

INSTITUTO FEDERAL
Goiás
Câmpus Anápolis

Fabricia Silva Machado

IDENTIFICAÇÃO E CLASSIFICAÇÃO DA PELAGEM DE EQUINOS UTILIZANDO TÉCNICAS DE DEEP LEARNING

Anápolis-GO

2026

Fabricia Silva Machado

IDENTIFICAÇÃO E CLASSIFICAÇÃO DA PELAGEM DE EQUINOS UTILIZANDO TÉCNICAS DE DEEP LEARNING

Projeto de Pesquisa apresentado ao curso de Bacharelado em Ciências da Computação, como requisito para obtenção do grau final na disciplina de Projeto Final de Curso.

Instituto Federal de Goiás - Câmpus Anápolis

Orientador: Prof. Me. Alexandre Bellezi José

Anápolis-GO
2026

Dados Internacionais de Catalogação na Publicação (CIP)

M538a Machado, Fabricia Silva.

Identificação e classificação da pelagem de equinos utilizando técnicas de *Deep Learning / Fabricia Silva Machado.* – Anápolis: IFG, 2026.

59 f. il. color.

Orientador: Prof. Me. Alexandre Bellezi José

Trabalho de Conclusão de Curso – Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Campus Anápolis – Bacharelado em Ciência da Computação, 2026

1. Visão computacional. 2. YOLOv11. 3. ResNet-50. 4. Pelagens equinas.
I. José, Alexandre Belezzi (orien.).
II. Título.

CDD 006.37

Ficha catalográfica elaborada pelo bibliotecário Matheus Rocha Piacenti CRB1/02992
IFG Campus Anápolis

TERMO DE AUTORIZAÇÃO PARA DISPONIBILIZAÇÃO NO REPOSITÓRIO DIGITAL DO IFG - ReDi IFG

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia de Goiás, a disponibilizar gratuitamente o documento no Repositório Digital (ReDi IFG), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IFG.

Identificação da Produção Técnico-Científica

[ ] Tese	[ ] Artigo Científico
[ ] Dissertação	[ ] Capítulo de Livro
[ ] Monografia – Especialização	[ ] Livro
[X] TCC - Graduação	[ ] Trabalho Apresentado em Evento
[ ] Produto Técnico e Educacional - Tipo: _____	

Nome Completo do Autor: Fabricia Silva Machado
Matrícula: 20221060140113
Título do Trabalho: IDENTIFICAÇÃO E CLASSIFICAÇÃO DA PELAGEM DE EQUINOS UTILIZANDO TÉCNICAS DE DEEP LEARNING

Autorização - Marque uma das opções

1. (X) Autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso aberto);
2. ( ) Autorizo disponibilizar meu trabalho no Repositório Digital do IFG somente após a data ___/___/_____ (Embargo);
3. ( ) Não autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso restrito).

Ao indicar a opção **2 ou 3**, marque a justificativa:

( ) O documento está sujeito a registro de patente.
( ) O documento pode vir a ser publicado como livro, capítulo de livro ou artigo.
( ) Outra justificativa: _____

DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O/A referido/a autor/a declara que:

i. o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
ii. obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia de Goiás os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
iii. cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia de Goiás.

Anápolis/GO, 27/08/2027.
Local Data

Documento assinado digitalmente
gov.br
FABRICIA SILVA MACHADO
Data: 31/08/2026 11:06:52-0300
Verifique em https://validar.iti.gov.br

Assinatura do Autor e/ou Detentor dos Direitos Autorais

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

Na presente data realizou-se a sessão pública de defesa do Trabalho de Conclusão de Curso intitulado **IDENTIFICAÇÃO E CLASSIFICAÇÃO DA PELAGEM DE EQUINOS UTILIZANDO TÉCNICAS DE DEEP LEARNING**, sob orientação de Alexandre Bellezi Jose , apresentado pela aluna **Fabricia Silva Machado (20221060140113)** do Curso **Bacharelado em Ciência da Computação (Câmpus Anápolis)**. Os trabalhos foram iniciados às **15:00** do dia **17/07/2026** pelo Professor presidente da banca examinadora, constituída pelos seguintes membros:

- **Alexandre Bellezi Jose** (Presidente)
- **Sergio Daniel Carvalho Canuto** (Examinador Interno)
- **Alessandro Rodrigues e Silva** (Examinador Interno)

A banca examinadora, tendo terminado a apresentação do conteúdo do Trabalho de Conclusão de Curso, passou à arguição da candidata. Em seguida, os examinadores reuniram-se para avaliação e deram o parecer final sobre o trabalho apresentado pela aluna, tendo sido atribuído o seguinte resultado:

[X] Aprovado [ ] Reprovado Nota : 7,4

Observação / Apreciações:

Proclamados os resultados pelo presidente da banca examinadora, foram encerrados os trabalhos e, para constar, eu **Alexandre Bellezi Jose** lavrei a presente ata que assino juntamente com os demais membros da banca examinadora.

Documento assinado eletronicamente por:

- **Alexandre Bellezi Jose**, em **17/07/2026 17:12:22** com chave **d20b42b8821b11f1ba4e005056a537a4**.
- **Sergio Daniel Carvalho Canuto**, em **17/07/2026 17:13:45** com chave **033b26e6821c11f1b19f005056a537a4**.
- **Alessandro Rodrigues e Silva**, em **17/07/2026 17:33:59** com chave **d6e53818821e11f1855d005056a537a4**.

Este documento foi emitido pelo SUAP. Para comprovar sua autenticidade, faça a leitura do QrCode ao lado ou acesse https://suap.ifg.edu.br/comum/autenticar_documento/ e informe os dados a seguir.

Tipo de Documento: Ata de Projeto Final

Data da Emissão: 31/07/2026

Código de Autenticação: 1ed43c

Lista de ilustrações

Figura 1 – Exemplo do Preenchimento da Resenha de Equinos 10
Figura 2 – Fluxograma da Visão Computacional 21
Figura 3 – Estrutura de Exemplo das Redes Neurais 22
Figura 4 – Exemplo da Arquitetura CNN 24
Figura 5 – Exemplo de Imagem Adequada para Projeto 31
Figura 6 – Comparação entre Pelagem Branca e Tordilha 32
Figura 7 – Exemplo do resultado da segmentação por instâncias das quatro regiões anatômicas. 33
Figura 8 – Arquitetura Multi-Input proposta para classificação de pelagens equinas. 36
Figura 9 – Matriz Confusão Classificação sem Aumento do Volume Amostral . . . 43
Figura 10 – Matriz Confusão da Avaliação com Aumento do Volume Amostral . . . 45
Figura 11 – Matriz Confusão da Avaliação Refinada 47
Figura 12 – Comparação entre Pelagens: Baio e Castanho 50
Figura 13 – Comparação entre Pelagens: Rosilha e Leopardo 51
Figura 14 – Comparação entre Pelagens: Lobuna e Alazão 52
Figura 15 – Comparação entre Pelagens: Oveira e Castanho 52

Lista de tabelas

Tabela 1 – Exemplos de Pelagens Brancas de Equinos e Muares 13
Tabela 2 – Exemplos de Pelagens Pretas de Equinos e Muares 13
Tabela 3 – Exemplos de Pelagens Alazões de Equinos e Muares 14
Tabela 4 – Exemplos de Pelagens Castanhos de Equinos e Muares 15
Tabela 5 – Exemplos de Pelagens Baias de Equinos e Muares 16
Tabela 6 – Exemplos de Pelagens Tordilhas de Equinos e Muares 17
Tabela 7 – Exemplos de Pelagens Rosilha de Equinos e Muares 18
Tabela 8 – Exemplos de Pelagens Diversas de Equinos e Muares 19
Tabela 9 – Exemplos de Pelagens Manchadas de Equinos e Muares 20
Tabela 10 – Parâmetros de treinamento do modelo de segmentação (YOLOv11-seg). 39
Tabela 11 – Parâmetros de treinamento do classificador Multi-Input ResNet-50. . . 40
Tabela 12 – Acurácia da Classificação sem Aumento do Volume Amostral 42
Tabela 13 – Métricas da Classificação sem Aumento do Volume Amostral 42
Tabela 14 – Avaliação com Aumento do Volume Amostral 45
Tabela 15 – Métricas da Classificação com Aumento do Volume Amostral 45
Tabela 16 – Acurácia da Classificação com Pré-processamento 47
Tabela 17 – Métricas da Classificação com Pré-processamento 47
Tabela 18 – Comparativo global de desempenho por classe e métricas de avaliação sob diferentes cenários experimentais. 48

Lista de abreviaturas e siglas

ACC	Acurácia
CNNs	Redes Neurais Convolucionais (*Convolutional Neural Networks*)
DNNs	Redes Neurais Profundas (*Deep Neural Networks*)
HSI	*Hue, Saturation, Brightness* (Matiz, Saturação e Brilho)
HSV	*Hue, Saturation, Value* (Matiz, Saturação e Valor)
K-Fold	Validação Cruzada K-Fold
K-NN	*K-Nearest Neighbors* (K-Vizinhos Mais Próximos)
LAB	Espaço de Cores CIELAB (Luminosidade, a e b)
mAP	*mean Average Precision* (Precisão Média Média)
SGD	Gradiente Descendente Estocástico (*Stochastic Gradient Descent*)
VC	Visão Computacional
YOLO	*You Only Look Once*

Agradecimentos

Agradeço primeiramente a Deus, pois todas as minhas conquistas e a força para superar os desafios, dentro e fora da instituição de ensino, vêm Dele.

Aos meus familiares, que sempre apoiaram as minhas escolhas, em especial ao meu núcleo familiar, presente em cada etapa desta jornada. Aos meus pais, Diane e Fabricio, e aos meus irmãos, Monique e Vinicius, pelo amor incondicional, carinho e por serem a base fundamental de quem sou hoje.

Aos amigos que construí ao longo dessa trajetória, pelos momentos compartilhados, pelas risadas, pelo apoio mútuo nas dificuldades e pelo incentivo constante. Principalmente ao Genêsis, meu namorado e melhor amigo, que tive a honra de conhecer ao longo do curso, na qual compartilhamos tudo de alegria e tristeza na graduação.

Aos professores do curso de Bacharelado em Ciência da Computação do Instituto Federal de Goiás (IFG), pelos ensinamentos compartilhados. Em especial, ao meu orientador, Prof. Me. Alexandre Bellezi, pela paciência, pelas valiosas orientações técnicas e pelo apoio contínuo dentro e fora de sala de aula.

A todos que de alguma forma contribuíram para a realização deste trabalho, registro minha profunda gratidão!

Resumo

O agronegócio, em especial o segmento da equinocultura, é um setor de expressivo impacto econômico no Brasil. Para a fiscalização e controle zootécnico desses animais, o preenchimento da Resenha Equina constitui um procedimento essencial em âmbito nacional. Embora esse documento seja emitido exclusivamente por médicos veterinários, ainda ocorrem inconsistências visuais no registro das pelagens devido à complexidade fenotípica e à subjetividade da inspeção humana. O presente trabalho propõe uma abordagem em dois estágios baseada em Visão Computacional e Aprendizado Profundo para a identificação e classificação de 11 tipos de pelagens equinas. Inicialmente, desenvolveu-se um conjunto customizado de anotações poligonais manuais em 200 imagens na plataforma Roboflow para treinar a rede YOLOv11-seg, responsável pela segmentação precisa de quatro regiões anatômicas de interesse (cabeça, dorso, membros e crina/cauda). Em seguida, essas partições isoladas alimentam uma arquitetura convolucional *Multi-Input* composta por quatro extratores de atributos ResNet-50 independentes, cujas saídas convergem em uma etapa de fusão tardia (*late fusion*) por concatenação vetorial (8192 dimensões) conectada a uma camada densa final de classificação. O modelo alcançou uma acurácia global de 71,17% no melhor cenário experimental (com expansão amostral), destacando-se com taxas de acerto superiores a 89% em pelagens frequentes como Castanho e Tordilha e viabilizando a predição de classes minoritárias anteriormente negligenciadas.Os resultados obtidos foram significativos para a maioria das classes avaliadas, demonstrando um avanço tecnológico promissor na identificação automatizada dessas características fenotípicas.

Palavras-chave: Visão Computacional. Segmentação por Instâncias. YOLOv11. ResNet-50. Fusão de Características. Pelagens Equinas. Anotação Poligonal.

Sumário

Introdução **9**

1 OBJETIVO **11**
1.1 Objetivo Geral **11**
1.2 Objetivos Específicos **11**

2 REFERENCIAL TEÓRICO **12**
2.1 Pelagem **12**
2.2 Visão Computacional **21**
2.3 Aprendizado Profundo (*Deep Learning*) **22**
2.4 Redes Neurais Convolucionais **23**
2.5 Redes Residuais Profundas **24**
2.5.1 ResNet-50 25
2.6 You Only Look Once (YOLO) **26**

3 TRABALHOS RELACIONADOS **28**
3.1 Estratégias Classificação de Pelagens **28**

4 METODOLOGIA **30**
4.1 Aquisição de Imagens **30**
4.2 Segmentação por Anotação de Polígonos **31**
4.3 Pré-processamento **33**
4.4 Reconhecimento de Padrões **34**
4.4.1 Configuração Experimental e Parametrização do Treinamento 37
4.5 Resultados e Avaliações **37**
4.5.1 Partição dos Dados 37
4.5.2 Métricas de Desempenho Computacional 38
4.5.3 Parametrização e Configuração dos Modelos 39
4.5.4 Validação Comparativa 40

5 RESULTADOS E DISCUSSÕES **42**
5.1 Avaliação sem Aumento do Volume Amostral **42**
5.2 Avaliação com Aumento do Volume Amostral **44**
5.3 Avaliação de Dados com Pré-processamento **46**
5.4 Análise Comparativa e Discussão Global dos Cenários **48**
5.5 Análise de Desempenho por Pelagem **49**
5.5.1 Alazão 49

5.5.2 Baio 49
5.5.3 Branco 50
5.5.4 Castanho 50
5.5.5 Leopardo 51
5.5.6 Lobuna 51
5.5.7 Oveira 51
5.5.8 Pampa 52
5.5.9 Preta 53
5.5.10 Rosilha 53
5.5.11 Tordilha 53

6 CONSIDERAÇÕES FINAIS 54
6.1 Objetivos Alcançados 54
6.2 Trabalhos Futuros 54

REFERÊNCIAS 56

Introdução

A pelagem é a coloração dada ao revestimento do conjunto de pele, pelo, crina, cauda e extremidade dos equídeos que acabam sendo o principal meio para sua identificação precisa. Para fins de notificação obrigatória, transporte, participação de competições utiliza-se a resenha, documento essencial que formaliza essas características para o registro em Associações de Criadores (SANTOS; KREBS; GODOI, 2021).

O preenchimento do documento é necessário especificar vários campos como nome, raça, sexo, idade, gestação, pelagem do animal, resenha gráfica do animal (desenhos na figura do cavalo) com o objetivo de que a identidade do cavalo seja verificada facilmente e única para cada animal, portando é necessário que os elementos característicos do animal para seu reconhecimento seja assegurado (EQUALLI, 2018), como mostra na Figura 1.

