

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DAINF - DEPARTAMENTO ACADÊMICO DE INFORMÁTICA
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

JORGE LUIZ DOS SANTOS RAMOS JUNIOR

**ESTUDO COMPARATIVO DE TÉCNICAS DE DETECÇÃO
DO ESPELHO NASAL DE BOVINOS**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA
2019

JORGE LUIZ DOS SANTOS RAMOS JUNIOR

**ESTUDO COMPARATIVO DE TÉCNICAS DE DETECÇÃO
DO ESPELHO NASAL DE BOVINOS**

Proposta de Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Sistemas de Informação da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do título de Bacharel.

Orientadora: Profa. Dra. Leyza Baldo Dorini
DAINF - Departamento Acadêmico de Informática - UTFPR

Coorientador: Prof. Dr. Ricardo Dutra da Silva
DAINF - Departamento Acadêmico de Informática - UTFPR

CURITIBA
2019

TERMO DE APROVAÇÃO

“ESTUDO COMPARATIVO DE TÉCNICAS DE DETECÇÃO DO ESPELHO NASAL DE BOVINOS”

por

“Jorge Luiz dos Santos Ramos Junior”

Este Trabalho de Conclusão de Curso foi apresentado no dia **16 de dezembro de 2019** como requisito parcial à obtenção do grau de Bacharel em Sistemas de Informação na Universidade Tecnológica Federal do Paraná - UTFPR - Câmpus Curitiba. O(a)s aluno(a)s foi(ram) arguido(a)s pelos membros da Banca de Avaliação abaixo assinados. Após deliberação a Banca de Avaliação considerou o trabalho

<p>_____</p> <p>Profa. Leyza Baldo Dorini (Presidente - UTFPR/Curitiba)</p>	<p>_____</p> <p>João Ari Gualberto Hill (Avaliador(a) 1 - IAPAR)</p>
<p>_____</p> <p>Guilherme Alves Rosa (Avaliador 2(a) - Bradesco)</p>	<p>_____</p> <p>Fabio Antonio Dorini (Avaliador 3(a) – UTFPR/Curitiba)</p>
<p>_____</p> <p>Profa. Leyza Baldo Dorini (Professora Responsável pelo TCC – UTFPR/Curitiba)</p>	
<p>_____</p> <p>Prof. Marcelo Mikosz Gonçalves (Coordenador do curso de Bacharelado em Sistemas de Informação – UTFPR/Curitiba)</p>	

“A Folha de Aprovação assinada encontra-se na Coordenação do Curso.

AGRADECIMENTOS

Agradeço ao Instituto Agronômico do Paraná (IAPAR) por providenciar esta oportunidade de pesquisa, grandes quantidades de imagens e oportunidade de estágio em boas condições.

Aos meus orientadores, Profa. Dra. Leyza Baldo Dorini e Prof. Dr. Ricardo Dutra da Silva, pela paciência, confiança, revisões e por sempre responder minhas dúvidas. Quaisquer erros ainda presentes neste texto são de responsabilidade do autor.

Ao Prof. Dr. Rodrigo Minetto por compartilhar recursos de hardware, sem os quais a execução das propostas experimentais sobre amostras de milhares de imagens seria quase impossível.

A todos os colegas do IAPAR, especialmente a Wyverson Bonasoli de Oliveira e Lucas da Silva Nolasco, pelas discussões construtivas e trechos de código-fonte.

Ao Laboratório de Física Aplicada e Computacional da Universidade de São Paulo (LAFAC - USP) por disponibilizar suas imagens de bovinos.

