PIGMENTAÇÃO OCULAR BOVINA USANDO DEEP LEARNING

Dissertação apresentada ao programa de Pós-graduação Stricto sensu em Computação Aplicada da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Computação Aplicada, área de concentração em Tecnologias para a produção agropecuária.

Orientador: Prof. Dr. Sandro da Silva Camargo

Coorientador: Prof. Dr. Fernando Flores Cardoso

Bagé
2026

Ficha catalográfica elaborada automaticamente com os dados fornecidos
pelo(a) autor(a) através do Módulo de Biblioteca do
Sistema GURI (Gestão Unificada de Recursos Institucionais) .

S237c Santos, Diego Silva dos
Classificação automática de escore de pigmentação ocular bovina usando deep learning / Diego Silva dos Santos.
98 p.

Dissertação(Mestrado)-- Universidade Federal do Pampa, MESTRADO EM COMPUTAÇÃO APLICADA, 2024.
"Orientação: Sandro da Silva Camargo".

1. Redes neurais artificiais. 2. Aprendizagem profunda. 3. Pecuária de precisão. 4. Análise fenotípica. 5. Processamento de imagens. I. Título.

DIEGO SILVA DOS SANTOS

Classificação automática de escore de pigmentação ocular bovina usando deep learning

Dissertação apresentada ao Programa de Pós-Graduação em Computação Aplicada da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Computação Aplicada.

Dissertação defendida e aprovada em: 21 de agosto de 2024.

Banca examinadora:

_____

Prof. Dr. Sandro da Silva Camargo
Orientador

(UNIPAMPA)

_____

Prof. Dr. Gabriel Soares Campos
(Purdue University, USA)

_____

Prof. Dr. Carlos Amaral Hölbig
(UPF)

Prof. Dr. João Ricardo Rebouças Dórea
(University of Wisconsin - Madison)

Assinado eletronicamente por **SANDRO DA SILVA CAMARGO**, **PROFESSOR DO MAGISTERIO SUPERIOR**, em 01/07/2026, às 10:00, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.

A autenticidade deste documento pode ser conferida no site https://sei.unipampa.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1520480** e o código CRC **3BA24776**.

Dedico este trabalho a meus pais e meu irmão, que me apoiaram durante todas as dificuldades logísticas enfrentadas não apenas no período do mestrado, mas por toda minha vida.

AGRADECIMENTOS

Agradeço primeiramente aos meus orientadores pela imensa paciência e apoio na redação desta dissertação.
Agradeço também aos meus amigos pelo apoio emocional dado em momentos de dificuldade.
Um agradecimento especial ao Professor Doutor Engenheiro Ivan Paulo Marques Alves, por todos os anos me auxiliando e me incentivando a buscar voos mais altos na minha carreira acadêmica.
E por último, mas não menos importante, agradeço a todas as pessoas que algum dia tenham participado positivamente da minha vida.

“It is not the critic who counts; not the man who points out how the strong man stumbles, or where the doer of deeds could have done them better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood; who strives valiantly; who errs, who comes short again and again, because there is no effort without error and shortcoming; but who does actually strive to do the deeds; who knows great enthusiasms, the great devotions; who spends himself in a worthy cause; who at the best knows in the end the triumph of high achievement, and who at the worst, if he fails, at least fails while daring greatly, so that his place shall never be with those cold and timid souls who neither know victory nor defeat.”

— Theodore Roosevelt

RESUMO

A crescente demanda mundial por proteína animal reforça a necessidade de aumentar a produtividade pecuária. Nesse contexto, o melhoramento genético contribui para a seleção de animais com características de interesse. Entre essas características, o escore de pigmentação ocular é utilizado na avaliação de bovinos da raça Hereford, mas sua atribuição é realizada visualmente por técnicos treinados, o que torna a avaliação sujeita à variabilidade entre avaliadores. Este estudo teve como objetivo automatizar a identificação e a classificação desse escore por meio de métodos de *deep learning*. Foram capturadas fotografias de 215 animais e avaliadas diferentes arquiteturas de redes neurais, incluindo as famílias *Residual Network* (ResNet) e *Visual Geometry Group* (VGG), o modelo *Faster Region-based Convolutional Neural Network* (Faster R-CNN) e a versão 8 da arquitetura *You Only Look Once* (YOLOv8), além de técnicas de aprendizado por transferência. O modelo YOLOv8 apresentou precisão média superior a 96% na detecção dos olhos, enquanto o desempenho na classificação do escore variou entre as classes avaliadas. Posteriormente, o modelo foi convertido para o formato TensorFlow Lite (TFLite) e integrado a um aplicativo Android, possibilitando sua execução em dispositivos móveis e sua utilização como ferramenta de apoio em condições de campo. Os resultados indicam a viabilidade técnica da aplicação de redes neurais profundas a essa tarefa e mostram que a automação pode contribuir para avaliações mais rápidas e padronizadas, auxiliando produtores e técnicos nos processos de seleção animal.

Palavras-chave: redes neurais artificiais; aprendizagem profunda; bovinos; pecuária de precisão; análise fenotípica; processamento de imagens.

ABSTRACT

The growing global demand for animal protein reinforces the need to increase livestock productivity. In this context, genetic improvement contributes to the selection of animals with desirable traits. Among these traits, the eye pigmentation score is used in the evaluation of Hereford cattle, but it is assigned visually by trained technicians, which makes the assessment subject to variability among evaluators. This study aimed to automate the identification and classification of this score using deep learning methods. Photographs of 215 animals were captured, and different neural network architectures were evaluated, including the Residual Network (ResNet) and Visual Geometry Group (VGG) families, the Faster Region-based Convolutional Neural Network (Faster R-CNN) model, and version 8 of the You Only Look Once architecture (YOLOv8), as well as transfer learning techniques. The YOLOv8 model achieved an average precision above 96% in eye detection, whereas performance in score classification varied among the evaluated classes. The model was subsequently converted to the TensorFlow Lite (TFLite) format and integrated into an Android application, enabling it to run on mobile devices and to be used as a support tool under field conditions. The results indicate the technical feasibility of applying deep neural networks to this task and show that automation may contribute to faster and more standardized assessments, assisting producers and technicians in animal selection processes.

Keywords: artificial neural networks; deep learning; cattle; precision livestock farming; phenotypic analysis; image processing.

LISTA DE FIGURAS

Figura 1 – Imagem capturada para a composição do conjunto de dados. 25
Figura 2 – Guia para análise de escore de pigmentação. 34
Figura 3 – Exemplo de escore de pigmentação. 35
Figura 4 – Modelo simplificado de neurônio animal. 39
Figura 5 – Modelo simplificado de neurônio artificial. 40
Figura 6 – Gráfico da função de limiar. 41
Figura 7 – Gráfico da função identidade ou linear. 42
Figura 8 – Gráfico da função tangente hiperbólica. 42
Figura 9 – Gráfico da função sigmoide. 43
Figura 10 – Gráfico da função sigmoide bipolar. 43
Figura 11 – Gráfico da função ReLU. 44
Figura 12 – Gráfico da função Leaky ReLU. 45
Figura 13 – Gráfico da função Softplus. 45
Figura 14 – Gráfico da função SoftMax. 46
Figura 15 – Diagrama de RNA totalmente conectada do tipo *feedforward*. 48
Figura 16 – Simulação de uma imagem após reconhecimento por rede neural. 50
Figura 17 – Diagrama da LeNet-5. 53
Figura 18 – Arquitetura da rede VGG-19 à esquerda, a rede plana com 34 camadas no centro e a ResNet com 34 camadas à direita. 55
Figura 19 – Linha temporal das arquiteturas YOLO. 57
Figura 20 – Visualização da arquitetura YOLOv8. 58
Figura 21 – Diagrama da revisão sistemática. 66
Figura 22 – Animal escondendo o rosto. 70
Figura 23 – Imagem capturada em ângulo frontal. 70
Figura 24 – Imagem borrada em razão do movimento do animal. 71
Figura 25 – Métricas de validação do YOLOv8 após 15 épocas. 74
Figura 26 – Conjunto de imagens com suas anotações de referência. 75
Figura 27 – Inferências do YOLOv8 para o conjunto de imagens da Figura 26. 76
Figura 28 – Matriz de confusão da classificação realizada pelo YOLOv8. 77

Figura 29 – Gráfico de precisão por confiança do YOLOv8. 78
Figura 30 – Gráfico de *recall* por confiança do YOLOv8. 78
Figura 31 – Gráfico de precisão por *recall* do YOLOv8. 79
Figura 32 – Gráfico de F1 por confiança do YOLOv8. 79
Figura 33 – Gráficos de métricas gerados pelo modelo YOLOv8. 80
Figura 34 – Predição correta do escore 4, com 68% de confiança. 81

LISTA DE QUADROS

Quadro 1 – Escore Hereford. 33
Quadro 2 – Matriz de confusão para duas classes. 59
Quadro 3 – Trabalhos correlatos selecionados. 65

LISTA DE SIGLAS

AHDB	Agriculture and Horticulture Development Board
ALICE	Acesso Livre à Informação Científica da Embrapa
AMD	Advanced Micro Devices
BRT	Brasília Time
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Networks
CPU	Central Processing Unit
DDR4	Double Data Rate 4
DEP	Diferença Esperada de Progênie
DNN	Deep Neural Networks
FN	Falso Negativo
FP	Falso Positivo
GPT-4	Generative Pre-trained Transformer 4
GPU	Graphics Processing Unit
IA	Inteligência Artificial
IEEE	Institute of Electrical and Electronics Engineers
IoU	Intersection over Union
IRNet	InceptionResNetV2
LOESS	Locally Estimated Scatterplot Smoothing
mAP	mean Average Precision
MNet	MobileNetV2
MNIST	Modified National Institute of Standards and Technology
MS COCO	Microsoft Common Objects in Context
PMGZ	Programa de Melhoramento Genético de Zebuínos
RAM	Random Access Memory
R-CNN	Region-based Convolutional Neural Network
ReLU	Rectified Linear Unit
ResNet	Residual Network
RNA	Redes Neurais Artificiais
RPN	Region Proposal Network
SORT	Simple Online and Realtime Tracking
STARE	Structured Analysis of the Retina
TFLite	TensorFlow Lite

VGG	Visual Geometry Group
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
YOLO	You Only Look Once

SUMÁRIO

1 INTRODUÇÃO **16**
1.1 Problema de Pesquisa **16**
1.2 Objetivos **17**
1.2.1 Objetivo Geral **17**
1.2.2 Objetivos Específicos **17**
1.3 Organização do Trabalho **18**
2 METODOLOGIA **19**
2.1 Caracterização **19**
2.2 Delimitação do Problema **19**
2.3 Protocolo da Revisão Sistemática de Literatura **20**
2.4 Delineamento Experimental **22**
2.5 Coleta de Dados **23**
2.6 Análise de Dados **24**
2.7 Riscos e Dificuldades **26**
2.8 Resultados Esperados **27**
3 REVISÃO DE LITERATURA **29**
3.1 Melhoramento Genético **29**
3.1.1 Programas de Avaliação Genética **30**
3.1.2 Diferença Esperada de Progênie **31**
3.1.3 Escores de Avaliação Subjetiva **32**
3.1.4 Escore de Pigmentação Ocular **32**
3.1.5 Usos de *deep learning* na pecuária bovina **35**
3.2 Tecnologia **36**
3.2.1 Inteligência Artificial **37**
3.2.2 Visão Computacional **37**
3.2.2.1 Segmentação *37*
3.2.3 Neurônios artificiais **38**
3.2.3.1 Funções de ativação *40*
3.2.3.2 Aprendizado de Máquina *47*
3.2.3.3 Redes Neurais Artificiais *47*
3.2.3.4 Aprendizado Supervisionado *49*
3.2.3.5 Aprendizado Profundo *51*

3.2.3.6 Redes Neurais Convolucionais *52*
3.2.3.7 Aprendizado por transferência *53*
3.2.3.8 VGG *54*
3.2.3.9 ResNet *54*
3.2.3.10 Faster R-CNN *56*
3.2.3.11 YOLOv8 ..... *56*
3.2.4 Avaliação de modelos de classificação **57**
3.2.4.1 Matriz de confusão *59*
3.2.4.2 Acurácia *60*
3.2.4.3 Precisão e Revocação (recall) *60*
3.2.4.4 F1-Score *60*
3.2.4.5 Perda Logarítmica *60*
3.3 Ferramentas **61**
3.3.1 Python **62**
3.3.2 TensorFlow **62**
3.3.3 PyTorch **62**
3.3.4 Albumentations **63**
3.3.5 Bibliotecas adicionais **63**
3.4 Trabalhos Correlatos **64**
4 ABORDAGEM PROPOSTA **69**
4.1 Preparação da base de dados **69**
4.2 Implementação **72**
4.3 Análise e comparação dos resultados **73**
5 CONCLUSÃO **82**
5.1 Trabalhos Futuros **83**
REFERÊNCIAS **84**

1 INTRODUÇÃO

A evolução da espécie humana, caracterizada por sua dieta onívora, foi marcada por transformações influenciadas por mudanças climáticas, que impulsionaram os ancestrais do *Homo sapiens sapiens* a aumentar o consumo de proteína animal. Essa adaptação é considerada um marco no desenvolvimento cerebral humano, conforme estudos nas áreas de nutrição e evolução (James *et al.*, 2019; O'Keefe *et al.*, 2022). Com a transição para um estilo de vida sedentário, há aproximadamente dez mil anos, a humanidade adotou a agricultura e desenvolveu a pecuária, o que gerou a necessidade de domesticar animais para suprir a demanda alimentar diante da diminuição dos recursos provenientes da caça. Assim, a pecuária consolidou-se como um componente essencial do desenvolvimento socioeconômico, embora também tenha levantado questões ecológicas significativas devido ao uso intensivo de recursos naturais.

No cenário contemporâneo, caracterizado pelo crescimento populacional e pelo aumento do acesso ao consumo de carne, a pressão sobre a produção animal intensificou-se, assim como as preocupações ambientais relacionadas ao desmatamento e à sustentabilidade da produção pecuária. Esses desafios evidenciam a necessidade de estratégias sustentáveis e eficientes, entre as quais o melhoramento genético desempenha um papel crucial, especialmente por meio da avaliação fenotípica, entendida como a análise das características físicas observáveis do animal (Abadias; Fonseca; Barbosa, 2020; Carvalho; Vieira, 2023; Malafaia; Biscola; Dias, 2020; Malafaia *et al.*, 2019; Silveira; Araújo, 2023; Souza, 2023).

Entre as características fenotípicas, o escore de pigmentação ocular constitui um critério subjetivo de seleção, cuja avaliação humana apresenta limitações. A automação dessa avaliação, por meio da aplicação de tecnologias de inteligência artificial (IA), particularmente do aprendizado profundo (*deep learning*) e das redes neurais convolucionais (CNNs, do inglês *Convolutional Neural Networks*), tem-se consolidado como uma alternativa viável, com potencial para proporcionar maior precisão, redução de custos e diminuição dos vieses associados à avaliação humana (Dong; Wang; Abbas, 2021; Freitas *et al.*, 2019; Lorensi; Camargo; Genro, 2022; Patel *et al.*, 2023; Rony *et al.*, 2021; Silva *et al.*, 2023).

1.1 Problema de Pesquisa

Este estudo investiga a viabilidade de empregar a IA, com foco em aprendizado profundo, para realizar a avaliação automática do escore de pigmentação ocular em bovinos e verificar se o desempenho do método se equipara ou supera o das avaliações realizadas por

especialistas humanos.

1.2 Objetivos

Esta seção apresenta os objetivos que orientaram o desenvolvimento desta pesquisa, os quais foram definidos com base no problema investigado e na proposta de avaliar a viabilidade da classificação automática do escore de pigmentação ocular bovina. O objetivo geral sintetiza a finalidade central do trabalho, enquanto os objetivos específicos organizam as etapas necessárias para alcançá-la.

1.2.1 Objetivo Geral

Desenvolver um método baseado em IA, com ênfase em aprendizado profundo e CNNs, para a classificação automática da pigmentação ocular em bovinos, a fim de fornecer subsídios ao processo de melhoramento genético.

1.2.2 Objetivos Específicos

Os objetivos específicos são:

a) construir um conjunto de dados de imagens rotuladas para o treinamento das CNNs;
b) implementar e avaliar diversas arquiteturas de CNN para a tarefa de classificação;
c) comparar o desempenho da arquitetura de CNN mais eficaz com as avaliações realizadas por especialistas humanos.

Quanto à metodologia, esta pesquisa é de natureza aplicada, com objetivos explicativos, abordagem quantitativa e procedimento experimental. Além de uma revisão sistemática da literatura, foram coletadas imagens de 215 bovinos da raça Hereford em Dom Pedrito, posteriormente catalogadas e anotadas com base nos escores atribuídos por especialistas. Os dados foram organizados em conjuntos de treinamento, validação e teste para a implementação e avaliação de diferentes arquiteturas de redes neurais, cujos resultados foram comparados às avaliações humanas por meio de métricas estatísticas.

1.3 Organização do Trabalho

Esta dissertação está organizada da seguinte forma: a Seção 2 descreve a metodologia adotada no estudo. A Seção 3 apresenta uma revisão da literatura pertinente, abordando os aspectos biológicos e tecnológicos relacionados ao melhoramento genético bovino e às aplicações de IA na pecuária. A Seção 4 detalha os resultados obtidos durante a pesquisa. Na Seção 5, são apresentadas as conclusões sobre os resultados obtidos, bem como uma breve discussão acerca de possíveis aplicações e melhorias a serem exploradas em trabalhos futuros.

2 METODOLOGIA

Esta seção aborda os métodos e os materiais empregados na condução da presente pesquisa. São detalhadas as etapas metodológicas, os materiais utilizados e os procedimentos experimentais adotados para alcançar os objetivos propostos.

2.1 Caracterização

De acordo com as diretrizes apresentadas por Prodanov e Freitas (2013), este estudo é de natureza aplicada, concentrando-se na solução de um problema real que envolve a criação de um método computacional para análise e classificação do escore de pigmentação ocular em bovinos. O escore, utilizado em programas de melhoramento genético, é atualmente atribuído por especialistas, demandando recursos e tempo significativos. Ao automatizar esse processo, espera-se reduzir custos e tornar a avaliação mais eficiente.

Quanto aos objetivos, esta pesquisa é explicativa, pois busca elucidar como abordagens de inteligência artificial baseadas em redes neurais artificiais podem ser aplicadas à classificação automatizada do escore de pigmentação ocular bovina. Essas abordagens foram aplicadas, otimizadas e avaliadas nesse contexto.

A pesquisa adota uma abordagem quantitativa, baseada em medidas numéricas para avaliar o desempenho dos métodos empregados. Para isso, é realizada uma análise estatística dos escores de pigmentação ocular, expressos em valores de 1 a 5, a fim de estimar a eficácia da técnica proposta na classificação da pigmentação ocular dos bovinos.

O procedimento adotado é de natureza experimental e envolve a implementação e a avaliação de diferentes abordagens baseadas em redes neurais artificiais. Os resultados obtidos são analisados com o objetivo de identificar o método mais adequado para automatizar a avaliação do escore de pigmentação ocular em bovinos, considerando sua precisão e confiabilidade.

2.2 Delimitação do Problema

A interação com a equipe do Laboratório de Bioinformática e Estatística Genômica (Labegen), da Embrapa Pecuária Sul, possibilitou compreender os problemas relacionados aos programas de melhoramento genético animal dos quais a equipe participa. Nesse contexto, surgiu a necessidade de desenvolver uma ferramenta capaz de avaliar o escore de pigmentação ocular de

maneira mais rápida e econômica, uma vez que essa tarefa é realizada manualmente por especialistas. Esse processo demanda tempo e recursos financeiros e está sujeito a variações decorrentes da subjetividade dos avaliadores. O escore de pigmentação ocular é um critério relevante em programas de melhoramento genético, como os conduzidos pela Embrapa, por contribuir para a seleção de animais destinados à reprodução e para o aprimoramento do rebanho bovino.

O problema central desta pesquisa, portanto, reside na necessidade de otimizar o processo de avaliação do escore de pigmentação ocular, buscando uma solução que não apenas acelere o processo, mas também aumente sua precisão e reduza os custos associados. A aplicação de uma abordagem automatizada, utilizando inteligência artificial, tem potencial para abordar essas questões ao fornecer uma ferramenta capaz de realizar as avaliações de maneira consistente e objetiva, com base em imagens digitais dos olhos dos animais.

2.3 Protocolo da Revisão Sistemática de Literatura

Durante a realização desta pesquisa, foi executada uma revisão sistemática para buscar trabalhos que realizassem algo semelhante ao proposto nesta dissertação: encontrar um ponto de interesse em uma imagem de bovinos, que, no caso deste trabalho, corresponde à área dos olhos, e, após sua localização, realizar a classificação dessa área. Preferencialmente, ambas as tarefas deveriam ser executadas por meio de redes neurais artificiais. Como nenhum trabalho com essas características foi encontrado, decidiu-se realizar uma nova revisão sistemática para buscar estudos que utilizassem redes neurais para detectar e/ou classificar imagens de forma mais ampla, com a expectativa de adaptar ao objeto desta pesquisa uma solução desenvolvida para outra tarefa.

Após a definição da questão de pesquisa, iniciou-se a busca por palavras-chave que fornecessem os resultados mais relevantes em repositórios especializados. Foram realizadas diversas tentativas com diferentes *strings* de busca. Em alguns casos, os repositórios retornavam um número excessivo de resultados com pouca ou nenhuma relação com o tema da pesquisa; em outros, não retornavam resultados. Entre as combinações testadas, a *(“neural network” AND “feature detection” AND image AND (eye OR ocular))* e sua versão em português, *(“rede neural” AND “detecção de característica” AND imagem AND (olho OR ocular))*, proporcionaram os resultados mais relevantes. Utilizou-se um filtro para descartar trabalhos publicados antes de 2021. Também se optou por limitar os resultados a trabalhos de livre acesso, para que outros pesquisadores possam consultar as mesmas fontes sem necessitar de acesso institucional ou pago.

Com o apoio da ferramenta online Parsifal[1], foram utilizados como meios de consulta os repositórios Engineering Village[2], Xplore[3], mantido pelo *Institute of Electrical and Electronics Engineers* (IEEE), ScienceDirect[4], Scopus[5] e Acesso Livre à Informação Científica da Embrapa (ALICE)[6].

Como critérios de inclusão, foram utilizados:

a) *publicações que tratem de detecção e/ou classificação em imagens*: esta dissertação busca criar uma ferramenta capaz de detectar os olhos dos animais e classificar a pigmentação em torno da região ocular. Considerando esse objetivo, os trabalhos incluídos na revisão deveriam tratar da detecção e/ou da classificação de imagens;
b) *trabalhos dos últimos três anos*: devido ao rápido avanço das técnicas de inteligência artificial e com o objetivo de evitar o uso de informações desatualizadas ou obsoletas, optou-se por limitar os resultados aos trabalhos publicados entre 2021 e janeiro de 2024;
c) *publicações que utilizem redes neurais para classificação de características fenotípicas em imagens de animais*: devido ao poder de generalização das redes neurais e ao fato de que esta dissertação busca testar especificamente métodos de redes neurais artificiais na resolução do problema de pesquisa, foram buscados trabalhos que utilizassem métodos semelhantes em situações próximas ao objetivo deste estudo;
d) *publicações de livre acesso*: para que esta dissertação possa servir de base a novas pesquisas e para ampliar o número de pesquisadores capazes de reproduzir e testar seus resultados, considerou-se preferível utilizar artigos disponíveis gratuitamente, mesmo para pessoas sem acesso institucional ou pago aos repositórios científicos consultados;
e) *publicações nos idiomas português ou inglês*: como o autor da pesquisa possui familiaridade com esses dois idiomas, foram utilizados apenas trabalhos escritos em português ou inglês.

Foram adotados os seguintes critérios de exclusão:

a) *pesquisas que não utilizem imagens no espectro visível*: este estudo visa desenvolver uma rede neural para integração em um aplicativo móvel, considerando que a maioria dos dispositivos móveis possui apenas câmeras capazes de captar o espectro visível da radiação eletromagnética. Portanto, foram excluídas pesquisas que envolvessem imagens multies-

[1]https://parsif.al/
[2]https://www.engineeringvillage.com/search/quick.url
[3]https://ieeexplore.ieee.org/
[4]https://www.sciencedirect.com
[5]https://www.scopus.com
[6]https://www.alice.cnptia.embrapa.br/

pectrais, de radar, sonar, lidar, raios X, ressonância magnética ou outras modalidades fora do espectro visível;

b) *trabalhos que não detalhem as ferramentas utilizadas e suas características de funcionamento*: esta revisão pretende encontrar estudos que tenham utilizado ferramentas passíveis de adaptação ao problema de pesquisa. Caso um trabalho não apresentasse detalhes como características da rede neural, linguagem de programação, ferramentas utilizadas e outras informações necessárias à adaptação da solução, tornava-se inviável utilizá-lo como base para a pesquisa;

c) *trabalhos que não utilizem redes neurais para a análise e a classificação de imagens*: como disposto na seção 2.8, o objetivo desta pesquisa é buscar um método eficiente, baseado em redes neurais artificiais, para classificar o escore de pigmentação ocular bovina e analisar sua utilidade em comparação com os métodos tradicionais;

d) *trabalhos duplicados*: trabalhos duplicados não acrescentam novas informações à pesquisa; portanto, apenas uma versão de cada estudo foi utilizada, e as duplicatas foram eliminadas.

2.4 Delineamento Experimental

O delineamento experimental da pesquisa visa avaliar a eficácia de redes neurais artificiais na classificação do escore de pigmentação ocular em bovinos. As variáveis essenciais incluem fotografias dos animais, brincos identificadores e avaliações de escore realizadas pelos especialistas e registradas no programa de melhoramento genético. Esses elementos são importantes para o desenvolvimento e a avaliação dos modelos de inteligência artificial propostos.

A amostra foi composta por animais de uma fazenda que participa ativamente de um programa de melhoramento genético. A inclusão de todos os animais do lote em estudo contribuiu para a coleta de um conjunto de dados representativo, necessário ao treinamento dos modelos de IA e à captura de uma diversidade significativa nas características de pigmentação ocular.

Para avaliar a robustez e a capacidade de generalização da solução criada, o conjunto de dados foi dividido em conjuntos distintos de treinamento, validação e teste. Essa separação contribui para reduzir o risco de sobreajuste e permite avaliar o desempenho dos modelos em imagens não utilizadas durante o treinamento.

O processo de treinamento de redes neurais no contexto da detecção e da classificação de imagens exige que o conjunto de imagens seja distribuído entre os conjuntos de treinamento, va-

lidação e teste. Como diferentes proporções podem ser utilizadas nessa divisão, cabe ao pesquisador definir a quantidade de imagens destinada a cada conjunto. Durante os testes preliminares, foram avaliadas diferentes proporções. A divisão de 80% das imagens para treinamento, 10% para validação e 10% para teste gerou resultados ligeiramente superiores aos das outras divisões testadas, em um particionamento semelhante ao utilizado por Kaliappan *et al.* (2023). Quantidades menores de imagens para o treinamento da rede geraram resultados de menor qualidade durante o processo de teste.

A etapa de análise estatística compara a acurácia dos modelos treinados com as avaliações realizadas pelos especialistas do programa de melhoramento genético. Essa comparação permite determinar a concordância entre os métodos de classificação automatizada e as avaliações humanas, fornecendo informações sobre a utilidade das redes neurais como ferramentas de apoio à avaliação fenotípica.

Por meio desse delineamento, a pesquisa busca demonstrar como as tecnologias de inteligência artificial podem complementar as abordagens tradicionais de avaliação no melhoramento genético bovino. Enfatiza-se a contribuição potencial das redes neurais para o aprimoramento das práticas existentes, promovendo avanços na precisão e na eficiência das avaliações fenotípicas.