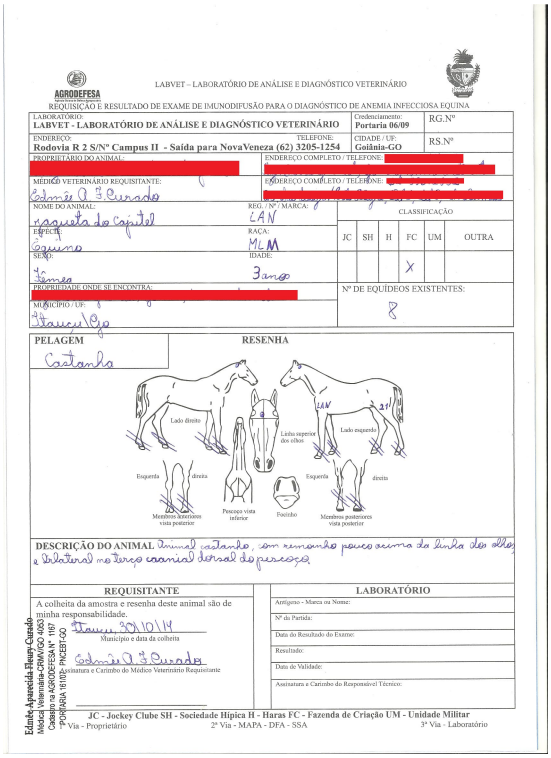
Para o preenchimento da resenha é formalizado, somente por médicos veterinários, as características do equino tanto de maneira descritiva, com particularidades como raça, idade e nome, quanto graficamente indicando manchas que individualizam o animal como manchas específicas indicando como são e onde está localizada no cavalo. Mas o modo como a resenha é feita atualmente acarreta várias falhas significativas. A resenha é um documento que obrigatoriamente causa a duplicidade de informações, inserindo as características do animal de forma gráfica e de forma textual, em decorrência a essa duplicidade pode-se acarretar outros erros como a inconsistência de dados. (NEVES, 2018).

As diversas pelagens existentes são classificadas em quatro categorias, sendo em que cada categoria possui vários tipos de pelagem, com coloração do pelagem externo distinto e cada tipo tem diversas variedades que são identificadas pelas diferentes tonalidades da coloração do pelo (AGRODEFESA, 2017). Dessa maneira, criasse-se um cenário na qual se tem uma grande diversidade de pelagens de equídeos tornando uma tarefa complexa não apenas para leigos, mas também para médicos veterinários, uma vez que a diferenciação exige a análise de características visuais muito específicas.

Na literatura recente, o estudo pioneiro de Neves (SILVA, 2019) propôs a automação da identificação de pelagens em equídeos aplicando o algoritmo K-NN sobre 145 imagens. A metodologia baseou-se na extração de médias cromáticas nos espaços RGB e HSV a partir de recortes anatômicos manuais (pescoço, espádua, costado e coxa) após a normalização da iluminação. Contudo, o método evidenciou limitações críticas na diferenciação de pelagens compostas e conjugadas — cuja identificação depende do arranjo espacial e de contrastes locais, e não de valores médios homogêneos de cor —, além de sofrer com o viés do desbalanceamento amostral severo entre fenótipos comuns e raros (SILVA, 2019).

Com base nessa problemática identificada, a fim de aprimorar a confiabilidade das

descrições das resenhas e avançar em relação ao estado da arte sobre as diversas pelagens desses animais a finalidade desse trabalho será criar ferramentas de visão computacional para recomendar aos profissionais possíveis pelagens por meio das imagens fotógraficas de equinos. Nesse trabalho será retratado o protocolo de levatamento dos trabalhos relacionados que inspira a criação da metodologia desse trabalho, obtenção das imagens que serão explorados com modo da captura de imagem dos animais com fotos padronizadas para uma maior eficiência do algoritmo utilizado, a abordagem do pré-processamento, os modelos escolhidos de segmentação e classificação e seus funcionamentos e por fim a validação dos resultados encontrados pelos os modelos escolhidos para o trabalho.



(a) Resenha Preencida



(b) Exemplar equino avaliado

Figura 1 – Exemplo do Preenchimento da Resenha de Equinos

Fonte: Próprio Autor

1 OBJETIVO

1.1 Objetivo Geral

O objetivo geral deste projeto é desenvolver e avaliar uma metodologia baseada em Deep Learning e Transfer Learning utilizando as Redes Neurais Convolucionais (CNNs) para a detecção e classificação das pelagens de equinos.

1.2 Objetivos Específicos

- **Base de Dados:** A príncipio, deve-se coletar e consolidar um conjunto de dados de uma abrangente quantidade de pelagens de cavalos disitintas, para garantir que não haja um treinamento limitado e que sofra de overtffiting por possuir uma pequena quantidade de amostras.
- **Implementação e Treinamento (Transfer Learning):**É fundamental implementar o modelo utilizando a abordagem de Transfer Learning para treinar e parametrizar modelos com e sem transferência de conhecimento.
- **Detecção de Características:** Neste projeto também deve-se analisar a capacidade do modelo de extrair e utilizar características visuais específicas da pelagem, como cor e marcações, que são cruciais para a diferenciação de pelagem.
- **Resultados e Avaliações:** E por fim, avaliar o desempenho do modelo final treinado utilizando métricas de classificação, como acurácia e mAP (mean Average Precision), e realizar uma análise comparativa dos resultados obtidos com diferentes metódos de segmentação.

2 Referencial Teórico

Nesse capítulo serão ressaltados alguns temas que fundamentaram a ideia desse projeto, como o conceito de classificação de pelagem dos equinos e muares, visão computacional e algoritmos relevantes para classificação da pelagem.

2.1 Pelagem

De acordo com a Saúde Animal, foram definidas 76 pelagens diferentes de cavalos, que compõem diversas subdivisões. As principais pelagens encontradas em equídeos são a branca, preta, alazã, castanha, baia, tordilha, rosilha, overo, tobiano, toveira, lobuna, pampa.(CARDOSO; WORTHINGTON, 2018).

Os ou cremelos podem ser branca ou creme bem claro, com crina e cauda brancas, pele cor-de-rosa por tudo e olhos azuis ou com ausência de pigmentação. (Tabela 1). Os pretos são aqueles que têm pêlos da cabeça, pescoço, troncos, membros, crina e cauda na coloração preta. (Tabela 2). Já os alazões são animais de pelos da cabeça, pescoço, troncos, membros, crina e cauda na coloração vermelha, que pode variar do escuro ao amarelado. A crina e/ou cauda podem ser de tonalidade mais clara. (Tabela 3). Em seguida, os castanhos têm pelos da cabeça, pescoço, troncos, membros vermelhos, com crina, cauda e extremidades pretas. (Tabela 4). Os baios têm pelos amarelos que variam do claro ao bronzeado na cabeça, pescoço e tronco, com crina, cauda e extremidades pretas. (Tabela 5). Os tordilhos têm a interpolação de pelos brancos em todo o corpo do animal. Todo equino de pelagem tordilha é produto de um acasalamento em que pelo menos um dos pais é também de pelagem tordilha.(Tabela 6). Os rosilhas possuem interpolação de pelos brancos nas diversas pelagens, com predomínio da pelagem de fundo na cabeça 7). Além desses possuem alguns com as classificações mas específicas pelo seu modelo único de pelagem como é o caso dos lobunas, pampa, tobino, oveira e toveira e Leopardo/Persa (Tabela 8).

Pelagem	Descrição	Imagem
Pombo/Gázeo	Pelos em pele com ausência quase total de pigmentação. Geralmente apenas os olhos se apresentam coloridos.	

Tabela 1 – Exemplos de Pelagens Brancas de Equinos e Muares

Fonte: (IAGRO, 2017)

Pelagem	Descrição	Imagem
Azeviche	Pelos pretos de tonalidade bem forte, que apresenta reflexos azulados.	
Maltinto	Pelagem preta com reflexos avermelhados nas regiões dos flancos e axilas, mas a cabeça sempre predomínio de pelos pretos. Muitas vezes animais que apresentam pelagem preta nascem lobuno e se tonam pretos na desmama.	

Tabela 2 – Exemplos de Pelagens Pretas de Equinos e Muares

Fonte: (IAGRO, 2017) (AGRODEFESA, 2017)

Pelagem	Descrição	Imagem
Amarilha/Palomina	Pelos de tonalidade amarela que pode variar da clara a escura com crina e cauda branca ou creme. Também chamada por algumas associações de Palomina ou Baia amarilha.	
Cereja	Pelos com tonalidade vermelha lembrando cor da cereja.	
Sobre Baia	Cabeça, pescoço, tronco amarelos, com crina, cauda e extremidades avermelhadas.	
Tostada	Pelos do corpo, crina e cauda de tonalidade vermelha escura lembrando café torrado.	

Tabela 3 – Exemplos de Pelagens Alazões de Equinos e Muares

Fonte: (IAGRO, 2017) (AGRODEFESA, 2017)

Pelagem	Descrição	Imagem
Clara	O vermelho da pelagem é de tonalidade clara com crina, cauda e membros pretos sendo que a tonalidade preta dos membros normalmente não atinge toda a canela.	
Escura	Pelos com tonalidade vermelho escuro com crina, cauda e extremidades pretas.	
Pinhão	Pelagem vermelha bem escura, quase preta. Pode ser diferenciada da preta maltinta avaliando-se a cabeça, que na castanha pinhão tem predominância de pelos vermelhos e na preta maltinta é coberta por maioria de pelos pretos.	
Zaina	Pelagem castanha escura ou pinhão que não apresenta particularidades na cabeça e nos membros.	

Tabela 4 – Exemplos de Pelagens Castanhos de Equinos e Muares

Fonte: (IAGRO, 2017)

Pelagem	Descrição	Imagem
Clara	Pelos amarelos de tonalidade clara com crina, cauda e os membros pretos, sendo que a canela não será totalmente clara preta pois terá pelos amarelos.	
Palha	Pelos amarelos bem claros, lembrando a coloração da palha de milho.	
Encerada	Pelos amarelos escuro lembrando cor de cera natural.	
Escura	A cabeça, o pescoço e o tronco apresentam a tonalidade amarela com intensa pigmentação. Crina, cauda e membros são pretos.	

Tabela 5 – Exemplos de Pelagens Baias de Equinos e Muares

Fonte: (IAGRO, 2017)

Pelagem	Descrição	Imagem
Cardã	Pelagem tordilha que apresenta reflexos avermelhados ou amarelados. Comum naqueles animais que nasceram castanhos, alazões ou baios. É uma variedade transitória, pois acontece em uma das fases de clareamento dos animais.	
Clara	Predomínio de pelos brancos na pelagem tordilha.	
Escura/Negra	Tordilho que apresenta pelagem preta com poucos pelos brancos. Acontece no início do clareamento.	
Pedrês/Mosqueado	Quando os pelos vermelhos ou pretos formam pequenos tufos no fundo branco.	
Ruça	Quando não mais se observar no tordilho os pelos pretos da pelagem de origem. O animal terá o corpo recoberto por pelos brancos e sua pele será excessivamente pigmentada nas extremidades, em virtude da migração do pigmento melânico que se acumula nessas regiões.	

Tabela 6 – Exemplos de Pelagens Tordilhas de Equinos e Muares

Fonte: (IAGRO, 2017)

Pelagem	Descrição	Imagem
Rosilha Alazã	Pelagem toda vermelha com interpolação de pelos brancos no pescoço e tronco.	
Rosilha Castanha	Pelagem castanha com interpolação de pelos brancos no pescoço e tronco.	
Rosilha Preta	Pelagem preta com interpolação de pelos brancos no pescoço e no tronco.	

Tabela 7 – Exemplos de Pelagens Rosilha de Equinos e Muares

Fonte: (IAGRO, 2017)

Pelagem	Descrição	Imagem
Lobuna	A pelagem lobuna é também caracterizada pelo predomínio de pelos pretos na cabeça.	
Pampa	É a conjugação de malhas brancas despigmentadas, bem delimitadas, em qualquer outra pelagem. A designação Pampa precede o nome da pelagem de fundo, se a proporção de malhas brancas for maior, ou deve vir depois do nome da pelagem de fundo, se as malhas brancas estiverem em menor proporção. Ocorre em muitas raças de equinos e pôneis	
Tobiana	É a conjugação de malhas brancas despigmentadas, bem delimitadas, em qualquer outra pelagem.	

Tabela 8 – Exemplos de Pelagens Diversas de Equinos e Muares

Fonte: (IAGRO, 2017)

Pelagem	Descrição	Imagem
Leopardo	Leopardo é a variedade da apaloosa quando a malha atinge todo o corpo do animal. Essa malha poderá se estender atingindo outras regiões do tronco e apresenta pintas na pelagem básica.	
Oveira	Composta de malhas despigmentação em fundo de qualquer pelagem. Essas malhas apresentam contorno irregular e não são bem delimitadas como as da pelagem pampa/tobiana pois podem se misturar com a pelagem de fundo.	
Toveira	A pelagem toveira é uma variedade da oveira, sendo uma mistura entre a pelagem oveira e a tobiana. Sua caracterização pode ser feita pela observação das malhas que são irregulares como na pelagem oveira e também grande parte da parte da cabeça do animal apresenta malha despigmentada (malacara) assim como acontece na pelagem oveira, porém as malhas do tronco ultrapassam a região dorsal e normalmente a área despigmentada é maior que a pigmentada.	

Tabela 9 – Exemplos de Pelagens Manchadas de Equinos e Muares

Fonte: (IAGRO, 2017)

2.2 Visão Computacional

A Visão Computacional (VC) estuda e implementa sistemas capazes de "enxergar"por meio de processos artificiais, com a capacidade de reconhecer e extrair informações de objetos que podem parecer tarefas complexas que necessitam de fortes conhecimentos matemáticos e de programação (BARELLI, 2017).

Atualmente, a VC abrange diversas aplicações, como reconhecimento facial, identificação de objetos e análise de padrões em produtos e na medicina. Assim, essa tecnologia unificada a técnicas de aprendizado de máquina, que permitem ao computador aprender e aperfeiçoar seu desempenho em alguma tarefa.

Para a criação de um sistema complexo que extraia informações específicas, como a variação de cores dos equinos e o padrão que cada pelagem possui demanda técnicas sofisticadas. Como por exemplo, o uso de segmentação das imagens é crucial para delimitar o objeto de interesse (o animal) e focar a análise e classificação e identificação de cores da pelagem dos animais. A diversidade de características visuais dos animais, como formas, texturas e cores, chama a atenção dos estudiosos dessa área buscando explorar essas características de forma eficaz (BARROS; PEDRINI, 2023).

Projetos que utilizam técnicas de visão computacional, em geral, estruturam seu funcionamento em etapas sequenciais para garantir a eficiência na extração de informações, conforme ilustrado na Figura 2 em que representa o fluxo de processamento de imagens, no qual a imagem é adquirida, pré-processada, segmentada, tem suas características extraídas, passa pelo reconhecimento de padrões e, por fim, gera um resultado. Outra parte que se torna essencial também é o aprendizado de máquina (Deep Learning) que na visão computacional tem sido um "divisor de águas", proporcionando avanços significativos na capacidade computacionais entenderem e interpretarem informações visuais (MATOS, 2023). É por esse segmento que a abordagem mais adequada para este tipo de trabalho será o uso de redes neurais convolucionais (CNNs), que será usado para a análise de pelagem.