RESUMO

RAMOS, Jorge. Estudo Comparativo de Técnicas de Detecção do Espelho Nasal de Bovinos. 2019. 41 f. Trabalho de Conclusão de Curso – Curso de Bacharelado em Sistemas de Informação, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

A identificação de bovinos é de suma importância em vários contextos, como o controle de surtos de doenças e gerenciamento da produção dos rebanhos. Um dos métodos que podem ser utilizados para realizar a identificação é a análise de características do espelho nasal, através de uma foto do animal. Neste contexto, um dos sub-problemas é a detecção da região do espelho nasal em uma foto que contém vários outros objetos. Este trabalho realiza uma comparação entre duas técnicas que podem ser utilizadas para a detecção sobre três grupos de fotos de bovinos: Classificador em Cascata usando características MB-LBP Cascade, e Classificador baseado na Rede Neural Convolucional *You Only Look Once* (YOLO). Em geral, o detector utilizando YOLO obteve melhores resultados em todos os cenários e obteve uma taxa de falsas detecções da região de interesse de 0.00%.

Palavras-chave: Detecção de objetos. Biometria animal. Rastreabilidade animal.

ABSTRACT

RAMOS, Jorge. Comparative Study of Bovine Muzzle Detection Techniques. 2019. 41 f. Trabalho de Conclusão de Curso – Curso de Bacharelado em Sistemas de Informação, Universidade Tecnológica Federal do Paraná. Curitiba, 2019.

Bovine identification is very important in many circumstances, such as controlling disease outbreaks and managing cattle production. One of the means by which identification can be done is by analyzing an animal's muzzle print features, as extracted out of a photo containing the animal. Under this context, one of the subproblems is detecting the muzzle region in a photo that also has many other objects. This work compares two techniques that can be used for object detection in this context: Cascade Classifier using MB-LBP Features and YOLO (You Only Look Once) CNN based Classifier. In general, the YOLO detector had better results in all scenarios and managed a region of interest false positive detection rate of 0.00%.

Keywords: Object detection. Animal biometrics. Animal traceability.

LISTA DE FIGURAS

Figura 1 – Exemplo de segmentação da região do espelho nasal (quadrilátero central).	12
Figura 2 – Representação do funcionamento de uma cascata.	13
Figura 3 – Exemplos de características <i>Haar-like</i>	14
Figura 4 – Demonstrações de conjuntos de vizinhos para algumas combinações de <i>p</i> e <i>r</i>	15
Figura 5 – Representação decimal do funcionamento de um LBP simples ($r=1, p=8$).	16
Figura 6 – Aplicação de MB-LBP 9×6	16
Figura 7 – Demonstração da aplicação de convolução sobre um <i>pixel</i> de uma imagem.	17
Figura 8 – Arquitetura Darknet-53 (YOLOv3).	18
Figura 9 – Modelo YOLO.	18
Figura 10 – Pontos de referência para Barry et al. (2007) (encontrados matematicamente).	20
Figura 11 – Ilustração do procedimento utilizado por Noviyanto e Arymurthy (2013).	20
Figura 12 – Abordagem proposta por Barry et al. (2007): (a) pontos de referência definidos manualmente e (b) região detectada com base em tais pontos.	21
Figura 13 – Amostra de imagens de dois focinhos bovinos utilizados por Hadad, Mahmoud e Mousa (2015).	21
Figura 14 – Interface do <i>software</i> de processamento desenvolvido por Gimenez (2015).	22
Figura 15 – Localização do focinho por Leick (2016).	23
Figura 16 – Localização do espelho nasal por Leick (2016) (retângulo central).	23
Figura 17 – Amostras da base Jersey_1.	25
Figura 18 – Amostras da base Jersey_2.	25
Figura 19 – Amostras da base Jersey_3.	25
Figura 20 – Amostras de fotos inviáveis nas bases Purunã.	26
Figura 21 – Amostras aprovadas da base Puruna_S1.	27
Figura 22 – Amostras aprovadas da base Puruna_S2.	27
Figura 23 – Amostras aprovadas da base Puruna_S3.	27
Figura 24 – Amostras da base USP.	27
Figura 25 – Fluxo dos algoritmos.	28
Figura 26 – Exemplos de anotação de narinas.	29
Figura 27 – Exemplos de amostras positivas de Focinho (retângulo maior).	29
Figura 28 – Exemplos de amostras negativas de Narina Direita (acima) e Narina Esquerda (abaixo)	30
Figura 29 – Exemplos de amostras negativas de Focinho.	30
Figura 30 – Exemplos de amostras negativas de Focinho geradas pelos modelos simples.	31