Para agilizar o processo de revisão e correção do texto, utilizou-se a ferramenta Generative Pre-trained Transformer 4 (GPT-4). Alguns parágrafos foram escritos pelo autor e, em seguida, submetidos à ferramenta para a correção de possíveis erros de digitação, ortografia e gramática. Também foram solicitadas sugestões de melhoria na fluência e na formalidade do texto. Essas sugestões foram analisadas pelo autor e utilizadas quando consideradas úteis para aumentar a qualidade e a legibilidade da escrita.

2.5 Coleta de Dados

A captura das imagens ocorreu em 20 de abril de 2023, abrangendo 215 bovinos da raça Hereford. O local selecionado para a coleta situa-se em Dom Pedrito, Rio Grande do Sul, Brasil, com coordenadas aproximadas de latitude 31°05′12,8″ S e longitude 54°16′38,9″ O. Ao todo, foram capturadas 5.656 imagens. Como não havia condições de infraestrutura para capturar imagens dos dois lados da cabeça dos bovinos, optou-se por registrar apenas o lado esquerdo. A escolha deveu-se ao fato de que os brincos utilizados para a identificação dos animais se encontravam nesse lado.

A captura das imagens foi realizada de modo a minimizar o desconforto dos animais,

utilizando-se o menor tempo possível para não prejudicar as atividades normais da fazenda e buscando-se maximizar o número de imagens obtidas. Para essa tarefa, os animais foram colocados em um brete, em grupos de dez a quinze indivíduos, no qual permaneceram confinados por alguns segundos. Durante esse período, foram utilizados dois *smartphones* para capturar o maior número possível de imagens de cada animal, sob diferentes ângulos e distâncias. Essa abordagem, embora não fosse ideal, foi a única considerada viável diante das limitações e restrições existentes no momento da coleta.

Dois dispositivos da marca Samsung, modelos SM-A736B e SM-A505GT, foram utilizados para capturar as imagens. As fotografias foram obtidas sob ângulos e distâncias variáveis, sem controle específico de iluminação, a fim de favorecer a robustez do modelo em desenvolvimento. A primeira imagem foi obtida às 8:01:05, no horário de Brasília (BRT, do inglês *Brasília Time*), e a última, às 10:13:39 BRT. As resoluções das imagens capturadas foram de 4000×3000, 2268 × 4032 e 12000 × 9000 *pixels*.

Essa metodologia visa assegurar ampla variabilidade nas imagens, necessária para que o modelo treinado generalize a partir de diferentes condições. Essa abordagem também é importante para a aplicabilidade do modelo em diversas situações práticas, aumentando sua utilidade em diferentes dispositivos móveis. Contudo, é necessário que esses dispositivos possuam capacidade de processamento suficiente para executar o modelo proposto.

As imagens coletadas foram armazenadas na plataforma em nuvem Google Photos, facilitando o acesso e a gestão dos dados para análise posterior. Um exemplo dessas imagens é apresentado na Figura 1, que ilustra o tipo de fotografia obtida durante o processo de coleta de dados.

2.6 Análise de Dados

Após a coleta das imagens, procedeu-se à catalogação manual dos códigos de identificação de cada bovino fotografado. Esses códigos foram utilizados para correlacionar as imagens capturadas com os respectivos escores de pigmentação ocular previamente atribuídos por especialistas. Essa etapa foi necessária para a preparação dos dados utilizados pelos modelos de redes neurais artificiais. Foram excluídas da análise as imagens que não exibiam claramente o rosto do animal ou que apresentavam qualidade insuficiente para a identificação e a classificação do escore de pigmentação ocular, a fim de não afetar negativamente o processo de treinamento da rede.

Durante o processo de reconhecimento e catalogação manual, foram encontrados 230

Figura 1 – Imagem capturada para a composição do conjunto de dados.



Fonte: Autor (2024).

códigos, dos quais apenas 215 eram compatíveis com os dados armazenados no programa de melhoramento genético. Essa diferença pode ter ocorrido porque muitos brincos estavam sujos ou desbotados, o que dificultava a leitura e o reconhecimento do código gravado. Além disso, os recursos disponíveis para capturar um grande número de imagens tornaram inviável a identificação prévia de cada animal fotografado.

O passo seguinte consistiu na anotação tanto da região dos olhos dos bovinos quanto do respectivo escore de pigmentação. Para essa tarefa, foi utilizada a ferramenta Label Studio[7]. Para obter resultados preliminares antes do recebimento dos escores registrados no banco de dados do programa de melhoramento, as primeiras avaliações e anotações foram realizadas pelo autor da dissertação. Após o recebimento dos escores atribuídos pelos especialistas, as anotações foram atualizadas para refletir a classificação de cada animal.

Das 5.656 imagens capturadas, apenas 4.100 foram consideradas adequadas para uso no treinamento, em razão das dificuldades apresentadas na seção 2.7. Entre essas imagens, não havia nenhuma com escore 1; havia 118 com escore 2, 2.031 com escore 3, 1.554 com escore 4 e 397 com escore 5.

Foi identificada uma possível falha no método inicialmente utilizado para separar as imagens entre os conjuntos de treinamento, validação e teste. Inicialmente, aplicava-se *data*

[7]https://labelstud.io/

augmentation a todas as imagens para equilibrar o número de exemplos entre as classes e, posteriormente, realizava-se a divisão aleatória de 80% das imagens de cada classe para treinamento, 10% para validação e 10% para teste. Esse método foi modificado para aumentar o rigor metodológico dos resultados. No método final, os animais foram primeiramente distribuídos entre os conjuntos de treinamento, validação e teste, de modo que as imagens de um mesmo animal permanecessem em apenas um desses conjuntos. Somente após essa separação foi realizada a etapa de *data augmentation*, mantendo-se o equilíbrio entre as classes. As imagens geradas artificialmente foram utilizadas em conjunto com as imagens originais no processo de treinamento da rede.

A análise estatística dos dados coletados e preparados concentrou-se principalmente na acurácia dos modelos de rede neural desenvolvidos, comparando os escores de pigmentação ocular gerados automaticamente pelos modelos testados com os escores originais atribuídos pelos especialistas. Essa comparação permitiu avaliar a precisão e a confiabilidade das classificações automatizadas, fornecendo um indicativo da capacidade da tecnologia de reproduzir as avaliações humanas. Métricas adicionais, como precisão, *recall* e medida F1, também foram calculadas para fornecer uma avaliação mais abrangente do desempenho dos modelos.

2.7 Riscos e Dificuldades

Durante a execução desta pesquisa, surgiram alguns desafios que afetaram a coleta de dados e a análise dos resultados. Um desafio significativo foi a quantidade limitada de imagens disponíveis para o treinamento da rede neural. A qualidade e a diversidade do conjunto de dados são fundamentais para desenvolver um modelo de inteligência artificial capaz de generalizar para novas instâncias.

A amostra restrita representou uma ameaça à capacidade do modelo de reconhecer variações não observadas durante o treinamento. Nesse contexto, após o recebimento dos dados referentes aos escores, constatou-se que nenhum dos animais possuía escore de pigmentação 1, o que limitou a capacidade de treinamento dos modelos. A ausência de animais com escore 1 no treinamento pode dificultar o reconhecimento dessa classe caso o modelo seja aplicado a novas imagens que contenham bovinos com esse escore. Além disso, havia apenas 4 animais com escore 2 e 18 com escore 5. Em comparação, havia 113 animais com escore 3 e 80 animais com escore 4.

A obtenção de fotografias adequadas também se mostrou desafiadora devido à inquietude dos bovinos. Muitos animais apresentaram comportamentos evasivos na presença da câmera,

dificultando a captura de imagens claras e bem enquadradas. A necessidade de várias tentativas para obter imagens utilizáveis prolongou o processo de coleta de dados e evidenciou a imprevisibilidade do comportamento animal em condições de campo.

Além disso, a resistência dos animais em manter posturas favoráveis para as fotografias impôs limitações. Como não era possível intervir diretamente no posicionamento dos animais, em razão da preocupação com o bem-estar e a minimização do estresse, algumas fotografias não apresentaram clareza suficiente nem mesmo para a realização de uma avaliação humana. Essa condição reduziu a quantidade de dados úteis disponíveis e evidenciou a necessidade de uma abordagem cuidadosa e persistente na coleta de dados.

A impossibilidade de registrar os dois olhos dos animais também constituiu uma limitação da pesquisa, pois a avaliação realizada pelos especialistas considera ambos os olhos. Por essa razão, o presente estudo deve ser considerado uma prova de conceito, que pode servir de base para novas pesquisas destinadas a aperfeiçoar os métodos de classificação do escore de pigmentação ocular em bovinos.

Esses desafios ilustram as dificuldades inerentes à coleta de dados em ambientes naturais e com animais em condições de campo. Embora não tenham sido adotadas estratégias específicas para superar todas essas barreiras, a experiência adquirida reforça a importância de um planejamento detalhado e da consideração das variáveis envolvidas em pesquisas futuras semelhantes.

2.8 Resultados Esperados

Esperava-se, ao final da pesquisa, desenvolver um modelo de rede neural profunda capaz de classificar com precisão o escore de pigmentação ocular em bovinos. Esse modelo, treinado e validado com um conjunto de dados representativo, deveria demonstrar capacidade de generalização e permitir a aplicação prática em diferentes contextos do melhoramento genético bovino.

Além da criação de um modelo eficaz, previa-se que a comparação dos resultados gerados pela inteligência artificial com as avaliações realizadas por especialistas forneceria informações sobre a viabilidade de utilizar essas tecnologias como ferramentas de apoio ao processo de seleção genética. Esperava-se que o modelo não apenas alcançasse alta acurácia na classificação da pigmentação ocular, mas também oferecesse uma abordagem mais rápida e econômica, sem comprometer a qualidade ou a confiabilidade das avaliações.

Outro resultado esperado era a demonstração da capacidade das redes neurais artificiais

de complementar as habilidades dos especialistas em melhoramento genético. Ao automatizar parte do processo de avaliação, o modelo proposto poderia permitir que os especialistas se concentrassem em tarefas mais complexas e decisivas, otimizando o uso de recursos humanos e de tempo na pesquisa e na prática do melhoramento genético.

Por fim, esperava-se que os resultados da pesquisa contribuíssem para o avanço do conhecimento na interseção entre inteligência artificial e genética animal. A implementação de um modelo de classificação automatizada do escore de pigmentação ocular em bovinos poderia servir como estudo de caso e incentivar a exploração de abordagens semelhantes em outras áreas da genética e da produção animal.

3 REVISÃO DE LITERATURA

Esta seção aborda o campo do melhoramento genético bovino, enfatizando a aplicação do escore de pigmentação ocular como um critério de seleção. Adicionalmente, explora-se a contribuição das técnicas de inteligência artificial, especificamente redes neurais artificiais, para a solução do problema de pesquisa proposto. Uma revisão sistemática da literatura fundamenta a discussão sobre trabalhos correlatos, estabelecendo o contexto no qual a presente pesquisa se insere e destacando a relevância das metodologias adotadas para o avanço do conhecimento na área.

3.1 Melhoramento Genético

A introdução de bovinos no território que hoje constitui o Brasil, conforme descrito por Alencar (2004), não foi apenas um marco na história da pecuária nacional, mas também o início de um processo evolutivo que conduziu ao desenvolvimento das primeiras raças brasileiras. Esse processo de adaptação e miscigenação, enriquecido posteriormente pela introdução do zebu e sua notável adaptação às condições brasileiras, representou os primórdios do melhoramento genético voltado à otimização do rebanho para as condições locais.

Cardoso (2007) salienta que o melhoramento genético, fundamentado na seleção e no cruzamento consciente de animais, permite aos criadores aprimorar a produtividade de forma sustentável. A genética superior dos animais não apenas promove o uso mais eficiente dos recursos disponíveis, mas também proporciona ganhos constantes e cumulativos em produtividade. Embora os resultados não sejam imediatos, em longo prazo observa-se um incremento significativo na eficiência produtiva do rebanho.

Os conceitos básicos destacados por Cardoso (2007) no campo do melhoramento genético incluem:

a) *fenótipo e genótipo*: o fenótipo, conjunto das características físicas observáveis do animal, resulta da interação entre o genótipo — o conjunto de genes do animal — e o ambiente. Enquanto o fenótipo pode ser influenciado pelo meio, o genótipo é determinado na concepção e permanece constante, servindo como o principal foco do melhoramento genético devido à sua estabilidade diante das variações ambientais;
b) *dominância e recessividade*: a expressão das características fenotípicas é influenciada pela relação entre alelos dominantes e recessivos. Características resultantes de alelos

dominantes prevalecem sobre as dos recessivos, enquanto a ausência de dominância completa entre dois alelos pode resultar em fenótipos intermediários;

c) *homozigotos e heterozigotos*: indivíduos homozigotos possuem dois alelos idênticos para determinada característica, enquanto heterozigotos apresentam alelos distintos. Em cruzamentos monohíbridos entre indivíduos heterozigotos para uma característica mendeliana simples, pode ocorrer uma proporção fenotípica de 3:1 entre características dominantes e recessivas na prole.

Conforme evidenciado por Euclides Filho (1999), a seleção e o cruzamento constituem ferramentas essenciais no processo de melhoramento genético. A seleção direciona a escolha dos indivíduos que contribuirão para a próxima geração, enquanto o cruzamento busca combinar características genéticas desejáveis dos progenitores na prole.

3.1.1 Programas de Avaliação Genética

Programas de avaliação genética têm sido fundamentais na criação de bovinos, com foco no aprimoramento de características vitais como a velocidade de ganho de peso, qualidade da carne e eficiência alimentar. Estas iniciativas visam incrementar a produtividade dos rebanhos por meio de intervenções genéticas precisas.

A genética, conforme discutido por Pereira (2004), desempenha um papel crucial ao definir o potencial produtivo dos animais, ao lado de fatores ambientais. Esse entendimento embasa os esforços de melhoramento e destaca a importância da seleção baseada em critérios genéticos rigorosos.

No Brasil, destacam-se programas como o Geneplus-Embrapa[1], o Programa de Melhoramento Genético de Zebuínos (PMGZ), da Associação Brasileira dos Criadores de Zebu[2], e o PampaPlus[3], conduzido pela Embrapa Pecuária Sul. Esses programas ilustram a diversidade de abordagens adotadas no país para otimizar a genética do rebanho nacional.

O Geneplus-Embrapa, especificamente, iniciou em 1996 como uma colaboração entre a Embrapa Gado de Corte e a Geneplus Consultoria Agropecuária Ltda. O programa tem como propósito auxiliar criadores de gado de corte na seleção dos animais, na elaboração de planos de acasalamento e nas atividades de comercialização, utilizando as avaliações genéticas de reprodutores, matrizes e seus produtos (Geneplus, [2024?]).

[1] https://geneplus.com.br/
[2] https://www.abcz.org.br/pmgz/apresentacao
[3] http://www.pampaplusnet.com.br/

Esses programas enfatizam a mensuração tanto de características genéticas quanto fenotípicas para a avaliação dos animais. Os resultados dessas avaliações, expressos em valores ou escores, fundamentam o cálculo da Diferença Esperada de Progênie (DEP), detalhada na seção 3.1.2.

3.1.2 Diferença Esperada de Progênie

A Diferença Esperada de Progênie (DEP), conforme descrita por Cardoso (2009), é um indicador crucial para avaliar o mérito genético de bovinos. Utilizando um modelo matemático que leva em consideração dados do animal, de seus parentes e de sua progênie, é possível estimar a média esperada para os descendentes de um reprodutor. Essa média serve como referência, com a DEP fornecendo valores que indicam a superioridade ou inferioridade genética do animal em relação a características específicas.

Valores negativos de DEP podem indicar que a progênie do animal tende a apresentar desempenho abaixo da média para as características avaliadas, enquanto valores positivos sugerem uma tendência de desempenho superior. Embora, em algumas circunstâncias específicas, uma DEP negativa possa ser desejável, esses casos estão fora do escopo deste trabalho e, portanto, não serão discutidos. Assim, a DEP é essencial para informar decisões de criação, orientando os criadores na seleção de reprodutores que possam melhorar geneticamente o rebanho.

O conceito de criação seletiva refere-se ao processo de escolher intencionalmente animais para reprodução com base em características genéticas específicas desejadas, visando aprimorar qualidades particulares no rebanho. Essa abordagem considera tanto o potencial genético quanto os resultados observados na progênie, promovendo ganhos genéticos transmissíveis às gerações futuras.

Avanços no campo da genômica têm aumentado a acurácia das avaliações genéticas, aumentando a confiabilidade da DEP. A integração de informações genômicas ao cálculo da DEP aprofunda o entendimento sobre o perfil genético dos animais, contribuindo para a seleção de reprodutores (Silva *et al.*, 2008; Zampar, 2007).

Portanto, a DEP constitui uma métrica dinâmica e indispensável na criação seletiva de bovinos, permitindo a integração de variadas informações genéticas. Com o suporte de tecnologias avançadas, a DEP viabiliza aos criadores o aprimoramento contínuo das características genéticas desejadas em seus rebanhos.

3.1.3 Escores de Avaliação Subjetiva

Avaliações subjetivas, que incluem a análise visual de características como conformação corporal, qualidade da pelagem e expressão geral, desempenham um papel fundamental na composição da Diferença Esperada de Progênie (DEP). Embora a DEP seja informada tanto por critérios objetivos quanto subjetivos, os escores de avaliação subjetiva fornecem informações relevantes que complementam as medidas quantitativas (Faria *et al.*, 2007).

A definição de escores de avaliação subjetiva exige critérios claros e padronizados para assegurar uma contribuição objetiva e confiável ao cálculo da DEP. A precisão e a consistência na aplicação desses escores são vitais para minimizar a variabilidade e garantir que as avaliações visuais sejam um componente eficaz na seleção genética (Oliveira, 2012).

Além disso, o treinamento dos avaliadores é fundamental para a eficácia dos escores de avaliação subjetiva. Avaliadores bem treinados são capazes de aplicar os critérios de forma consistente, o que é essencial para manter a objetividade e a confiabilidade das avaliações visuais (Teixeira *et al.*, 2015).

A integração de escores de avaliação subjetiva na DEP reflete a compreensão de que a seleção genética deve ser informada por uma gama completa de características, abrangendo não apenas dados quantitativos, mas também qualidades fenotípicas que são mais bem avaliadas visualmente. Essa abordagem holística permite uma seleção mais informada e direcionada, contribuindo para o aprimoramento contínuo das qualidades genéticas e fenotípicas do rebanho.

Em resumo, os escores de avaliação subjetiva são componentes indispensáveis nos programas de melhoramento genético de bovinos. Eles permitem uma avaliação mais completa do potencial genético dos animais, combinando dados objetivos com avaliações visuais para formar uma DEP robusta e informativa. A padronização dos escores e o treinamento rigoroso dos avaliadores contribuem para a melhor aplicação dessa metodologia, o que pode favorecer avanços na qualidade e produtividade dos rebanhos.

3.1.4 Escore de Pigmentação Ocular

O escore de pigmentação ocular é um valor qualitativo atribuído com base na pigmentação da pelagem ao redor dos olhos de cada animal avaliado. Segundo Curso Internacional de Melhoramento de Bovinos de Corte do PampaPlus (2016), essa pigmentação é desejável em animais da raça Hereford, pois, quanto mais escura for a região próxima aos olhos, maior será a proteção contra os raios solares que podem causar desconforto e doenças oculares.

De acordo com Davis *et al.* (2015) e Pausch *et al.* (2012), entre bovinos de raças com a face branca, aqueles que apresentam pigmentação escura ao redor dos olhos possuem menor incidência de carcinoma espinocelular ocular, comumente chamado de “tumor ocular” ou “olho de câncer”, e maior resistência à ceratoconjuntivite infecciosa (Jara *et al.*, 2022).

A avaliação do escore é realizada de forma específica para cada raça, variando de 1, para animais com baixa qualidade de pigmentação, a 5, para animais com alta qualidade de pigmentação. No caso da raça Hereford, Curso Internacional de Melhoramento de Bovinos de Corte do PampaPlus (2016) fornece alguns parâmetros para a avaliação:

Quadro 1 – Escore Hereford.

Escore	Descrição	Na raça Hereford
1	Ausência	Sem pigmentação
2	Traços	Um olho parcialmente pigmentado
3	Razoável	Um olho totalmente pigmentado
4	Boa	Dois olhos parcialmente pigmentados
5	Excelente	Dois olhos totalmente pigmentados

Fonte: Adaptado de Curso Internacional de Melhoramento de Bovinos de Corte do PampaPlus (2016).

Como mostra o Quadro 1, o escore é atribuído com base na pigmentação de ambos os olhos de cada animal. Entretanto, devido à inviabilidade logística de obter fotografias dos dois olhos, conforme explicado na seção 2.5, neste trabalho foi possível analisar apenas o lado esquerdo e considerar que o outro olho apresentava a mesma pigmentação. Caso o sistema proposto nesta dissertação funcione adequadamente para um olho, estudos futuros poderão testar uma nova abordagem na qual ambos os olhos sejam fotografados e o escore seja atribuído de acordo com o quadro.

Embora seja desafiador encontrar materiais que tratem especificamente de o escore de pigmentação ocular, algumas técnicas digitais já foram propostas para esta classificação, como no estudo de Davis *et al.* (2015). Neste estudo, a quantificação da pigmentação ocular em bovinos foi conduzida por meio da captura de imagens digitais dos olhos dos animais, seguida por uma análise detalhada de *pixels*. As imagens foram obtidas utilizando câmeras digitais específicas e, após a captura, foram processadas para isolar visualmente as regiões das pálpebras e da esclera corneana.

A edição das imagens envolveu a conversão delas para escala de cinza, facilitando a identificação e contagem dos *pixels* com base em sua intensidade, que variava de completamente preto (0) a completamente branco (255). Utilizou-se o software ImageJ para essa análise, permitindo a quantificação precisa da área pigmentada em relação à área total observada nas

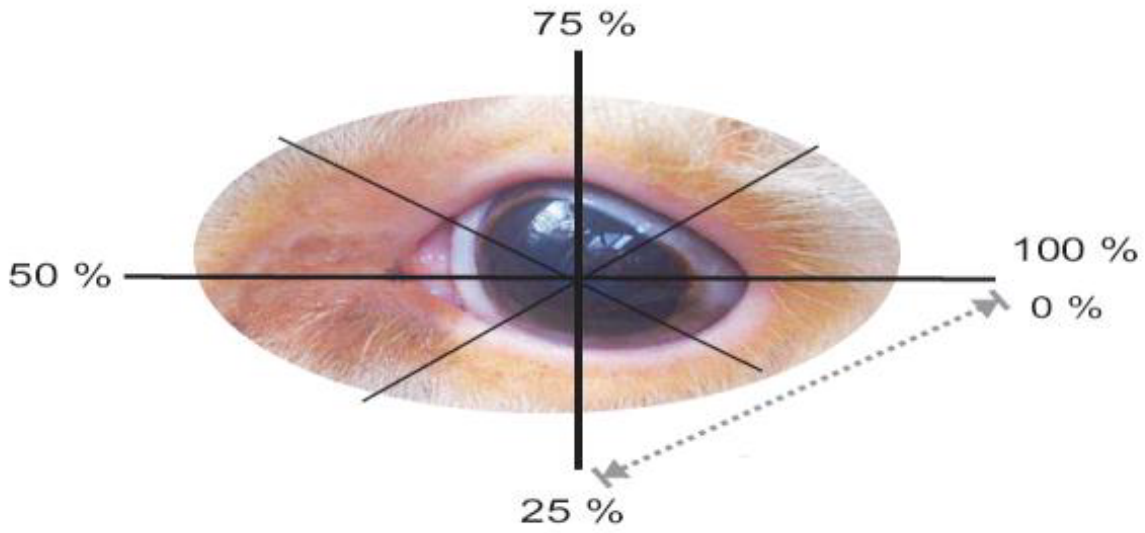
regiões de interesse.

Para diferenciar entre as áreas pigmentadas e não pigmentadas nas imagens, aplicou-se a regressão *Locally Estimated Scatterplot Smoothing* (LOESS) aos dados dos *pixels*. Essa técnica estatística possibilitou a identificação de dois picos principais nas distribuições de intensidade dos *pixels*, correspondendo às áreas pigmentadas e não pigmentadas, e um vale (*trough*) entre esses picos para delimitar as categorias.

Os resultados demonstraram que a proporção de pigmentação variou significativamente entre diferentes grupos raciais, com cruzamentos de Hereford e raças de pigmentação escura apresentando maiores níveis de pigmentação ocular. Além disso, uma maior pigmentação estava associada a uma redução na incidência de lesões oculares, sugerindo um efeito protetor. Esse método proporcionou uma abordagem objetiva e quantitativa para a avaliação da pigmentação ocular, superando as limitações das avaliações visuais subjetivas e proporcionando uma base sólida para futuras pesquisas sobre a relação entre pigmentação ocular e saúde em bovinos.

Apesar do exposto, durante a revisão sistemática realizada neste trabalho, não foi encontrado nenhum estudo que utilizasse métodos de *deep learning* para classificar o escore de pigmentação ocular bovina. Esse escore é um dos parâmetros utilizados no programa PampaPlus para o melhoramento genético bovino, conforme mostram Teixeira *et al.* (2015) e Vidal (2021). As Figuras 2 e 3 apresentam, respectivamente, o guia utilizado para relacionar a proporção de pigmentação ao escore e exemplos visuais dos diferentes níveis de pigmentação ocular.

Figura 2 – Guia para análise de escore de pigmentação.



Fonte: Sociedad Criadores de Hereford del Uruguay (2009), *apud* Curso Internacional de Melhoramento de Bovinos de Corte do PampaPlus (2016).

Figura 3 – Exemplo de escore de pigmentação.



Fonte: Sociedad Criadores de Hereford del Uruguay (2009), *apud* Curso Internacional de Melhoramento de Bovinos de Corte do PampaPlus (2016).

3.1.5 Usos de *deep learning* na pecuária bovina

Considera-se relevante citar alguns usos de *deep learning* como ferramenta de apoio para a resolução de problemas na pecuária de gado bovino. Abaixo são resumidos alguns artigos nos quais pesquisadores utilizaram redes neurais na classificação e na detecção de imagens de bovinos para aplicações em problemas diversos.

Li, Erickson e Xiong (2022) compararam 59 modelos de redes neurais em 4.923 imagens de 268 animais, com o objetivo de reconhecer os indivíduos por meio dos padrões encontrados em seus focinhos, que são únicos e comparáveis às impressões digitais humanas. A melhor acurácia obtida pelos pesquisadores foi de 98,7%, e eles concluíram que, com *data augmentation* e melhorias na função de perda de entropia cruzada, seria possível aumentar a acurácia com um número menor de imagens, facilitando o processo de identificação e o gerenciamento preciso da produção.

Porter, Wieland e Basran (2021) utilizaram a arquitetura GoogLeNet com uma abordagem de *transfer learning* para detectar mastite em vacas leiteiras. Com um total de 398 imagens dos úberes de vacas de duas fazendas diferentes, os pesquisadores alcançaram bons resultados na detecção de hiperqueratose no tecido mamário utilizando um escore de quatro níveis: normal, liso, áspero e muito áspero. Os pesquisadores concluíram que o uso de redes neurais pode ser útil como ferramenta de detecção da condição, que normalmente é feita por especialistas.

No estudo de Barney *et al.* (2023), os pesquisadores desenvolveram um sistema para detectar claudicação em vários bovinos em tempo real, utilizando uma versão adaptada do modelo *Mask Region-based Convolutional Neural Network* (Mask R-CNN) e o algoritmo *Simple Online and Realtime Tracking* (SORT) para analisar a postura e o deslocamento dos animais. A precisão do sistema foi avaliada por especialistas do *Agriculture and Horticulture Development Board* (AHDB) e validada por meio de validação cruzada e eliminação recursiva de características. O sistema alcançou 100% na detecção de claudicação e 94% na classificação de sua gravidade, com coeficiente kappa de Cohen de 0,8782, precisão de 0,8650 e *recall* de 0,9209.