Figura 2 – Fluxograma da Visão Computacional

Fonte: (BARELLI, 2017)

2.3 Aprendizado Profundo (*Deep Learning*)

O *Deep Learning* (Aprendizado Profundo) é um subcampo do aprendizado de máquina fundamentado em redes neurais artificiais com múltiplas camadas de processamento, denominadas redes neurais profundas (*Deep Neural Networks* – DNNs), cuja concepção é inspirada na organização estrutural dos neurônios biológicos (HOLDSWORTH; SCAPICCHIO, 2025). Essas arquiteturas são treinadas sobre expressivos volumes de dados, capacitando-se a extrair representações latentes e a realizar predições de forma automatizada sem a necessidade de engenharia manual de atributos (*handcrafted features*).

A unidade elementar de processamento é o neurônio artificial (ou nó computacional), organizado em camadas interconectadas. As redes profundas distinguem-se das arquiteturas rasas pelo empilhamento de múltiplas camadas ocultas intermediárias (DATABRICKS, 2025), viabilizando uma hierarquia de abstração na qual as camadas iniciais capturam padrões primitivos de baixo nível (bordas, texturas) e as camadas subsequentes compõem representações semânticas de alto nível.

O processamento em cada unidade compreende a soma ponderada dos sinais de entrada multiplicados por seus respectivos pesos sinápticos, acrescida de um termo de polarização (*bias*). Em seguida, esse somatório é submetido a uma função de ativação não linear, cuja finalidade é permitir que a rede modele relações matemáticas complexas não lineares e mapeie o sinal para um intervalo de saturação específico. Esse fluxo de computação propaga-se sequencialmente da camada de entrada até a camada de saída, constituindo a etapa de propagação direta (*feedforward*), conforme ilustrado na Figura 3 (OLIVEIRA, 2024).

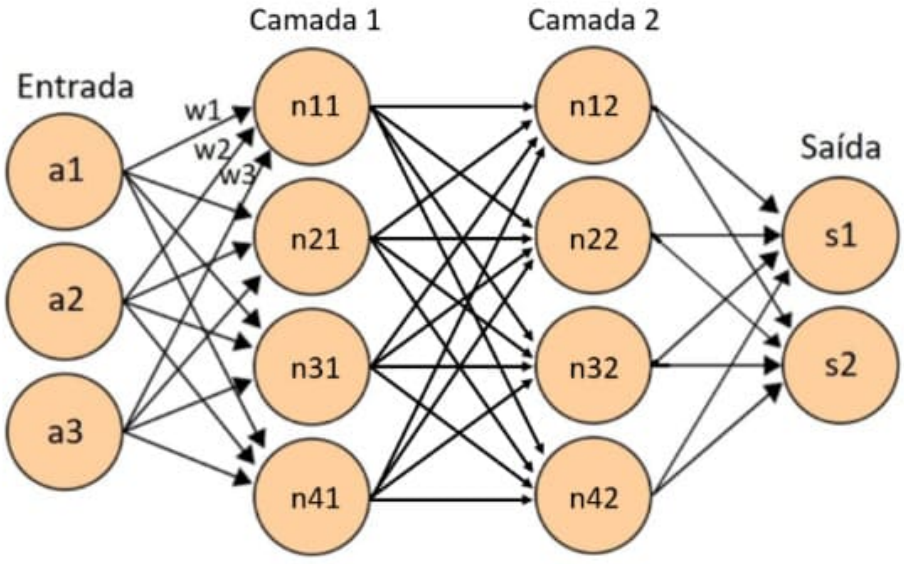


Figura 3 – Estrutura de Exemplo das Redes Neurais

Fonte: (DATABRICKS, 2025)

2.4 Redes Neurais Convolucionais

As redes neurais convolucionais (*Convolutional Neural Networks* ou CNNs) são um subconjunto do aprendizado de máquina utilizado com frequência para tarefas de classificação e visão computacional (IBM, 2025). As CNNs são modelos de *Deep Learning* que têm como foco o processamento visual, aprendendo automaticamente representações de características (*features*) complexas e eliminando a necessidade de extração manual de atributos (*hand-engineered features*).

Uma particularidade importante desta arquitetura, como ilustrado na Figura 8a, está relacionada à importância das camadas convolucionais (*Convolutional Layers* - CL), que consistem em conjuntos de filtros que, combinados a funções de ativação não lineares, percorrem sequencialmente os dados de entrada (ou da camada anterior) para produzir matrizes denominadas mapas de características (*feature maps*) (NEVES, 2017). Esses filtros convolucionais são empregados para extrair informações visuais, aprendendo os parâmetros do *kernel* para aquele conjunto específico de dados e compreendendo, por exemplo, padrões de cores e bordas circulares ou horizontais (SILVA, 2021).

Outra característica fundamental das CNNs é a utilização de camadas de agrupamento (*pooling*). Visto que processar uma imagem em sua resolução original ao longo de todas as camadas é computacionalmente oneroso, o *pooling* reduz a dimensionalidade espacial dos dados. Esse processo não apenas diminui o custo computacional, mas também permite que a rede identifique características em diferentes escalas, conferindo maior robustez e abstração aos padrões extraídos (SILVA, 2021).

O avanço da Visão Computacional tem sido impulsionado pelo desenvolvimento de diversas arquiteturas de CNNs (BARROS; PEDRINI, 2023). Para a classificação de imagens de objetos, destacam-se estudos que utilizam essas redes para a identificação de raças de cavalos, nos quais a arquitetura ResNet-50 alcançou acurácia de 95,90% (ATABAY, 2017).

Contudo, conforme discutido por (CHEN et al., 2024), o avanço contínuo do *Deep Learning* em visão computacional impõe um dilema fundamental entre a profundidade das redes e a sua eficiência paramétrica. Modelos densos exigem a estimativa de milhões de pesos sinápticos, demandando volumes massivos de dados para treinamento e limitando sua aplicabilidade direta em cenários computacionalmente restritos.

Para mitigar a necessidade de grandes bases de treinamento e o custo computacional de redes profundas, adota-se amplamente a técnica de *Transfer Learning* (Aprendizado por Transferência) (ATABAY, 2017). O *Transfer Learning* permite que os pesos de uma rede pré-treinada em um conjunto de dados abrangente e de propósito geral sejam reaproveitados para uma tarefa específica com menor disponibilidade amostral. Isso é realizado por meio da modificação e do retreinamento das camadas finais da arquitetura. Quando apenas as

camadas densas totalmente conectadas são ajustadas, o processo atua como um extrator fixo de características; por outro lado, quando camadas convolucionais intermediárias também são atualizadas, o procedimento é denominado *fine-tuning* (ajuste fino) (ATABAY, 2017).

Um padrão complexo que deve ser considerado é a classificação cromática dos animais, apontada como uma das propriedades determinantes na identificação de uma pelagem. Em contexto análogo, um estudo desenvolvido pelo Departamento de Engenharia de Software da Universidade Cibernética de Sejong aplicou CNNs com foco estrito na detecção de cores em veículos automotores (KIM, 2024).

Embora as CNNs sejam convencionalmente empregadas para o reconhecimento de contornos e formas geométricas, as camadas convolucionais iniciais são responsáveis por capturar características de baixo nível, como arestas, texturas e as próprias cores. À medida que os dados avançam pelas camadas mais profundas, a rede aprende a sintetizar essas informações para estruturar padrões abstratos e de alto nível, os quais apresentam maior robustez a pequenas variações de iluminação ao correlacionar textura, cor e morfologia (KIM, 2024).

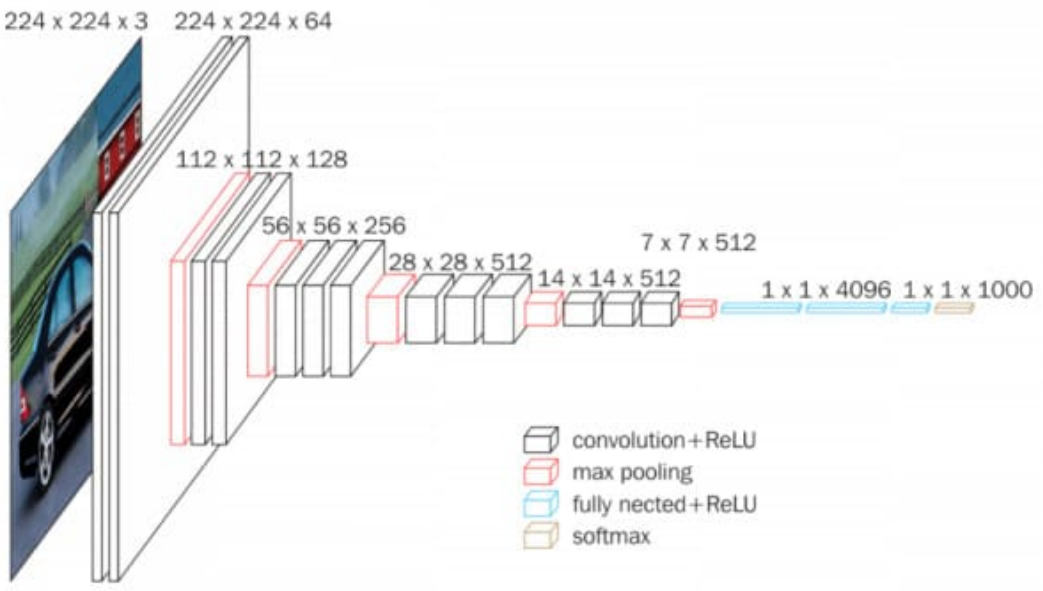


Figura 4 – Exemplo da Arquitetura CNN

Fonte: (SILVA, 2021)

2.5 Redes Residuais Profundas

As Redes Neurais Residuais (*Residual Networks* ou ResNets) constituem uma classe especializada de Redes Neurais Convolucionais projetada para viabilizar o treinamento eficiente de modelos extremamente profundos (ULTRALYTICS, 2026). Tradicionalmente,

o empilhamento sucessivo de camadas convolucionais enfrenta o problema da degradação do desempenho e o desvanecimento do gradiente (*vanishing gradient*), fenômeno no qual os sinais de erro atenuam-se excessivamente durante a etapa de retropropagação (*backpropagation*), impedindo a convergência da rede (COSTA, 2024).

A robustez dos blocos residuais torna essa família de arquiteturas especialmente vantajosa para tarefas complexas de visão computacional, como a extração hierárquica de características morfológicas e texturais em imagens de alta resolução. No contexto da identificação de pelagens equinas, como em modelos que precisam distinguir desde padrões cromáticos uniformes até transições e manchas intrincadas, a capacidade de reter detalhes espaciais em diferentes níveis de abstração é determinante. Dentre as configurações existentes dessa arquitetura, a variante **ResNet-50** destaca-se por incorporar blocos residuais com estrutura de gargalo (*bottleneck*), alcançando um equilíbrio ótimo entre profundidade representacional, capacidade de generalização e custo computacional (ULTRALYTICS, 2026).

2.5.1 ResNet-50

A ResNet-50 é uma arquitetura de rede neural convolucional caracterizada pela sua estrutura composta por 50 camadas, projetada para atingir um equilíbrio ótimo entre custo computacional e capacidade de representação e precisão preditiva (ULTRALYTICS, 2026). Devido à sua eficácia e estabilidade no treinamento, essa variante é amplamente adotada em tarefas complexas de visão computacional, incluindo detecção de objetos e classificação fina de imagens (KIM, 2024).

A aplicabilidade da ResNet-50 na extração de atributos cromáticos é evidenciada no estudo desenvolvido por KIM (2024) (KIM, 2024). A proposta do autor consiste em um sistema capaz de identificar e classificar, em tempo real, os tipos e cores dos veículos detectados, fornecendo suporte tecnológico baseado em uma maior densidade de informações (KIM, 2024). A metodologia empregada utiliza uma CNN ResNet-50 pré-treinada como rede base para o treinamento do YOLO v4. Após a detecção do veículo na sequência de imagens, a *bounding box* (caixa delimitadora) gerada é utilizada para a classificação subsequente do tipo e da cor do objeto

Para isso, o estudo utilizou um conjunto de dados composto por 32.500 imagens para detecção e 156.000 imagens para a classificação de cores, adquiridas via DashCam em ambientes rodoviários coreanos. As amostras foram categorizadas em quatro tipos de veículos (sedã, SUV, caminhão e ônibus) e oito categorias de cores (branco, preto, prata, amarelo, vermelho, laranja, azul e verde) (KIM, 2024).

A arquitetura foi adaptada por meio de *transfer learning*, em que as camadas finais (*fully connected* e *softmax*) foram reconfiguradas para as classes específicas do estudo

(KIM, 2024). Como resultados, o método alcançou acurácias de 91,5% na detecção, 93,9% na classificação de tipos e 94,2% na classificação de cores (KIM, 2024).

Destaca-se que, enquanto Kim (KIM, 2024) foca na identificação de veículos para navegação autônoma, este trabalho adapta a lógica de extração de características cromáticas para a identificação de pelagens em equídeos. Dada a eficácia do uso de YOLO v4 e Redes Neurais Convolucionais (CNNs), em especifico a ResNet-50, com *transfer learning* demonstrada pelo autor, a proposta deste trabalho é seguir uma metodologia similar, aproveitando a robustez dessas arquiteturas e adaptando para o contexto biológico.

2.6 You Only Look Once (YOLO)

Como foi explorado na seção da visão computacional, depois da aquisição das imagens, uma etapa fundamental para o identificação das características é o pré-processamente para isso é de suma importância que seja feita a segmentação das regiões de interesse da imagem. Neste sentido, a segmentação de imagens é a tarefa de agrupar partes de uma imagem que pertencem à mesma classe de objeto. Esse processo também é chamado de classificação em nível de pixel (COMPUTACIONAL, 2022).

E as regiões de interesse os elementos presentes na imagem que queremos identificar e obter informações. Com os objetos de interesse definidos, algumas técnicas de pré-processamento para destacar bordas e formas geométricas e tratar ruídos são aplicadas na imagem para realçá-la, facilitando a obtenção de suas informações pelo sistema de Visão Computacional (BARELLI, 2017).

Enquanto a classificação lida com um único objeto e a imagem como um todo, a Detecção e a Segmentação se concentram em cenários mais complexos com múltiplos objeto, localizando os objetos de diferentes classes em uma cena (OLIVEIRA, 2024). Para tal tarefa, um dos algoritmos que tem mais destaque é o You Only Look Once (YOLO) é uma CNN de "estado da arte" na detecção de objetos em tempo real. Seu diferencial é processar a imagem apenas uma vez (You Only Look Once) para realizar todas as detecções, tornando-o muito mais rápido que métodos anteriores.

No artigo, Detecção e Classificação de derramamento de Óleo na superfície oceânica baseada em aprendizagem profunda via algoritmo Yolo a autora apresenta como objetivo central a proposição de uma metodologia baseada em aprendizagem profunda para o monitoramento e mitigação de impactos socioambientais causados por desastres petrolíferos. O estudo utilizou um conjunto de dados proveniente da missão SENTINEL-1, composto por 1.112 imagens de Radar de Abertura Sintética (SAR) em tons de cinza, com a metodologia baseada na família de algoritmos YOLOv8 (modelos nano, small e medium), utilizando a arquitetura composta pelas camadas backbone, neck e head (SILVA, 2024).