Figura 31 – Exemplos de amostras negativas de Narinas geradas pelos modelos simples.	31
Figura 32 – Exemplos de ROIs detectadas pelos algoritmos e anotação manual. . .	34
Figura 33 – Amostras de falha do detector YOLO nas imagens Jersey_2 + Jersey_3	34
Figura 34 – Detecção aceitável e um falso positivo do detector MB-LBP.	35
Figura 35 – Falhas de detecção do detector YOLO em imagens das Bases Purunã. .	36
Figura 36 – Exemplos onde MB-LBP detectou narinas e YOLO não detectou. . . .	36
Figura 37 – Exemplos aleatórios de falsos positivos do detector MB-LBP nas Bases Purunã.	36
Figura 38 – Falhas de detecção do detector YOLO em imagens da base Nelore_USP.	37
Figura 39 – Exemplos aleatórios de falsos positivos do detector MB-LBP nas Bases Purunã.	37

LISTA DE QUADROS

Quadro 1 – Bases de imagens.	24
--------------------------------------	----

LISTA DE TABELAS

Tabela 1 – Imagens aprovadas nas Bases Purunã	26
Tabela 2 – Resultados para os Cenários Jersey (1, 2 e 3).	33
Tabela 3 – Resultados para os Cenários Purunã (4 e 5).	35
Tabela 4 – Resultados para os Cenários Nelore (6 e 7).	37

SUMÁRIO

1 – INTRODUÇÃO	11
2 – REVISÃO DE LITERATURA	13
2.1 Técnicas de detecção da ROI em imagens digitais	13
2.1.1 Haar Cascade	13
2.1.2 MB-LBP Cascade (<i>Multi-scale Block Local Binary Pattern</i>)	15
2.1.3 YOLO (<i>You Only Look Once</i>)	15
2.2 Trabalhos anteriores	20
2.2.1 Impressões de tinta do focinho digitalizadas	20
2.2.2 Extração manual da ROI	21
2.2.3 Detecção automática usando <i>MB-LBP Cascade</i>	22
3 – MATERIAIS E MÉTODOS	24
3.1 Bases de imagens utilizadas	24
3.1.1 Bases Jersey	24
3.1.2 Bases Purunã	25
3.1.3 Base Nelore_USP	27
3.2 Implementação dos algoritmos	28
3.2.1 MB-LBP Cascade	29
3.2.2 YOLO (<i>You Only Look Once</i>)	31
3.3 Experimento	32
3.3.1 Métricas utilizadas	32
3.3.2 Fluxo experimental	32
4 – ANÁLISE E DISCUSSÃO DOS RESULTADOS	33
4.1 Cenários 1-3: Bases Jersey_2 + Jersey_3	33
4.2 Cenários 4-5: Bases Puruna_S1 + Puruna_S2 + Puruna_S3	35
4.3 Cenários 6-7: Base Nelore_USP	37
5 – CONCLUSÃO	38
Referências	39

1 INTRODUÇÃO

A identificação de bovinos é algo que possui implicações tanto para a saúde humana quanto no gerenciamento da produção dos rebanhos. Alguns dos motivos para identificar os animais são o controle de surtos de doenças, identificação dos proprietários e requerimentos para exportação (BOWLING et al., 2008).

A carne brasileira sofre com constantes limitações na exportação devido à aplicação de barreiras não-tarifárias por parte de várias regiões. O caso de febre aftosa ocorrido em 2004 no Pará, por exemplo, é usado como motivo para restringir as importações de carne bovina por parte de países da União Européia, e levou ao estabelecimento de sanções por parte da Rússia e países da América do Norte (VALE, 2016). As crescentes exigências do mercado internacional sobre os atributos de segurança do alimento, bem-estar e sustentabilidade animal obrigam o Brasil a implementar rigorosos protocolos de produção para manter a competitividade (SILVA; TRICHES; MALAFAIA, 2011).