Já Bai *et al.* (2023) desenvolveram uma versão modificada do modelo *Residual Network* com 50 camadas (ResNet50) para analisar imagens de células somáticas do leite. Esse modelo integra técnicas que permitem uma identificação detalhada de células, como macrófagos, linfócitos, células epiteliais e neutrófilos. A pesquisa utilizou um conjunto de dados preparado por meio de pré-processamento para aumentar a variedade de amostras e aplicou o PolyLoss para lidar com amostras desbalanceadas e complexas, enquanto adotou estratégias para prevenir o sobreajuste e melhorar a capacidade de generalização do modelo. Os testes indicaram que o modelo aprimorado alcançou resultados superiores aos de modelos anteriores, superando modelos anteriores em precisão e eficácia na classificação de células em amostras de leite.

Em Nye, Zingaretti e Pérez-Enciso (2020), foi apresentado o método DeepAPS, uma abordagem que combina dois algoritmos para gerar máscaras precisas em imagens de bovinos, facilitando a classificação automática e acurada de características fenotípicas. A metodologia demonstrou alta precisão quando comparada às classificações manuais da proporção de cor da pelagem, alcançando coeficiente R^2 ajustado de 0,926, e possibilitou a extração automática de informações referentes a outras 14 características morfológicas, evidenciando herdabilidades significativas para esses traços. Essa abordagem contribui para aumentar a eficiência e a precisão da avaliação fenotípica em pesquisas com bovinos.

Esses estudos ilustram a aplicabilidade e eficácia do *deep learning* na resolução de desafios complexos na pecuária bovina, desde a identificação individual até a detecção de condições de saúde. A tecnologia, ao facilitar análises precisas e rápidas, demonstra potencial para transformar práticas agrícolas ao aumentar a precisão diagnóstica e operacional, reduzindo custos e melhorando o bem-estar animal. Esses avanços sugerem que métodos avançados de análise de dados têm o potencial de se expandir dentro do setor agrícola, à medida que as ferramentas de aprendizado de máquina se tornam mais adaptadas às necessidades específicas do setor e mais acessíveis.

3.2 Tecnologia

Esta seção contextualiza alguns dos conceitos computacionais utilizados na construção da ferramenta de análise automática do escore de pigmentação. São discutidos conceitos básicos de visão computacional, além de aspectos relacionados às técnicas de inteligência artificial e à sua utilização na detecção e na classificação de imagens.

3.2.1 Inteligência Artificial

Russell e Norvig (2022) apresentam duas formas de definir a inteligência artificial (IA). Uma linha de pensamento caracteriza a IA em termos de fidelidade ao desempenho humano, enquanto outra prioriza uma definição abstrata e formal da inteligência como racionalidade, isto é, como a capacidade de realizar a ação correta. A história da IA pode ser traçada desde Aristóteles, que formulou um conjunto de leis destinado a descrever a parte racional da mente.

Wooldridge (2021) prefere citar o matemático Alan Turing como uma figura-chave no início do conceito contemporâneo de IA. Wooldridge considera o teste de Turing — experimento no qual um interrogador humano conversa por escrito com algo ou alguém e, ao final, tenta identificar se a outra parte era uma pessoa ou uma máquina — um marco do conceito moderno de IA. Para Turing, uma máquina capaz de convencer o interrogador de que era humana demonstraria comportamento inteligente. Entretanto, Wooldridge observa que, apesar de ser um conceito interessante, muitos programadores passaram a utilizar estratégias voltadas apenas a enganar o interrogador, em vez de desenvolver sistemas verdadeiramente capazes de responder de forma inteligente.

A IA é um termo amplo para definir diversas técnicas de resolução automatizada de problemas. Podem ser citadas a resolução por meio de busca informada (heurística) e não informada, agentes lógicos, raciocínio e programação probabilística, aprendizado de máquina, processamento de linguagem natural, visão computacional e robótica (Russell; Norvig, 2022).

3.2.2 Visão Computacional

Utilizando a definição de Marengoni e Stringhini (2009), a visão computacional busca emular a visão humana, utilizando uma imagem como entrada e tendo como saída uma interpretação de uma parte ou de toda a imagem. A visão computacional consiste em um conjunto de técnicas que visam extrair informações de imagens digitais com o objetivo de emular, no computador, a visão animal, buscando identificar ou classificar quaisquer objetos ou seres que sejam retratados na imagem (Klette, 2014).

3.2.2.1 Segmentação

No contexto de visão computacional, a segmentação de imagens busca dividir uma imagem digital em partes menores e mais simples de lidar, com o objetivo de facilitar o proces-

samento e a análise do que está retratado (Shapiro; Stockman, 2001). Já Saldanha e Freitas (2009) utilizam a definição de Jain (1989), que apresenta a segmentação como o processo de decomposição de uma imagem digital em diversos segmentos (regiões) que a formam.

Com a segmentação busca-se detectar regiões da imagem onde os *pixels* possuam informações como brilho, cor e textura semelhantes, fato que geralmente ocorre em situações onde se encontra um objeto contínuo na imagem. Nas regiões de fronteira entre objetos, existe uma tendência de que as informações dos *pixels* possuam uma variação maior (Russell; Norvig, 2022).

Entre os métodos de segmentação encontram-se a limiarização, o crescimento de regiões, a divisão e fusão e a segmentação *watershed*. Cada método apresenta vantagens e limitações, e diferentes situações podem responder melhor a técnicas distintas. Tradicionalmente, a segmentação de imagens utiliza algoritmos desenvolvidos por especialistas. Entretanto, redes neurais artificiais (RNAs), uma das abordagens da inteligência artificial (IA), também podem ser utilizadas nessa tarefa.

3.2.3 Neurônios artificiais

O primeiro modelo de neurônio artificial, entendido como uma aproximação matemática de uma célula neuronal, surgiu na década de 1940 (McCulloch; Pitts, 1943). Para facilitar a compreensão dos neurônios artificiais, apresenta-se inicialmente uma explicação simplificada do funcionamento de uma célula nervosa animal e, em seguida, descreve-se o neurônio artificial, destacando-se as semelhanças entre os dois modelos.

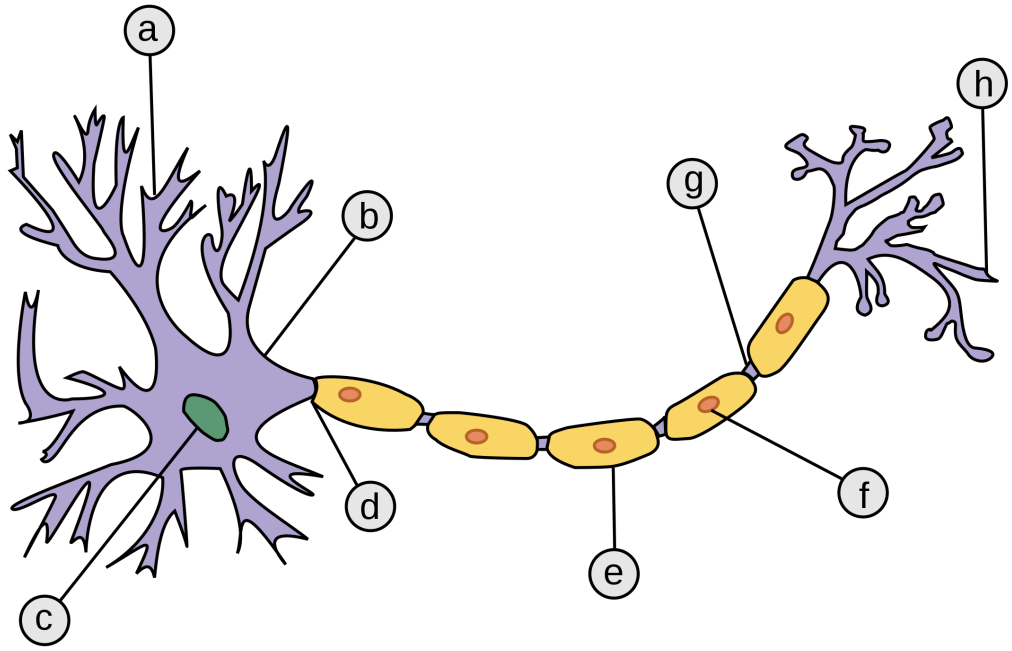
Embora existam diferentes tipos de neurônios biológicos, a Figura 4 apresenta um modelo simplificado de uma célula nervosa. Na imagem, em (a), encontra-se um dendrito, responsável por receber informações na forma de sinais eletroquímicos. Esses sinais chegam ao corpo celular (b), aumentando ou diminuindo o potencial elétrico do neurônio. Em (c), encontra-se o núcleo da célula.

Dependendo do tipo de neurônio e de outros fatores fisiológicos e bioquímicos, quando um potencial elétrico específico é atingido, o neurônio dispara uma descarga eletroquímica por meio de seu axônio (d), revestido pela bainha de mielina (e). Cada segmento da bainha de mielina é formado por uma célula de Schwann (f), e os segmentos são separados por nódulos de Ranvier (g). Ao final do axônio, encontram-se os terminais axonais (h), que transmitem sinais a outros neurônios ou células.

Uma explicação mais detalhada e precisa das células neuronais e do seu funcionamento

pode ser encontrada em livros de neurociência ou neurobiologia como os de Kandel *et al.* (2014) e Martin *et al.* (2021).

Figura 4 – Modelo simplificado de neurônio animal.



Fonte: Lokal Profil (2006).

A seguir, analisa-se o funcionamento de um neurônio artificial. Na Figura 5, as entradas são representadas por x_1, x_2 e x_3, podendo existir tantos valores de entrada quantos forem necessários. Essas entradas correspondem, de forma simplificada, aos sinais recebidos pela célula biológica. Cada valor é multiplicado por um peso, representado na figura por w_1, w_2 e w_3. O círculo interno representa o corpo do neurônio artificial, no qual se calcula o somatório das entradas ponderadas juntamente com o *bias* b. O resultado é utilizado como entrada da função de ativação f. O *bias* pode aumentar ou diminuir a entrada líquida da função de ativação, conforme seu valor seja positivo ou negativo, respectivamente (Haykin, 2001). Dependendo do valor de entrada e da função de ativação, obtém-se a saída y.

Matematicamente o neurônio artificial pode ser representado pela função:

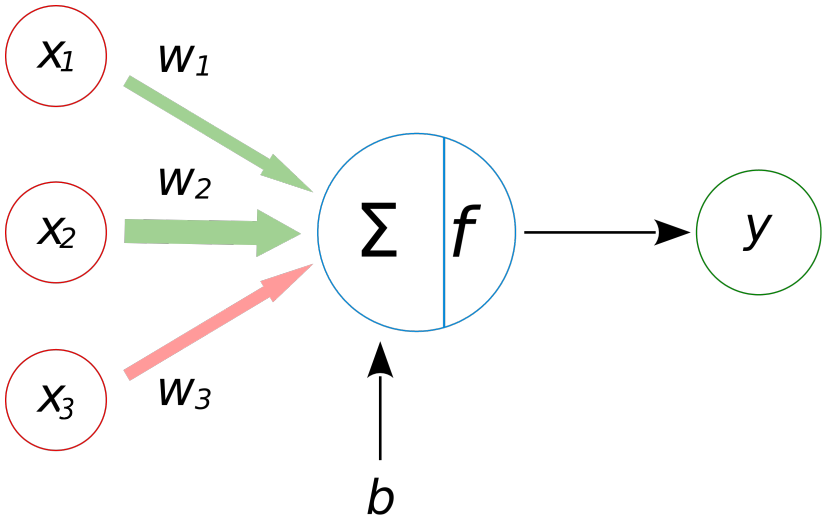
$$g(x_1, x_2, \ldots, x_n) = f\left(b + \sum_{i=1}^{n} x_i w_i\right) \tag{1}$$

em que $g(x_1, x_2, \ldots, x_n)$ representa o neurônio, $f(\cdot)$ é a função de ativação, b é o *bias*, x_i representa os valores de entrada e w_i, os pesos correspondentes.

Essa não é a única forma de representação possível de um neurônio artificial. Como mencionado anteriormente, o primeiro modelo de neurônio artificial foi descrito por McCulloch

e Pitts (1943). Mais detalhes sobre modelos de neurônios artificiais podem ser encontrados em Haykin (2001) e Russell e Norvig (2022), entre outros livros e artigos da área.

Figura 5 – Modelo simplificado de neurônio artificial.



Fonte: Gerazov (2023).

3.2.3.1 Funções de ativação

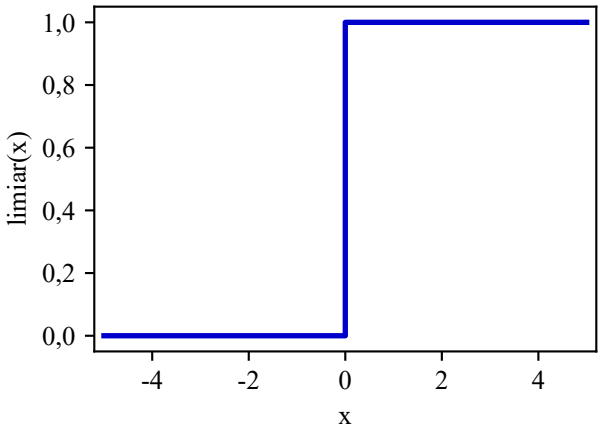
As funções de ativação desempenham um papel fundamental no funcionamento dos neurônios artificiais, pois introduzem não linearidades essenciais à capacidade de representar relações complexas. Sem essas funções, a combinação de diferentes camadas lineares permaneceria equivalente a uma única transformação linear, o que limitaria a variedade de problemas que poderiam ser aprendidos pela rede. As funções de ativação determinam como o valor produzido pelo somatório das entradas ponderadas e pelo *bias* será transformado antes de ser encaminhado aos neurônios seguintes. Dessa maneira, elas influenciam tanto a forma das respostas produzidas pelo modelo quanto o comportamento dos gradientes utilizados durante o treinamento (Rasamoelina; Adjailia; Sinčák, 2020; Sharma; Sharma; Athaiya, 2020). A escolha de uma função adequada depende da arquitetura, do tipo de problema e das características desejadas para a saída.

A primeira função de ativação, proposta por McCulloch e Pitts (1943), possuía um comportamento por vezes denominado “tudo ou nada” (Haykin, 2001). Essa denominação decorre do fato de que esse tipo de função, chamada de *função de limiar* ou modelo de McCulloch–Pitts, possui uma definição matemática semelhante à seguinte:

$$f(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \tag{2}$$

em que o argumento v é o valor gerado pelo somatório das entradas e pelo *bias*. A Figura 6 apresenta graficamente o comportamento dessa função, com saída 0 para valores negativos e 1 para valores iguais ou superiores a zero.

Figura 6 – Gráfico da função de limiar.



Fonte: Autor (2024).

Uma variação da função anterior é a chamada função bipolar (Apicella *et al.*, 2021), cuja diferença está no valor de saída para entradas menores do que zero. Na função bipolar, o resultado é -1 em vez de 0.

$$f(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ -1, & \text{se } v < 0 \end{cases} \tag{3}$$

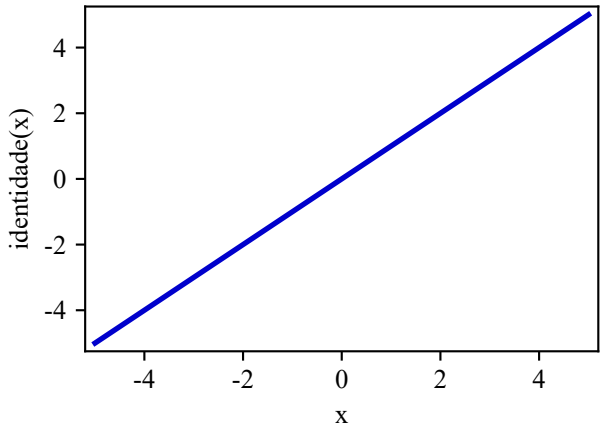
em que v é o valor de entrada da função.

A função identidade (Figura 7) também pode ser usada como uma função de ativação, assim como a função tangente hiperbólica (Figura 8). Outras duas funções que podem ser citadas são as funções sigmoide (Figura 9) e sigmoide bipolar (Figura 10) (Apicella *et al.*, 2021; Rasamoelina; Adjailia; Sinčák, 2020; Sharma; Sharma; Athaiya, 2020). A seguir, são definidas as funções aqui citadas.

$$identidade(x) = x \tag{4}$$

em que x é o valor de entrada da função.

Figura 7 – Gráfico da função identidade ou linear.

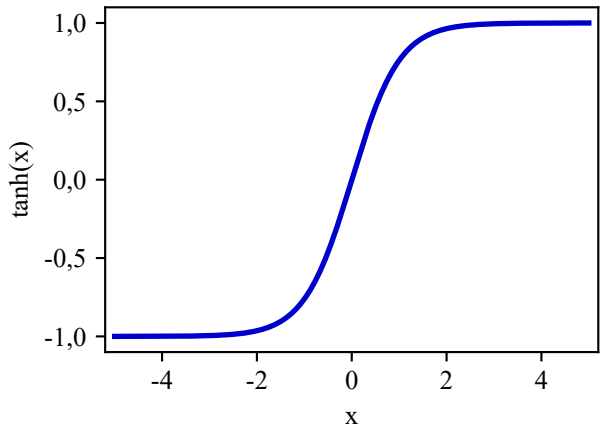


Fonte: Autor (2024).

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{5}$$

em que x é o valor de entrada da função.

Figura 8 – Gráfico da função tangente hiperbólica.



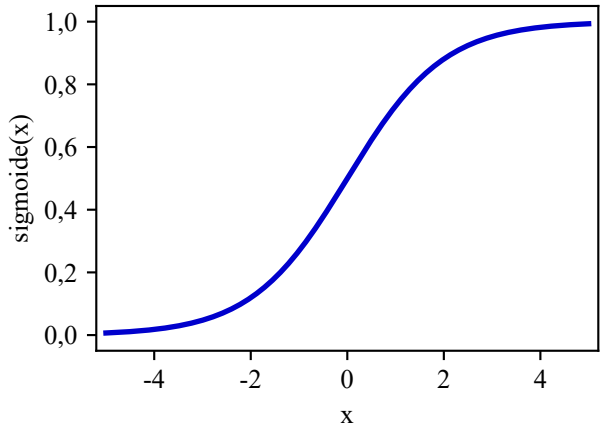
Fonte: Autor (2024).

$$\sigma(v) = \frac{1}{1 + e^{-v}} \tag{6}$$

em que v é o valor de entrada.

$$\sigma_B(v) = \frac{1 - e^{-v}}{1 + e^{-v}} \tag{7}$$

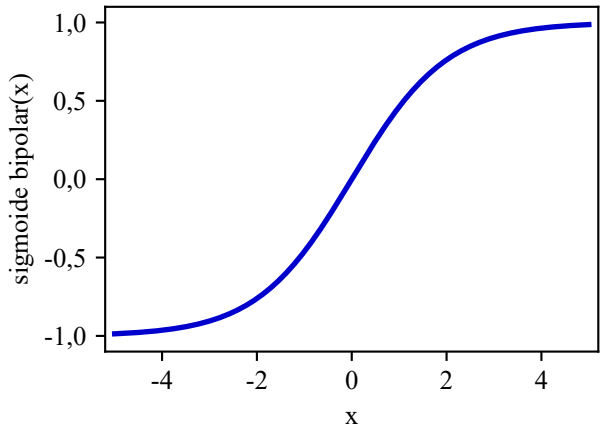
Figura 9 – Gráfico da função sigmoide.



Fonte: Autor (2024).

em que *v* é o valor de entrada.

Figura 10 – Gráfico da função sigmoide bipolar.



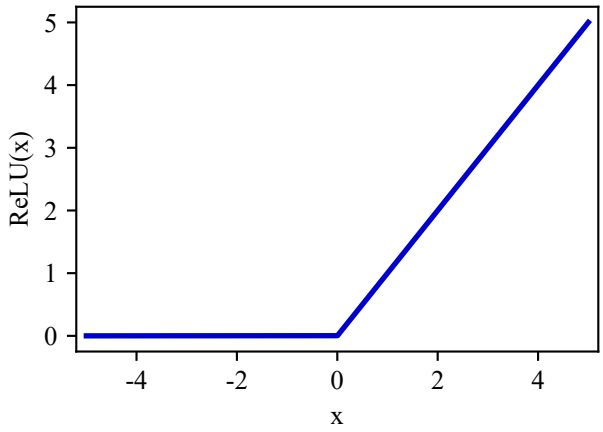
Fonte: Autor (2024).

A Figura 11 mostra a função *Rectified Linear Unit* (ReLU). Embora ela tenha sido introduzida por Fukushima (1969) em um contexto de neurociência visual, a função passou a ser usada em redes neurais artificiais após o trabalho de Nair e Hinton (2010). Segundo esses autores, ela apresentou melhores resultados do que a função sigmoide para tarefas de reconhecimento de imagens. Glorot, Bordes e Bengio (2011) também concluíram que a função ReLU produz melhores resultados do que funções como sigmoide e tangente hiperbólica. Matematicamente, a ReLU pode ser definida como:

$$f(x) = \begin{cases} 0, & \text{se } x < 0 \\ x, & \text{se } x \geq 0 \end{cases} \tag{8}$$

em que x é o valor de entrada.

Figura 11 – Gráfico da função ReLU.



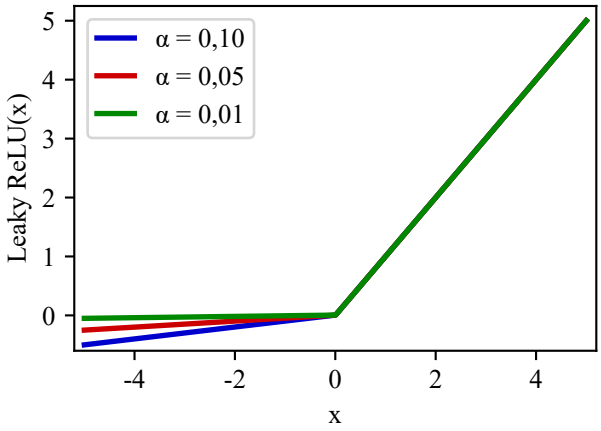
Fonte: Autor (2024).

Bengio, Simard e Frasconi (1994) descreveram o fenômeno conhecido como *vanishing gradient*, ou gradiente que se anula. Esse problema ocorre principalmente em redes neurais profundas (*Deep Neural Networks* — DNNs), discutidas na seção 3.2.3.5, quando os gradientes propagados pelas camadas se tornam muito pequenos e dificultam o ajuste dos pesos. A função ReLU reduz parte desse problema para entradas positivas, mas pode apresentar outra limitação, denominada *dying ReLU*, na qual determinados neurônios passam a produzir saída zero de forma persistente (Berner; Elbrächter; Grohs, 2019; Douglas; Yu, 2018; Ferlez; Shoukry, 2020; Goel; Karmalkar; Klivans, 2019; Hayou; Doucet; Rousseau, 2019; Lu *et al.*, 2020; Oostwal; Straat; Biehl, 2021; Petrosyan; Dereventsov; Webster, 2020). Como alternativa, Maas, Hannun e Ng (2013) propuseram uma variação denominada Leaky ReLU, cujo gráfico é apresentado na Figura 12. Para valores menores do que zero, aplica-se a multiplicação αx, sendo α um valor pequeno, geralmente 0,01 ou 0,001. A função Leaky ReLU é definida como:

$$f(x) = \begin{cases} \alpha x, & \text{se } x < 0 \\ x, & \text{se } x \geq 0 \end{cases} \tag{9}$$

Outra função que busca reduzir algumas limitações da ReLU é a Softplus. A função é definida como:

Figura 12 – Gráfico da função Leaky ReLU.



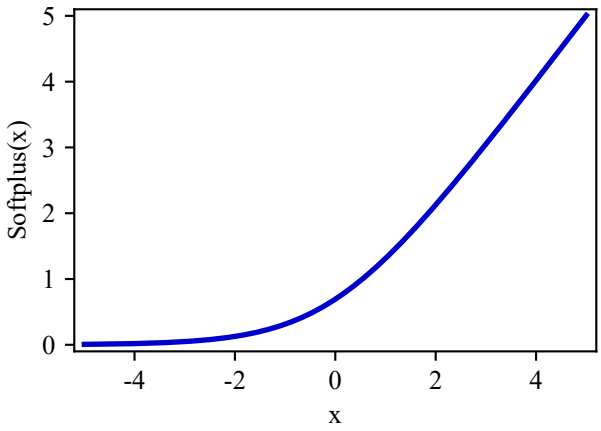
Fonte: Autor (2024).

$$\text{softplus}(x) = \log(1 + e^{x}) \tag{10}$$

em que x é o valor de entrada.

A Figura 13 apresenta o gráfico da função Softplus.

Figura 13 – Gráfico da função Softplus.



Fonte: Autor (2024).

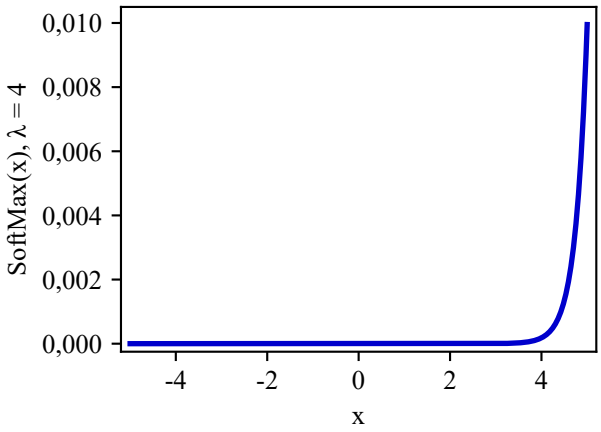
Uma função utilizada na última camada de classificadores multiclasse para fazer com que os resultados de saída sejam transformados em valores probabilísticos é a chamada SoftMax (Bridle, 1990a, 1990b). A função SoftMax é definida como:

$$\text{SoftMax}(z_i, \lambda) = \frac{e^{\lambda z_i}}{\sum_j e^{\lambda z_j}} \tag{11}$$

em que z_i representa os valores de entrada, λ é o parâmetro de ajuste de temperatura e

o denominador corresponde à soma de $e^{\lambda z_j}$ para todos os valores de j, garantindo que as saídas somem 1 e possam ser interpretadas como probabilidades. A Figura 14 apresenta o gráfico da função SoftMax com $\lambda = 4$.

Figura 14 – Gráfico da função SoftMax.



Fonte: Autor (2024).

No trabalho de Apicella *et al.* (2021), as funções de ativação apresentadas são classificadas como funções de forma fixa, pois suas características são predefinidas e não se modificam durante o treinamento das redes. Nesse mesmo estudo, os autores descrevem outro tipo, denominado função treinável, cujas características podem ser ajustadas juntamente com os pesos e os *biases* da rede durante o treinamento. O objetivo é melhorar os resultados e o desempenho dos modelos. As funções de ativação treináveis não serão detalhadas neste trabalho, pois não foram utilizadas na solução do problema de pesquisa.