Como resultados, o modelo YOLOv8 *medium* obteve o desempenho mais satisfatório, alcançando uma precisão de 89,3%, recall de 80,6%, mAP50 de 85% e mAP50-95 de 71,6% (SILVA, 2024). Com relação às principais diferenças entre o trabalho de (SILVA, 2024) e este estudo, destaca-se que, enquanto a autora foca em manchas de óleo em imagens de satélite, este trabalho adapta a lógica de detecção e remoção de fundo para a identificação de pelagens equinas. A inspiração no uso de segmentação por máscaras de (SILVA, 2024) é aqui aplicada para garantir que, após a detecção do animal, o ambiente externo ao animal seja removido, permitindo que a extração de características das cores se concentre exclusivamente no pelagem do equídeo, reduzindo ruídos ambientais na geração da resenha animal.

O projeto de classificação de pelagem, como a identificação de raças de cavalos ou padrões de cores, utiliza diretamente os princípios de Classificação e Detecção de Imagens baseados em CNNs. A classificação da pelagem depende da capacidade da rede neural de extrair características complexas: nas camadas iniciais, a CNN aprende a distinguir texturas e cores básicas que definem o padrão da pelagem. Para garantir que a classificação se concentre apenas no objeto de interesse (o animal ou a área da pelagem) e ignore o fundo, técnicas de Detecção de Objetos, como o algoritmo YOLO, e a utilização de Bounding Boxes são essenciais. Isso permite isolar a região da pelagem, aumentando a precisão da rede ao garantir que ela esteja classificando as características do objeto e não do ambiente circundante, um passo similar ao usado para localizar defeitos em tecidos ou manchas de óleo em imagens de satélite. (SILVA, 2024).

De maneira complementar, o estudo **End-to-end human parsing and detection optimized for resource-constrained devices** os autores abordam a tarefa de *human parsing*, que consiste na segmentação detalhada de partes do corpo e peças de roupa para associação individual. Para alcançar essa eficiência, a arquitetura baseia-se na estrutura da rede YOLO, um diferencial metodológico importante proposto no estudo é a utilização de anotações por polígonos delimitadores para realizar a detecção e a segmentação simultaneamente. Em vez de depender exclusivamente de máscaras densas e tradicionais baseadas em pixels, o modelo identifica os contornos morfológicos convertendo-os em polígonos representados por uma lista de vértices diretos.

A lógica por trás do método de segmentação por polígonos demonstra-se altamente pertinente para a identificação automatizada de equinos. Como a classificação de pelagens não deriva de um padrão visual global único, mas exige a análise combinada de regiões anatômicas essenciais e distintas do cavalo, especificamente dorso, cabeça, membros, crina e cauda (AGRODEFESA, 2017), a adaptação desse conceito de *parsing* permite que a rede neural isole e processe essas estruturas com precisão espacial. Dessa forma, a arquitetura computacional aproxima-se da avaliação morfológica segmentada que é realizada por especialistas humanos no domínio (HOSEN; ISLAM, 2025).

3 Trabalhos Relacionados

Este capítulo apresenta uma revisão de trabalhos que abordam metodologias e tecnologias fundamentais para a estruturação desta pesquisa. O objetivo é mapear diferentes abordagens presentes na literatura para fundamentar as escolhas metodológicas aqui adotadas, além de situar este estudo no contexto do estado da arte.

3.1 Estratégias Classificação de Pelagens

O trabalho Identificação de Cores de Equinos e Muares desenvolvida por (SILVA, 2019), cujo escopo central reside na automação do reconhecimento da pelagem de equídeos. O desenvolvimento do conceito foi motivado pela recorrência de inconsistências e erros humanos durante o preenchimento manual da Resenha Clínica Animal (NEVES, 2018). Para mitigar o problema, o autor implementa uma abordagem computacional utilizando o algoritmo de aprendizado supervisionado K-Nearest Neighbors (K-NN) (SILVA, 2019).

No pipeline metodológico do estudo de referência, estruturou-se um banco de dados composto por 145 imagens (SILVA, 2019). As amostras visuais foram submetidas a um algoritmo de normalização de iluminação (*Colour Invariants*) e, consecutivamente, a um processo de segmentação manual conduzido pelo autor, gerando até três recortes por espécime de regiões anatômicas específicas: pescoço, espádua, costado e coxa (SILVA, 2019). A classificação baseou-se no cálculo da média aritmética dos vetores de características extraídos dos espaços de cores RGB e HSV pertencentes a cada recorte, associando o padrão numérico à pelagem de maior probabilidade estatística (SILVA, 2019).

Contudo, a abordagem descrita evidenciou limitações severas, em especial na categorização de pelagens do tipo compostas e conjugadas, cuja identidade taxonômica não é definida puramente por uma cor sólida e uniforme, mas pela distribuição espacial de padrões cromáticos em áreas topográficas específicas do corpo do animal (SILVA, 2019). Soma-se a isso o viés gerado pelo desequilíbrio de classes decorrente da escassez de registros fotográficos para pelagens raras (SILVA, 2019). Diante desse cenário e considerando os avanços contemporâneos em arquiteturas de redes neurais convolucionais e segmentação de instâncias, a presente pesquisa propõe ser um artigo metodológico com a finalidade de transpor tais barreiras ao implementar métodos robustos para a classificação geral de pelagens equinas.

No estudo Deep Learning for Horse Breed Recognition, desenvolvido por (ATABAY, 2017), explora-se a identificação de raças de cavalos em imagens de cenários naturais por meio de métodos de Aprendizado Profundo (Deep Learning). Devido à indisponibilidade

de bases de dados públicas específicas para essa finalidade na época, o autor compilou um conjunto contendo 1.693 imagens distribuídas entre seis raças distintas. A metodologia baseou-se na técnica de Transfer Learning, empregando redes neurais convolucionais (CNNs) previamente treinadas no conjunto ImageNet. Foram avaliadas arquiteturas consolidadas como VGG16, VGG19, Inception V3, Xception e ResNet50.

Os resultados apontaram a superioridade da arquitetura ResNet50, que atingiu a maior acurácia média de classificação, com 95,90% (ATABAY, 2017). A pesquisa em questão fornece um excelente ponto de partida, pois valida empiricamente a eficácia da ResNet50 na extração de características visuais de equídeos, mesmo diante de desafios como a variação de iluminação e a estrutura não rígida do corpo do animal (ATABAY, 2017).

Contudo, a abordagem descrita na literatura utiliza a imagem global do cavalo para inferir a sua linhagem genética. A presente pesquisa utiliza esse conhecimento como base, mas propõe um avanço metodológico distinto ao focar na classificação de pelagens. Para lidar com as sutilezas visuais desse domínio, a estratégia aqui desenvolvida não processa a imagem como um bloco único, mas aplica algoritmos de segmentação para isolar áreas anatômicas críticas, alimentando essas características de forma independente em uma arquitetura de múltiplas entradas.

4 Metodologia

A metodologia empregada neste projeto baseia-se em um fluxo de trabalho semelhante ao adotado em pesquisas de Visão Computacional, conforme fundamentado na literatura teórica (BARELLI, 2017) e em estudos aplicados da área (KIM, 2024). Embora a abordagem do autor original trabalhe com a detecção de tipos e cores de veículos, a arquitetura baseada na detecção de objetos (carros) e suas cores é bem análogo para a detecção de pelagem equina. Esta transposição justifica-se pela robustez do projeto na identificação de objetos cromáticos e sua eficiência de o uso de redes-pré treinadas para tarefas de identificação visual.

4.1 Aquisição de Imagens

A aquisição de imagens é uma etapa fundamental que tem como foco principal garantir a qualidade e o volume adequado dos dados. Essa importância justifica-se pelo fato de que as Redes Neurais Convolucionais dependem estritamente das características extraídas desse conjunto; logo, o uso de informações de baixa qualidade comprometerá a tomada de decisão e não produzirá classificações precisas (OLIVEIRA, 2025).

Como já foi citado anteriormente, a pelagem é o revestimento externo do animal, caracterizado pela coloração do conjunto formado por pele, pelos, crina e cauda (EQUALLI, 2018). Por isso, o conjunto de dados de imagens das diversas pelagens dos equinos coletados, visa incluir o corpo inteiro do cavalo para que o conjunto de todo o animal seja analisado a fim de identificar por várias características de todo corpo do animal para determinar a pelagem pela rede neural. Com uma padronização das imagens para que não haja interferência no resultado como ângulo, iluminação, tamanho das imagens (KIM, 2024).

O *dataset* utilizado neste trabalho foi construído a partir do repositório Kaggle, uma plataforma que hospeda conjuntos de dados para análise e aprendizado de máquina (ICICT/FIOCRUZ, 2025), utilizando-se especificamente os conjuntos de dados (BELITSKAYA, 2021) e (NNEJI, 2022). Adicionalmente, integraram-se dados originários do RoboFlow, uma plataforma em nuvem voltada ao desenvolvimento simplificado de projetos de visão computacional (ROBOFLOW, 2026), de onde foram extraídos os bancos de imagens públicos (HARDWARE, 2026) e (DLABEQUICARE, 2024).

O processo de curadoria e seleção das amostras seguiu critérios metodológicos estritos: priorizaram-se fotografias do equídeo em vista lateral completa, garantindo que o animal fosse o único objeto de interesse na cena, com total ausência de outros elementos (como humanos ou múltiplos animais). Outrossim, foram selecionados espécimes sem

adereços que ocultassem partes significativas da pelagem. Essa exigência assegura a visibilidade desimpedida das regiões anatômicas essenciais do cavalo, como dorso, cabeça, membros, crina e cauda, mitigando ruídos visuais e otimizando o desempenho dos algoritmos de segmentação e classificação. Tais medidas garantem que, nas etapas subsequentes de segmentação e extração de coordenadas, as variáveis ambientais não interfiram na precisão do modelo, resultando em uma abordagem mais robusta e eficiente.



Figura 5 – Exemplo de Imagem Adequada para Projeto

Fonte: (BELITSKAYA, 2021)

4.2 Segmentação por Anotação de Polígonos

A anotação poligonal é uma técnica de rotulagem de dados usada em visão computacional para delinear objetos em imagens ou vídeos, desenhando formas com múltiplos lados, ou polígonos, ao redor deles. Conectando uma série de pontos (vértices) ao longo das bordas de um objeto, os anotadores criam uma forma fechada que captura com precisão os contornos do objeto (BANKAS, 2025). Portanto, a anotação poligonal torna-se indispensável na preparação dos dados de treinamento, permitindo que as Redes Neurais identifiquem com exatidão os objetos de interesse presentes na imagem.

A classificação precisa da pelagem exige a análise conjunta de regiões específicas do animal, que serão denominadas como os objetos de interesse nesse artigo, que irá englobar a cabeça, a região dorsal, a crina, a cauda e as extremidades dos membros (AGRODEFESA, 2017).

Para a localização dos animais, utiliza-se uma abordagem de Detecção de Objetos com a arquitetura YOLO (*You Only Look Once*) para identificar a *bounding box* do animal e isolar a área da pelagem, removendo o fundo da imagem. Esse passo impede

que ruídos do ambiente afetem negativamente o reconhecimento, de forma semelhante ao trabalho de (SILVA, 2024), no qual a identificação de manchas de petróleo obteve resultados expressivamente melhores após a supressão do fundo.

Complementada a esse metódo, a aplicação da anotação por polígonos consolidou-se como um pilar fundamental para treinar modelos na interpretação de dados visuais complexos (BANKAS, 2025). Nesse contexto, a rede YOLO não apenas localiza o equino em sua totalidade, mas também delimita suas regiões anatômicas de interesse para a classificação. Analisar o animal inteiro sem essa separação dificulta o processo, pois as distinções entre certas pelagens baseiam-se em detalhes sutis. Por exemplo, equinos tordilhos e brancos possuem a região dorsal visualmente semelhante; no entanto, o tordilho apresenta focinho e patas com pigmentação escura, além de crina e cauda frequentemente amareladas ou acinzentadas. Em contrapartida, o cavalo branco possui extremidades rosadas (despigmentadas) e pelos da crina e cauda totalmente brancos, conforme ilustrado na Figura 6.



(a) Exemplar Cavalo Pelagem Branca (b) Exemplar Cavalo Pelagem Tordilha

Figura 6 – Comparação entre Pelagem Branca e Tordilha

Fonte: (IAGRO, 2017) (AGRODEFESA, 2017)

A fim de viabilizar esse nível de detalhamento, optou-se pela arquitetura YOLOv11 na sua variante de segmentação de instâncias (*YOLOv11-seg*). Devido à ausência de conjuntos de dados públicos pré-anotados especificamente para as partes anatômicas de equídeos e à grande variação morfológica entre as raças, desenvolveu-se um conjunto customizado de anotações poligonais. O processo de rotulagem manual foi conduzido na plataforma *Roboflow* sobre um conjunto de 200 fotografias de equinos em vista lateral, delimitando com precisão os vértices de quatro regiões de interesse anatômico: *horse body* (dorso/tronco), *horse head* (cabeça), *horse legs* (membros) e *horse hair* (crina e cauda), conforme ilustrado na Figura 7.

O modelo foi treinado ao longo de 105 épocas, monitorando-se as funções de perda de localização (*Box Loss*), classificação (*Class Loss*) e presença de objeto (*Object Loss*). No conjunto de validação, o modelo alcançou um desempenho global satisfatório, registrando

uma Precisão Média de mAP_{50} de 88,3% (com mAP_{50-95} atingindo aproximadamente 57,0%), Precisão de 90,2%, Revocação (*Recall*) de 82,4% e F_1-Score de 86,1%.

Na avaliação individualizada por classe anatômica, a rede apresentou alta capacidade discriminativa para estruturas volumétricas bem definidas, atingindo mAP_{50} de 99,0% para a cabeça (*horse head*) e 95,0% para o dorso (*horse body*). Para as regiões mais finas e morfologicamente variáveis, como membros (*horse legs*) e crina/cauda (*horse hair*), o classificador obteve mAP_{50} de 82,0% e 78,0%, respectivamente. Esses resultados garantem a extração de máscaras e coordenadas precisas para o recorte automatizado (*crop*) e supressão de fundo das regiões anatômicas, viabilizando a alimentação consistente da arquitetura *Multi-Input* na etapa de reconhecimento de padrões.

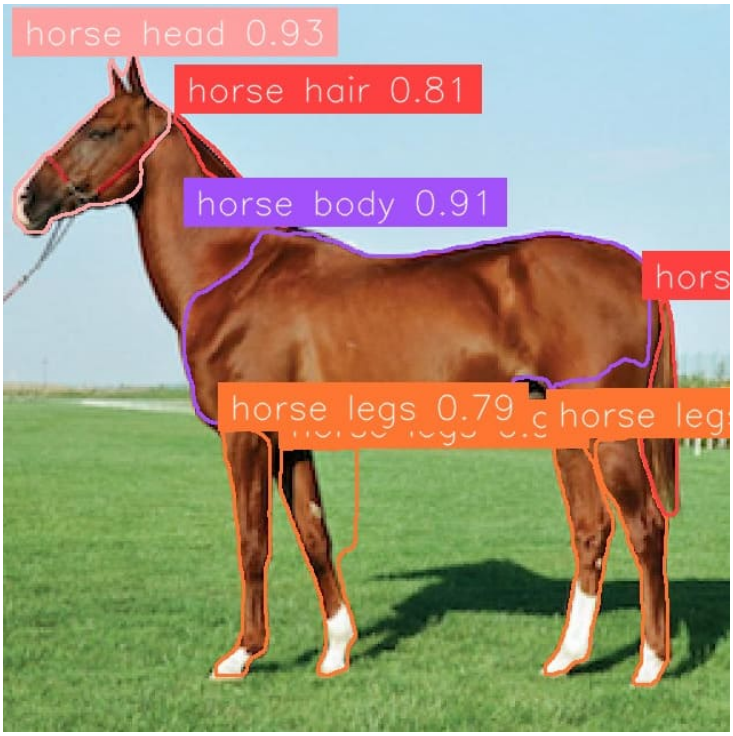


Figura 7 – Exemplo do resultado da segmentação por instâncias das quatro regiões anatômicas.