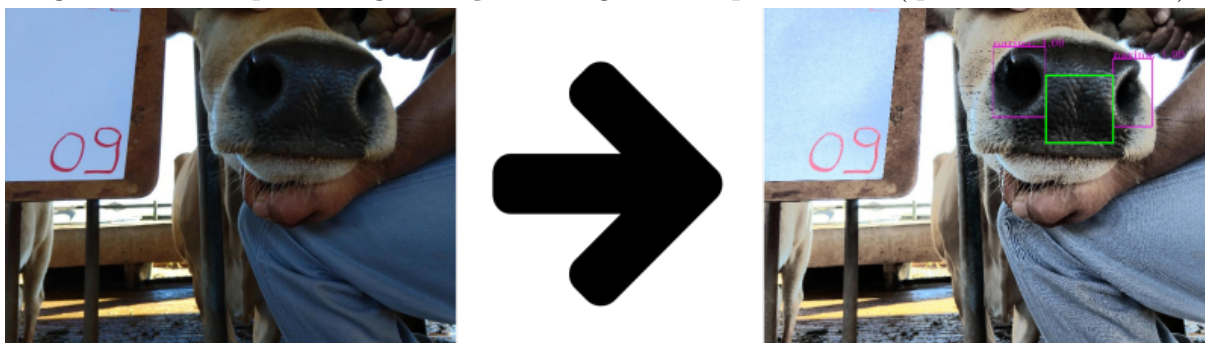
Alguns dos problemas inerentes aos métodos clássicos de identificação (temporários, elétricos, ou permanentes — não-biométricos) são a perda de marcadores, vulnerabilidade a deformações ou fraude, e a infligência de dor e estresse nos animais (AWAD, 2016). Dados os problemas com os métodos clássicos, abordagens biométricas apresentam-se como uma alternativa, pois atendem aos requisitos necessários de universalidade (todos os animais possuem o traço sendo medido), unicidade (o traço utilizado é diferente para cada animal), desempenho (o uso do traço alcança boa precisão com uma velocidade aceitável) e circunvenção (é difícil imitar o traço) (AWAD, 2016).

Alguns dos possíveis identificadores biométricos para bovinos são padrões dos vasos da retina (ALLEN et al., 2008), padrões da íris (SUN; YANG; ZHAO, 2013), perfis de DNA (HEATON et al., 2002) e padrões do espelho nasal (BARRY et al., 2007; NOVIYANTO; ARYMURTHY, 2013; HADAD; MAHMOUD; MOUSA, 2015; GIMENEZ, 2015; LEICK, 2016; KUMAR; SINGH, 2017). A abordagem através do espelho nasal foi utilizada neste trabalho por sua maior conveniência para a obtenção de imagens.

Para identificar unicamente um bovino a partir de uma foto contendo muitos outros objetos além de seu espelho nasal, faz-se útil primeiramente detectar e isolar a Região de Interesse (do inglês *Region of Interest*, abreviado como ROI no restante deste trabalho), conforme ilustrado na Figura 1.

Tal procedimento reduz a região a ser processada, destacando o que é mais significativo e resultando em uma melhoria no desempenho dos algoritmos de classificação (SZELISKI, 2010).

Figura 1 – Exemplo de segmentação da região do espelho nasal (quadrilátero central).



Fonte: Autoria própria.

Este trabalho visa comparar os resultados de detecção da Região de Interesse (ROI) em imagens de espelhos nasais de bovinos considerando duas abordagens. A primeira utiliza um classificador em cascata com características *Multi-scale Block Local Binary Pattern* (MB-LBP) e a outra utiliza um classificador baseado na Rede Neural Convolutiva denominada *You Only Look Once* (YOLO).