Em resumo, as funções de ativação são componentes críticos na modelagem de neurônios artificiais, cada uma trazendo características únicas que as tornam adequadas para diferentes contextos dentro da inteligência artificial. A evolução das funções de ativação desde a função de limiar, passando pela sigmoide, tangente hiperbólica, até as mais recentes como ReLU e suas variantes, Leaky ReLU e SoftPlus, reflete a busca contínua por eficiência e eficácia no treinamento de modelos. A função de limiar e sua variante bipolar são fundamentais para tarefas que requerem decisões binárias simples. Funções contínuas e suaves, como a sigmoide e tangente hiperbólica, são preferidas para problemas onde a suavidade do gradiente é crucial. A ReLU, por sua simplicidade e eficácia, tornou-se a escolha dominante para redes neurais profundas, embora suas variantes, como a Leaky ReLU e a SoftPlus, tenham sido desenvolvidas para mitigar limitações específicas, como o efeito de neurônios mortos e *vanishing gradient*. A função SoftMax é particularmente útil em contextos de classificação multiclasse, pois ela converte os valores de saída brutos e não normalizados de uma rede neural em probabilidades claramente

interpretáveis. Esse panorama das funções de ativação ilustra como a escolha depende intrinsecamente do tipo de problema, da arquitetura do modelo e do comportamento desejado durante o aprendizado, orientando assim a seleção para otimizar o desempenho e a convergência do modelo. Os conceitos de redes neurais e redes neurais profundas serão apresentados nas seções seguintes.

3.2.3.2 Aprendizado de Máquina

Considera-se agente toda entidade que age ou interage com algo ou com determinada situação. Um agente racional busca agir de modo a alcançar o melhor resultado possível. Seu desempenho pode melhorar quando realiza observações sobre a tarefa e o ambiente em que ela ocorre. Quando o agente que aprende é um computador, esse processo é denominado aprendizado de máquina (Russell; Norvig, 2022).

De acordo com Monard e Baranauskas (2003), o aprendizado de máquina é um campo da inteligência artificial voltado ao desenvolvimento de técnicas de aprendizagem computacional e de sistemas capazes de adquirir conhecimento automaticamente. O aprendizado de máquina é uma tecnologia que está presente em diversos aspectos da sociedade atual, incluindo, entre outros, mecanismos de busca na internet, filtros de conteúdos em redes sociais, sistemas de recomendação em *websites* de lojas, e diversos produtos de consumo como câmeras e telefones celulares (LeCun; Bengio; Hinton, 2015).

3.2.3.3 Redes Neurais Artificiais

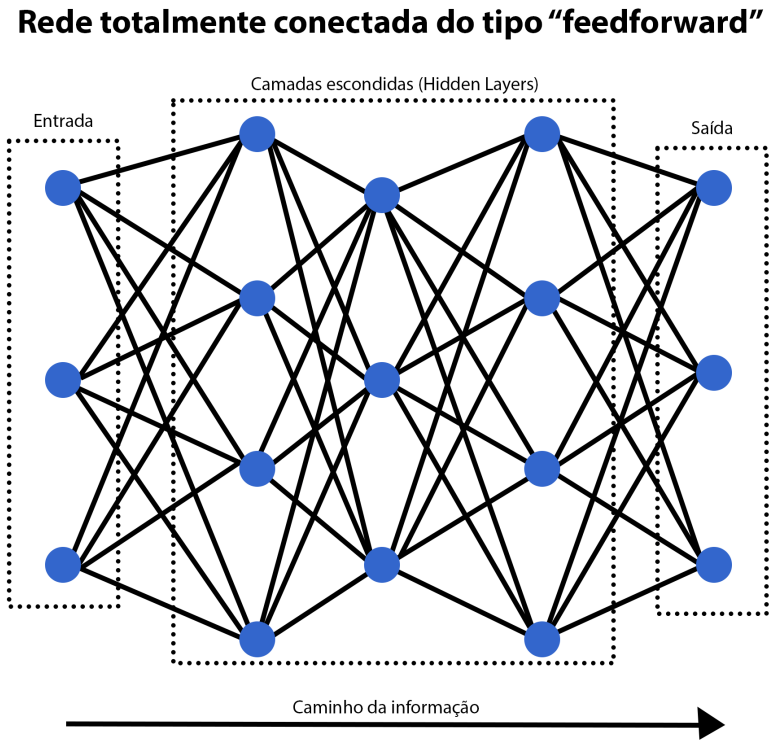
Com base nos conceitos de neurônios artificiais e funções de ativação, pode-se definir uma rede neural artificial (RNA). Em sua forma mais simples, uma RNA consiste em uma organização de neurônios artificiais capaz de realizar tarefas mais complexas do que aquelas processadas por um único neurônio. Um dos primeiros modelos foi o perceptron, proposto por Rosenblatt (1962). Por possuir uma única camada de processamento, esse modelo é limitado à solução de problemas linearmente separáveis.

Devido à característica acima descrita, a definição do perceptron de camada única pode ser usada de forma intercambiável com a definição de neurônio artificial mostrada em 3.2.3. Quando são unidos vários neurônios em uma camada, forma-se de modelo de múltiplos perceptrons ou rede neural de camada única (Haykin, 2001). Uma rede também pode ter várias camadas, com quantidades diferentes de neurônios em cada uma. As camadas que não correspondem à entrada nem à saída são denominadas de *hidden layers*, ou camadas ocultas. O

ajuste de quantas camadas uma rede terá, quantos neurônios em cada camada e quais funções de ativação utilizar são exemplos de hiperparâmetros, e devem ser ajustados de acordo com quais conjuntos de hiperparâmetros geram o melhor resultado final para o problema a ser solucionado.

Quando todos os neurônios de uma camada estão conectados a todos os neurônios da camada seguinte, a rede é denominada *fully connected* ou *densely connected*. Em uma arquitetura do tipo *feedforward*, os sinais se propagam em uma única direção durante a etapa de inferência: partem da entrada, atravessam as camadas ocultas e chegam à camada de saída, como demonstrado na Figura 15. Na Figura 15, o diagrama apresenta três camadas ocultas, duas delas com quatro neurônios e uma com três neurônios. Para que os parâmetros das camadas sejam ajustados durante o treinamento, utiliza-se o algoritmo de *backpropagation*, desenvolvido em meados da década de 1970 (Linnainmaa, 1976; Werbos, 1982, 1974). Esse algoritmo calcula os gradientes do erro em relação aos pesos e aos *biases*, propagando essas informações da saída em direção às camadas anteriores e permitindo o ajuste dos parâmetros das camadas ocultas.

Figura 15 – Diagrama de RNA totalmente conectada do tipo *feedforward*.



Fonte: Autor (2024).

Com o *backpropagation*, o erro calculado pela função de perda é propagado no sentido

inverso ao fluxo de saída, e seus gradientes são utilizados por um método de otimização, como o gradiente descendente estocástico, para ajustar os pesos e os *biases* dos neurônios (Haykin, 2001). A passagem completa do conjunto de dados pela rede é denominada *epoch*.

À medida que o número de *epochs* aumenta, a rede pode aproximar seus pesos de valores mais adequados à tarefa, mas também cresce o risco de *overfitting*. No *overfitting*, o modelo se ajusta excessivamente aos dados de treinamento, podendo memorizar padrões específicos e apresentar desempenho inferior em dados não utilizados durante o treinamento. Por outro lado, um número insuficiente de *epochs* pode causar *underfitting*, situação em que a rede não aprende adequadamente os padrões necessários à tarefa e permanece incapaz de produzir resultados satisfatórios.

Para evitar tanto o risco de *overfitting* quanto de *underfitting* é essencial que o número de dados utilizados para treinamento seja o maior e mais variado possível, sejam eles imagens, vídeos, valores, áudios ou qualquer outra informação que se queira utilizar a rede para resolver um problema. Entretanto, um número maior de dados também aumenta o tempo de treinamento da rede (Goodfellow; Bengio; Courville, 2016). Existem algumas outras estratégias para tentar amenizar o *overfitting*, como o *dropout*, que desliga um ou mais neurônios durante o treinamento, para forçar a rede a utilizar outros neurônios, e reduzir a dependência de combinações específicas de neurônios.

Entretanto, durante muito tempo não havia poder computacional suficiente para se realizar tarefas complexas em tempo hábil utilizando-se RNAs. Mais recentemente, com o aumento do poder computacional e a criação de hardware especializado para o cálculo necessário em uma RNA os computadores passaram a dispor de velocidade suficiente para realizar o processamento de informações de maior complexidade em um tempo viável para a utilização deste método de forma mais intensa.

3.2.3.4 Aprendizado Supervisionado

O aprendizado supervisionado é uma das maneiras de se treinar uma RNA. No aprendizado supervisionado, um conjunto de dados previamente classificado por um especialista humano é utilizado no treinamento de uma rede neural. Inicialmente, os pesos utilizados em cada nó da rede podem ser inicializados com valores aleatórios, e após cada ciclo de treinamento, o resultado gerado pela rede é comparado com a classificação já existente, e, utilizando uma função de perda sobre os resultados obtidos, os pesos da rede são reajustados buscando minimizar o erro (Haykin, 2001).

Redes neurais artificiais podem ser utilizadas em diferentes tipos de dados e informações.

Figura 16 – Simulação de uma imagem após reconhecimento por rede neural.



Fonte: Autor (2024).

No caso específico em que uma rede neural é utilizada para reconhecer ou extrair informações de imagens, pode-se compreender o processo de aprendizado supervisionado utilizando a explicação dada por LeCun, Bengio e Hinton (2015): primeiro um grande conjunto de imagens é coletado, podendo mostrar diversos objetos ou seres diferentes. Essas imagens são categorizadas e recebem uma *label*, ou rótulo, uma descrição do que a imagem mostra. Durante o processo de treinamento a máquina recebe as imagens e gera como saída um vetor de *scores*, um para cada categoria.

Após esse processo, o resultado gerado pela rede neural é então comparado com as informações reais de cada imagem, e por meio de uma função de erro calcula-se o quão distante a rede está de produzir as respostas corretas. Os pesos e *biases* da rede então são modificados na tentativa de diminuir a distância entre seus resultados e a referência correta. O objetivo é diminuir o erro ao mínimo possível, ou seja, fazer com que o resultado gerado pela rede neural aproxime-se o máximo possível da realidade, até que eventualmente a rede consiga identificar o que está na imagem com uma precisão próxima ou mesmo maior do que um observador humano teria.

A Figura 16 simula a saída de uma RNA treinada para reconhecer gatos e cachorros. Na imagem, os animais são circundados por *bounding boxes*, acompanhadas de uma *label* e de um valor entre 0 e 1. Esse valor representa o nível de confiança da classificação, sendo 0 equivalente a 0% e 1 equivalente a 100%. No aprendizado supervisionado, a imagem possui um arquivo separado com as coordenadas das *bounding boxes* e os respectivos rótulos, que, nesse exemplo, poderiam ser “gato” e “cachorro”. A rede recebe a imagem como entrada e tenta identificar seu

conteúdo. A menos que tenha sido inicializada com os pesos exatos para encontrar os animais, as chances de realizar corretamente a detecção e a classificação já na primeira *epoch* são muito pequenas. Entretanto, à medida que o treinamento avança e os pesos são ajustados, espera-se que a rede aprenda a tarefa e alcance níveis elevados de precisão e acurácia.

Com base no que foi explicado sobre a Figura 16, pode-se perceber que RNAs apresentam resultados probabilísticos. Uma rede bem treinada recebendo como entrada uma imagem clara e de boa qualidade do objeto a ser reconhecido conseguirá atingir uma confiança maior na sua classificação. Em contraste, uma rede que não foi corretamente treinada ou que recebe uma imagem pouco clara do objeto, pode não conseguir reconhecê-lo, pode reconhecer com baixo nível de confiança, ou ainda pode realizar uma classificação errada, como no caso da Figura 16 onde uma rede mal treinada poderia confundir o gato com um segundo cachorro, ou o inverso.

3.2.3.5 Aprendizado Profundo

Métodos convencionais de aprendizado de máquina eram limitados em sua habilidade de processar dados naturais em sua forma pura. O conceito de aprendizado profundo é dado a redes nas quais existe mais de uma camada oculta, entretanto redes neurais profundas possuem frequentemente dezenas ou centenas de camadas ocultas (Goodfellow; Bengio; Courville, 2016). Cada camada é responsável por detectar características cada vez mais abstratas, tornando-as capazes de aprender funções de grande complexidade (LeCun; Bengio; Hinton, 2015).

De acordo com Russell e Norvig (2022), o aprendizado profundo compreende uma ampla família de técnicas de aprendizado de máquina nas quais as hipóteses assumem a forma de circuitos algébricos complexos, cujas intensidades de conexão podem ser ajustadas. Como já dito, tipicamente uma rede neural profunda pode possuir dezenas, centenas ou mesmo milhares de camadas, como, por exemplo, a arquitetura *Residual Network* (ResNet), que chegou a ser testada com um total de 1.202 camadas (He *et al.*, 2016), porém possui versões com 50 camadas (ResNet50), 101 camadas (ResNet101), entre outras.

O aumento do número de camadas traz vantagens e desvantagens. Entre as vantagens, as redes profundas possuem potencial para resolver problemas de elevada complexidade, que seriam difíceis para redes com apenas uma camada oculta. Entre as desvantagens, podem ser citados o aumento do tempo de processamento, a necessidade de maior capacidade computacional e o risco de *overfitting* (Dong; Wang; Abbas, 2021).

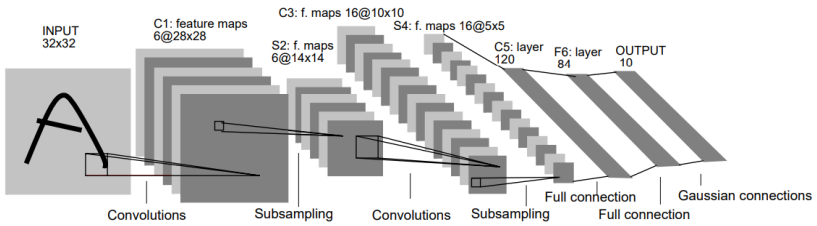
3.2.3.6 Redes Neurais Convolucionais

As redes neurais convolucionais (CNNs, do inglês *Convolutional Neural Networks*) tiveram como precursor o neocognitron proposto por Fukushima (1980) e foram posteriormente desenvolvidas em aplicações como a apresentada por LeCun *et al.* (1998). No trabalho de LeCun *et al.* (1998), os autores buscaram uma forma mais eficiente de reconhecer caracteres escritos à mão. O funcionamento das CNNs foi inspirado em estudos sobre o córtex visual de gatos (Hubel; Wiesel, 1962).

A rede proposta por LeCun *et al.* (1998), denominada LeNet-5, recebia imagens em escala de cinza com dimensão de 32×32 *pixels*, contendo caracteres escritos à mão. A arquitetura possui sete camadas: duas convolucionais, duas de *subsampling*, duas totalmente conectadas e uma camada de saída. A Figura 17 apresenta um diagrama da LeNet-5.

a) *camadas de convolução*: o objetivo dessas camadas é extrair características da imagem de entrada ou da camada anterior. A extração ocorre pela aplicação de uma pequena matriz de pesos, denominada *kernel* ou filtro, geralmente com dimensão de 3 × 3, embora possa ser maior. Ao multiplicar os valores dos *pixels*, o filtro gera uma versão modificada da imagem, na qual determinadas características podem ser acentuadas ou suavizadas. Dependendo dos valores do *kernel*, podem ser detectados detalhes como curvaturas, linhas verticais ou horizontais e bordas. Os *kernels* também podem produzir efeitos como aumento de nitidez, desfoque e relevo. Os *kernels* deslocam-se pela imagem de acordo com um passo denominado *stride*. Quanto maior o *stride*, menor tende a ser o mapa de características resultante. A escolha desse valor depende dos objetivos da modelagem, das características do conjunto de dados e do equilíbrio desejado entre eficiência computacional e preservação de informações. Testes sistemáticos são importantes para definir um valor adequado ao problema;
b) *camadas de pooling ou subsampling*: a função do *pooling* é reduzir a dimensionalidade dos mapas de características da camada anterior, geralmente uma camada convolucional. Em vez de aplicar matrizes como *kernels*, no *pooling* pequenas janelas são extraídas do mapa de características anterior. Com uma operação de *max pooling*, preserva-se apenas o maior valor da janela; com o *average pooling*, calcula-se a média de seus valores. Essa janela também se desloca por meio de um *stride* predefinido. Quanto maiores forem a janela e o *stride*, menor será o mapa de características produzido. Com essa redução, diminuem o número de parâmetros da rede, o risco de *overfitting* e a sensibilidade a pequenas translações e deformações. Contudo, é importante não reduzir excessivamente a

Figura 17 – Diagrama da LeNet-5.



Fonte: LeCun *et al.* (1998).

dimensionalidade, pois isso pode causar a perda de informações relevantes para os processos de detecção e classificação;

c) *camadas totalmente conectadas*: após as etapas de convolução e *subsampling*, as informações são encaminhadas a camadas totalmente conectadas, nas quais ocorre a classificação. Conforme LeCun *et al.* (1998), seria possível executar todo o processo com uma rede totalmente conectada, mas o custo computacional seria maior e os resultados não necessariamente teriam a mesma qualidade. As camadas de convolução e *pooling* reduzem significativamente o número de parâmetros, sobretudo em dados como imagens e vídeos, que exigiriam grande quantidade de neurônios de entrada.

3.2.3.7 Aprendizado por transferência

O aprendizado por transferência, ou *transfer learning*, é um processo que vem sendo desenvolvido ao longo do tempo (Lim; Salakhutdinov; Torralba, 2011; Pan; Yang, 2010), cujo objetivo é reutilizar uma rede neural previamente treinada em uma tarefa para executar outra tarefa semelhante, reduzindo os recursos computacionais e o tempo necessários. Essa técnica se beneficia do fato de que os pesos do modelo já foram ajustados durante um treinamento anterior, em vez de partirem de valores aleatórios, que inicialmente são incapazes de produzir resultados úteis. Para exemplificar, pode-se imaginar uma rede treinada para detectar cachorros e gatos em imagens que será reutilizada para detectar cavalos. Durante o treinamento original, os pesos da rede foram ajustados a partir de uma grande quantidade de imagens, permitindo que o modelo aprendesse características gerais relevantes para tarefas de reconhecimento visual. Dessa forma, espera-se que o novo treinamento seja mais eficiente do que o treinamento de uma rede iniciada com pesos aleatórios. Ao utilizar o *transfer learning*, é possível modificar a camada de saída da rede, em geral substituindo-a por uma ou mais camadas, e também congelar os pesos das primeiras camadas para preservar sua capacidade de detectar características mais gerais das

imagens. Após esses ajustes, a rede pode ser treinada com um número menor de imagens do que seria necessário para treinar um modelo desde o início (Goodfellow; Bengio; Courville, 2016; Sarkar; Bali; Ghosh, 2018; Shanmugamani, 2018). Como será demonstrado na subseção 3.4, algumas arquiteturas de rede são amplamente utilizadas em processos de aprendizado por transferência. Nas seções 3.2.3.8, 3.2.3.9 e 3.2.3.10, são apresentadas algumas arquiteturas descritas na literatura que se destacam nesse tipo de aplicação devido ao desempenho alcançado em tarefas e competições de detecção e classificação de imagens.

3.2.3.8 VGG

Visual Geometry Group (VGG) é um grupo de pesquisa do Departamento de Engenharia da Universidade de Oxford[4]. No artigo *Very Deep Convolutional Networks for Large-Scale Image Recognition* (Simonyan; Zisserman, 2015), os autores propuseram uma arquitetura de rede neural convolucional com configurações de até 19 camadas, sendo que a versão de 16 camadas também é utilizada com sucesso em diversas tarefas de classificação em imagens (Ahmed *et al.*, 2023; Geng *et al.*, 2019; Guan *et al.*, 2019; Islam *et al.*, 2019; Sitaula; Hossain, 2021). Essas redes são hoje conhecidas como VGG-16 e VGG-19, sendo a numeração no nome um indicativo do número de camadas que cada uma delas possui.

A rede VGG é configurada para receber imagens coloridas com dimensão de 224 × 224 *pixels* e, diferentemente de modelos anteriores, utiliza filtros convolucionais pequenos, de 3 × 3. A arquitetura alcançou o primeiro e o segundo lugares nas tarefas de localização e classificação, respectivamente, na competição *ImageNet Challenge 2014*.

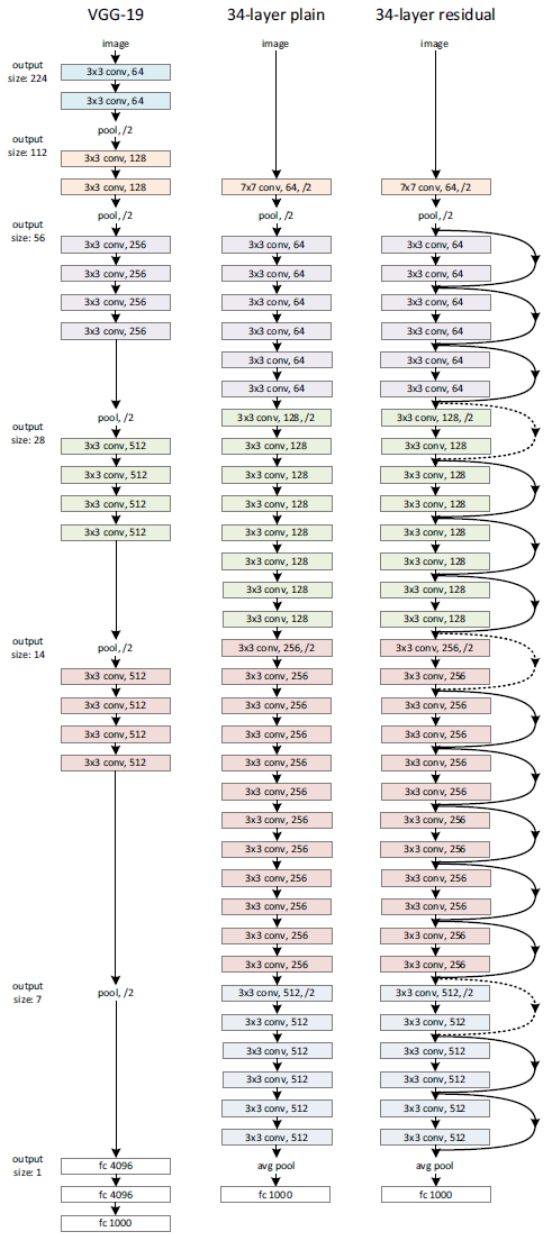
3.2.3.9 ResNet

Residual Network (ResNet) é uma arquitetura profunda que utiliza conexões de atalho entre camadas para transmitir informações residuais e reduzir problemas de treinamento em redes muito profundas. A arquitetura foi proposta por He *et al.* (2016) e teve como ponto de partida conceitos empregados na VGG, como filtros convolucionais pequenos. Os autores construíram inicialmente uma rede plana de 34 camadas e, em seguida, acrescentaram conexões residuais para formar uma ResNet de mesma profundidade. A Figura 18 apresenta um diagrama dessas arquiteturas.

Os autores testaram diferentes números de camadas, entre elas 34, 50, 101, 152 e até mesmo 1.202 camadas. Embora a rede de 1.202 camadas tenha se saído relativamente bem,

[4]https://www.robots.ox.ac.uk/~vgg/

Figura 18 – Arquitetura da rede VGG-19 à esquerda, a rede plana com 34 camadas no centro e a ResNet com 34 camadas à direita.



Fonte: He *et al.* (2016).

ela não demonstrou melhora significativa em comparação com modelos com quantidades bem menores de camadas. Com a ResNet, os autores conseguiram vencer várias categorias em competições no ano de 2015, superando o modelo VGG.

3.2.3.10 Faster R-CNN

A Faster R-CNN é um modelo proposto por Ren *et al.* (2015) como uma evolução da Fast R-CNN (Girshick, 2015), que, por sua vez, se baseou em trabalhos anteriores (Girshick *et al.*, 2014). Na Faster R-CNN, uma rede neural convolucional denominada *Region Proposal Network* (RPN) é utilizada para localizar possíveis regiões da imagem nas quais um objeto de interesse possa estar presente. As propostas produzidas por essa rede são encaminhadas à etapa de detecção e classificação. Os autores integraram as duas redes em uma única arquitetura, tornando o processo significativamente mais rápido do que no modelo Fast R-CNN. A versão anterior dependia de algoritmos externos para gerar as regiões de interesse, enquanto a Faster R-CNN passou a aprender essa etapa juntamente com o restante do modelo.

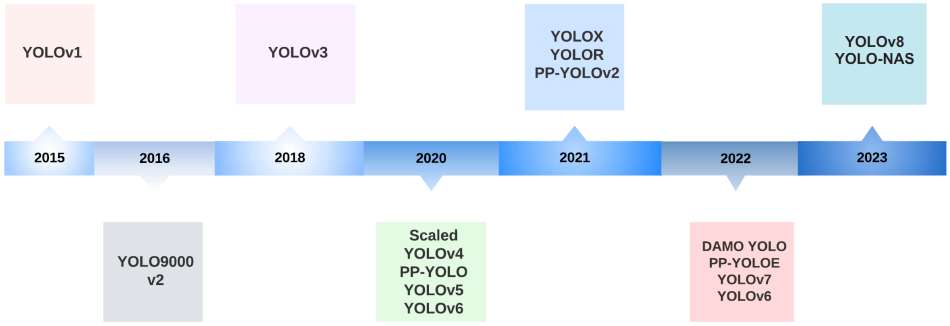
3.2.3.11 YOLOv8

A arquitetura *You Only Look Once* (YOLO) foi introduzida no trabalho de Redmon *et al.* (2016), com o objetivo de criar uma rede neural capaz de detectar e classificar objetos em uma imagem em uma única etapa. Após a introdução da versão inicial, outras versões com melhorias foram geradas por diferentes pesquisadores, sendo chamadas, em geral, de YOLOv2, YOLOv3, até chegar à YOLOv8, utilizada neste trabalho. Por ser uma rede com código-fonte aberto, outras versões com nomes levemente diferentes também foram criadas para uso em tarefas específicas (Terven; Córdova-Esparza; Romero-González, 2023). Na Figura 19 é possível ver uma linha temporal das versões da arquitetura YOLO.

A YOLOv8, lançada pela Ultralytics em janeiro de 2023, constitui uma família de modelos para diferentes tarefas de visão computacional e sucede a YOLOv5, também desenvolvida pela empresa. Os modelos podem executar tarefas como detecção de objetos, segmentação, estimativa de pose, rastreamento e classificação.

A arquitetura do YOLOv8 foi aprimorada para integrar recursos de alto nível com informações contextuais, aumentando a precisão e a eficiência na detecção. A YOLOv8 apresenta avanços consideráveis em relação aos seus predecessores, consolidando-se como uma ferramenta essencial para aplicações de visão computacional em tempo real. Avaliações no conjunto de dados *Microsoft Common Objects in Context* (MS COCO) confirmaram sua eficácia,

Figura 19 – Linha temporal das arquiteturas YOLO.



Fonte: Adaptado de Terven, Córdova-Esparza e Romero-González (2023).

demonstrando seu potencial para diversas aplicações práticas (Jocher; Chaurasia; Qiu, 2023; Terven; Córdova-Esparza; Romero-González, 2023). A Figura 20 mostra uma visualização da arquitetura YOLOv8.

A YOLOv8 possui cinco versões principais: YOLOv8x, YOLOv8l, YOLOv8m, YOLOv8s e YOLOv8n. Essas configurações variam da YOLOv8x, que possui maior número de camadas e parâmetros, à YOLOv8n, a menor entre elas e a utilizada neste trabalho. As versões maiores apresentam maior capacidade de representação e podem alcançar melhor desempenho em determinadas tarefas, mas exigem mais memória, capacidade de processamento e tempo de treinamento e inferência. A escolha da YOLOv8n foi realizada com o objetivo de reduzir a necessidade desses recursos e facilitar a posterior execução do modelo em dispositivos móveis.