Fonte: Próprio Autor.

4.3 Pré-processamento

Em fluxos de Visão Computacional, a etapa de pré-processamento é imprescindível para garantir a robustez da arquitetura, conforme abordado por (KIM, 2024). Uma vez que as imagens utilizadas no projeto são provenientes de diferentes fontes, parte delas apresenta alto nível de pixelização, contornos mal definidos ou variações de iluminação que interferem na percepção das cores dos equinos, sendo necessária uma padronização prévia dos dados.

Inicialmente, as imagens são normalizadas para o intervalo $[0, 1]$ (ponto flutuante). Para destacar os contornos do animal e a textura dos pelos, aplica-se a técnica de máscara de nitidez (*unsharp masking*). Esta etapa consiste na subtração de uma versão suavizada da imagem original para gerar uma máscara de detalhes, que é posteriormente somada à imagem base, ressaltando as bordas sem a introdução de artefatos digitais excessivos (KIM, 2024). Tal procedimento foi adotado com o intuito de destacar não apenas a coloração original dos pelos, mas, principalmente, para evidenciar as marcações características nas pelagens dos equídeos.

Com o objetivo de evitar interferências negativas na etapa de segmentação, os métodos de realce citados foram aplicados exclusivamente após a conclusão da segmentação das imagens. Dessa forma, a avaliação da influência desses filtros foi isolada e direcionada especificamente para o desempenho da Rede Neural na etapa de classificação. Adicionalmente, as imagens foram redimensionadas para atender às dimensões de entrada da CNN, conforme a metodologia de classificação de imagens de veículos descrita por (KIM, 2024).

4.4 Reconhecimento de Padrões

A extração de características visuais será conduzida pela arquitetura ResNet-50, uma Rede Neural Convolucional reconhecida por sua capacidade de mitigar o problema do desvanecimento do gradiente por meio de conexões residuais (VINA, 2025). Essa robustez arquitetural a torna altamente eficaz na extração de padrões complexos, como demonstrado em trabalhos de classificação de veículos (KIM, 2024) e na identificação de raças equinas (ATABAY, 2017).

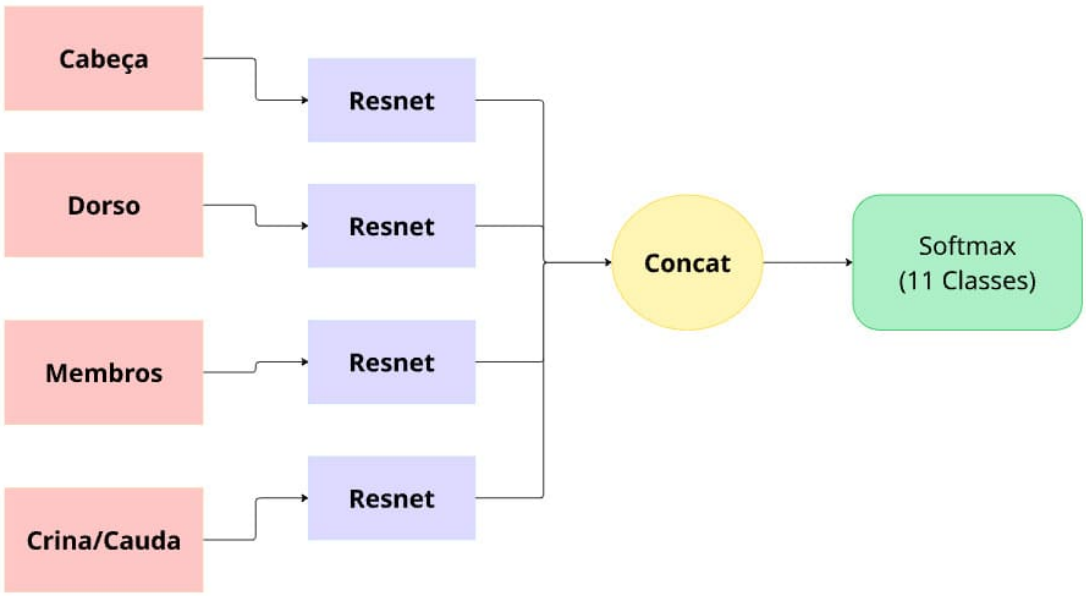
Devido à vasta diversidade fenotípica de pelagens equinas, o escopo deste projeto foi delimitado a 11 categorias principais: Alazão, Baia, Branca, Castanho, Leopardo, Lobuna, Oveira, Pampa, Preta, Rosilha, Tordilha. Para atender às especificidades morfológicas das pelagens, o modelo adotado não processará o cavalo de forma global, mas sim através de uma arquitetura de múltiplas entradas (Multi-Input). A rede extrairá características de forma independente a partir de quatro regiões anatômicas isoladas (cabeça, dorso, membros, crina e cauda). Os vetores de características profundas de cada região serão concatenados na camada final, permitindo uma classificação conjunta e baseada em evidências locais, essa arquitetura proposta é mostrada na Figura 8a.

O processo de aprendizado fundamenta-se na técnica de Transfer Learning, aproveitando os pesos pré-treinados no conjunto ImageNet (VINA, 2025). As camadas convolucionais iniciais, responsáveis por extrair formas geométricas elementares, serão mantidas congeladas (VRBANčIč, 2020). O ajuste fino (Fine-Tuning) ocorrerá na última camada totalmente conectada (Fully Connected), a qual foi redimensionada para receber a fusão das quatro regiões anatômicas e adaptada para o número exato de classes de pelagens

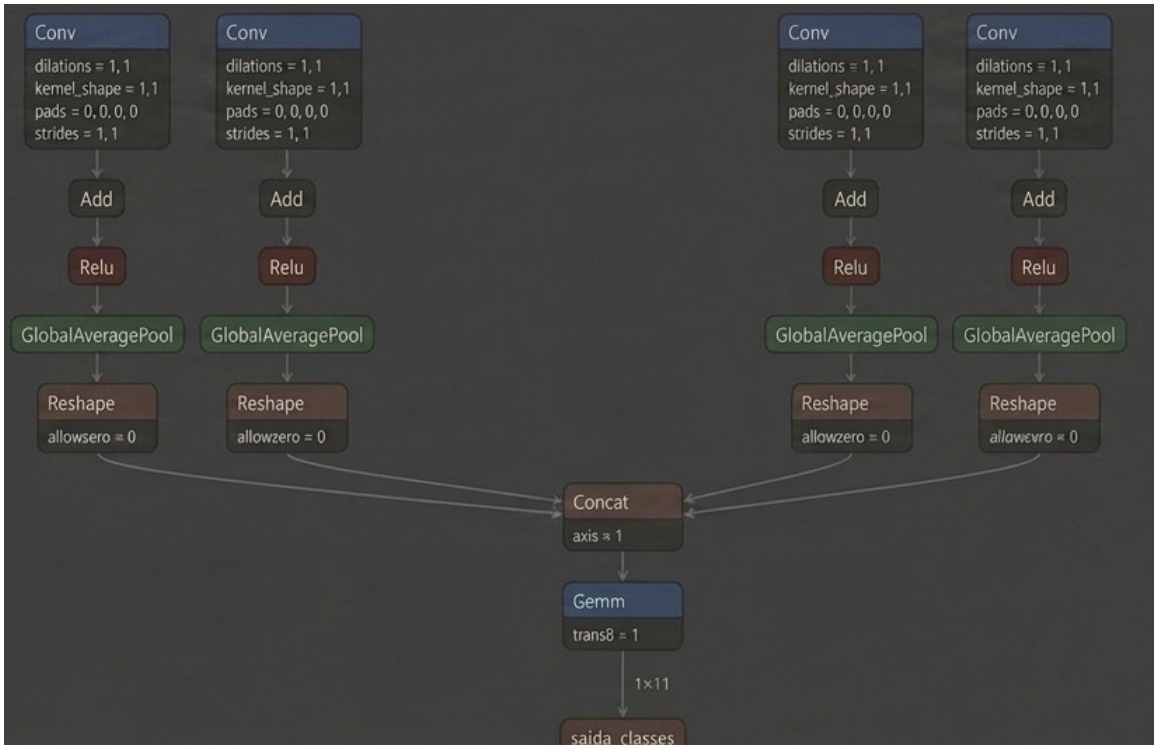
mapeadas no projeto. Para otimizar o aprendizado e mitigar o impacto do desbalanceamento natural do banco de dados, a função de perda será penalizada ativamente utilizando pesos inversamente proporcionais à frequência de cada classe. Essa estratégia assegura que pelagens com menor volume amostral obtenham a devida representatividade na atualização dos pesos da rede, reduzindo o custo computacional e maximizando a eficácia em conjuntos de dados limitados (VRBANčIč, 2020).

A etapa de fusão de atributos foi implementada por meio de uma estratégia de fusão tardia (*late fusion*), integrando as representações visuais profundas extraídas de cada uma das quatro regiões anatômicas isoladas (cabeça, dorso, membros e crina/cauda). Conforme detalhado no grafo computacional da Figura 8b, cada partição imagética passa de forma independente pelas camadas convolucionais da sua respectiva sub-rede ResNet-50 até alcançar a camada de agrupamento médio global (*Global Average Pooling*). Essa operação reduz os mapas de ativação espaciais bidimensionais a um vetor unidimensional de características de 2048 dimensões por região, o qual é linearizado pela operação de *Reshape* para assegurar a compatibilidade dimensional.

Em seguida, os quatro vetores resultantes são combinados ao longo do eixo das características ($dim = 1$) por meio de um operador de concatenação (*Concat*), consolidando as informações locais em um vetor representativo unificado de 8192 dimensões (2048×4). Esse vetor concatenado sintetiza a morfologia cromática global do animal e alimenta uma camada totalmente conectada final (*Gemm/Linear*), responsável por realizar a projeção linear para as 11 classes de pelagens mapeadas. Por fim, a distribuição de probabilidades normalizada sobre o espaço amostral de fenótipos é obtida pela aplicação da função *Softmax*.



(a) Arquitetura Multi-Input: Extração de Características por Região Anatômica e Fusão de Dados.



(b) Grafo computacional detalhando a convergência dos 4 ramos no operador *Concat* e na camada densa final (1×11).

Figura 8 – Arquitetura Multi-Input proposta para classificação de pelagens equinas.

Fonte: Próprio Autor.

4.4.1 Configuração Experimental e Parametrização do Treinamento

O treinamento foi formulado como um problema de classificação multiclasse supervisionado. Para mitigar o impacto do desbalanceamento natural da base de dados entre pelagens comuns e raras, adotou-se a função de perda Entropia Cruzada Ponderada (*Weighted Cross-Entropy Loss*), na qual pesos penalizadores associados a cada classe são calculados dinamicamente de forma balanceada e inversamente proporcional à sua frequência de ocorrência no subconjunto de treino de cada partição (*fold*) (AURELIO, 2017).

A otimização dos pesos da camada totalmente conectada foi conduzida pelo algoritmo Gradiente Descendente Estocástico com Momento (*SGD with Momentum*). Nesse método, os parâmetros treináveis são atualizados iterativamente na direção oposta ao gradiente da função de custo, com a magnitude do ajuste controlada pela taxa de aprendizado (*learning rate*). Adicionalmente, o mecanismo de momento atua como um termo de inércia ou memória baseado no histórico das iterações anteriores, acelerando a convergência em trajetórias consistentes e amortecendo oscilações indesejadas (LIMA; MIRANDA; FARIAS, 2026). Para a parametrização do classificador, definiu-se a taxa de aprendizado inicial em $\eta = 0{,}01$ e o fator de momento em 0,9, restringindo a atualização de gradientes estritamente à camada linear de fusão.

As entradas correspondentes às quatro regiões anatômicas isoladas (dorso, membros, crina e cabeça) foram padronizadas na resolução de $224 \times 224 \times 3$ pixels e organizadas em lotes (*batch size*) de 4 amostras. O treinamento do classificador foi executado ao longo de 10 épocas por partição no *framework PyTorch* (versão 2.5.1), empregando aceleração por hardware via CUDA em uma GPU dedicada NVIDIA GeForce RTX 2050 (4 GB de VRAM).

4.5 Resultados e Avaliações

Para aferir a eficácia do modelo ResNet-50 na identificação das pelagens, estabeleceu-se um protocolo experimental rigoroso, dividido em fases de treinamento, validação e teste final. O desempenho foi quantificado através de métricas estatísticas de classificação multiclasse, processadas em ambiente computacional utilizando a biblioteca *PyTorch*.

4.5.1 Partição dos Dados

Nesta seção, detalha-se a estratégia de partição do conjunto de dados aplicada no desenvolvimento da arquitetura proposta, conforme ilustrado na Figura 8a. A separação estruturada das imagens de treinamento e validação é uma etapa essencial para garantir a capacidade de generalização do modelo, impedindo a mera memorização dos padrões visuais,

um fenômeno prejudicial conhecido como *overfitting* (GOOGLE, 2022). Para assegurar a robustez da avaliação, adotou-se a técnica de Validação Cruzada Estratificada com 5 partições (*5-Fold Stratified Cross-Validation*). Essa abordagem, amplamente estabelecida em pesquisas da área, garante que todos os subconjuntos gerados mantenham a mesma distribuição estatística de classes do conjunto original (LEARN, 2017).

O conjunto total de imagens pré-processadas foi segmentado matematicamente em 5 blocos distintos. Durante a execução, o algoritmo realiza um rodízio dinâmico onde, a cada iteração, 4 blocos (80% dos dados) são alocados para o treinamento e ajuste de pesos da rede, enquanto 1 bloco (20%) é isolado para a validação. Esse ciclo repete-se 5 vezes, garantindo que todas as imagens sejam utilizadas como dados inéditos de teste em algum momento do processo.

A estratificação dessa partição é o critério mais crítico para a confiabilidade metodológica do classificador. Devido ao desbalanceamento natural existente entre pelagens equinas comuns e raras no banco de imagens, a divisão dos blocos ocorre de forma estritamente proporcional. O algoritmo calcula a representatividade global de cada classe e assegura que essa exata mesma porcentagem seja replicada dentro de cada um dos 5 blocos, evitando que o modelo seja prejudicado pela ausência de classes minoritárias durante o treinamento.

4.5.2 Métricas de Desempenho Computacional

Seguindo a metódologia de avaliação do classificador de (KIM, 2024), fundamentou-se em seis pilares principais, extraídos diretamente do comportamento da rede neural durante o teste:

1. **Função de Perda (*Cross-Entropy Loss*):** As funções de perda são utilizadas para calcular o erro cometido pelo modelo CNN, em relação ao resultado desejado, sendo a mais utilizida a *Cross-Entropy* (SILVA, 2022). A *Final Cross-Entropy Loss* mensura o quão distante a predição probabilística do modelo está do rótulo real, permitindo identificar a confiança da rede nas classificações de pelagens visualmente similares.

2. **Acurácia Geral (*Overall Accuracy*):** Calculada pela razão entre o total de predições corretas e o número total de amostras no conjunto de teste. Esta métrica fornece uma visão holística da robustez do modelo (FILHO, 2023).