Este projeto tem como objetivos específicos:

1. Auxiliar no controle do manejo de bovinos, minimizando erros e problemas causados por limitações com os métodos de identificação atuais.
2. Propor uma metodologia de identificação que preze pelo bem-estar do animal, evitando procedimentos que causem infligência de dor e estresse.
3. Comparar o desempenho dos detectores em termos de quantidade de ROIs detectadas, qualidade da detecção (avaliada qualitativamente) e capacidade de identificação das ROIs detectadas (avaliada através de um classificador VGG16).

É importante destacar que esse projeto foi resultado de uma parceria entre o Instituto Agrônomo do Paraná (IAPAR) e a Universidade Tecnológica Federal do Paraná (UTFPR) durante os anos de 2018 e 2019, o que possibilitou a coleta de bases de imagens e a interação com especialistas das áreas de medicina veterinária e zootecnia.

Este trabalho está estruturado como descrito a seguir. No Capítulo 2 são apresentadas técnicas de detecção da ROI e abordagens utilizadas em trabalhos anteriores. No Capítulo 3 são descritas as bases de imagens utilizadas, detalhes sobre a implementação dos algoritmos e informações sobre as condições do experimento e métricas de comparação. No Capítulo 4 são apresentados os resultados experimentais. Por fim, no Capítulo 5 são expostas as conclusões do estudo e sugestões para trabalhos futuros.

2 REVISÃO DE LITERATURA

A Seção 2.1 discorre sobre algumas técnicas de detecção da ROI em imagens digitais, enquanto a Seção 2.2 discute brevemente alguns trabalhos anteriores sobre identificação de bovinos através do espelho nasal sob a ótica de detecção da ROI.

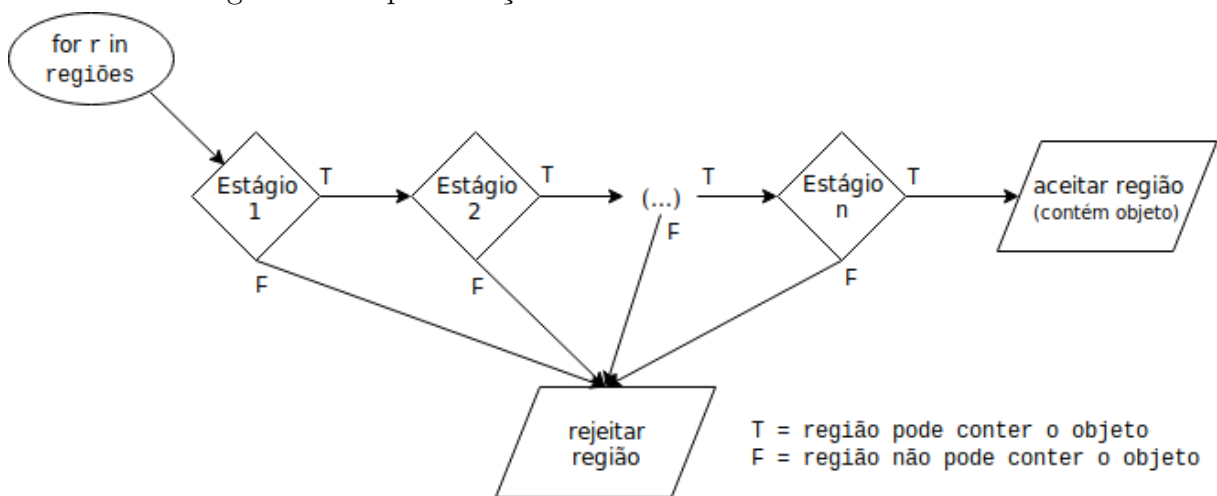
2.1 Técnicas de detecção da ROI em imagens digitais

É possível executar o processo de extração da ROI manualmente com um editor de imagens, mas tal processo é extremamente laborioso quando se faz necessária sua aplicação sobre centenas de amostras. A utilização de técnicas computacionais de detecção automática evita a necessidade de supervisão de um ser humano, viabilizando a utilização de mais imagens e diminuindo a subjetividade do processo.