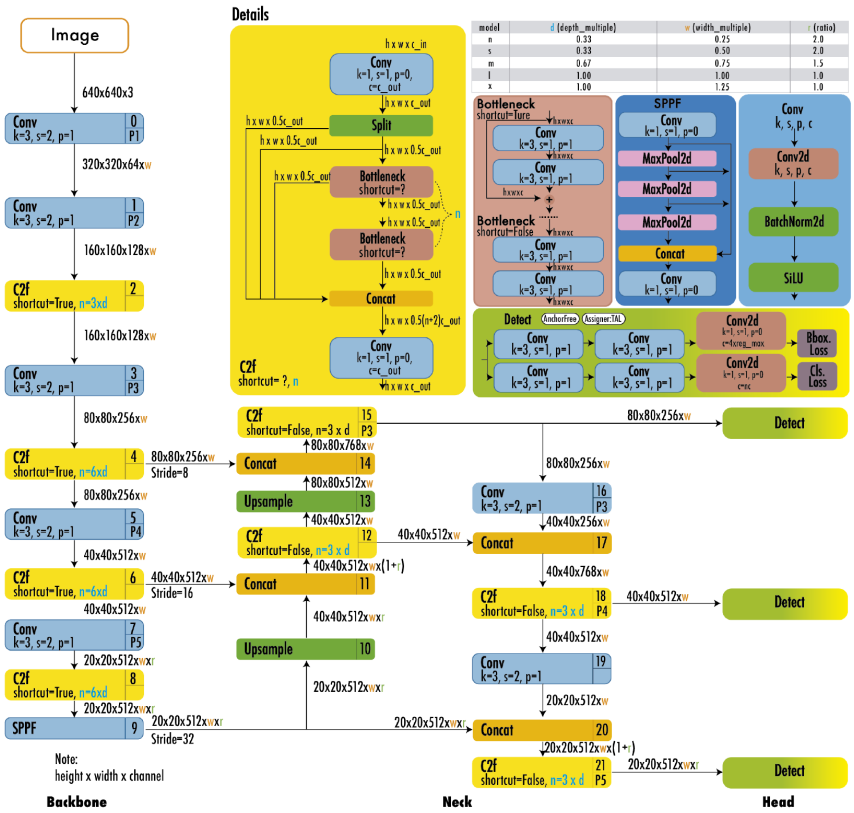
3.2.4 Avaliação de modelos de classificação

No contexto da inteligência artificial, um modelo de classificação busca prever resultados de uma determinada entrada (conjunto de valores) e indicar em qual categoria, entre duas ou mais, a informação pertence. Um exemplo de um classificador simples seria uma rede neural que recebe imagens de cães e gatos e responde se o animal representado na imagem é um cão ou um gato.

Outro exemplo bastante utilizado no ensino de redes neurais utiliza o conjunto de dados *Modified National Institute of Standards and Technology* (MNIST), que consiste de 70.000 imagens de números escritos à mão, variando entre zero e nove (Deng, 2012). O objetivo nesse caso é treinar a rede para detectar qual dígito está representado na imagem.

Porém, é necessário utilizar métodos para avaliar o desempenho desses modelos de classificação. Existem diversas métricas, algumas mais simples e outras mais complexas, e cada

Figura 20 – Visualização da arquitetura YOLOv8.



Fonte: Adaptado de Terven, Córdova-Esparza e Romero-González (2023).

uma delas é mais adequada para uma situação diferente.

A seguir, apresenta-se uma breve explicação de algumas métricas comuns de avaliação de modelos de classificação. A lista não é exaustiva, e determinadas métricas são calculadas de forma diferente em problemas binários e multiclasse.

3.2.4.1 Matriz de confusão

Uma matriz de confusão, também chamada de tabela de contingência (Powers, 2011), cruza as classes reais com as classes previstas. Em um problema de classificação binária, cada resultado é enquadrado em uma de quatro possibilidades:

a) *verdadeiro positivo*: ocorre quando o resultado previsto é positivo e coincide com a classe real. Nesse caso, o modelo identifica corretamente como positiva uma ocorrência que realmente pertence à classe positiva;
b) *verdadeiro negativo*: ocorre quando o resultado previsto é negativo e também coincide com a classe real. Nesse caso, o modelo identifica corretamente como negativa uma ocorrência que pertence à classe negativa;
c) *falso positivo*: ocorre quando o modelo classifica como positiva uma ocorrência que, na realidade, pertence à classe negativa;
d) *falso negativo*: ocorre quando o modelo classifica como negativa uma ocorrência que, na realidade, pertence à classe positiva.

Quadro 2 – Matriz de confusão para duas classes.

Realidade x Predição	**Sim**	**Não**
Sim	Verdadeiro Positivo (VP)	Falso Positivo (FP)
Não	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Fonte: Adaptado de Vilela Junior *et al.* (2022).

No exemplo do Quadro 2, é mostrada uma situação em que existem apenas duas classes de saída, sim e não. A matriz pode ser expandida para quantas classes forem necessárias, bastando adicionar mais uma linha e uma coluna para cada classe extra. Idealmente, espera-se que todos os resultados previstos recaiam sobre a diagonal principal, o que indicaria 100% de acertos nas previsões.

3.2.4.2 Acurácia

A acurácia é calculada diretamente a partir da matriz de confusão, dividindo-se a soma de VP e VN pela soma de VP, VN, FP e FN (Hossin; Sulaiman, 2015; Vujović, 2021).

$$\text{Acurácia} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{VN} + \text{FP} + \text{FN}} \tag{12}$$

A acurácia representa a proporção de previsões corretas em relação ao total de previsões realizadas. Embora seja uma medida de interpretação simples, sua análise deve considerar a distribuição das classes, especialmente em conjuntos de dados desbalanceados.

3.2.4.3 Precisão e Revocação (recall)

A precisão corresponde à proporção de verdadeiros positivos entre todos os elementos previstos como positivos. Essa métrica indica o grau de confiabilidade das previsões positivas do modelo.

$$\text{Precisão} = \frac{\text{VP}}{\text{VP} + \text{FP}} \tag{13}$$

A revocação, ou *recall*, é semelhante à precisão, mas utiliza os falsos negativos no denominador em vez dos falsos positivos (Flach; Kull, 2015). A revocação mede a capacidade do modelo de identificar os elementos positivos presentes no conjunto de dados.

$$\text{Recall} = \frac{\text{VP}}{\text{VP} + \text{FN}} \tag{14}$$

3.2.4.4 F1-Score

O F1-Score corresponde à média harmônica entre a precisão e o *recall*. Seu melhor valor é 1, e o pior, 0.

$$\text{F1-Score} = 2 \cdot \left(\frac{\text{Precisão} \cdot \text{Recall}}{\text{Precisão} + \text{Recall}} \right) \tag{15}$$

3.2.4.5 Perda Logarítmica

A perda logarítmica, também chamada de entropia cruzada, é geralmente utilizada quando o modelo produz uma probabilidade para cada classe. Essa métrica avalia a qualidade das previsões e penaliza mais intensamente aquelas que atribuem elevada confiança a resultados

incorretos. Quanto mais próxima de zero estiver a perda, melhor será o desempenho do modelo segundo essa medida (Janocha; Czarnecki, 2016).

Para saídas não binárias, utiliza-se a entropia cruzada categórica, na qual as *labels* podem ser codificadas no formato *one-hot*. Nesse formato, cada categoria é representada por um vetor com número de elementos igual ao total de classes. Em um problema com três classes, por exemplo, os vetores poderiam ser $[1, 0, 0]$, $[0, 1, 0]$ e $[0, 0, 1]$ (Aggarwal, 2018).

Esses vetores podem ser associados a categorias compreensíveis para o usuário quando se deseja apresentar os resultados. Em uma situação com as classes cachorro, gato e pássaro, por exemplo, o programador pode associar o vetor $[1, 0, 0]$ a uma das classes e os demais vetores às outras categorias. A escolha da correspondência é arbitrária, mas deve ser mantida de forma consistente durante o treinamento e a utilização do modelo (Brownlee, 2020; Li *et al.*, 2018). Além da codificação *one-hot*, as *labels* podem ser representadas no formato esparso (*sparse*), no qual, em vez de um vetor, cada categoria recebe um único valor inteiro, como $0, 1, 2, 3, \ldots$ (Srivastava *et al.*, 2014; Sutskever *et al.*, 2013). Como o problema de classificação deste trabalho possui cinco classes, correspondentes aos escores de 1 a 5, a função de perda logarítmica é definida conforme apresentado a seguir:

a) n: número de observações;
b) m: número de classes possíveis;
c) y_{ij}: indicador binário que assume valor 1 quando a observação i pertence à classe j e 0 nos demais casos;
d) p_{ij}: probabilidade estimada pelo modelo de que a observação i pertença à classe j.

$$\text{Perda Logarítmica} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij}\log(p_{ij}) \tag{16}$$

Na equação, y_{ij} indica se a observação i pertence à classe j;
p_{ij} indica a probabilidade estimada de que a observação i pertença à classe j.

3.3 Ferramentas

Para a codificação e os testes do sistema, foi utilizado um computador com 16 gigabytes de memória de acesso aleatório (*Random Access Memory* — RAM) do tipo *Double Data Rate 4* (DDR4) de 3.200 MHz, processador Ryzen 5 3600, da Advanced Micro Devices (AMD), e placa de vídeo Gigabyte NVIDIA GeForce RTX 3070. A ferramenta Google Colab também foi

considerada como alternativa em situações que demandassem recursos adicionais.

3.3.1 Python

Para o desenvolvimento do sistema, foi escolhida a linguagem Python. A escolha deveu-se ao fato de Python ser uma linguagem simples, fácil de aprender e compreender e capaz de expressar conceitos em um número menor de linhas do que linguagens tradicionais, como C (Srinath, 2017). Além disso, a linguagem possui bibliotecas que facilitam a criação de modelos de redes neurais artificiais, como NumPy, scikit-learn, PyTorch e TensorFlow.

3.3.2 TensorFlow

TensorFlow é uma plataforma livre e de código aberto desenvolvida pelo Google e amplamente utilizada na implementação de redes neurais (Pang; Nijkamp; Wu, 2020). A ferramenta oferece operações para manipulação de tensores, diferenciação automática e construção e treinamento de modelos. Entre suas vantagens está a possibilidade de paralelizar o processamento e utilizar unidades de processamento gráfico (GPUs), o que pode reduzir o tempo necessário para o treinamento. Em placas compatíveis, também podem ser empregados recursos especializados para cálculos com tensores, como os *Tensor Cores*. Esses recursos são particularmente úteis em modelos que exigem grande quantidade de operações matriciais.

3.3.3 PyTorch

PyTorch é uma biblioteca de aprendizado profundo que alia usabilidade e desempenho. Lançada em 2016, destaca-se pelo estilo de programação imperativo e integrado ao Python, o que facilita a depuração e se integra bem com outras bibliotecas científicas. PyTorch oferece execução dinâmica de cálculos de tensores, diferenciação automática e aceleração por GPU, mantendo um desempenho semelhante às bibliotecas mais rápidas de aprendizado profundo. Essas características tornaram a biblioteca uma escolha popular na comunidade de pesquisa (Paszke *et al.*, 2019).

A versão 2 do PyTorch introduziu recursos como TorchDynamo e TorchInductor, destinados à compilação e à otimização dos grafos de execução. O TorchDynamo atua em nível de Python e analisa dinamicamente o *bytecode* para capturar partes do programa que podem ser

compiladas. O TorchInductor, por sua vez, traduz programas PyTorch para Triton, no caso de GPUs, ou para C++, no caso de unidades centrais de processamento (CPUs). Esses recursos permitem aplicar otimizações sem abandonar o estilo de programação imperativo característico do PyTorch e produziram melhorias de desempenho tanto na inferência quanto no treinamento (Ansel *et al.*, 2024).

3.3.4 Albumentations

Ao realizar o treinamento de uma rede neural, busca-se utilizar o maior número possível de informações para reduzir o *overfitting* e também para favorecer a generalização da rede para situações em que não foi treinada. Considerando-se que muitas vezes é difícil obter uma quantidade suficiente de imagens para o treinamento adequado de uma rede neural, os pesquisadores utilizam técnicas de *data augmentation*, que consistem na aplicação de filtros e modificações às imagens originais para gerar artificialmente um conjunto maior de imagens para o treinamento das redes.

A Albumentations é uma biblioteca que facilita a geração de novas imagens (Buslaev *et al.*, 2020). Por meio de filtros e modificações definidos pelo pesquisador, ela pode gerar um número teoricamente ilimitado de versões a partir de cada imagem original. Isso é realizado pela aplicação e pela combinação de pequenas modificações aleatórias ao conjunto original de dados.

3.3.5 Bibliotecas adicionais

Além das bibliotecas principais mencionadas, foram consideradas outras ferramentas voltadas à programação científica e à análise de dados. Entre elas destaca-se a NumPy, uma das bibliotecas mais utilizadas para computação numérica em Python. A biblioteca oferece uma estrutura de *arrays* multidimensionais, além de um amplo conjunto de operações matemáticas que facilitam a realização de cálculos numéricos com eficiência e precisão. Essas estruturas também servem de base para o funcionamento de diversas bibliotecas utilizadas em aprendizado de máquina e processamento de imagens.

Outra biblioteca relevante é a scikit-learn, amplamente reconhecida por sua vasta gama de algoritmos para *machine learning* e ferramentas de pré-processamento de dados. Essa biblioteca é essencial para a implementação de modelos de aprendizado supervisionado e não super-

visionado, oferecendo também recursos para avaliação de modelos, seleção de características e ajuste de hiperparâmetros, todos fundamentais para o desenvolvimento de soluções robustas e eficazes.

A biblioteca pandas também foi utilizada ocasionalmente, devido à sua utilidade no tratamento e na análise de dados estruturados. A pandas fornece estruturas de dados flexíveis e de alto desempenho, como DataFrame e Series, que são úteis para a manipulação eficiente de grandes volumes de dados e facilitam operações como filtragem, agregação e visualização.

3.4 Trabalhos Correlatos

O processo de seleção dos trabalhos correlatos ocorreu em etapas: inicialmente, foram lidos apenas os títulos e os *abstracts* dos trabalhos e, nessa etapa, foram eliminados aqueles que não possuíam relevância para a pesquisa proposta. Em seguida, foi realizada a leitura completa dos trabalhos remanescentes para determinar quais deveriam ser utilizados como referencial teórico.

Foram encontrados 32 trabalhos nos repositórios pesquisados: 14 no ScienceDirect[5], 17 no Scopus[6], um no IEEE Xplore[7], nenhum no Engineering Village[8] e nenhum no Acesso Livre à Informação Científica da Embrapa (ALICE)[9]. Na etapa de leitura dos títulos e dos *abstracts*, alguns trabalhos foram descartados por tratarem de assuntos sem relação com esta dissertação ou por não atenderem aos critérios de inclusão e exclusão da revisão. Após essa etapa, restaram nove artigos para leitura integral. Entre os nove trabalhos lidos completamente, apenas quatro cumpriram todos os critérios estabelecidos para a presente pesquisa.

O Quadro 3 apresenta uma visão geral dos trabalhos analisados, incluindo o contexto de aplicação, o problema abordado e o modelo empregado na solução. A Figura 21 apresenta um diagrama do processo de revisão.

Mehedi Shamrat *et al.* (2024) utilizaram 15 arquiteturas pré-treinadas e compararam seus resultados com os de um modelo próprio para classificar imagens de retina e reconhecer o nível de danos causados pela retinopatia diabética. Segundo os autores, foi aplicada uma abordagem metodológica mista, combinando informações qualitativas provenientes da pesquisa médica com a capacidade quantitativa dos algoritmos de aprendizagem profunda.

[5]https://www.sciencedirect.com
[6]https://www.scopus.com
[7]https://ieeexplore.ieee.org/
[8]https://www.engineeringvillage.com/search/quick.url
[9]https://www.alice.cnptia.embrapa.br/

Quadro 3 – Trabalhos correlatos selecionados.

Autor	Aplicação	Modelo
Mehedi Shamrat *et al.* (2024)	Reconhecimento de danos em imagens de fundo de retina	Baseado em redes neurais (DarkNet13)
Ben Abdallah, Elleuch e Guermazi (2021)	Identificar expressões faciais e classificar a situação de alunos	Baseado em redes neurais (VGG-16 modificado)
García-García *et al.* (2023)	Identificar pontos orofaciais	Duas redes neurais modificadas: InceptionResNetV2 (IRNet) e MobileNetV2 (MNet)
Sheet *et al.* (2022)	Classificar doenças de retina	Filtro de equalização adaptativa de histograma com limitação de contraste (CLAHE) e aprendizagem por transferência em rede neural (ResNet50)

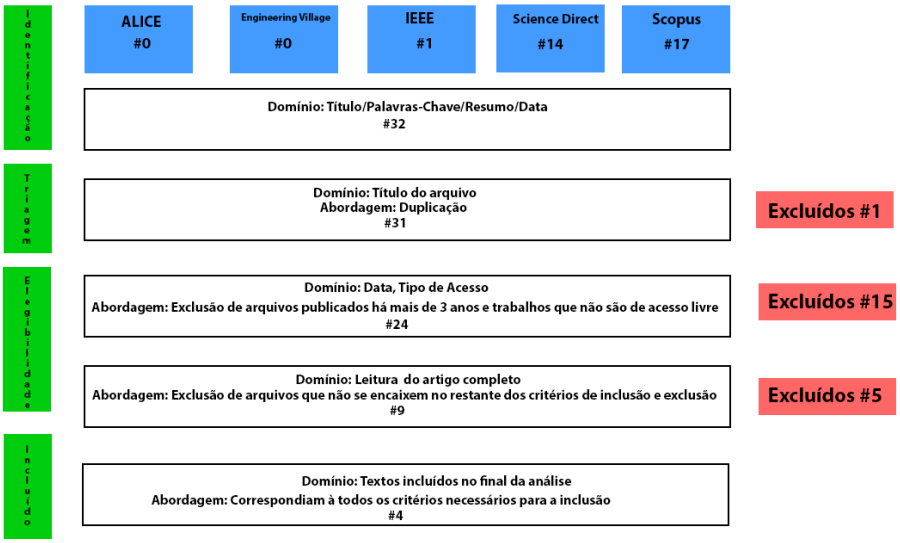
Fonte: Autor (2024).

O uso dessa abordagem mista permitiu superar limitações metodológicas observadas em pesquisas anteriores, caracterizadas pela dependência da extração manual de detalhes e informações e pelo uso de algoritmos de aprendizado de máquina com baixa capacidade de adaptação. Durante os testes, os pesquisadores desenvolveram uma RNA denominada DarkNet13, com o objetivo de melhorar os resultados de pesquisas anteriores na área e alcançar um equilíbrio mais adequado entre acurácia e demanda de recursos computacionais.

Após a realização dos testes, os pesquisadores concluíram que a DarkNet13 apresentou resultados superiores aos dos modelos utilizados como referência no estudo. A arquitetura demonstrou maior capacidade de reconhecer detalhes nas imagens de validação e apresentou melhora de acurácia, simultaneamente à redução do tempo de processamento. Segundo os autores, a combinação desses fatores indica maior eficiência no uso dos recursos computacionais disponíveis. Esses resultados evidenciam o potencial da DarkNet13 para aplicações nas quais a precisão e a economia de recursos computacionais sejam relevantes, especialmente em tarefas de visão computacional e aprendizado profundo.

Ben Abdallah, Elleuch e Guermazi (2021) modificaram a arquitetura VGG-16 para identificar expressões faciais de estudantes e, a partir delas, classificar se os alunos compreendiam ou não o conteúdo apresentado em aula. Os autores organizaram a abordagem em três etapas e utilizaram *transfer learning*. Inicialmente, foi empregado um conjunto de dados composto por vídeos de estudantes, no qual uma rede neural, em conjunto com a biblioteca dlib, detectou e rastreou os rostos em cada *frame*. Em seguida, os rostos detectados foram recortados dos *frames*, redimensionados para 48 × 48 *pixels* e convertidos para escala de cinza. Nessa etapa, as

Figura 21 – Diagrama da revisão sistemática.



Fonte: Autor (2024).

expressões faciais foram reconhecidas por meio de uma versão modificada da VGG-16.

Após o treinamento da rede para reconhecer as expressões, utilizou-se o *transfer learning* para adaptar o modelo a um conjunto menor de vídeos, por meio de um procedimento semelhante ao da primeira etapa. As imagens dos rostos passaram novamente pela rede, que foi modificada para classificar as expressões nas categorias *entendendo* e *não entendendo*.

De acordo com os autores, a análise comparativa revelou que a versão modificada da VGG-16, quando combinada com técnicas de *transfer learning*, apresentou desempenho superior ao das configurações originais da VGG-16 e da VGG-19. Esse resultado foi obtido mesmo diante de um conjunto de dados de entrada consideravelmente limitado. A comparação sugere que a abordagem adotada contribuiu para melhorar o processo de aprendizado e a capacidade de generalização do modelo. Segundo os pesquisadores, a integração de modificações arquitetônicas específicas à rede, juntamente com o reaproveitamento de pesos previamente treinados, permitiu obter melhores resultados sem a necessidade de um conjunto de dados de grande volume. Assim, a estratégia pode constituir uma alternativa útil para problemas semelhantes em que a disponibilidade de imagens rotuladas seja limitada.

García-García *et al.* (2023) utilizaram versões modificadas de duas redes neurais convolucionais profundas, InceptionResNetV2 (IRNet) e MobileNetV2 (MNet), para identificar pontos orofaciais em fotografias pré-operatórias, determinar sua visibilidade e estimar suas coordenadas bidimensionais. Esses modelos foram treinados de maneira supervisionada com o consenso entre duas anotações de anestesiologistas humanos utilizado como referência. A pes-

quisa realizou etapas sucessivas de *transfer learning* em ordem crescente de similaridade com relação ao conjunto de dados DiffAirW.

Os modelos foram avaliados por meio de validação cruzada estratificada por sexo, a fim de garantir uma avaliação justa e robusta do desempenho em imagens não vistas durante o treinamento. Os resultados foram comparados às discrepâncias entre os anotadores, adotadas como padrão de referência, e ao desempenho de cinco modelos deformáveis considerados de estado da arte. De maneira geral, o modelo IRNet apresentou capacidade de generalização satisfatória, alcançando desempenho comparável ao de especialistas humanos, enquanto o modelo MNet apresentou apenas uma ligeira redução de desempenho. Na vista lateral, ambas as redes tiveram um desempenho estatisticamente inferior aos humanos, embora com tamanhos de efeito pequenos (não significativos para IRNet). Esse desempenho inferior nas imagens em vista lateral já havia sido descrito na literatura por outros autores, que utilizaram técnicas distintas, e pode estar relacionado à dificuldade intrínseca de distinguir visualmente determinadas características, mesmo por especialistas treinados. Para o melhor modelo, não foi encontrado viés de desempenho em relação ao sexo.

Sheet *et al.* (2022) apresentam um estudo voltado à detecção de doenças retinianas por meio de imagens de fundo de olho. A pesquisa propõe um modelo automatizado para a identificação de doenças da retina, com o objetivo de reduzir a intervenção manual e manter altos níveis de precisão na classificação das imagens. O estudo utilizou imagens de fundo de olho obtidas do site *Structured Analysis of the Retina* (STARE), classificadas em cinco categorias com base em anotações de especialistas. As imagens foram processadas por meio de um filtro de equalização adaptativa de histograma com limitação de contraste (CLAHE), com o objetivo de realçar as imagens, reduzir o ruído e normalizar o espectro de intensidade. O modelo proposto foi desenvolvido por meio de aprendizado por transferência, utilizando como base a rede pré-treinada ResNet50, que foi ajustada para a detecção de doenças retinianas. Os resultados da avaliação do modelo indicaram 100% de sensibilidade, 100% de especificidade e 100% de precisão. Os autores concluem que o modelo proposto apresenta desempenho significativamente superior ao de estudos anteriores e pode constituir uma ferramenta útil para auxiliar oftalmologistas na identificação rápida e precisa de doenças retinianas.

A análise dos trabalhos correlatos fornece uma base teórica e metodológica para o desenvolvimento desta pesquisa, ao mesmo tempo que permite identificar os aspectos que a distinguem dos estudos anteriores. Os esforços empregados na seleção de metodologias e tecnologias relevantes, como o uso de redes neurais convolucionais e de técnicas de aprendizado por transferência, destacam-se como componentes importantes para o avanço da pesquisa na área de

computação aplicada. A escolha de modelos específicos, como DarkNet13 e ResNet50, e de arquiteturas utilizadas como base nos trabalhos analisados, como VGG-16, VGG-19, Inception-ResNetV2 e MobileNetV2, ilustra a importância de adaptar e aperfeiçoar abordagens existentes para explorar seu potencial em novos contextos. A revisão também evidencia a necessidade de considerar limitações como a dependência de conjuntos de dados reduzidos e a escolha de modelos que não estejam suficientemente alinhados aos objetivos específicos da pesquisa. Dessa forma, as lições extraídas dos trabalhos correlatos orientam as escolhas metodológicas deste estudo e contribuem para evitar abordagens inadequadas, mantendo o desenvolvimento da solução fundamentado em práticas já avaliadas na literatura.

4 ABORDAGEM PROPOSTA

Esta seção apresenta a abordagem utilizada para tratar o problema de pesquisa, detalhando as etapas percorridas ao longo do estudo. Esse processo incluiu a preparação da base de imagens, a implementação dos métodos e, por fim, a análise e a comparação dos resultados obtidos. A preparação compreendeu a seleção, a organização e o aumento artificial das imagens; a implementação envolveu o teste das arquiteturas previstas; e a etapa final reuniu a avaliação das métricas, a discussão das limitações e a construção do protótipo móvel.

4.1 Preparação da base de dados

Como já explicitado na Seção 2, inicialmente foram realizadas a captura das imagens e a catalogação manual dos códigos de identificação presentes nos brincos dos animais. As informações obtidas por meio da análise das imagens foram organizadas em uma tabela e enviadas à equipe do programa de melhoramento genético, que retornou os escores dos animais identificados.

Nos testes preliminares, foram identificados alguns fatores que dificultavam o treinamento adequado dos modelos de redes neurais. Entre eles, destacaram-se o desequilíbrio entre as classes de escore, decorrente do pequeno número de animais em algumas delas, e a impossibilidade de localizar o olho em imagens com desfoque excessivo. Também houve casos em que o animal apresentou comportamento evasivo durante a captura e teve a região ocular encoberta pelo corpo de outro bovino. Além disso, alguns modelos exigiam dimensões específicas de entrada, enquanto as fotografias haviam sido obtidas em diferentes resoluções. Esses fatores evidenciaram a necessidade de selecionar, redimensionar e organizar adequadamente os dados destinados ao treinamento, à validação e ao teste.

Para mitigar os problemas decorrentes das limitações do estudo, foram adotadas estratégias de tratamento e seleção das imagens. Excluíram-se aquelas nas quais nem mesmo um avaliador humano conseguiria localizar o olho ou atribuir uma classificação, bem como as imagens em que não foi possível identificar o animal, por exemplo, devido à ilegibilidade ou à ausência de visibilidade do brinco. As Figuras 22, 23 e 24 apresentam exemplos de imagens consideradas inadequadas para o treinamento da rede neural. Os exemplos representam problemas distintos: ocultação da região ocular, ângulo que dificulta a visualização lateral do olho e perda de nitidez provocada pelo movimento do animal durante a captura.

Também foi realizada uma etapa de *data augmentation* com a biblioteca Albumentati-

Figura 22 – Animal escondendo o rosto.



Fonte: Autor (2024).

Figura 23 – Imagem capturada em ângulo frontal.



Fonte: Autor (2024).

ons, com o objetivo de equilibrar a quantidade de dados disponível em cada classe da base de imagens. Inicialmente, os animais foram distribuídos entre os conjuntos de treinamento, vali-

Figura 24 – Imagem borrada em razão do movimento do animal.