3. **Acurácia por Classe:** Essencial devido ao desequilíbrio amostral entre certas pelagens (como o desafio entre Brancos e Tordilhos). Esta métrica permite identificar quais fenótipos são mais facilmente distinguíveis pelo extrator de características da ResNet-50.

4. **Matriz de Confusão:** Uma matriz de confusão (ou matriz de erro) é um método de visualização para resultados de algoritmo classificador. Mais especificamente, é uma tabela que decompõe o número de instâncias da verdade fundamental de uma classe específica em relação ao número de instâncias de classes previstas (MUREL, 2024). No contexto deste trabalho, a matriz é crucial para analisar a dispersão entre pelagens com sobreposição cromática, como a confusão entre Alazão, Baio e Castanho, por exemplo.

5. **F1-Score Micro:** Agrega os verdadeiros positivos, falsos positivos e falsos negativos de todas as classes para calcular a métrica de forma global. Por tratar cada amostra individual com o mesmo peso, o resultado final reflete predominantemente o desempenho do modelo nas classes mais frequentes (majoritárias) (SEFIDIAN, 2019).

6. **F1-Score Macro:** Calcula o F1-Score para cada classe de maneira isolada e, em seguida, extrai a média aritmética simples entre elas. Essa abordagem trata todas as categorias com peso igual, independentemente do volume de dados de cada uma. Trata-se de uma métrica fundamental para revelar a eficácia do classificador em conjuntos de dados desbalanceados (SEFIDIAN, 2019).

4.5.3 Parametrização e Configuração dos Modelos

A reprodutibilidade dos experimentos e a estabilidade de convergência durante as etapas de segmentação anatômica e classificação de pelagens fundamentam-se na definição criteriosa dos hiperparâmetros de treinamento. As escolhas arquiteturais e de otimização foram estruturadas para equilibrar o custo computacional, a capacidade de generalização e a mitigação do desbalanceamento amostral severo inerente ao conjunto de dados.

Na etapa de segmentação por instâncias, utilizou-se o modelo YOLOv11-seg treinado na plataforma Ultralytics/Roboflow. Os hiperparâmetros configurados para a extração das quatro regiões de interesse anatômico estão detalhados na Tabela 10.

Tabela 10 – Parâmetros de treinamento do modelo de segmentação (YOLOv11-seg).

Parâmetro / Configuração	**Especificação Adotada**
Arquitetura base	YOLOv11-seg
Framework / Ambiente	Ultralytics / Roboflow
Hardware	GPU NVIDIA RTX 2050 (4 GB)
Resolução da entrada	640×640 pixels
Tamanho do lote (*Batch*)	16
Número de épocas	105
Critério de parada	*Early Stopping* (patience = 20)
Classes anatômicas	4 (corpo, membros, crina, cabeça)

Fonte: Próprio Autor.

Para a etapa de classificação de pelagens com a arquitetura *Multi-Input*, empregou-se a ResNet-50 pré-treinada na base ImageNet-1k como extrator de atributos profundos. O treinamento da camada linear de fusão foi conduzido com o otimizador SGD com Momento ($\mu = 0{,}9$), taxa de aprendizado $\eta = 0{,}01$ e a função de custo Entropia Cruzada Ponderada (*Weighted Cross-Entropy*), cujas configurações operacionais estão sumarizadas na Tabela 11.

Tabela 11 – Parâmetros de treinamento do classificador Multi-Input ResNet-50.

Parâmetro / Configuração	**Especificação Adotada**
Arquitetura base	ResNet-50 (*Pretrained ImageNet-1k*)
Framework / Ambiente	PyTorch 2.5.1 / CUDA 12.x
Hardware	GPU NVIDIA RTX 2050 (4 GB)
Resolução da entrada	$4 \times (224 \times 224 \times 3)$ pixels
Tamanho do lote (*Batch*)	16
Número de épocas	50
Otimizador	SGD com Momento ($\mu = 0{,}9$)
Taxa de aprendizado (η)	1×10^{-2}
Decaimento de peso (*WD*)	1×10^{-4}
Função de perda (*Loss*)	*Weighted Cross-Entropy*
Critério de parada	*Early Stopping* (patience = 10)
Classes fenotípicas	11 pelagens equinas

Fonte: Próprio Autor.

A estratégia de congelamento parcial de pesos (*feature extraction*) na ResNet-50 preservou as representações visuais generalistas pré-treinadas, permitindo que apenas os pesos da camada densa de fusão fossem atualizados. Essa decisão reduziu a quantidade de parâmetros treináveis, acelerou a convergência e mitigou riscos de sobreajuste (*overfitting*) diante da escala amostral por classe.

4.5.4 Validação Comparativa

O experimento foi estruturado a partir de diferentes configurações do conjunto de dados, permitindo uma análise comparativa rigorosa do desempenho da rede neural. Essa abordagem metodológica possibilita isolar o impacto de cada técnica aplicada ao longo do projeto, abrindo espaço para discussões fundamentadas sobre as melhores abordagens para a classificação. Os cenários de validação foram divididos em:

- **Avaliação sem Aumento do Volume Amostral:** Treinamento realizado estritamente com o volume original de imagens coletadas e segmentadas, estabelecendo uma linha de base para o desempenho inicial do classificador.
- **Avaliação com Aumento do Volume Amostral:** Análise do comportamento da rede após a inserção de novas imagens com variações de pelagens no conjunto de

treinamento, visando verificar o ganho de capacidade de generalização do modelo.

- **Avaliação de Dados com Pré-processamento:** Medição do impacto da aplicação do *pipeline* de filtros de imagem nos recortes isolados. O objetivo é validar se o realce das texturas da pelagem facilita a extração de características pelas camadas convolucionais.

5 Resultados e Discussões

Este capítulo apresenta os resultados obtidos por meio da aplicação dos fundamentos teóricos e da metodologia proposta para a identificação e classificação de pelagens equinas. Apresenta-se análise comparativa explicitadas no capítulo de metodologia.

5.1 Avaliação sem Aumento do Volume Amostral

Nesta seção, analisa-se o comportamento e o desempenho inicial da arquitetura de classificação utilizando o conjunto de dados em seu volume primário reduzido, antes da etapa de expansão do *dataset*. Este cenário, definido como *Baseline*, estabelece a linha de base empírica para avaliar a necessidade e o real impacto da coleta e inserção de novas imagens originais no treinamento do classificador.

Tabela 12 – Acurácia da Classificação sem Aumento do Volume Amostral

Pelagem	**Acurácia (%)**
Alazão	100,00%
Baia	100,00%
Branca	0,00%
Castanho	65,79%
Leopardo	0,00%
Lobuna	0,00%
Oveira	0,00%
Pampa	0,00%
Preta	7,69%
Rosilha	0,00%
Tordilha	100,00%
ACC Geral	**63,27%**

Tabela 13 – Métricas da Classificação sem Aumento do Volume Amostral

Métricas	**Valores**
Cross-Entropy Loss	13,60
F1-Score Micro	0,63
F1-Score Macro	0,25

Conforme sumarizado na Tabela 12, o modelo apresentou uma Acurácia Geral de 63,27%. Do ponto de vista global analisando as métricas na Tabela 13, a arquitetura alcançou uma *CrossEntropy Loss* de 13,60, um F1-Score Micro de 0,63 e um F1-Score Macro de apenas 0,25. A interpretação cruzada destas métricas globais com a distribuição

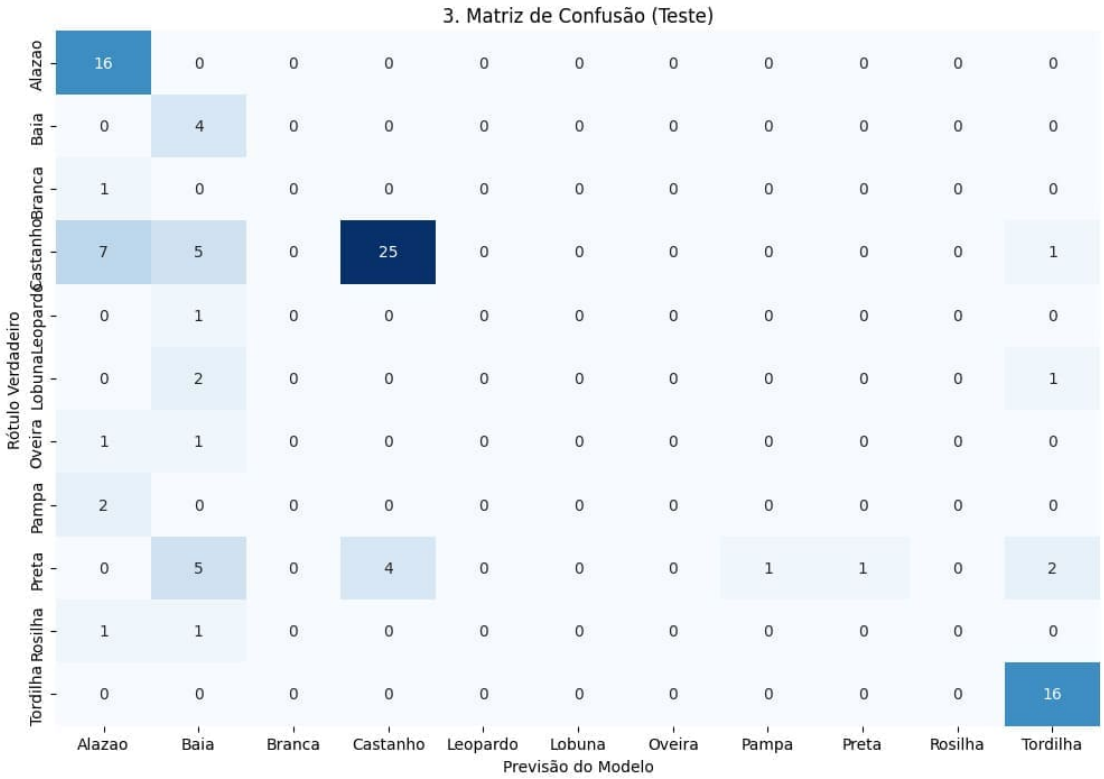


Figura 9 – Matriz Confusão Classificação sem Aumento do Volume Amostral

de acertos por classe revela um forte viés de predição, comportamento característico de redes neurais treinadas em espaços de características (*feature spaces*) severamente desbalanceados.

O F1-Score Micro de 0,63 e a acurácia global de 63,27% sugerem um modelo funcional à primeira vista. No entanto, esses valores são sustentados quase que exclusivamente pelo excelente desempenho nas pelagens mais frequentes do conjunto de treinamento. As classes Alazão, Baia e Tordilha alcançaram 100% de acurácia, não registrando falsos negativos, enquanto a classe Castanho obteve 65,79%. Como a métrica Micro trata cada amostra individual com o mesmo peso, o alto volume de acertos nessas classes dominantes inflou os resultados, mascarando as falhas estruturais do classificador em lidar com a diversidade do banco de imagens.

A fragilidade do modelo é matematicamente exposta pelo F1-Score Macro de apenas 0,25. Ao extrair a média tratando todas as categorias com peso igual (independentemente do volume de dados), o F1-Macro despenca, provando que a rede negligenciou por completo as pelagens raras. Classes como Branca, Leopardo, Lobuna, Oveira, Pampa e Rosilha registraram 0% de acurácia. O algoritmo, na tentativa de reduzir a função de perda global pelo caminho de menor resistência, adotou o viés de prever amostras desconhecidas sempre como uma das pelagens majoritárias (diluindo os erros entre Alazão e Baia).

A *CrossEntropy Loss* extremamente elevada (13,60) é o principal sintoma matemá-

tico da instabilidade do aprendizado. Na mecânica do modelo, uma perda desta magnitude indica que a rede não está apenas errando, mas operando com predições incorretas e com um alto nível de incerteza. Até mesmo para pelagens com características sólidas de cor, como a Preta, o índice de acerto foi de ínfimos 7,69%, demonstrando que a ausência de variações de imagens impediu o extrator de características de construir fronteiras de decisão nítidas.

A análise da Matriz de Confusão do cenário na Figura 9 evidencia o impacto severo do desbalanceamento original dos dados, revelando um classificador fortemente enviesado. A dispersão das predições demonstra que a rede neural adotou uma estratégia conservadora de minimização de perda, recorrendo sistematicamente à predição das classes majoritárias (como Alazão e Baia), que atuaram como atratores para a grande maioria dos erros. Embora o modelo tenha consolidado o aprendizado parcial de pelagens abundantes como Castanho e Tordilha, a sua acurácia global é estatisticamente ilusória, sustentada apenas pela alta frequência natural dessas categorias no conjunto de testes.

Como consequência direta desse viés quantitativo, o modelo falhou criticamente na identificação das pelagens minoritárias (Branca, Leopardo, Lobuna, Oveira, Pampa e Rosilha), registrando total ausência de acertos. Além disso, demonstrou uma confusão estrutural com a pelagem Preta, erroneamente associada a tons terrosos ou reflexos claros devido à deficiência de exemplos no treinamento. Esse diagnóstico visual corrobora que a Rede Neural foi incapaz de construir fronteiras de decisão precisas no volume amostral primário, justificando metodologicamente a necessidade da expansão quantitativa do *dataset* com novas imagens reais.

5.2 Avaliação com Aumento do Volume Amostral

Nesta etapa, o modelo foi submetido a um novo ciclo de treinamento, incorporando a expansão quantitativa do conjunto de dados. A base original, composta por 491 imagens, foi ampliada com a coleta de 620 novas amostras, totalizando 1.111 imagens, um aumento de 126,27 % no volume de dados, conforme detalhado na metodologia. O objetivo foi mensurar o impacto direto da diversidade amostral na capacidade de generalização da rede. Os resultados obtidos, sumarizados nas Tabelas 14 e 15, demonstram uma evolução estrutural no aprendizado da arquitetura:

Tabela 14 – Avaliação com Aumento do Volume Amostral

Pelagem	*Acurácia(%)*
Alazão	62,50%
Baia	16,67%
Branco	50%
Castanho	95%
Leopardo	33,33%
Lobuna	0,00%
Oveira	0,00%
Pampa	33,33%
Preta	65%
Rosilha	50%
Tordilha	89,36%
ACC Geral	**71,17%**

Tabela 15 – Métricas da Classificação com Aumento do Volume Amostral

Métricas	**Valores**
Cross-Entropy Loss	8,74
F1-Score Micro	0,71
F1-Score Macro	0,44

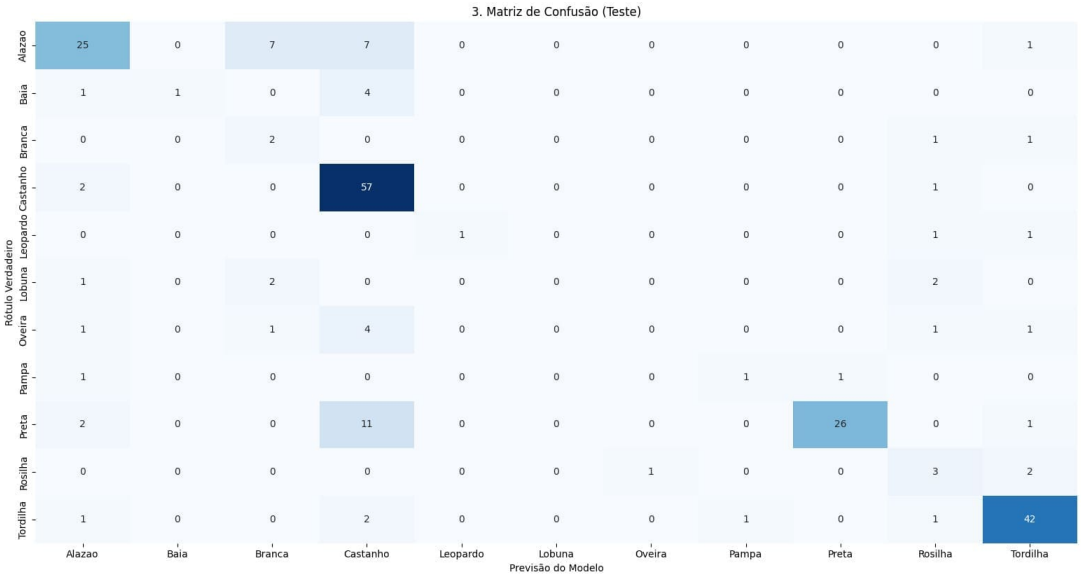


Figura 10 – Matriz Confusão da Avaliação com Aumento do Volume Amostral

A *Cross-Entropy Loss* reduziu-se de 13,60 (cenário *Baseline*) para 8,74. Esta redução de quase 36% na métrica de erro indica que o modelo convergiu para um estado de maior confiança estatística. O aumento do volume amostral permitiu que a ResNet-50 mapeasse variações morfológicas que, no conjunto original, eram interpretadas apenas como ruído.