2.1.1 Haar Cascade

Proposta por Viola, Jones et al. (2001), esta técnica consiste em utilizar uma cascata com N estágios, sendo cada estágio um classificador binário baseado em características *Haar-like*. Cada estágio da cascata tem o objetivo de descartar regiões que não contém o objeto, mas sem rejeitar regiões que possam conter o objeto. Se uma região passar por todos os estágios, assume-se que ela contém o objeto. Deste modo, é diminuída a quantidade de processamento utilizada em áreas que não são de interesse. O funcionamento de uma cascata é representado na Figura 2.

Figura 2 – Representação do funcionamento de uma cascata.



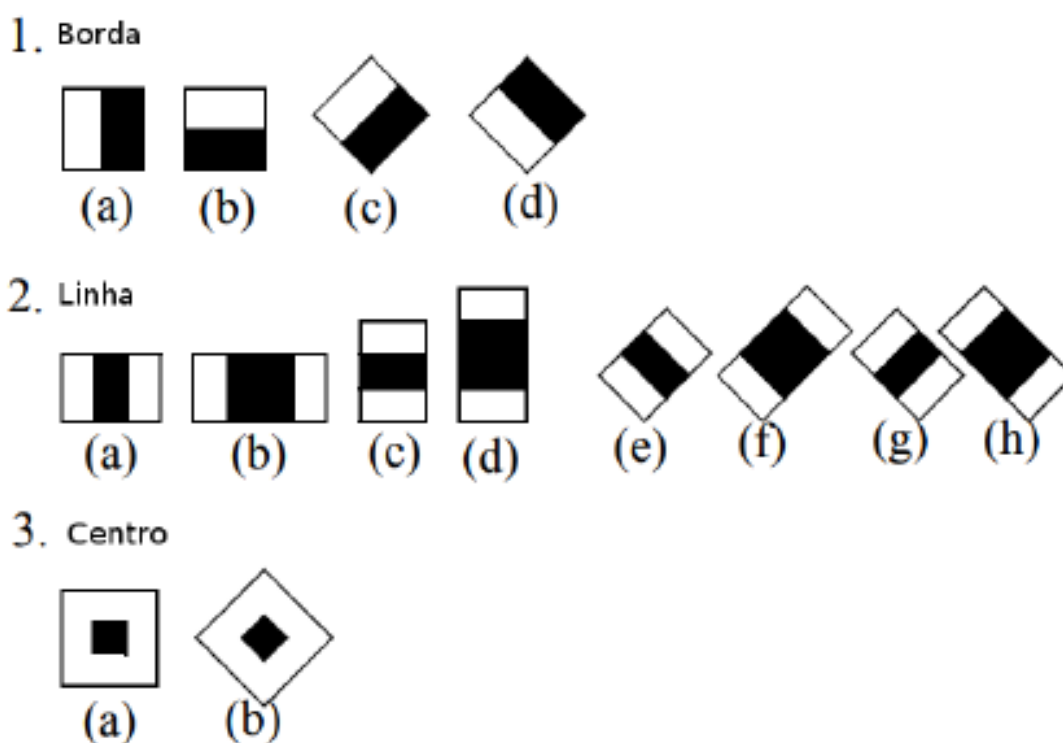
Fonte: A autoria própria.

Classificadores utilizam características para determinar a qual classe um objeto pertence. No caso de classificadores binários, há apenas duas classes, que neste contexto

são “pode conter o objeto” e “não contém o objeto”. O adjetivo “forte” ou “fraco” diz respeito à taxa de acertos do classificador, sendo um classificador fraco aquele que é apenas levemente melhor comparado a um completamente aleatório.

No Haar Cascade, cada estágio é um classificador forte composto por uma combinação de classificadores fracos binários, computacionalmente baratos e que utilizam características *Haar-like*. Tais características são baseadas na diferença de intensidade entre regiões retangulares de uma imagem e representam formas como bordas e linhas, conforme ilustrado na Figura 3.

Figura 3 – Exemplos de características *Haar-like*.