Fonte: Autor (2024).

dação e teste, evitando-se que imagens de um mesmo animal fossem incluídas em mais de um conjunto. Essa ordem foi importante para impedir que versões transformadas de uma mesma fotografia original fossem distribuídas entre treinamento, validação e teste. Após essa divisão, utilizou-se *data augmentation* para gerar aproximadamente 50 mil imagens de cada classe disponível no conjunto de treinamento, totalizando cerca de 200 mil imagens. Para os conjuntos de validação e teste, foram geradas as quantidades necessárias para que cada classe tivesse pouco mais de 5 mil imagens, totalizando aproximadamente 20 mil imagens em cada conjunto.

Como já explicado nesta dissertação, não havia animais com escore 1 entre os fotografados e havia apenas 4 animais com escore 2, o que pode limitar a capacidade do modelo de reconhecer essas classes. O aumento artificial do número de imagens não cria, por si só, a diversidade biológica que seria obtida com novos animais. Para aprimorar o modelo de modo a incluir a classe ausente e melhorar o desempenho na classe sub-representada, será necessário obter novas imagens, priorizando as classes extremas, e realizar uma nova etapa de treinamento. Para isso, pode-se utilizar o método YOLO.train() da biblioteca Ultralytics, tomando os pesos previamente aprendidos como ponto de partida. Esse procedimento permite continuar o treinamento com os novos dados, em vez de descartar integralmente o aprendizado obtido na etapa anterior.

O número aproximado de 50 mil imagens por classe de escore foi escolhido na tentativa de equilibrar os recursos computacionais disponíveis e a necessidade de fornecer uma quantidade adequada de dados para o treinamento. Um volume excessivo exige maior poder computacional, mais espaço de armazenamento e maior tempo de processamento, enquanto um volume insuficiente pode não oferecer variação suficiente para o aprendizado do modelo. Considerando

os recursos e o tempo disponíveis durante o estudo, o total aproximado de 200 mil imagens foi considerado um compromisso adequado para a realização dos experimentos, sem a pretensão de que essa quantidade eliminasse as limitações decorrentes do pequeno número de animais em algumas classes.

4.2 Implementação

A implementação foi planejada em duas fases principais: a detecção do olho do animal e a classificação do escore de pigmentação. A revisão de literatura indicou que o aprendizado por transferência (*transfer learning*) é frequentemente empregado em tarefas de detecção e classificação de imagens, razão pela qual essa estratégia foi adotada para reduzir o tempo necessário ao treinamento. Para a detecção dos olhos, foram testados modelos pré-treinados como ResNet50, VGG16, Faster R-CNN e YOLOv8, além do método utilizado por Lorensi (2024) em sua dissertação. Pretendia-se avaliar o desempenho dessas abordagens e determinar qual delas seria mais adequada à finalidade específica deste estudo. A comparação planejada consideraria tanto a capacidade de localizar a região ocular quanto a viabilidade de executar o treinamento com os recursos disponíveis. A seleção dos modelos e métodos baseou-se em sua utilização prévia em tarefas de detecção e no objetivo de explorar diferentes aplicações de redes neurais artificiais.

Na fase seguinte, a região identificada como correspondente ao olho seria recortada e processada por uma segunda rede neural artificial (RNA), responsável especificamente pela classificação do escore de pigmentação. Essa divisão permitiria avaliar separadamente o desempenho do detector e do classificador, além de substituir um dos componentes sem necessariamente alterar o outro. Nessa etapa, também se pretendia utilizar aprendizado por transferência e modelos das famílias VGG e ResNet, entre outros anteriormente citados. No entanto, como detalhado na Seção 4.3, somente o modelo YOLOv8 foi empregado na solução final. Como esse modelo realiza a localização e a classificação em uma única execução, a separação inicialmente planejada em duas redes deixou de ser necessária. Essa decisão decorreu das limitações encontradas durante a implementação e dos resultados obtidos nos testes preliminares.

Reconhecendo que o processo de pesquisa científica é dinâmico e pode exigir ajustes durante sua execução (Lakatos; Marconi, 2021; Marconi; Lakatos, 2021a, 2021b; Pereira, 2011), manteve-se a possibilidade de incorporar ou excluir modelos dos testes finais. Essa flexibilidade permitiu a inclusão do YOLOv8, que não estava previsto nas etapas iniciais. Também se optou por não prosseguir com os modelos que não puderam ser treinados adequadamente ou que não produziram resultados utilizáveis no ambiente e no período disponíveis para a pesquisa.

4.3 Análise e comparação dos resultados

Após diversas tentativas, não foram obtidos resultados utilizáveis na tarefa de detecção dos olhos com ResNet50, VGG16, Faster R-CNN ou com o método empregado por Lorensi (2024). Os problemas observados variaram desde a incapacidade de localizar a área de interesse até falhas de execução em momentos aparentemente aleatórios do treinamento. Como os processos não foram concluídos de forma equivalente, seus resultados parciais não permitiam uma comparação estatística válida com o treinamento concluído do YOLOv8.

Alguns modelos não puderam ser treinados devido à dependência de combinações específicas de versões de diferentes bibliotecas. Em determinados casos, uma versão necessária para uma biblioteca entrava em conflito com a versão exigida por outro componente do ambiente. Em outros, ocorreram falhas de execução em etapas distintas do treinamento, sem que fosse possível reproduzi-las de maneira consistente. No ambiente utilizado, esses problemas não foram solucionados durante o período disponível para a pesquisa. A investigação das causas de cada intercorrência exigiria a reconstrução dos ambientes, testes controlados e estudos específicos para cada modelo; por isso, evitou-se formular explicações especulativas no escopo desta dissertação.

Com o modelo pré-treinado YOLOv8, foi possível localizar a área de interesse em mais de 96% das instâncias do conjunto de validação após 15 épocas, com tempo de processamento aproximado de 7,6 horas. O método de treinamento estava configurado inicialmente para 100 épocas, mas o comportamento das métricas foi acompanhado durante a execução. A partir de aproximadamente oito épocas, as perdas do conjunto de treinamento continuaram diminuindo, enquanto algumas perdas de validação passaram a aumentar. Esse afastamento entre os resultados dos dois conjuntos sugeriu a ocorrência de sobreajuste, isto é, uma melhora no ajuste aos dados de treinamento sem ganho correspondente na validação. Por essa razão, considerou-se desnecessário e potencialmente prejudicial prolongar o treinamento até o número inicialmente configurado.

A utilização de pesos pré-treinados pode ter contribuído para a obtenção dos resultados em um número reduzido de épocas. Além disso, o YOLOv8 realiza a localização da área de interesse e a atribuição da classe em uma única etapa, eliminando a necessidade de empregar uma segunda rede e simplificando o fluxo de processamento.

Os conjuntos de treinamento, validação e teste eram independentes quanto aos animais neles representados, de modo que imagens de um mesmo bovino não fossem distribuídas entre conjuntos diferentes. Os pesos do modelo foram ajustados com base nas imagens de treina-

mento. O conjunto de validação, não utilizado diretamente para esse ajuste, serviu para acompanhar o comportamento das métricas durante o treinamento, identificar indícios de sobreajuste e auxiliar na escolha dos pesos a serem utilizados. O conjunto de teste foi empregado somente após a conclusão do treinamento, para avaliar a inferência em imagens derivadas de animais não utilizados nas etapas anteriores. Essa separação buscou reduzir o risco de que o modelo fosse avaliado com imagens excessivamente semelhantes às empregadas no ajuste dos pesos.

Na Figura 25, são apresentados os resultados obtidos pelo YOLOv8 após 15 épocas. Considerando todas as classes, o modelo alcançou precisão de 0,424, *recall* de 0,385, *mean Average Precision* (mAP) de 0,318 no limiar de 0,50 (mAP50) e de 0,230 na média dos limiares de 0,50 a 0,95 (mAP50–95). Na classe 2, todas essas métricas foram iguais a zero, indicando que nenhuma instância dessa classe foi corretamente reconhecida. A classe 1 não aparece nos resultados porque não havia animais com esse escore no conjunto de dados. Entre as demais classes, o escore 3 apresentou precisão de 0,294 e *recall* de 0,554; o escore 4, precisão de 0,438 e *recall* de 0,828; e o escore 5, precisão de 0,962 e *recall* de 0,159. Esses valores mostram que uma classe pode apresentar poucas previsões incorretas e, simultaneamente, deixar de recuperar grande parte das instâncias existentes.

Figura 25 – Métricas de validação do YOLOv8 após 15 épocas.

```
15 epochs completed in 7.585 hours.
Optimizer stripped from results\train\weights\last.pt, 6.2MB
Optimizer stripped from results\train\weights\best.pt, 6.2MB

Validating results\train\weights\best.pt...
Ultralytics YOLOv8.2.83 🚀 Python-3.12.5 torch-2.4.0+cu124 CUDA:0 (NVIDIA GeForce RTX 3070, 8192MiB)
Model summary (fused): 168 layers, 3,006,428 parameters, 0 gradients, 8.1 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 628/628 [02:15<00:00,  4.62it/s]
                   all      20082      20082      0.424      0.385      0.318       0.23
                     2       5005       5005          0          0          0          0
                     3       5019       5019      0.294      0.554       0.26      0.185
                     4       5018       5018      0.438      0.828      0.452      0.297
                     5       5040       5040      0.962      0.159      0.561      0.438
Speed: 0.1ms preprocess, 1.2ms inference, 0.0ms loss, 1.5ms postprocess per image
Results saved to results\train
2024-08-30 20:48:02,411 - INFO - Treinamento concluído.

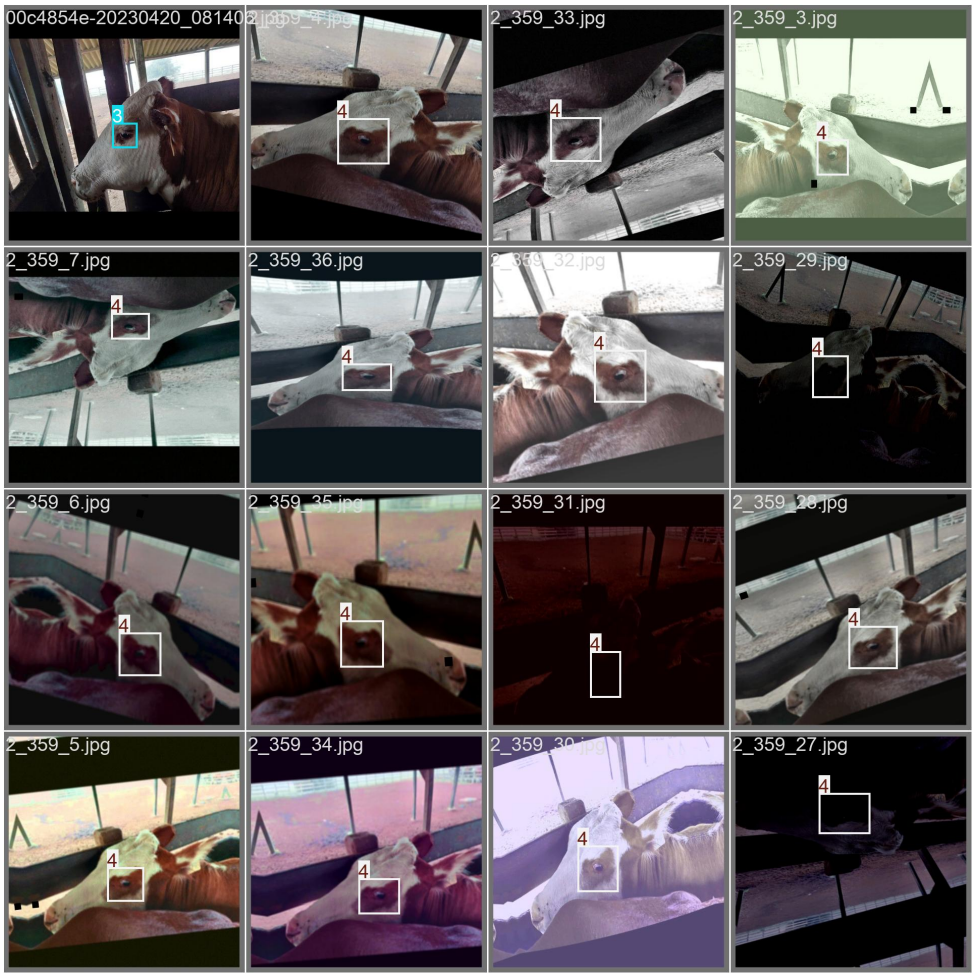
Process finished with exit code 0
```

Fonte: Autor (2024).

As Figuras 26 e 27 foram geradas automaticamente pelo YOLOv8 a partir do mesmo lote de imagens de validação. A Figura 26 apresenta as caixas delimitadoras e os escores anotados como referência, enquanto a Figura 27 apresenta as caixas e as classes previstas pelo modelo, acompanhadas dos respectivos níveis de confiança. A comparação visual permite verificar se a região ocular foi localizada e observar as diferenças entre as anotações de referência e as previsões. Nesse lote específico, a maior parte dos olhos foi localizada, mas também é possível observar classificações divergentes do escore anotado, o que é compatível com as limitações

indicadas pelas métricas agregadas.

Figura 26 – Conjunto de imagens com suas anotações de referência.

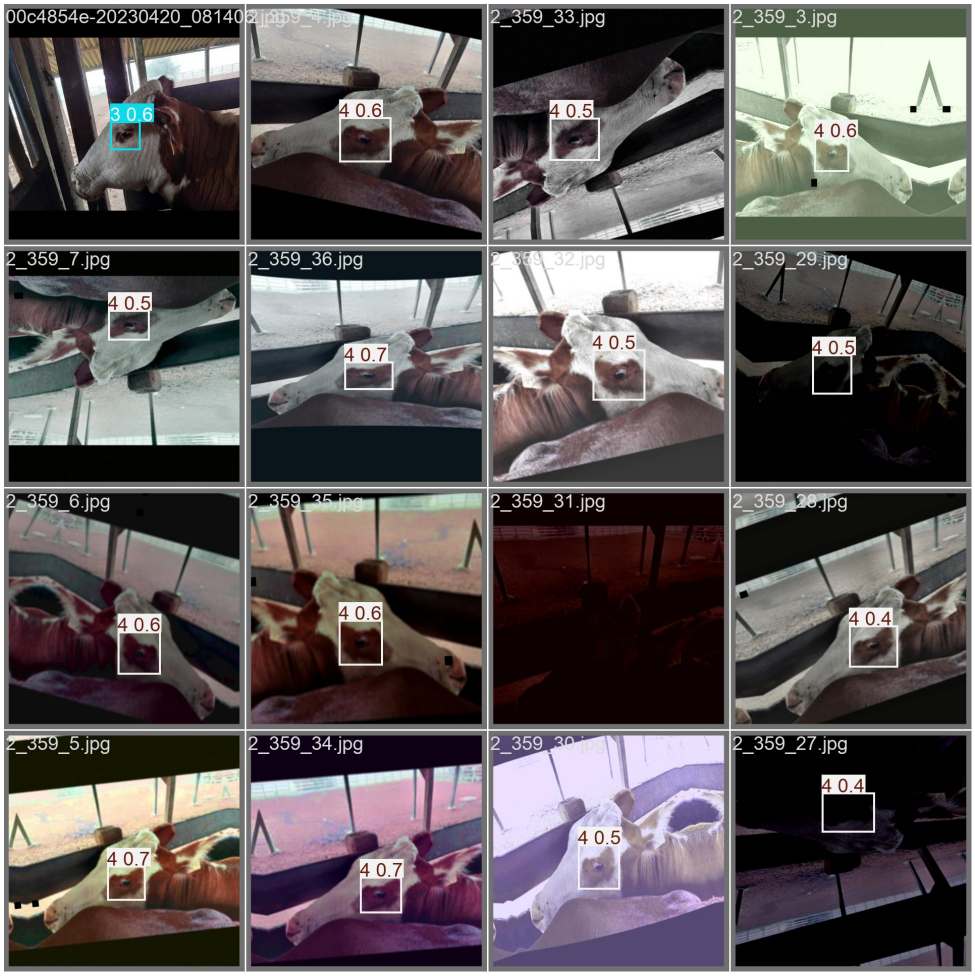


Fonte: Autor (2024).

A Figura 28 apresenta a matriz de confusão gerada com os dados de validação ao final do treinamento. A linha e a coluna denominadas *background* representam anotações ou detecções sem correspondência e não constituem uma classe existente nas anotações originais. A matriz indica que pouco mais de 3,6% das instâncias anotadas não tiveram a área de interesse localizada. Também evidencia a incapacidade do modelo de reconhecer corretamente o escore 2: aproximadamente 96% das instâncias dessa classe foram classificadas como escore 3. Em contraste, 79,8% das instâncias com escore 4 foram corretamente classificadas.

Esses resultados evidenciam a importância de um conjunto de treinamento mais robusto e da utilização de imagens dos dois olhos do animal para a inferência do escore. Havia apenas 4 animais com escore 2: 2 foram destinados ao treinamento, 1 à validação e 1 ao teste. Embora tenha sido aplicado *data augmentation*, essa técnica não substitui a diversidade proporcionada por um número maior de animais em cada classe. Observou-se ainda que os animais com escore

Figura 27 – Inferências do YOLOv8 para o conjunto de imagens da Figura 26.

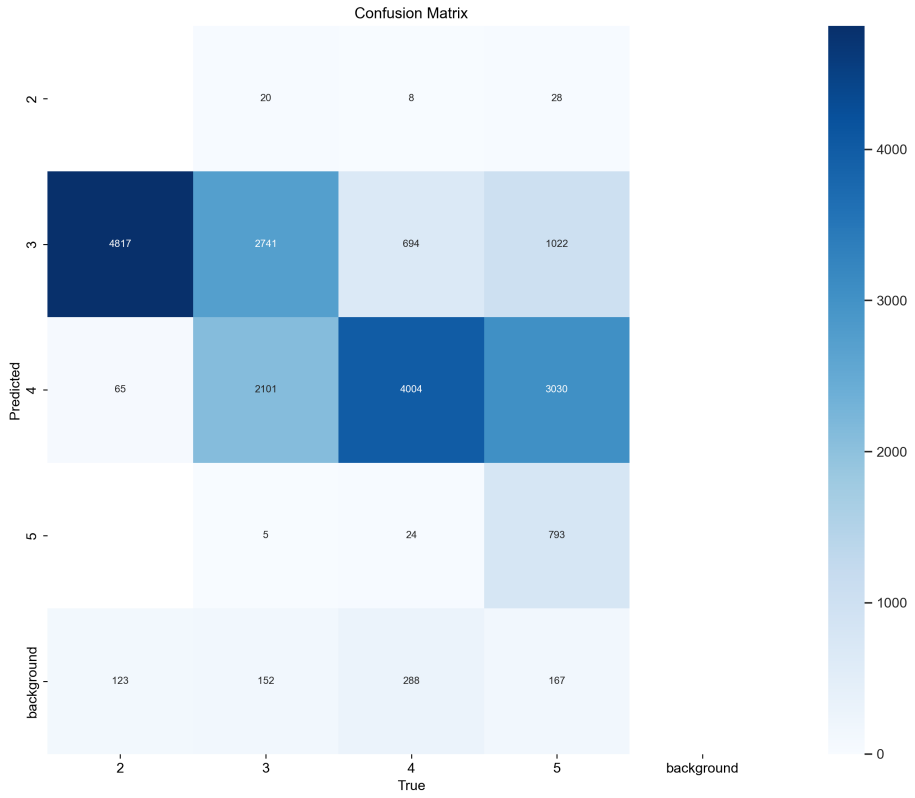


Fonte: Autor (2024).

4 foram mais bem reconhecidos do que aqueles com escore 3. Dos 80 animais com escore 4, 64 foram utilizados no treinamento; entre os 113 animais com escore 3, 91 foram destinados a esse conjunto.

Uma possível explicação é que animais com escore 3 podem apresentar níveis distintos de pigmentação entre os olhos. Assim, uma imagem que mostre apenas um deles pode se assemelhar visualmente a um escore maior ou menor, dependendo da pigmentação do olho fotografado. Essa condição introduz uma diferença entre o escore atribuído pela avaliação conjunta do animal e a informação efetivamente disponível em uma fotografia unilateral. Caso essa hipótese seja confirmada, a utilização de um conjunto mais amplo, com maior número de animais em todas as classes e imagens de ambos os olhos, poderá melhorar os resultados de inferência. Essa interpretação é compatível com o desempenho do escore 4, corretamente reconhecido em 79,8% das instâncias, pois os animais dessa classe apresentam níveis elevados de pigmentação nos dois

Figura 28 – Matriz de confusão da classificação realizada pelo YOLOv8.



Fonte: Autor (2024).

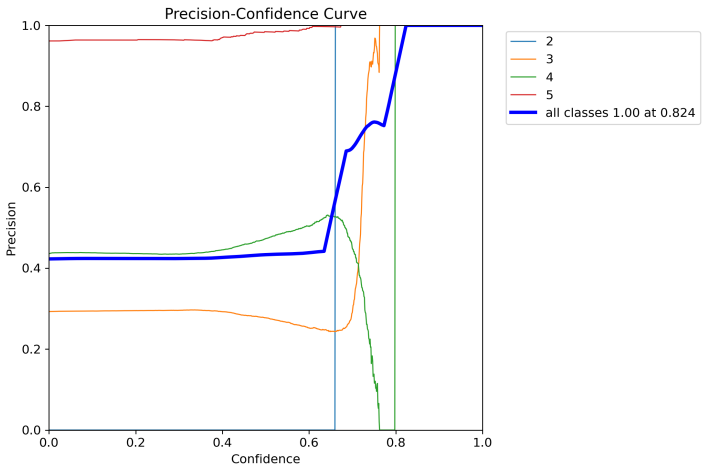
olhos, embora a pigmentação total em ambos corresponda ao escore 5. Ainda assim, a hipótese não pode ser confirmada apenas pelos resultados deste experimento e deverá ser avaliada em trabalhos futuros com uma coleta especificamente planejada para essa finalidade.

A Figura 29 apresenta a relação entre a precisão e o limiar de confiança no conjunto de validação. Em limiares baixos, a precisão agregada permaneceu próxima de 0,42. À medida que o limiar aumentou, a precisão das previsões remanescentes também aumentou, alcançando 1,00 no limiar de 0,824, embora com redução correspondente no número de detecções mantidas.

A Figura 30 apresenta a relação entre o *recall* e o limiar de confiança. A curva agregada atingiu *recall* máximo de aproximadamente 0,39 no limiar de confiança 0 e diminuiu à medida que esse limiar foi elevado. Entre as classes, o escore 4 apresentou o maior *recall*, aproximadamente 0,83, seguido do escore 3, com cerca de 0,55. O escore 5 apresentou aproximadamente 0,16, enquanto o escore 2 permaneceu em zero. Esses resultados reforçam a influência do desequilíbrio e da baixa diversidade de animais nas classes menos representadas.

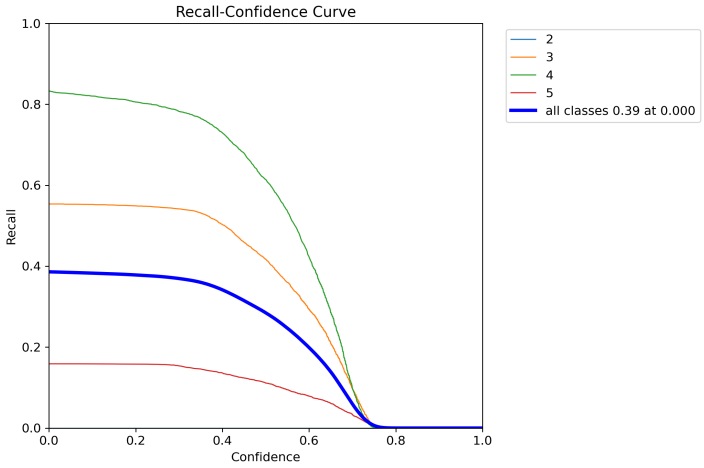
A Figura 31 apresenta a relação entre precisão e *recall*. Considerando todas as classes, o valor de mAP50 foi 0,318. A classe 5 obteve o maior valor, 0,561, seguida pelas classes 4, com 0,452, e 3, com 0,260. A classe 2 apresentou valor igual a zero.

Figura 29 – Gráfico de precisão por confiança do YOLOv8.



Fonte: Autor (2024).

Figura 30 – Gráfico de *recall* por confiança do YOLOv8.

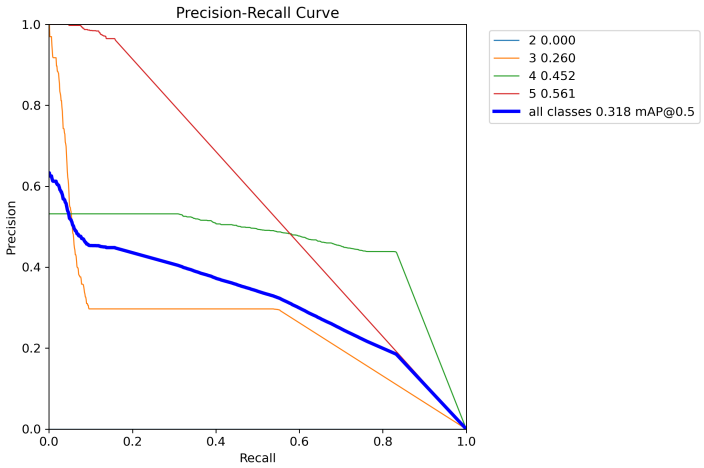


Fonte: Autor (2024).

A Figura 32 apresenta a métrica F1, correspondente à média harmônica entre precisão e *recall*. O valor agregado máximo foi de aproximadamente 0,31, obtido com limiar de confiança de 0,032. A classe 4 apresentou os maiores valores ao longo da maior parte da curva, enquanto a classe 2 permaneceu sem resultados. O valor agregado diminuiu progressivamente com o aumento do limiar de confiança e aproximou-se de zero em torno de 0,75.

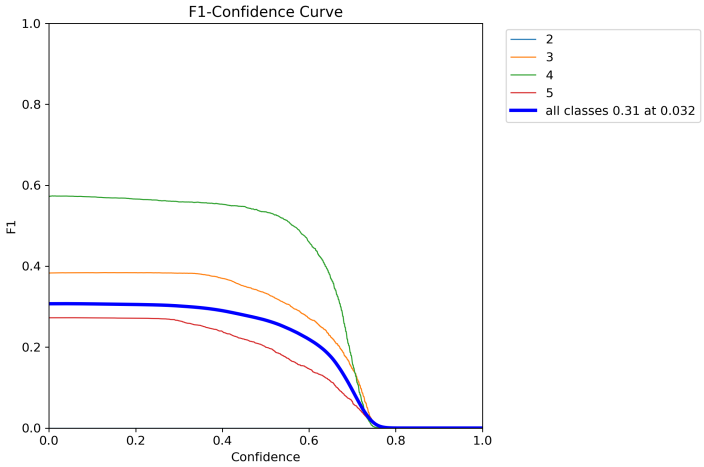
A Figura 33 reúne dez gráficos gerados automaticamente pelo YOLOv8 ao longo das 15 épocas. São apresentadas as perdas relacionadas às caixas delimitadoras, à classificação e à distribuição focal, tanto no treinamento quanto na validação, além das curvas de precisão, *recall*, mAP50 e mAP50–95. A métrica mAP50 considera um limiar de interseção sobre união (IoU)

Figura 31 – Gráfico de precisão por *recall* do YOLOv8.



Fonte: Autor (2024).

Figura 32 – Gráfico de F1 por confiança do YOLOv8.