Complementando a alteração da *Cross-Entropy Loss*, a métrica F1-Score Micro elevou-se para 0,71, refletindo um ganho de precisão global. Contudo, o indicador mais significativo é o F1-Score Macro, que saltou de 0,25 para 0,44. Esse crescimento expressivo valida que a rede neural deixou de depender exclusivamente das classes majoritárias, passando a construir fronteiras de decisão robustas para categorias que antes apresentavam acurácia nula.

A partir da Matriz de Confusão (Figura 10), ilustra o desdobramento do aprendizado. Observa-se uma desconcentração das predições: o "atrator"de erros que anteriormente convergia para Alazão e Baia foi dissipado, dando lugar a uma distribuição de acertos mais orgânica. A classe Castanho atingiu uma acurácia de 95%, enquanto a classe Preta, anteriormente negligenciada, alcançou 65% de precisão. Pelagens que apresentavam 0% de acerto, como Branca (50%), Leopardo (33,33%) e Pampa (33,33%), registraram suas primeiras identificações bem-sucedidas.

Apesar do avanço qualitativo, a redução da acurácia absoluta em classes como Alazão (de 100% para 62,50%) e Baia (de 100% para 16,67%) não representa uma falha, mas sim uma correção de viés. Na Avaliação Inicial, esses valores eram inflados por predições enviesadas. Com o aumento da diversidade de dados, o modelo foi forçado a ser mais criterioso. A persistência de 0% de acurácia em Lobuna e Oveira indica que, embora o volume amostral tenha crescido, a complexidade visual e a sutil diferença cromática dessas pelagens ainda exigem estratégias de extração de características mais refinadas.

5.3 Avaliação de Dados com Pré-processamento

A aplicação de rotinas de pré-processamento, incluindo ajustes de contraste, brilho e realce de nitidez nas imagens segmentadas, resultou em um desempenho geral de 46,85% de acurácia, representando uma queda significativa em relação aos cenários anteriores. Esse resultado, acompanhado por uma *Cross-Entropy Loss* elevada de 21,64 e uma retração nos indicadores F1-Score Micro (0,46) e Macro (0,29), evidencia que o pré-processamento, na configuração testada, comprometeu a integridade das características visuais essenciais para a identificação das pelagens.

A análise da Matriz de Confusão (Figura 11) ilustra um fenômeno de polarização acentuada, no qual a classe Alazão passou a atuar como um "atrator"global para as predições do modelo. O ajuste artificial de contraste homogeneizou a representação cromática das amostras, induzindo a rede neural a classificar incorretamente uma vasta gama de pelagens distintas — como Castanho, Tordilha e Preta — sob o rótulo de Alazão. Esse comportamento sugere que a equalização das imagens, embora tivesse como objetivo destacar padrões, acabou por suprimir a variabilidade cromática natural que a ResNet-50 utilizava anteriormente para diferenciar os subtipos de pelagem.

Tabela 16 – Acurácia da Classificação com Pré-processamento

Pelagem	*Acurácia(%)*
Alazão	100%
Baia	16,67%
Branco	0%
Castanho	20%
Leopardo	33,33%
Lobuna	20%
Oveira	0%
Pampa	0%
Preta	70%
Rosilha	33,33%
Tordilha	40,43%
ACC Geral	**46,85%**

Tabela 17 – Métricas da Classificação com Pré-processamento

Métricas	**Valores**
Cross-Entropy Loss	21,64
F1-Score Micro	0,46
F1-Score Macro	0,29

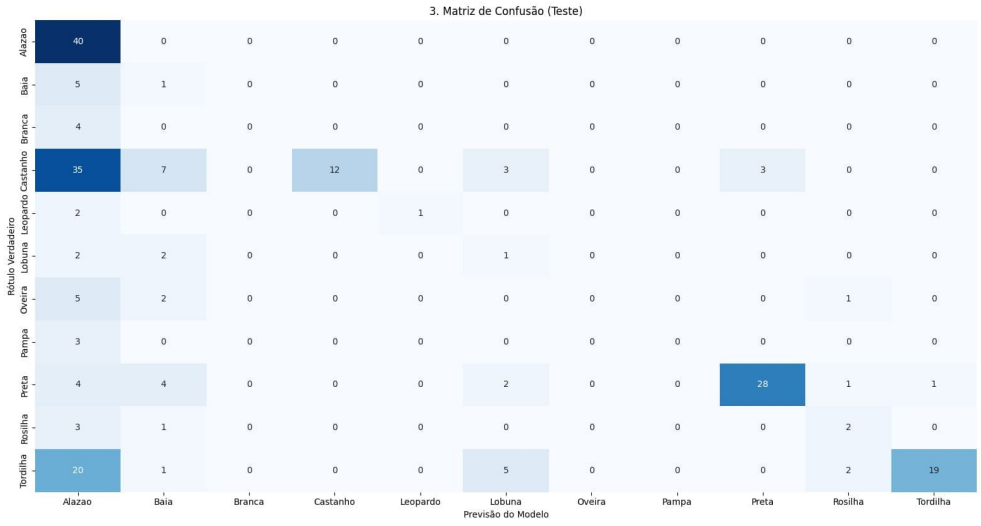


Figura 11 – Matriz Confusão da Avaliação Refinada

Consequentemente, fronteiras de decisão que haviam sido consolidadas na etapa de expansão do volume amostral foram severamente degradadas. Pelagens como a Castanho e a Tordilha, que apresentavam indicadores de performance robustos, sofreram colapsos em sua acurácia, indicando que os filtros aplicados provavelmente distorceram as texturas finas e os tons de transição necessários para a classificação correta. Por outro lado, observa-se

que o pré-processamento trouxe benefícios marginais para classes de difícil detecção; a pelagem Preta registrou uma acurácia de 70%, indicando que o realce de bordas pode ser útil apenas para casos específicos de baixa luminosidade ou texturas muito densas.

Em suma, esta etapa do protocolo de validação comparativa revelou que, para o domínio da equinocultura, a fidelidade das propriedades cromáticas originais é um fator crítico superior ao realce sintético de bordas. O experimento conclui que a manipulação agressiva dos pixels distorceu a distribuição original de cores, tornando o modelo mais instável e menos capaz de generalizar o aprendizado. Portanto, os resultados obtidos neste cenário ratificam a decisão metodológica de priorizar a integridade do dado original como base para o classificador, evidenciando que nem toda técnica de refinamento computacional produz ganhos práticos em cenários onde a cor é o principal preditor semântico.

5.4 Análise Comparativa e Discussão Global dos Cenários

A consolidação dos três cenários experimentais permite traçar um panorama a respeito do comportamento da arquitetura Multi-Input ResNet-50 diante de diferentes tratamentos do conjunto de dados. A Tabela 18 sumariza os desempenhos de acurácia por classe fenotípica, além dos indicadores globais de erro e convergência estatística.

Tabela 18 – Comparativo global de desempenho por classe e métricas de avaliação sob diferentes cenários experimentais.

Pelagem	**Metódo 1**	**Metódo 2**	**Metódo 3**
Alazão	100,00%	62,50%	100,00%
Baia	100,00%	16,67%	16,67%
Branca	0,00%	50,00%	0,00%
Castanho	65,79%	95,00%	20,00%
Leopardo	0,00%	33,33%	33,33%
Lobuna	0,00%	0,00%	20,00%
Oveira	0,00%	0,00%	0,00%
Pampa	0,00%	33,33%	0,00%
Preta	7,69%	65,00%	70,00%
Rosilha	0,00%	50,00%	33,33%
Tordilha	100,00%	89,36%	40,43%
Acurácia Geral	**63,27%**	**71,17%**	**46,85%**
F1-Score Micro	0,63	0,71	0,46
F1-Score Macro	0,25	0,44	0,29
Cross-Entropy Loss	13,60	8,74	21,64

Fonte: Próprio Autor.

O cenário *Baseline*, embora registrasse acurácia global aparente de 63,27%, revelou-se estruturalmente enviesado devido ao F_1-Macro de apenas 0,25, com total cegueira

preditiva (0,00%) sobre seis das onze pelagens. A expansão amostral proporcionou o ganho de F_1-Macro para 0,44, viabilizando a classificação satisfatória de classes raras como Branca (50,00%), Leopardo (33,33%), Pampa (33,33%), Rosilha (50,00%) e uma subida substancial na pelagem Preta (de 7,69% para 65,00%). A *Cross-Entropy Loss* caiu de 13,60 no *Baseline* para o mínimo que foi obtido de 8,74 no cenário expandido. Em contrapartida, as transformações de realce de contraste e nitidez no cenário com pré-processamento elevaram a perda para 21,64, confirmando que a alteração forçada dos canais cromáticos introduziu ruído nas camadas convolucionais. A retração da acurácia para 46,85% no cenário de pré-processamento comprova que técnicas genéricas de filtragem e equalização espacial prejudicam tarefas em que o tom e a saturação natural da pelagem são os principais vetores discriminativos. A perda de fronteiras outrora sólidas nas pelagens Castanho (de 95,00% para 20,00%) e Tordilha (de 89,36% para 40,43%) corrobora que a integridade fotométrica original dos dados é imprescindível para o extrator ResNet-50.

Conclui-se que a coleta de novos dados reais agregada à arquitetura *Multi-Input* com segmentação automatizada representa o arranjo de melhor equilíbrio e generalização para a automação da identificação de pelagens equinas.

5.5 Análise de Desempenho por Pelagem

Esta seção dedica-se à análise detalhada do desempenho da arquitetura de classificação por classe, utilizando como base os resultados obtidos no cenário de Expansão do Volume Amostral. O objetivo é discutir como a rede neural interpretou as características semânticas de cada pelagem.

5.5.1 Alazão

A pelagem Alazão caracteriza-se por uma tonalidade avermelhada distribuída de forma homogênea pelo corpo, incluindo crinas, cauda, cabeça e membros. O modelo obteve um desempenho satisfatório, com acurácia de 62,50%. Os erros de classificação concentraram-se na confusão com a pelagem Castanha, cujas características cromáticas apresentam alta similaridade, dificultando a diferenciação em condições de iluminação variáveis.

5.5.2 Baio

A pelagem Baia destaca-se pelos pelos amarelados — que variam para um marrom claro — em contraste com crina e cauda pretas. Apesar de ser um fenótipo bem definido, a rede neural apresentou dificuldade de processamento, registrando acurácia de 16,67%. A confusão com exemplares da classe Castanha ocorre devido à proximidade tonal entre o

marrom claro da Baia e os tons de castanho de luminosidade elevada, como ilustrado na Figura 12.



(a) Exemplar Baio (b) Exemplar Castanho

Figura 12 – Comparação entre Pelagens: Baio e Castanho

Fonte: (IAGRO, 2017)

5.5.3 Branco

A pelagem Branca é uma das mais raras, exigindo frequentemente a análise de extremidades não pigmentadas para a correta identificação. Embora a escassez de dados para treinamento represente um desafio, o modelo alcançou uma acurácia de 50%. A confusão observada com a pelagem Tordilha é um fenômeno esperado, visto que ambas compartilham tons claros que frequentemente se sobrepõem no espaço de características da rede.

5.5.4 Castanho

A pelagem Castanho apresentou a maior representatividade na base de dados, constituindo a classe majoritária do conjunto. Fenotipicamente, esta pelagem caracteriza-se pela coloração marrom do corpo — que pode variar de tons claros aos mais escuros — em contraste com a coloração preta nas crinas, cauda e extremidades dos membros. O modelo obteve um desempenho superior nesta categoria, alcançando uma acurácia de 95%. Este resultado corrobora a premissa de que a robustez da rede neural é diretamente proporcional à frequência amostral e à consistência visual da classe, permitindo que a arquitetura estabeleça fronteiras de decisão claras e eficazes, servindo como padrão de referência para o desempenho do modelo em relação às demais pelagens.

5.5.5 Leopardo

O padrão Leopardo é caracterizado por manchas arredondadas dispersas pelo dorso. O modelo obteve acurácia de 33,33%. A dificuldade de classificação reside na complexidade do padrão visual, que frequentemente é confundido com a sobreposição de tons em pelagens Rosilhas ou Tordilhas (Figura 13). A dispersão dos erros na Matriz de Confusão (Figura 10) reforça que o modelo não convergiu para um padrão único, o que é agravado pelo volume amostral reduzido desta classe.



(a) Exemplar Leopardo



(b) Exemplar Rosilha

Figura 13 – Comparação entre Pelagens: Rosilha e Leopardo

Fonte: (BELITSKAYA, 2021), (NNEJI, 2022)

5.5.6 Lobuna

A pelagem Lobuna apresenta uma variação cromática peculiar, com um tom acinzentado/terroso que se estende pelo corpo. Com acurácia de 0%, os resultados indicam que a rede interpretou a coloração da Lobuna como um tom "intermediário". A confusão recorrente com as pelagens Branca e Rosilha justifica-se pelo fato de o modelo ter priorizado a luminosidade do dorso em detrimento das características cromáticas específicas, conforme exemplificado na Figura 14.

5.5.7 Oveira

Conforme abordado no referencial teórico, a pelagem Oveira destaca-se pela distribuição irregular de manchas na região dorsal, sendo o padrão espacial das manchas mais relevante do que a cor base. No entanto, a maioria das amostras do banco de dados apresenta tons castanhos em regiões como cabeça e membros. Como resultado, o modelo obteve acurácia de 0%, classificando a maioria dos exemplares como Castanhos. Isso



(a) Exemplar Lobuna (b) Exemplar Alazão

Figura 14 – Comparação entre Pelagens: Lobuna e Alazão

Fonte: (BELITSKAYA, 2021), (NNEJI, 2022)

demonstra a dificuldade da arquitetura em compreender a sutileza das manchas em relação à coloração de fundo, como observado na Figura 15.