Fonte: Adaptado de Lienhart e Maydt (2002)

A combinação de classificadores é feita por meio do algoritmo *AdaBoost* (FREUND; SCHAPIRE, 1997), cuja ideia básica é escolher apenas os classificadores fracos mais relevantes e então calcular uma média ponderada dos mesmos, o que resulta em um classificador forte.

Como características *Haar-like* são baseadas em regiões retangulares, se faz útil converter a imagem de entrada em uma imagem integral. Uma imagem integral permite calcular a soma de qualquer região retangular em tempo constante ($O(1)$) (CROW, 1984).

Para treinar o classificador, são necessários um conjunto de imagens positivas (que contém o objeto que se visa detectar) e um conjunto de imagens negativas (que não contém o objeto que se visa detectar). Em geral, o objetivo do treinamento é encontrar conjuntos de características que aceitem todas as amostras positivas, mas rejeitem a maior

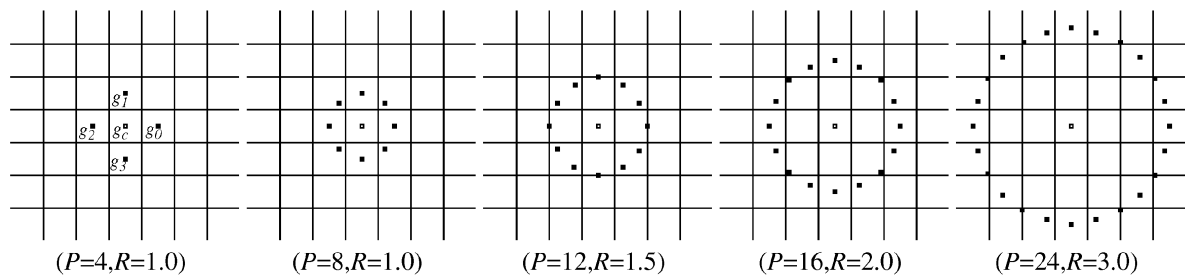
proporção de amostras negativas possível.

2.1.2 MB-LBP Cascade (*Multi-scale Block Local Binary Pattern*)

Assim como *Haar Cascade*, esta técnica também utiliza uma cascata com N estágios de classificadores binários fortes, que são compostos pela aplicação do algoritmo *AdaBoost* sobre conjuntos de classificadores binários fracos. A diferença é que a característica usada nos classificadores fracos é o MB-LBP.

O *LBP* é um descritor popularizado por Ojala, Pietikäinen e Mäenpää (2002), cuja ideia é extrair informação de textura local comparando a intensidade de um *pixel* com a de um número p de *pixels* em uma vizinhança circularmente simétrica a um raio r . A Figura 4 mostra alguns exemplos de combinações de p e r .

Figura 4 – Demonstrações de conjuntos de vizinhos para algumas combinações de p e r .



Fonte: Ojala, Pietikäinen e Mäenpää (2002)

Cada componente da vizinhança é associado a um *bit* que é definido como 1 se possuir intensidade maior ou igual que a do centro, ou 0 se menor. O resultado é um número que representa aquela textura. O procedimento é demonstrado na Figura 5.

Na versão MB-LBP, em vez de a vizinhança ser composta por *pixels* individuais, são usados blocos compostos por vários *pixels*. O bloco é dividido em N sub-regiões com uma igual quantidade P de *pixels* cada, e o valor de cada sub-região é definido como a média aritmética da soma da intensidade de seus componentes. A figura 6 apresenta um exemplo de aplicação do MB-LBP.