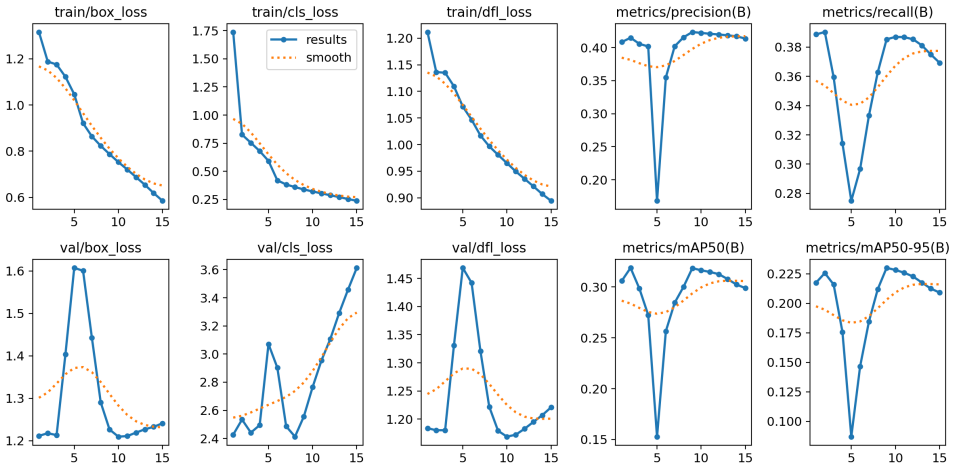


Fonte: Autor (2024).

de 0,50, enquanto a mAP50–95 corresponde à média calculada em limiares de 0,50 a 0,95.

Para avaliar a qualidade da inferência, foram utilizadas aproximadamente 10% das imagens no conjunto de teste. Ao todo, o teste incluiu 20.093 imagens obtidas por meio de *data augmentation* aplicado às imagens previamente destinadas a esse conjunto. Como os animais haviam sido separados antes da geração das imagens aumentadas, nenhum bovino do teste estava presente nos conjuntos de treinamento ou validação. Devido ao volume de amostras, não seria viável apresentar individualmente todas as imagens e previsões no texto da dissertação. A Figura 34 apresenta uma dessas imagens após a localização do olho, o desenho da caixa delimitadora, a classificação do escore e a indicação do nível de confiança. A inspeção dos resultados indicou que a rede localizou corretamente o olho em mais de 90% das imagens de teste. Os

Figura 33 – Gráficos de métricas gerados pelo modelo YOLOv8.



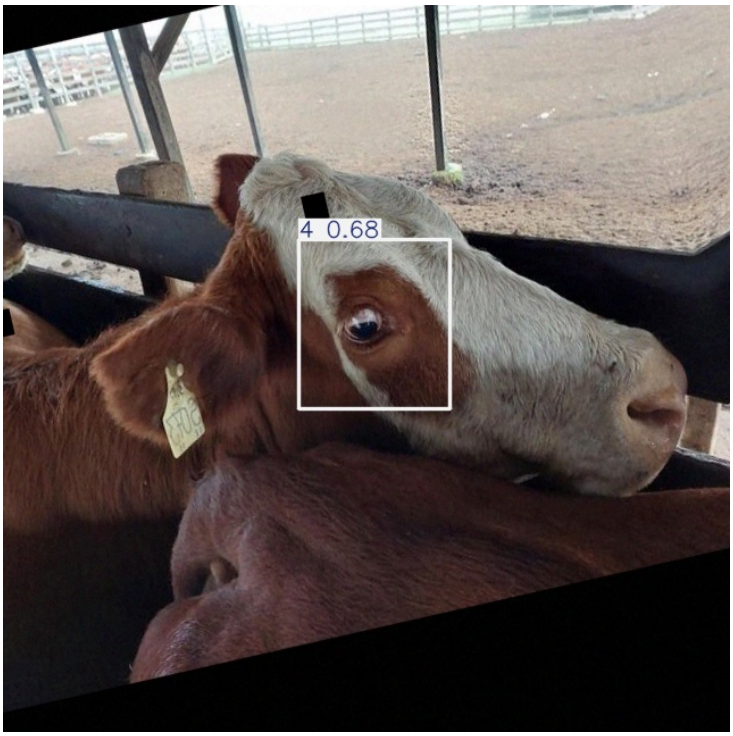
Fonte: Autor (2024).

erros de localização ocorreram principalmente em imagens excessivamente distorcidas pelas transformações de *data augmentation*, algumas das quais dificultavam a identificação do olho até mesmo por um observador humano. O desempenho da classificação permaneceu próximo ao observado na validação, em torno de 40% quando consideradas conjuntamente as classes, com resultados mais favoráveis para as classes 3, 4 e 5. Esse comportamento reforça que a localização da região ocular foi mais consistente do que a atribuição correta do escore.

O objetivo inicial do trabalho era testar a viabilidade de diferentes redes neurais na detecção e na classificação da região ocular de bovinos e comparar estatisticamente seus resultados. Entretanto, como apenas o YOLOv8 produziu resultados utilizáveis no ambiente e no período disponíveis, não foi possível realizar a comparação entre modelos originalmente planejada. Em vez de apresentar uma comparação baseada em execuções incompletas ou em resultados não equivalentes, optou-se por concentrar a etapa final na demonstração prática do modelo treinado. Diante disso, foi desenvolvido um protótipo de aplicativo para dispositivos móveis Android, utilizando a rede treinada para a classificação automatizada do escore de pigmentação ocular em bovinos.

Para a construção do protótipo, utilizou-se o Android Studio em ambiente amd64, com ambiente de execução 17.0.10+0–11572160. O modelo YOLOv8n foi convertido para o formato TensorFlow Lite (TFLite) por meio do método YOLO.export() da biblioteca Ultralytics e incorporado ao aplicativo para realizar inferências sobre as imagens capturadas. Ao ser iniciado, o aplicativo aciona a câmera do dispositivo; ao apontá-la para o rosto de um animal, o sistema desenha uma caixa delimitadora sobre o olho detectado e apresenta a classificação do escore. O protótipo demonstra a execução do modelo em um dispositivo móvel e mantém a possibili-

Figura 34 – Predição correta do escore 4, com 68% de confiança.



Fonte: Autor (2024).

dade de desenvolvimento de uma versão ampliada, com funções adicionais e uma interface mais completa.

5 CONCLUSÃO

Esta dissertação teve como objetivo desenvolver um método baseado em inteligência artificial (IA), com ênfase em aprendizado profundo (*deep learning*) e redes neurais convolucionais (CNNs), para a classificação automática do escore de pigmentação ocular em bovinos, buscando fornecer subsídios ao processo de melhoramento genético.

O objetivo geral foi alcançado em caráter de prova de conceito. Foi desenvolvida uma solução baseada no modelo YOLOv8 capaz de localizar a região ocular e atribuir uma classe de escore de pigmentação, posteriormente integrada a um protótipo de aplicativo Android. Os resultados demonstraram a viabilidade técnica da abordagem, mas também evidenciaram limitações que ainda impedem sua utilização como substituto da avaliação realizada por especialistas.

Quanto aos objetivos específicos, foi construído um conjunto de dados de imagens rotuladas para o treinamento dos modelos. Entretanto, a base foi formada por animais de um único produtor, não continha indivíduos com escore 1 e apresentava apenas 4 animais com escore 2 e 18 com escore 5. A aplicação de técnicas de *data augmentation* permitiu equilibrar numericamente as classes disponíveis, mas não substituiu a diversidade biológica proporcionada por um número maior de animais em cada classe.

A implementação e a avaliação de diferentes arquiteturas foram parcialmente realizadas. ResNet50, VGG16, Faster R-CNN e outras abordagens previstas foram testadas, mas não produziram resultados utilizáveis no ambiente e no período disponíveis para a pesquisa. Por essa razão, não foi possível realizar a comparação estatística entre modelos inicialmente planejada. O YOLOv8 foi o único modelo que concluiu o treinamento com resultados aproveitáveis.

A comparação das previsões do YOLOv8 com os escores atribuídos pelos especialistas mostrou desempenho distinto entre a localização dos olhos e a classificação da pigmentação. O modelo localizou a região ocular em mais de 96% das instâncias do conjunto de validação e em mais de 90% das imagens inspecionadas no conjunto de teste. Por outro lado, o desempenho agregado da classificação permaneceu em torno de 40%, com resultados mais favoráveis para os escores 3, 4 e 5 e incapacidade de reconhecer corretamente o escore 2. A principal limitação esteve, portanto, na atribuição correta do escore, especialmente nas classes com poucos animais.

Entre as contribuições do trabalho, destacam-se a criação do conjunto de imagens rotuladas, o treinamento do modelo YOLOv8, sua conversão para o formato TFLite e o desenvolvimento do protótipo móvel. O estudo também evidenciou que o aumento artificial de imagens não elimina a necessidade de uma coleta ampla e equilibrada e que a utilização de apenas um dos olhos pode introduzir divergências entre a informação presente na fotografia e o escore atribuído

pela avaliação conjunta dos dois olhos.

As principais limitações foram a origem restrita dos dados, o reduzido número de animais nas classes extremas, a ausência do escore 1, a captura de apenas um lado da cabeça, a impossibilidade de comparar estatisticamente os modelos e a falta de validação do aplicativo com produtores e técnicos em condições reais de campo. Assim, os resultados devem ser interpretados como uma prova de conceito e como base para pesquisas posteriores.

5.1 Trabalhos Futuros

Recomenda-se ampliar o conjunto de dados com imagens provenientes de diferentes propriedades e com maior número de animais dos escores 1, 2 e 5. A nova coleta deve incluir imagens dos dois olhos e preservar a separação dos animais entre os conjuntos de treinamento, validação e teste. Também se recomenda avaliar outras arquiteturas e estratégias de treinamento em ambientes computacionais controlados, permitindo a comparação entre modelos executados sob as mesmas condições.

A coleta poderá ser aprimorada por meio do uso de câmeras de vídeo, das quais poderão ser extraídos *frames* adequados, desde que seja possível identificar previamente cada animal e associar corretamente as imagens ao respectivo escore. A captura deverá priorizar os escores extremos e incluir variações de iluminação, distância e posicionamento, sem comprometer o bem-estar dos animais.

Por fim, o aplicativo deverá ser validado com produtores e técnicos responsáveis por avaliações de campo. Essa etapa permitirá verificar o desempenho do modelo em condições reais, a facilidade de uso do sistema e sua utilidade como ferramenta de apoio. Os resultados deste estudo poderão servir de ponto de partida para novas pesquisas voltadas à integração entre inteligência artificial, avaliação fenotípica e melhoramento genético bovino.

REFERÊNCIAS

ABADIAS, Ivalmir Mota; FONSECA, Paulo Roberto Beltramin da; BARBOSA, Cleison Hugo. Manejo da pecuária: uma análise sobre impactos ambientais. **Educamazônia - Educação, Sociedade e Meio Ambiente**, Humaitá, AM, v. 25, n. 1, p. 113–125, jan./jun. 2020. Disponível em: https://periodicos.ufam.edu.br/index.php/educamazonia/article/view/7669. Acesso em: 22 maio 2024.

AGGARWAL, Charu C. **Neural Networks and Deep Learning**: A Textbook. Cham, Switzerland: Springer, 2018. E-book. DOI: 10.1007/978-3-319-94463-0. Disponível em: https://doi.org/10.1007/978-3-319-94463-0. Acesso em: 12 maio 2024.

AHMED, Fahad *et al.* Identification and Prediction of Brain Tumor Using VGG-16 Empowered with Explainable Artificial Intelligence. **International Journal of Computational and Innovative Sciences**, Pakistan, v. 2, n. 2, p. 24–33, june 2023. Disponível em: https://ijcis.com/index.php/IJCIS/article/view/69. Acesso em: 22 maio 2024.

ALENCAR, Maurício Mello de. Perspectivas para o melhoramento genético de bovinos de corte no Brasil. *In:* REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 41., 2004, Campo Grande, MS. **A produção animal e a segurança alimentar**: anais dos simpósios. Campo Grande, MS: Sociedade Brasileira de Zootecnia: Embrapa Gado de Corte, 2004. p. 358–367. 1 CD-ROM. Disponível em: https://www.alice.cnptia.embrapa.br/alice/handle/doc/45707. Acesso em: 24 maio 2024.

ANSEL, Jason *et al.* PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Transformation and Graph Compilation. *In:* ACM INTERNATIONAL CONFERENCE ON ARCHITECTURAL SUPPORT FOR PROGRAMMING LANGUAGES AND OPERATING SYSTEMS, 29., 2024, La Jolla. **Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2**. New York: Association for Computing Machinery, 2024. p. 929–947. DOI: 10.1145/3620665.3640366. Disponível em: https://doi.org/10.1145/3620665.3640366. Acesso em: 17 maio 2024.

APICELLA, Andrea *et al.* A survey on modern trainable activation functions. **Neural Networks**, [Oxford], v. 138, p. 14–32, june 2021. DOI: 10.1016/j.neunet.2021.01.026. Disponível em: https://doi.org/10.1016/j.neunet.2021.01.026. Acesso em: 23 maio 2024.

BAI, Jie *et al.* Classification and recognition of milk somatic cell images based on PolyLoss and PCAM-Reset50. **Mathematical Biosciences and Engineering**, Springfield, MO, v. 20, n. 5, p. 9423–9442, 20 mar. 2023. DOI: 10.3934/mbe.2023414. Disponível em: https://doi.org/10.3934/mbe.2023414. Acesso em: 18 maio 2024.

BARNEY, Shaun *et al.* Deep learning pose estimation for multi-cattle lameness detection. **Scientific Reports**, London, v. 13, n. 1, 4499, 18 mar. 2023. DOI: 10.1038/s41598-023-31297-1. Disponível em: https://doi.org/10.1038/s41598-023-31297-1. Acesso em: 15 maio 2024.

BEN ABDALLAH, Taoufik; ELLEUCH, Islam; GUERMAZI, Radhouane. Student Behavior Recognition in Classroom using Deep Transfer Learning with VGG-16. **Procedia Computer Science**, Amsterdam, Netherlands, v. 192, p. 951–960, 2021. DOI: 10.1016/j.procs.2021.08.098. Disponível em: https://doi.org/10.1016/j.procs.2021.08.098. Acesso em: 14 maio 2024.

BENGIO, Yoshua; SIMARD, Patrice; FRASCONI, Paolo. Learning long-term dependencies with gradient descent is difficult. **IEEE Transactions on Neural Networks**, New York, v. 5, n. 2, p. 157–166, mar. 1994. DOI: 10.1109/72.279181. Disponível em: https://doi.org/10.1109/72.279181. Acesso em: 25 maio 2024.

BERNER, Julius; ELBRÄCHTER, Dennis Maximilian; GROHS, Philipp. How degenerate is the parametrization of neural networks with the ReLU activation function? *In:* CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 33., 2019, Vancouver, Canada. **Advances in Neural Information Processing Systems**. Red Hook, NY: Curran Associates, Inc., 2019. v. 32, p. 7790–7801. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2019/hash/04115ec378e476c56d19d827bcf8db56-Abstract.html. Acesso em: 18 maio 2024.

BRIDLE, John S. Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition. *In:* NATO ADVANCED RESEARCH WORKSHOP ON NEUROCOMPUTING ALGORITHMS, ARCHITECTURES AND APPLICATIONS, 1989, Les Arcs, France. **Neurocomputing**: Algorithms, Architectures and Applications. Berlin: Springer, 1990a. v. 68, p. 227–236. DOI: 10.1007/978-3-642-76153-9_28. Disponível em: https://doi.org/10.1007/978-3-642-76153-9_28. Acesso em: 19 maio 2024.

BRIDLE, John S. Training Stochastic Model Recognition Algorithms as Networks can Lead to Maximum Mutual Information Estimation of Parameters. *In:* IEEE CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS – NATURAL AND SYNTHETIC, 3., 1989, Denver, CO. **Advances in Neural Information Processing Systems**. San Mateo, CA: Morgan Kaufmann Publishers, 1990b. v. 2, p. 211–217. Disponível em: https://proceedings.neurips.cc/paper/1989/hash/0336dcbab05b9d5ad24f4333c7658a0e-Abstract.html. Acesso em: 15 maio 2024.

BROWNLEE, Jason. **A Gentle Introduction to Cross-Entropy for Machine Learning**. Machine Learning Mastery. 22 dec. 2020. Disponível em: https://machinelearningmastery.com/cross-entropy-for-machine-learning/. Acesso em: 28 maio 2024.

BUSLAEV, Alexander *et al.* Albumentations: Fast and Flexible Image Augmentations. **Information**, Basel, v. 11, n. 2, 125, 24 feb. 2020. DOI: 10.3390/info11020125. Disponível em: https://doi.org/10.3390/info11020125. Acesso em: 26 maio 2024.

CARDOSO, Fernando Flores. **Ferramentas e estratégias para o melhoramento genético de bovinos de corte**. Bagé: Embrapa Pecuária Sul, 2009. Disponível em: http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/657470. Acesso em: 19 maio 2024.

CARDOSO, Fernando Flores. Melhoramento genético de bovinos de corte. *In:* BARCELLOS, Júlio Otávio Jardim; CHRISTOFARI, Luciana Fagundes; LOPA, Thaís Maria Bento Pires (ed.). **Caderno de atualização técnica e julgamento de Hereford e Braford**. Porto Alegre: Associação Brasileira de Hereford e Braford, 2007. p. 114–135. Disponível em: https://www.alice.cnptia.embrapa.br/alice/handle/doc/217790. Acesso em: 27 maio 2024.

CARVALHO, Victor Heitor Garcia de; VIEIRA, Paulo Rufino Paulino. Melhoramento genético como estratégia de avanço da produção e da produtividade durante o confinamento bovino. **Scientia Generalis**, Patos de Minas, v. 4, n. 1, p. 61–77, 2023. DOI: 10.22289/sg.V4N1A7. Disponível em: https://doi.org/10.22289/sg.V4N1A7. Acesso em: 26 maio 2024.

CURSO INTERNACIONAL DE MELHORAMENTO DE BOVINOS DE CORTE DO PAMPAPLUS, 8., 2016, Bagé. **Evoluir é democratizar o conhecimento**: manual. Organização: Fernando Flores Cardoso, Thaís Maria Bento Pires Lopa e Bruno Borges Teixeira. Bagé: Embrapa Pecuária Sul: ABHB, 2016. [43] p. Disponível em: http://www.alice.cnptia.embrapa.br/alice/handle/doc/1042308. Acesso em: 20 maio 2024.

DAVIS, Kaycee Marie *et al.* Technical note: Digital quantification of eye pigmentation of cattle with white faces. **Journal of Animal Science**, Champaign, v. 93, n. 7, p. 3654–3660, july 2015. DOI: 10.2527/jas.2014-8786. Disponível em: https://doi.org/10.2527/jas.2014-8786. Acesso em: 15 maio 2024.

DENG, Li. The MNIST database of handwritten digit images for machine learning research. [Best of the Web]. **IEEE Signal Processing Magazine**, [New York], v. 29, n. 6, p. 141–142, nov. 2012. DOI: 10.1109/MSP.2012.2211477. Disponível em: https://doi.org/10.1109/MSP.2012.2211477. Acesso em: 18 maio 2024.

DONG, Shi; WANG, Ping; ABBAS, Khushnood. A survey on deep learning and its applications. **Computer Science Review**, [Amsterdam], v. 40, 100379, may 2021. DOI: 10.1016/j.cosrev.2021.100379. Disponível em: https://doi.org/10.1016/j.cosrev.2021.100379. Acesso em: 20 maio 2024.

DOUGLAS, Scott C.; YU, Jiutian. Why ReLU units sometimes die: analysis of single-unit error backpropagation in neural networks. *In:* ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS, AND COMPUTERS, 52., 2018, Pacific Grove. **2018 52nd Asilomar Conference**

on Signals, Systems, and Computers. Piscataway, NJ: IEEE, oct. 2018. p. 864–868. DOI: 10.1109/ACSSC.2018.8645556. Disponível em: https://doi.org/10.1109/ACSSC.2018.8645556. Acesso em: 27 maio 2024.

EUCLIDES FILHO, Kepler. **Melhoramento genético animal no Brasil**: fundamentos, história e importância. Campo Grande, MS: Embrapa Gado de Corte, 1999. Disponível em: https://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/323391. Acesso em: 29 maio 2024.

FARIA, Carina Ubirajara de *et al.* **Utilização de escores visuais de características morfológicas de bovinos nelore como ferramenta para o melhoramento genético animal**. Planaltina, DF: Embrapa Cerrados, 2007. Disponível em: http://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/571967. Acesso em: 18 maio 2024.

FERLEZ, James; SHOUKRY, Yasser. AReN: assured ReLU NN architecture for model predictive control of LTI systems. *In:* ACM INTERNATIONAL CONFERENCE ON HYBRID SYSTEMS: COMPUTATION AND CONTROL, 23., 2020, Sydney. **Proceedings of the 23rd International Conference on Hybrid Systems**: Computation and Control. New York: Association for Computing Machinery, 22 apr. 2020. p. 1–11. DOI: 10.1145/3365365.3382213. Disponível em: https://doi.org/10.1145/3365365.3382213. Acesso em: 12 maio 2024.

FLACH, Peter A.; KULL, Meelis. Precision-Recall-Gain Curves: PR Analysis Done Right. *In:* ANNUAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 29., 2015, Montréal, Canada. **Advances in Neural Information Processing Systems**. Red Hook, NY, USA: Curran Associates, Inc., 2015. v. 28, p. 838–846. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2015/file/33e8075e9970de0cfea955afd4644bb2-Paper.pdf. Acesso em: 20 maio 2024.

FREITAS, Dhyonatan Santos de *et al.* Reconhecimento da ceratoconjuntivite infecciosa bovina utilizando imagens termográficas e redes neurais convolucionais. **Revista Brasileira de Computação Aplicada**, Passo Fundo, v. 11, n. 3, p. 133–145, 16 out. 2019. DOI: 10.5335/rbca.v11i3.9210. Disponível em: https://doi.org/10.5335/rbca.v11i3.9210. Acesso em: 23 maio 2024.

FUKUSHIMA, Kunihiko. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. **Biological Cybernetics**, Berlin, v. 36, n. 4, p. 193–202, apr. 1980. DOI: 10.1007/BF00344251. Disponível em: https://doi.org/10.1007/BF00344251. Acesso em: 15 maio 2024.

FUKUSHIMA, Kunihiko. Visual Feature Extraction by a Multilayered Network of Analog Threshold Elements. **IEEE Transactions on Systems Science and Cybernetics**, New York, v. 5, n. 4, p. 322–333, oct. 1969. DOI: 10.1109/TSSC.1969.300225. Disponível em: https://doi.org/10.1109/TSSC.1969.300225. Acesso em: 11 maio 2024.

GARCÍA-GARCÍA, Fernando *et al.* Automated location of orofacial landmarks to characterize airway morphology in anaesthesia via deep convolutional neural networks. **Computer Methods and Programs in Biomedicine**, Amsterdam, Netherlands, v. 232, 107428, apr. 2023. DOI: 10.1016/j.cmpb.2023.107428. Disponível em: https://doi.org/10.1016/j.cmpb.2023.107428. Acesso em: 15 maio 2024.

GENEPLUS. **Programa Embrapa de Melhoramento de Gado de Corte – Geneplus**. [2024?]. Disponível em: https://geneplus.com.br/sobre. Acesso em: 20 maio 2024.

GENG, Lei *et al.* Lung segmentation method with dilated convolution based on VGG-16 network. **Computer Assisted Surgery**, Abingdon, England, v. 24, p. 27–33, oct. 2019. Suplemento 2: Advances in Minimally Invasive Surgery and Clinical Measurement. DOI: 10.1080/24699322.2019.1649071. Disponível em: https://doi.org/10.1080/24699322.2019.1649071. Acesso em: 19 maio 2024.

GERAZOV, Branislav. **Artificial neuron**. 3 feb. 2023. 1 ilustração, SVG. Disponível em: https://commons.wikimedia.org/wiki/File:Artificial_neuron.svg. Acesso em: 23 maio 2024.

GIRSHICK, Ross. Fast R-CNN. *In:* IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION, 15., 2015, Santiago, Chile. **Proceedings of the IEEE International Conference on Computer Vision (ICCV)**. Piscataway, NJ: IEEE, dec. 2015. p. 1440–1448. DOI: 10.1109/ICCV.2015.169. Disponível em: https://doi.org/10.1109/ICCV.2015.169. Acesso em: 19 maio 2024.

GIRSHICK, Ross *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. *In:* IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2014, Columbus, OH. **2014 IEEE Conference on Computer Vision and Pattern Recognition**. Los Alamitos: IEEE Computer Society, 2014. p. 580–587. DOI: 10.1109/CVPR.2014.81. Disponível em: https://doi.org/10.1109/CVPR.2014.81. Acesso em: 16 maio 2024.

GLOROT, Xavier; BORDES, Antoine; BENGIO, Yoshua. Deep Sparse Rectifier Neural Networks. *In:* INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS, 14., 2011, Fort Lauderdale, FL, USA. **Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics**. [*S. l.*]: PMLR, 14 june 2011. v. 15, p. 315–323. Disponível em: https://proceedings.mlr.press/v15/glorot11a.html. Acesso em: 18 maio 2024.

GOEL, Surbhi; KARMALKAR, Sushrut; KLIVANS, Adam R. Time/Accuracy Tradeoffs for Learning a ReLU with respect to Gaussian Marginals. *In:* CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 33., 2019, Vancouver, Canada. **Advances in Neural Information Processing Systems**. Red Hook, NY: Curran Associates, 2019. v. 32, p. 8584–8593. Disponível em: https://papers.nips.cc/paper_files/paper/2019/hash/067a26d87265ea39030f5bd82408ce7c-Abstract.html. Acesso em: 23 maio 2024.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. Cambridge, MA: The MIT Press, 2016. E-book. Disponível em: https://www.deeplearningbook.org. Acesso em: 30 maio 2024.

GUAN, Qing *et al.* Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study. **Journal of Cancer**, Sydney, NSW, v. 10, n. 20, p. 4876–4882, aug. 2019. DOI: 10.7150/jca.28769. Disponível em: https://doi.org/10.7150/jca.28769. Acesso em: 15 maio 2024.

HAYKIN, Simon. **Redes neurais**: princípios e prática. Tradução: Paulo Martins Engel. 2. ed. Porto Alegre: Bookman, 2001.

HAYOU, Soufiane; DOUCET, Arnaud; ROUSSEAU, Judith. On the Impact of the Activation Function on Deep Neural Networks Training. *In:* INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 36., 2019, Long Beach, California. **Proceedings of the 36th International Conference on Machine Learning**. [*S. l.*]: PMLR, 24 may 2019. v. 97, p. 2672–2680. Disponível em: https://proceedings.mlr.press/v97/hayou19a.html. Acesso em: 16 maio 2024.

HE, Kaiming *et al.* Deep residual learning for image recognition. *In:* IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2016, Las Vegas, NV, USA. **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Piscataway: IEEE, june 2016. p. 770–778. DOI: 10.1109/CVPR.2016.90. Disponível em: https://doi.org/10.1109/CVPR.2016.90. Acesso em: 24 maio 2024.

HOSSIN, Mohammad bin; SULAIMAN, Md. Nasir. A Review on Evaluation Metrics for Data Classification Evaluations. **International Journal of Data Mining & Knowledge Management Process**, Chennai, India, v. 5, n. 2, p. 1–11, mar. 2015. DOI: 10.5121/ijdkp.2015.5201. Disponível em: https://doi.org/10.5121/ijdkp.2015.5201. Acesso em: 16 maio 2024.

HUBEL, David Hunter; WIESEL, Torsten Nils. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. **The Journal of Physiology**, London, v. 160, n. 1, p. 106–154, jan. 1962. DOI: 10.1113/jphysiol.1962.sp006837. Disponível em: https://doi.org/10.1113/jphysiol.1962.sp006837. Acesso em: 18 maio 2024.

ISLAM, Shazzadul *et al.* Bird species classification from an image using VGG-16 network. *In:* INTERNATIONAL CONFERENCE ON COMPUTER AND COMMUNICATIONS MANAGEMENT, 7., 2019, Bangkok. **Proceedings of the 2019 7th International Conference on Computer and Communications Management**. New York, NY: Association for Computing Machinery, 2019. p. 38–42. DOI: 10.1145/3348445.3348480. Disponível em: https://doi.org/10.1145/3348445.3348480. Acesso em: 19 maio 2024.