(a) Exemplar Oveira (b) Exemplar Castanho

Figura 15 – Comparação entre Pelagens: Oveira e Castanho

Fonte: (DLABEQUICARE, 2024), (IAGRO, 2017)

5.5.8 Pampa

A pelagem Pampa enfrenta desafios similares à Leopardo, sendo o reconhecimento dependente da identificação de manchas brancas extensas. A acurácia de 33,33% reflete a dificuldade do modelo em segmentar a estrutura das manchas em relação à cor de fundo,

tratando o padrão como uma variação cromática genérica em vez de uma característica morfológica estrutural.

5.5.9 Preta

Equinos com pelagem Preta possuem uma cor homogênea e característica. O modelo obteve acurácia de 65%, apresentando bom desempenho na identificação. A maior parte dos erros originou-se da confusão com exemplares da pelagem Castanha, especificamente em casos nos quais a iluminação mascarava subtons marrons, levando a rede a interpretar tons terrosos escuros como pelagem preta.

5.5.10 Rosilha

A pelagem Rosilha, caracterizada pela mistura de pelos brancos sobre uma base colorida, apresentou acurácia de 50%. Apesar da raridade, a rede neural demonstrou capacidade de reconhecer o padrão de "mescla"distintivo desta pelagem, ainda que tenha apresentado erros de classificação em relação a outras categorias cromáticas que também dependem da análise de tonalidade base.

5.5.11 Tordilha

A pelagem Tordilha foi uma das mais representativas na base de dados, caracterizada pela interpolação de pelos brancos em todo o corpo. Esse padrão foi bem assimilado pela rede neural, sendo a segunda classe com maior desempenho, atingindo 89,36% de acurácia.

6 Considerações Finais

6.1 Objetivos Alcançados

Os objetivos apresentados inicialmente se tornam satisfatórios com a criação do protótipo de metodologia baseada em Deep Learning e Transfer Learning utilizando as Redes Neurais Convolucionais (CNNs) para a detecção e classificação das pelagens de equinos, e em cenários diferentes justificando as estratégias propostas ao longo do trabalho.

O protocolo de validação comparativa realizado permitiu uma compreensão sistêmica do comportamento da rede neural ResNet-50 frente aos desafios específicos da classificação de pelagens equinas. A análise dos três cenários experimentais, da *Baseline*, Expansão do Volume Amostral e Pré-processamento, ofereceu evidências empíricas decisivas para a consolidação da metodologia proposta, demonstrando que o desempenho do modelo está intrinsecamente ligado à qualidade e à diversidade cromática dos dados, superando em importância as intervenções algorítmicas de realce visual.

O conjunto de experimentos conclui que a configuração ótima para a classificação de pelagens equinas consiste na utilização da arquitetura ResNet-50 alimentada por um volume diversificado de imagens originais e segmentação adequada, preservando a autenticidade das cores e texturas. Esta metodologia não apenas entrega resultados mais acurados, mas, fundamentalmente, assegura um modelo mais equânime e robusto, capaz de lidar com a complexidade da diversidade fenotípica equina de forma estatisticamente confiável.

6.2 Trabalhos Futuros

Embora este projeto tenha consolidado uma metodologia eficaz para a classificação de pelagens equinas, o domínio da visão computacional aplicado à equinocultura é vasto e permite contínuas evoluções. A exploração de novas estratégias e o refinamento das técnicas aqui empregadas são fundamentais para elevar a precisão, a interpretabilidade e a aplicabilidade da ferramenta em cenários reais. Nesse contexto, destacam-se como possibilidades para trabalhos futuros:

- **Árvores de Decisão:** Investigar o uso de árvores de decisão como classificador secundário, utilizando como entrada as *features* extraídas pela ResNet-50. Essa abordagem visa aumentar a interpretabilidade do modelo, permitindo a hierarquização das decisões. Por exemplo, ao identificar características morfológicas primárias, o sistema poderia descartar automaticamente classes incompatíveis, refinando a predição final

de maneira lógica, como diferenciar equinos de pelagens escuras (Preta/Castanho) de pelagens claras em um nível hierárquico.

- **Otimização para Dispositivos Móveis (MobileNet):** Migração da arquitetura atual para modelos mais leves, como a MobileNet, visando a portabilidade. O objetivo é reduzir o custo computacional e a latência, viabilizando a execução da inferência em tempo real diretamente em dispositivos móveis, além de possibilitar um estudo comparativo da eficiência energética do modelo, aspecto crítico para aplicações de campo.

- **Integração e Validação em Aplicativo para Médicos Veterinários:** Desenvolvimento de uma interface voltada ao suporte na elaboração de resenhas equinas. A ferramenta permitiria ao profissional manipular os polígonos de segmentação com classificação automática em tempo real. Além disso, propõe-se a implementação de um ciclo de *feedback* contínuo: as imagens coletadas pelos veterinários em ambiente real poderiam ser integradas à base de dados para o refinamento constante do modelo, replicando os ganhos de performance observados com a expansão do volume amostral neste trabalho.

- **Ampliação do Escopo e das Classes Equinas:** Expandir o banco de dados para incluir uma maior diversidade de raças, idades e padrões de pelagem ainda não contemplados. A escalabilidade do modelo para cobrir novas classes é um passo essencial para transformar esta ferramenta em um catálogo completo e universal de identificação e registro equino.

- **Aprimoramento da Arquitetura e Pipeline de Classificação:** Investigar novas técnicas de pré-processamento e normalização cromática menos invasivas, além de realizar estudos comparativos avaliando outros extratores de características profundas (como *Vision Transformers*, EfficientNet ou MobileNet) e diferentes estratégias de fusão de atributos (fusão precoce ou com mecanismos de atenção), visando elevar a capacidade discriminativa sobre classes raras.

Referências

AGRODEFESA. *AS PELAGENS DOS EQUIDEOS*. [S.l.]: GERÊNCIA DE SANIDADE ANIMAL, 2017. Citado 6 vezes nas páginas 9, 13, 14, 27, 31 e 32.

ATABAY, H. A. Deep learning for horse breed recognition. *Department of Computer Engineering*, 2017. Citado 5 vezes nas páginas 23, 24, 28, 29 e 34.

AURELIO, Y. S. Estudo de funções de custo para redes neurais com dados desbalanceados. Universidade Federal de Minas Gerais, 2017. Citado na página 37.

BANKAS, E. W. *What Is Polygon Annotation and How Does It Work?* Ayadata - Global Reach. World Class AI Solutions, 2025. Disponível em: <https://www.ayadata.ai/what-is-polygon-annotation-and-how-does-it-work/>. Citado 2 vezes nas páginas 31 e 32.

BARELLI, F. *Introdução à Visão Computacional*. [S.l.]: Casa do Código, 2017. Citado 3 vezes nas páginas 21, 26 e 30.

BARROS, G. P.; PEDRINI, H. Análise comparativa de redes neurais em classificação de imagens de animais. *UNICAMP*, 2023. Citado 2 vezes nas páginas 21 e 23.

BELITSKAYA, O. *Horse Breeds*. Kaggle, 2021. Disponível em: <https://www.kaggle.com/datasets/olgabelitskaya/horse-breeds/data>. Citado 4 vezes nas páginas 30, 31, 51 e 52.

CARDOSO, H. L. d. I.; WORTHINGTON, R. C. *MANUAL DE CONFECCÕES DE RESENHAS*. [S.l.]: Programa Estadual de Sanidade dos Equídeos, 2018. Citado na página 12.

CHEN, F. et al. Review of lightweight deep convolutional neural networks. *Archives of Computational Methods in Engineering*, Springer, v. 31, n. 4, p. 1915–1937, 2024. Citado na página 23.

COMPUTACIONAL, V. *IdentificaÇão, DetecÇão, Reconhecimento e SegmentaÇão de Imagem e Objetos*. 2022. Disponível em: <https://visaocomputacional.com.br/identificacao-deteccao-reconhecimento-e-segmentacao-de-imagem-e-objetos/>. Citado na página 26.

COSTA, L. J. P. D. Gans (generative adversarial network) na ampliação da base de dados para a detecção de defeitos em linhas de produção da indústria têxtil. 2024. Citado na página 25.

DATABRICKS. *Deep learning*. 2025. Disponível em: <https://www.databricks.com/br/glossary/deep-learning>. Citado na página 22.

DLABEQUICARE. Open Source Dataset, *Horse Face Dataset*. Roboflow, 2024. <https://universe.roboflow.com/dlabequicare/horse-face-nfeoi>. Visited on 2026-06-29. Disponível em: <https://universe.roboflow.com/dlabequicare/horse-face-nfeoi>. Citado 2 vezes nas páginas 30 e 52.

EQUALLI. *GUIA PRÁTICO PARA PREENCHIMENTO DE RESENHA*. [S.l.]: Equalli Laboratório – Patologia Clínica Veterinária, 2018. Citado 2 vezes nas páginas 9 e 30.

FILHO, M. *O Que é Acurácia Em Machine Learning?* Mario Filho, 2023. Disponível em: <https://mariofilho.com/o-que-e-acuracia-em-machine-learning/#:~:text=Ela%20%C3%A9%20calculada%20dividindo%20o,sua%20acur%C3%A1cia%20%C3%A9%20de%2085%25.> Citado na página 38.

GOOGLE. *Conjuntos de dados: como dividir o conjunto de dados original*. Google, 2022. Disponível em: <https://developers.google.com/machine-learning/crash-course/overfitting/dividing-datasets?hl=pt-br>. Citado na página 38.

HARDWARE. Open Source Dataset, *horse Dataset*. Roboflow, 2026. <https://universe.roboflow.com/hardware-biea3/horse-6gpg3>. Visited on 2026-06-29. Disponível em: <https://universe.roboflow.com/hardware-biea3/horse-6gpg3>. Citado na página 30.

HOLDSWORTH, J.; SCAPICCHIO, M. *O que é deep learning?* 2025. Disponível em: <https://www.ibm.com/br-pt/think/topics/deep-learning>. Citado na página 22.

HOSEN, T. A. M. I.; ISLAM, M. B. End-to-end human parsing and detection optimized for resource-constrained devices. *Scientific Reports*, 2025. Citado na página 27.

IAGRO. *Manual App – Resenha Versão 1.0*. [S.l.]: Programa Estadual de Sanidade dos Equídeos, 2017. Citado 11 vezes nas páginas 13, 14, 15, 16, 17, 18, 19, 20, 32, 50 e 52.

IBM. *O que são redes neurais convolucionais?* 2025. Disponível em: <https://www.ibm.com/topics/convolutional-neural-networks>. Citado na página 23.

ICICT/FIOCRUZ, I. de Comunicação e Informação Científica e Tecnológica em S. *Aprendizagem colaborativa - Kaggle*. Instituto de Comunicação e Informação Científica e Tecnológica em Saúde Icict/Fiocruz, 2025. Disponível em: <https://pcdas.icict.fiocruz.br/rede-de-parcerias/kaggle/#:~:text=O%20Kaggle%20%C3%A9%20uma%20plataforma,das%20competi%C3%A7%C3%B5es%20em%20v%C3%A1rias%20categorias.> Citado na página 30.

KIM, J. Deep learning-based vehicle type and color classification to support safe autonomous driving [cite: 5]. *Applied Sciences*, MDPI, v. 14, n. 4, p. 1600, 2024. Citado 7 vezes nas páginas 24, 25, 26, 30, 33, 34 e 38.

LEARN, S. *StratifiedKFold*. Scikit Learn Developers, 2017. Disponível em: <https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.StratifiedKFold.html>. Citado na página 38.

LIMA, G. G. d.; MIRANDA, G. C. d.; FARIAS, T. d. S. Introdução às redes neurais para físicos. *Revista Brasileira de Ensino de Física*, SciELO Brasil, v. 48, p. e20250183, 2026. Citado na página 37.

MATOS, G. *Deep Learning na Visão Computacional: Aplicações e Implementação Prática*. 2023. Disponível em: <https://share.atelie.software/deep-learning-na-vis%C3%A3o-computacional-aplica%C3%A7%C3%B5es-e-implementa%C3%A7%C3%A3o-pr%C3%A1tica-80192df11d3d>. Citado na página 21.

MUREL, E. K. J. *O Que é uma matriz de confusão?* IBM, 2024. Disponível em: <https://www.ibm.com/br-pt/think/topics/confusion-matrix>. Citado na página 39.

NEVES, C. E. d. S. Identificação de cores de equinos e muares. 2018. Citado 2 vezes nas páginas 9 e 28.

NEVES, C. E. da S. Redes neurais convolucionais para aprendizagem e reconhecimento de objetos 3d. *EESC/USP*, 2017. Citado na página 23.

NNEJI, I. *Donkeys, Horses & Zebra Images Dataset*. Kaggle, 2022. Disponível em: <https://www.kaggle.com/dsv/3984419>. Citado 3 vezes nas páginas 30, 51 e 52.

OLIVEIRA, L. A. A. J. N. O papel da qualidade da informação nos sistemas de informação. *Departamento de Sistemas de Informação Universidade do Minho*, 2025. Citado na página 30.

OLIVEIRA, T. T. d. Gans (generative adversarial network) na ampliação da base de dados para a detecção de defeitos em linhas de produção da indústria têxtil. *UFRN*, 2024. Citado 2 vezes nas páginas 22 e 26.

ROBOFLOW. *RoboFlow*. Roboflow, 2026. Visited on 2026-06-29. Disponível em: <https://roboflow.com>. Citado na página 30.

SANTOS, M. M. M.; KREBS, L. C.; GODOI, F. N. Por trás das pelagens dos equídeos – revisão de literatura. *Archivos de Zootecnia*, UFRRJ, v. 70, n. 271, p. 389–395, 2021. Citado na página 9.

SEFIDIAN, A. M. *Understanding Micro, Macro, and Weighted Averages for Scikit-Learn metrics in multi-class classification with example*. Lar Academia Sefidiana, 2019. Disponível em: <https://iamirmasoud.com/2022/06/19/understanding-micro-macro-and-weighted-averages-for-scikit-learn-metrics-in-multi-class-classification>. Citado na página 39.

SILVA, D. M. da. Identificação de cores de equinos e muares. *IFG - Inhumas*, 2019. Citado 2 vezes nas páginas 9 e 28.

SILVA, D. M. da. AplicaÇÃo de funcães de perda em modelos de verificacão de assinaturas. *UPE*, 2022. Citado na página 38.

SILVA, T. C. S. Detecção e classificação de derramamento de óleo na superfície oceânica baseada em aprendizagem profunda via algoritmo yolo. *UFMA*, 2024. Citado 3 vezes nas páginas 26, 27 e 32.

SILVA, T. M. B. da. Segmentação semântica de Areas de desmatamento na amazônia a partir de imagens orbitais. *PUC-Rio*, 2021. Citado 2 vezes nas páginas 23 e 24.

ULTRALYTICS. *Residual Networks (ResNet)*. Ultralytics, 2026. Disponível em: <https://www.ultralytics.com/pt/glossary/residual-networks-resnet>. Citado 2 vezes nas páginas 24 e 25.

VINA, A. *What is Resnet-50 and what is it's relevance in Computer Vision?* Ultralytics Inc., 2025. Disponível em: <https://www.ultralytics.com/pt/blog/what-is-resnet-50-and-what-is-its-relevance-in-computer-vision>. Citado na página 34.

VRBANčIč, V. P. G. Transfer learning with adaptive fine-tuning. *IEEE*, 2020. Citado 2 vezes nas páginas 34 e 35.