JAIN, Anil K. **Fundamentals of Digital Image Processing**. Englewood Cliffs: Prentice Hall, 1989.

JAMES, W. P. T. *et al.* Nutrition and its role in human evolution. **Journal of Internal Medicine**, Oxford, UK, v. 285, n. 5, p. 533–549, may 2019. DOI: 10.1111/joim.12878. Disponível em: https://doi.org/10.1111/joim.12878. Acesso em: 13 maio 2024.

JANOCHA, Katarzyna; CZARNECKI, Wojciech Marian. On Loss Functions for Deep Neural Networks in Classification. **Schedae Informaticae**, Kraków, v. 25, p. 49–59, 2016. DOI: 10.4467/20838476SI.16.004.6185. Disponível em: https://doi.org/10.4467/20838476SI.16.004.6185. Acesso em: 22 maio 2024.

JARA, Eugenio *et al.* Revealing the genetic basis of eyelid pigmentation in Hereford cattle. **Journal of Animal Science**, Oxford, England, v. 100, n. 5, skac110, may 2022. DOI: 10.1093/jas/skac110. Disponível em: https://doi.org/10.1093/jas/skac110. Acesso em: 28 maio 2024.

JOCHER, Glenn; CHAURASIA, Ayush; QIU, Jing. **Ultralytics YOLOv8**. [*S. l.*]: Ultralytics, 10 jan. 2023. Versão 8.0.0. Programa de computador. Disponível em: https://github.com/ultralytics/ultralytics. Acesso em: 14 maio 2024.

KALIAPPAN, Jayakumar *et al.* Impact of Cross-Validation on Machine Learning Models for Early Detection of Intrauterine Fetal Demise. **Diagnostics**, Basel, v. 13, n. 10, 1692, 10 may 2023. DOI: 10.3390/diagnostics13101692. Disponível em: https://doi.org/10.3390/diagnostics13101692. Acesso em: 19 maio 2024.

KANDEL, Eric R. *et al.* (ed.). **Princípios de neurociências**. 5. ed. Porto Alegre: AMGH, 2014. E-book.

KLETTE, Reinhard. **Concise Computer Vision**: An Introduction into Theory and Algorithms. London, UK: Springer, 2014. DOI: 10.1007/978-1-4471-6320-6. Disponível em: https://doi.org/10.1007/978-1-4471-6320-6. Acesso em: 17 maio 2024.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Técnicas de pesquisa**. 9. ed. São Paulo: Atlas, 2021.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **Nature**, London, UK, v. 521, n. 7553, p. 436–444, 28 may 2015. DOI: 10.1038/nature14539. Disponível em: https://doi.org/10.1038/nature14539. Acesso em: 24 maio 2024.

LECUN, Yann *et al.* Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, New York, v. 86, n. 11, p. 2278–2324, nov. 1998. DOI: 10.1109/5.726791. Disponível em: https://doi.org/10.1109/5.726791. Acesso em: 21 maio 2024.

LI, Guoming; ERICKSON, Galen E.; XIONG, Yijie. Individual Beef Cattle Identification Using Muzzle Images and Deep Learning Techniques. **Animals**, Basel, v. 12, n. 11, 1453, 4

june 2022. DOI: 10.3390/ani12111453. Disponível em: https://doi.org/10.3390/ani12111453. Acesso em: 13 maio 2024.

LI, Hao *et al.* Visualizing the Loss Landscape of Neural Nets. *In:* CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 32., 2018, Montréal, Canada. **Advances in Neural Information Processing Systems**. Red Hook, NY: Curran Associates, 2018. v. 31, p. 6389–6399. Disponível em: https://proceedings.neurips.cc/paper/2018/hash/a41b3bb3e6b050b6c9067c67f663b915-Abstract.html. Acesso em: 18 maio 2024.

LIM, Joseph J.; SALAKHUTDINOV, Ruslan; TORRALBA, Antonio. Transfer Learning by Borrowing Examples for Multiclass Object Detection. *In:* ANNUAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 25., 2011, Granada, Spain. **Advances in Neural Information Processing Systems 24**. Red Hook, NY: Curran Associates, Inc., 2011. p. 118–126. DOI: 10.5555/2986459.2986473. Disponível em: https://doi.org/10.5555/2986459.2986473. Acesso em: 30 maio 2024.

LINNAINMAA, Seppo. Taylor expansion of the accumulated rounding error. **BIT Numerical Mathematics**, Copenhagen, v. 16, n. 2, p. 146–160, june 1976. DOI: 10.1007/BF01931367. Disponível em: https://doi.org/10.1007/BF01931367. Acesso em: 28 maio 2024.

LOKAL PROFIL. **Neuron, LangNeutral**. 20 dec. 2006. 1 ilustração, SVG. Disponível em: https://commons.wikimedia.org/wiki/File:Neuron,_LangNeutral.svg. Acesso em: 16 maio 2024.

LORENSI, Rui Antonio da Rocha. **Simulação preliminar da predição do volume globular bovino**: um método não invasivo por meio de imagens RGB e redes neurais convolucionais. 2024. Dissertação (Mestrado em Computação Aplicada) – Universidade Federal do Pampa, Campus Bagé, Bagé, 30 nov. 2023. Disponível em: https://repositorio.unipampa.edu.br/handle/riu/8952. Acesso em: 15 maio 2024.

LORENSI, Rui Antonio da Rocha; CAMARGO, Sandro da Silva; GENRO, Teresa Cristina Moraes. Inteligência artificial para predição de doenças em bovinos através de imagens: uma revisão sistemática da literatura. **JAIIO, Jornadas Argentinas de Informática**, [Buenos Aires], v. 8, n. 4, p. 93–106, 26 dez. 2022. Trabalho apresentado no Congreso Argentino de AgroInformática (CAI 2022), integrante das 51 Jornadas Argentinas de Informática e Investigación Operativa, Buenos Aires, 2022. Disponível em: https://revistas.unlp.edu.ar/JAIIO/article/view/18279. Acesso em: 27 maio 2024.

LU, Lu *et al.* Dying ReLU and initialization: theory and numerical examples. **Communications in Computational Physics**, Hong Kong, v. 28, n. 5, p. 1671–1706, nov. 2020. DOI: 10.4208/cicp.OA-2020-0165. Disponível em: https://doi.org/10.4208/cicp.OA-2020-0165. Acesso em: 28 maio 2024.

MAAS, Andrew L.; HANNUN, Awni Y.; NG, Andrew Y. Rectifier nonlinearities improve neural network acoustic models. *In:* ICML WORKSHOP ON DEEP LEARNING FOR AUDIO, SPEECH AND LANGUAGE PROCESSING, 2013, Atlanta, GA. **Proceedings of the 30th International Conference on Machine Learning**. [*S. l.: s. n.*], 2013. v. 28. [6] p. Disponível em: https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf. Acesso em: 17 maio 2024.

MALAFAIA, Guilherme Cunha; BISCOLA, Paulo Henrique Nogueira; DIAS, Fernando Rodrigues Teixeira. **Projeções para o mercado mundial de carne bovina 2020-2029**. Análise da semana de 18 a 24 de julho. Campo Grande, MS: Embrapa Gado de Corte, 2020. Boletim CiCarne. Disponível em: https://www.infoteca.cnptia.embrapa.br/infoteca/handle/doc/1125596. Acesso em: 15 maio 2024.

MALAFAIA, Guilherme Cunha *et al.* A sustentabilidade na cadeia produtiva da pecuária de corte brasileira. *In:* BUNGENSTAB, Davi José *et al.* (ed.). **ILPF**: inovação com integração de lavoura, pecuária e floresta. Brasília: Embrapa, 2019. cap. 8, p. 117–130. Disponível em: http://www.alice.cnptia.embrapa.br/alice/handle/doc/1112915. Acesso em: 23 maio 2024.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**: métodos científicos, técnicas de pesquisa, elaboração de referências bibliográficas. Atualização João Bosco Medeiros. 9. ed. São Paulo: Atlas, 2021a.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Metodologia do trabalho científico**: projetos de pesquisa, pesquisa bibliográfica, teses de doutorado, dissertações de mestrado, trabalhos de conclusão de curso. Atualização João Bosco Medeiros. 9. ed. São Paulo: Atlas, 2021b.

MARENGONI, Maurício; STRINGHINI, Denise. Tutorial: introdução à visão computacional usando OpenCV. **Revista de Informática Teórica e Aplicada**, Porto Alegre, v. 16, n. 1, p. 125–160, 2009. DOI: 10.22456/2175-2745.11477. Disponível em: https://doi.org/10.22456/2175-2745.11477. Acesso em: 14 maio 2024.

MARTIN, A. Robert *et al.* **From Neuron to Brain**. 6th ed. Sunderland, Massachusetts: Oxford University Press/Sinauer, 2021.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, [Chicago], v. 5, n. 4, p. 115–133, dec. 1943. DOI: 10.1007/BF02478259. Disponível em: https://doi.org/10.1007/BF02478259. Acesso em: 30 maio 2024.

MEHEDI SHAMRAT, F. M. Javed *et al.* An advanced deep neural network for fundus image analysis and enhancing diabetic retinopathy detection. **Healthcare Analytics**, New York, NY, v. 5, 100303, p. 1–19, june 2024. DOI: 10.1016/j.health.2024.100303. Disponível em: https://doi.org/10.1016/j.health.2024.100303. Acesso em: 13 maio 2024.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. Conceitos sobre aprendizado de máquina. *In:* REZENDE, Solange Oliveira (org.). **Sistemas inteligentes**: fundamentos e aplicações. 1. ed. Barueri: Manole, 2003. cap. 4, p. 89–114.

NAIR, Vinod; HINTON, Geoffrey. Rectified Linear Units Improve Restricted Boltzmann Machines. *In:* INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 27., 2010, Haifa, Israel. **Proceedings of the 27th International Conference on Machine Learning (ICML-10)**. Madison, WI: Omnipress, 2010. p. 807–814. Disponível em: https://icml.cc/Conferences/2010/papers/432.pdf. Acesso em: 18 maio 2024.

NYE, Jessica; ZINGARETTI, Laura M.; PÉREZ-ENCISO, Miguel. Estimating Conformational Traits in Dairy Cattle With DeepAPS: A Two-Step Deep Learning Automated Phenotyping and Segmentation Approach. **Frontiers in Genetics**, Lausanne, v. 11, 513, 21 may 2020. DOI: 10.3389/fgene.2020.00513. Disponível em: https://doi.org/10.3389/fgene.2020.00513. Acesso em: 14 maio 2024.

O'KEEFE, James Henry *et al.* Debunking the vegan myth: The case for a plant-forward omnivorous whole-foods diet. **Progress in Cardiovascular Diseases**, Philadelphia, v. 74, p. 2–8, sept./oct. 2022. DOI: 10.1016/j.pcad.2022.08.001. Disponível em: https://doi.org/10.1016/j.pcad.2022.08.001. Acesso em: 18 maio 2024.

OLIVEIRA, Bruna Marcele Martins de. **Hemodinâmica uterina avaliada por ultrassom Doppler colorido e taxa de fertilidade de vacas submetidas à IATF com sêmen analisado por sondas fluorescentes**. 2012. Dissertação (Mestrado em Ciências, área de Reprodução Animal) – Universidade de São Paulo, Faculdade de Medicina Veterinária e Zootecnia, São Paulo, 31 jan. 2012. DOI: 10.11606/D.10.2012.tde-19072012-110306. Disponível em: https://doi.org/10.11606/D.10.2012.tde-19072012-110306. Acesso em: 11 maio 2024.

OOSTWAL, Elisa; STRAAT, Michiel; BIEHL, Michael. Hidden unit specialization in layered neural networks: ReLU vs. sigmoidal activation. **Physica A:** Statistical Mechanics and its Applications, Amsterdam, Netherlands, v. 564, 125517, feb. 2021. DOI: 10.1016/j.physa.2020.125517. Disponível em: https://doi.org/10.1016/j.physa.2020.125517. Acesso em: 29 maio 2024.

PAN, Sinno Jialin; YANG, Qiang. A Survey on Transfer Learning. **IEEE Transactions on Knowledge and Data Engineering**, New York, v. 22, n. 10, p. 1345–1359, oct. 2010. DOI: 10.1109/TKDE.2009.191. Disponível em: https://doi.org/10.1109/TKDE.2009.191. Acesso em: 17 maio 2024.

PANG, Bo; NIJKAMP, Erik; WU, Ying Nian. Deep Learning With TensorFlow: A Review. **Journal of Educational and Behavioral Statistics**, [Thousand Oaks, CA], v. 45, n. 2, p. 227–248, apr. 2020. DOI: 10.3102/1076998619872761. Disponível em: https://doi.org/10.3102/1076998619872761. Acesso em: 25 maio 2024.

PASZKE, Adam *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. *In:* CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 33., 2019, Vancouver, Canada. **Advances in Neural Information Processing Systems 32**. Red Hook, NY, USA: Curran Associates, Inc., 2019. p. 8024–8035. Disponível em: https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html. Acesso em: 26 maio 2024.

PATEL, Neel *et al.* Biometric-based Unique Identification for Bovine Animals — Comparative Study of Various Machine and Deep Learning Computer Vision Methods. *In:* SOMAIYA INTERNATIONAL CONFERENCE ON TECHNOLOGY AND INFORMATION MANAGEMENT, 9., 2023, Mumbai, India. **2023 Somaiya International Conference on Technology and Information Management (SICTIM)**. Piscataway, NJ, USA: IEEE, 2023. p. 1–5. DOI: 10.1109/SICTIM56495.2023.10105004. Disponível em: https://doi.org/10.1109/SICTIM56495.2023.10105004. Acesso em: 26 maio 2024.

PAUSCH, Hubert *et al.* Identification of QTL for UV-Protective Eye Area Pigmentation in Cattle by Progeny Phenotyping and Genome-Wide Association Analysis. **PLoS ONE**, San Francisco, CA, v. 7, n. 5, e36346, 2 may 2012. DOI: 10.1371/journal.pone.0036346. Disponível em: https://doi.org/10.1371/journal.pone.0036346. Acesso em: 19 maio 2024.

PEREIRA, Jonas Carlos Campos. **Melhoramento genético aplicado à produção animal**. 4. ed. Belo Horizonte: FEPMVZ Editora, 2004.

PEREIRA, Maurício Gomes. **Artigos científicos**: como redigir, publicar e avaliar. Rio de Janeiro: Guanabara Koogan, 2011.

PETROSYAN, Armenak; DEREVENTSOV, Anton; WEBSTER, Clayton G. Neural network integral representations with the ReLU activation function. *In:* MATHEMATICAL AND SCIENTIFIC MACHINE LEARNING CONFERENCE, 1., 2020, Princeton, NJ, USA. **Proceedings of the First Mathematical and Scientific Machine Learning Conference**. [*S. l.*]: PMLR, 16 aug. 2020. v. 107, p. 128–143. Disponível em: https://proceedings.mlr.press/v107/petrosyan20a.html. Acesso em: 14 maio 2024.

PORTER, Ian R.; WIELAND, Matthias; BASRAN, Parminder S. Feasibility of the use of deep learning classification of teat-end condition in Holstein cattle. **Journal of Dairy Science**, Champaign, v. 104, n. 4, p. 4529–4536, apr. 2021. DOI: 10.3168/jds.2020-19642. Disponível em: https://doi.org/10.3168/jds.2020-19642. Acesso em: 11 maio 2024.

POWERS, David Martin Ward. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. **Journal of Machine Learning Technologies**, [Pune], v. 2, n. 1, p. 37–63, 15 dec. 2011. Disponível em: https://bioinfopublication.org/pages/article.php?id=BIA0001114. Acesso em: 12 maio 2024.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do Trabalho Científico**: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico. 2. ed. Novo Hamburgo:

Feevale, 2013. E-book. Disponível em: https://www.feevale.br/Comum/midias/0163c988-1f5d-496f-b118-a6e009a7a2f9/E-book%20Metodologia%20do%20Trabalho%20Cientifico.pdf. Acesso em: 23 maio 2024.

RASAMOELINA, Andrinandrasana David; ADJAILIA, Fouzia; SINČÁK, Peter. A Review of Activation Function for Artificial Neural Network. *In:* IEEE WORLD SYMPOSIUM ON APPLIED MACHINE INTELLIGENCE AND INFORMATICS (SAMI), 18., 2020, Herl'any, Slovakia. **2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)**. Piscataway, NJ: IEEE, 2020. p. 281–286. DOI: 10.1109/SAMI48414.2020.9108717. Disponível em: https://doi.org/10.1109/SAMI48414.2020.9108717. Acesso em: 13 maio 2024.

REDMON, Joseph *et al.* You Only Look Once: Unified, Real-Time Object Detection. *In:* IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2016, Las Vegas, NV, USA. **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Las Vegas, NV, USA: IEEE, june 2016. p. 779–788. DOI: 10.1109/CVPR.2016.91. Disponível em: https://doi.org/10.1109/CVPR.2016.91. Acesso em: 20 maio 2024.

REN, Shaoqing *et al.* Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *In:* ANNUAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 29., 2015, Montreal, Quebec, Canada. **Advances in Neural Information Processing Systems 28**. [*S. l.*]: Neural Information Processing Systems Foundation, Inc., 2015. p. 91–99. Disponível em: https://papers.nips.cc/paper/2015/hash/14bfa6bb14875e45bba028a21ed38046-Abstract.html. Acesso em: 11 maio 2024.

RONY, Md. Awlad Hossen *et al.* Cattle External Disease Classification Using Deep Learning Techniques. *In:* INTERNATIONAL CONFERENCE ON COMPUTING COMMUNICATION AND NETWORKING TECHNOLOGIES, 12., 2021, Kharagpur, India. **2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)**. Piscataway, NJ, USA: IEEE, 2021. p. 1–7. DOI: 10.1109/ICCCNT51525.2021.9579662. Disponível em: https://doi.org/10.1109/ICCCNT51525.2021.9579662. Acesso em: 14 maio 2024.

ROSENBLATT, Frank. **Principles of Neurodynamics**: Perceptrons and the Theory of Brain Mechanisms. Washington, DC: Spartan Books, 1962.

RUSSELL, Stuart Jonathan; NORVIG, Peter. **Inteligência artificial**: uma abordagem moderna. Tradução: Daniel Vieira e Flávio Soares Corrêa da Silva. 4. ed. Rio de Janeiro: GEN LTC, 2022.

SALDANHA, Marcus Fabiano Silva; FREITAS, Corina da Costa. Segmentação de Imagens Digitais: Uma Revisão. *In:* WORKSHOP DOS CURSOS DE COMPUTAÇÃO APLICADA DO INPE (WORCAP), 9., 2009, São José dos Campos. **Anais**. São José dos Campos: Instituto

Nacional de Pesquisas Espaciais (INPE), 2009. [9] p. Disponível em: http://urlib.net/rep/8JMKD3MGP8W/37NC86S. Acesso em: 19 maio 2024.

SARKAR, Dipanjan; BALI, Raghav; GHOSH, Tamoghna. **Hands-On Transfer Learning with Python**: Implement advanced deep learning and neural network models using TensorFlow and Keras. 1st ed. Birmingham: Packt Publishing, 2018.

SHANMUGAMANI, Rajalingappaa. **Deep Learning for Computer Vision**: expert techniques to train advanced neural networks using TensorFlow and Keras. Birmingham: Packt Publishing, 2018.

SHAPIRO, Linda G.; STOCKMAN, George C. **Computer Vision**. Upper Saddle River, NJ: Prentice Hall, 2001.

SHARMA, Siddharth; SHARMA, Simone; ATHAIYA, Anidhya. Activation Functions in Neural Networks. **International Journal of Engineering Applied Sciences and Technology**, [Delhi], v. 4, n. 12, p. 310–316, apr. 2020. DOI: 10.33564/IJEAST.2020.v04i12.054. Disponível em: https://doi.org/10.33564/IJEAST.2020.v04i12.054. Acesso em: 16 maio 2024.

SHEET, Sinan S. Mohammed *et al.* Retinal disease identification using upgraded CLAHE filter and transfer convolution neural network. **ICT Express**, [Seoul], v. 8, n. 1, p. 142–150, mar. 2022. DOI: 10.1016/j.icte.2021.05.002. Disponível em: https://doi.org/10.1016/j.icte.2021.05.002. Acesso em: 27 maio 2024.

SILVA, Adriana Moreira da *et al.* Uso de inteligência artificial na pecuária: revisão de literatura. **Research, Society and Development**, Vargem Grande Paulista, SP, v. 12, n. 4, e6612440777, 27 mar. 2023. DOI: 10.33448/rsd-v12i4.40777. Disponível em: https://doi.org/10.33448/rsd-v12i4.40777. Acesso em: 22 maio 2024.

SILVA, Fabyano Fonseca e *et al.* Comparação bayesiana de modelos de previsão de diferenças esperadas nas progênies no melhoramento genético de gado Nelore. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 43, n. 1, p. 37–45, jan. 2008. DOI: 10.1590/S0100-204X2008000100006. Disponível em: https://doi.org/10.1590/S0100-204X2008000100006. Acesso em: 13 maio 2024.

SILVEIRA, Gabriela Cristina Neves; ARAÚJO, Flávia Ferreira. Melhoramento genético em bovinos de corte no Brasil. **Revista de Trabalhos Acadêmicos – Universo Belo Horizonte**, Belo Horizonte, v. 1, n. 8, 2023. [2] p. Trabalho apresentado na Jornada Acadêmica – Universo BH, 2023, Belo Horizonte. Disponível em: https://revista.universo.edu.br/index.php?journal=3universobelohorizonte3&op=view&page=article&path%5B%5D=11432. Acesso em: 18 maio 2024.

SIMONYAN, Karen; ZISSERMAN, Andrew. Very Deep Convolutional Networks for Large-Scale Image Recognition. *In:* INTERNATIONAL CONFERENCE ON LEARNING

REPRESENTATIONS, 3., 2015, San Diego, CA. **Conference Track Proceedings**. [*S. l.*]: Computational and Biological Learning Society, 2015. p. 1–14. Disponível em: https://arxiv.org/abs/1409.1556. Acesso em: 10 maio 2024.

SITAULA, Chiranjibi; HOSSAIN, Mohammad Belayet. Attention-based VGG-16 model for COVID-19 chest X-ray image classification. **Applied Intelligence**, Dordrecht, v. 51, n. 5, p. 2850–2863, may 2021. DOI: 10.1007/s10489-020-02055-x. Disponível em: https://doi.org/10.1007/s10489-020-02055-x. Acesso em: 28 maio 2024.

SOUZA, Igor Rocha Ribeiro de. **Pecuária de corte bovina brasileira e indústria de nutrição animal**: mapeamento de cenários. 2023. Dissertação (Mestrado Profissional em Administração de Empresas) – Fundação Getulio Vargas, Escola de Administração de Empresas de São Paulo, São Paulo, 30 jun. 2023. Disponível em: https://repositorio.fgv.br/bitstreams/d50eda05-f3b1-4be4-9099-2f013a68405c/download. Acesso em: 24 maio 2024.

SRINATH, K. R. Python – the fastest growing programming language. **International Research Journal of Engineering and Technology**, [Trichy], v. 4, n. 12, p. 354–357, dec. 2017. Disponível em: https://www.irjet.net/archives/V4/i12/IRJET-V4I1266.pdf. Acesso em: 28 maio 2024.

SRIVASTAVA, Nitish *et al.* Dropout: A Simple Way to Prevent Neural Networks from Overfitting. **Journal of Machine Learning Research**, Brookline, MA, v. 15, n. 56, p. 1929–1958, june 2014. Disponível em: https://jmlr.org/papers/v15/srivastava14a.html. Acesso em: 21 maio 2024.

SUTSKEVER, Ilya *et al.* On the importance of initialization and momentum in deep learning. *In:* INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 30., 2013, Atlanta, GA. **Proceedings of the 30th International Conference on Machine Learning**. Proceedings of Machine Learning Research, v. 28, n. 3. [*S. l.*]: PMLR, 26 may 2013. p. 1139–1147. Disponível em: https://proceedings.mlr.press/v28/sutskever13.html. Acesso em: 26 maio 2024.

TEIXEIRA, Bruno Borges Machado *et al.* Herdabilidades e correlações genéticas para critérios de seleção das raças Hereford e Braford. *In:* SIMPÓSIO BRASILEIRO DE MELHORAMENTO ANIMAL, 11., 2015, Santa Maria, RS. **Melhoramento animal da academia ao campo**: uma parceria em construção: anais. Santa Maria, RS: SBMA: UFSM, 2015. p. 1–3. Disponível em: https://www.alice.cnptia.embrapa.br/handle/doc/1024449. Acesso em: 27 maio 2024.

TERVEN, Juan; CÓRDOVA-ESPARZA, Diana-Margarita; ROMERO-GONZÁLEZ, Julio-Alejandro. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. **Machine Learning and Knowledge Extraction**, Basel, v. 5, n. 4, p. 1680–1716, 20 nov. 2023. DOI: 10.3390/make5040083. Disponível em: https://doi.org/10.3390/make5040083. Acesso em: 17 maio 2024.

VIDAL, Ezequiel Luís. **Otimização da seleção de acasalamentos para melhoramento genético animal**. 2021. Dissertação (Mestrado em Computação Aplicada) – Universidade Federal do Pampa, Campus Bagé, Programa de Pós-Graduação em Computação Aplicada, Bagé, 18 ago. 2021. Disponível em: https://repositorio.unipampa.edu.br/handle/riu/7809. Acesso em: 24 maio 2024.

VILELA JUNIOR, Guanis de Barros *et al.* Determinação das métricas usuais a partir da matriz de confusão de classificadores multiclasses em algoritmos inteligentes nas ciências do movimento humano. **Revista CPAQV - Centro de Pesquisas Avançadas em Qualidade de Vida**, Campinas, SP, v. 14, n. 2, p. 1–9, 2022. DOI: 10.36692/v14n2-01. Disponível em: https://doi.org/10.36692/v14n2-01. Acesso em: 12 maio 2024.

VUJOVIĆ, Željko Đ. Classification Model Evaluation Metrics. **International Journal of Advanced Computer Science and Applications**, West Yorkshire, UK, v. 12, n. 6, p. 599–606, 2021. DOI: 10.14569/IJACSA.2021.0120670. Disponível em: https://doi.org/10.14569/IJACSA.2021.0120670. Acesso em: 18 maio 2024.

WERBOS, Paul John. Applications of advances in nonlinear sensitivity analysis. *In:* IFIP CONFERENCE ON SYSTEM MODELING AND OPTIMIZATION, 10., 1981, New York, NY, USA. **System Modeling and Optimization**. Berlin, Germany: Springer, 30 june 1982. v. 38, p. 762–770. DOI: 10.1007/BFb0006203. Disponível em: https://doi.org/10.1007/BFb0006203. Acesso em: 24 maio 2024.

WERBOS, Paul John. **Beyond Regression**: New Tools for Prediction and Analysis in the Behavioral Sciences. 1974. Tese (Doutorado em Estatística) – Harvard University, Committee on Applied Mathematics, Cambridge, Massachusetts, aug. 1974.

WOOLDRIDGE, Michael John. **A Brief History of Artificial Intelligence**: What It Is, Where We Are, and Where We Are Going. New York: Flatiron Books, 2021.

ZAMPAR, Aline. **Comparação de estimativas de diferenças esperadas de progênie por três metodologias em touros da raça Nelore**. 2007. Dissertação (Mestrado em Zootecnia) – Universidade de São Paulo, Faculdade de Zootecnia e Engenharia de Alimentos, Pirassununga, 11 out. 2007. Área de concentração: Qualidade e Produtividade Animal. DOI: 10.11606/D.74.2007.tde-14112007-152946. Disponível em: https://doi.org/10.11606/D.74.2007.tde-14112007-152946. Acesso em: 12 maio 2024.