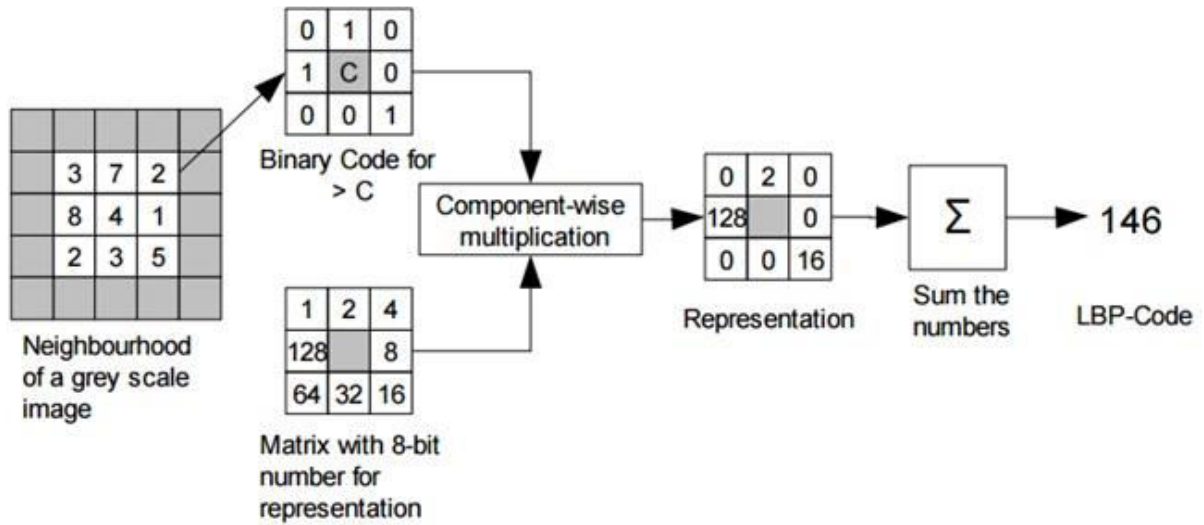
Algumas vantagens são maior robustez e melhor capacidade de representar estruturas de maior escala (LIAO et al., 2007). Com o uso de imagens integrais, o custo operacional é apenas levemente maior do que o do LBP original (LIAO et al., 2007).

Outra vantagem desta técnica em comparação com *Haar Cascade* é que o processo de treinamento é consideravelmente mais rápido (ZHANG et al., 2007).

2.1.3 YOLO (*You Only Look Once*)

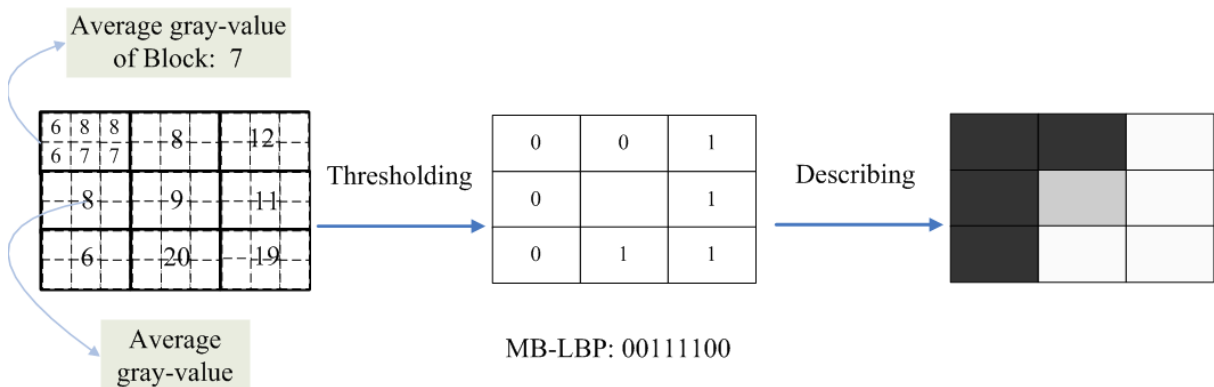
YOLO é um algoritmo de detecção de objetos baseado em Redes Neurais Convolucionais (do inglês CNN, *Convolutional Neural Networks*) idealizado por Redmon et

Figura 5 – Representação decimal do funcionamento de um LBP simples ($r=1, p=8$).



Fonte: Zeno, Yudin e Alkhatib (2016)

Figura 6 – Aplicação de MB-LBP 9×6 .



Fonte: Zhang et al. (2007)

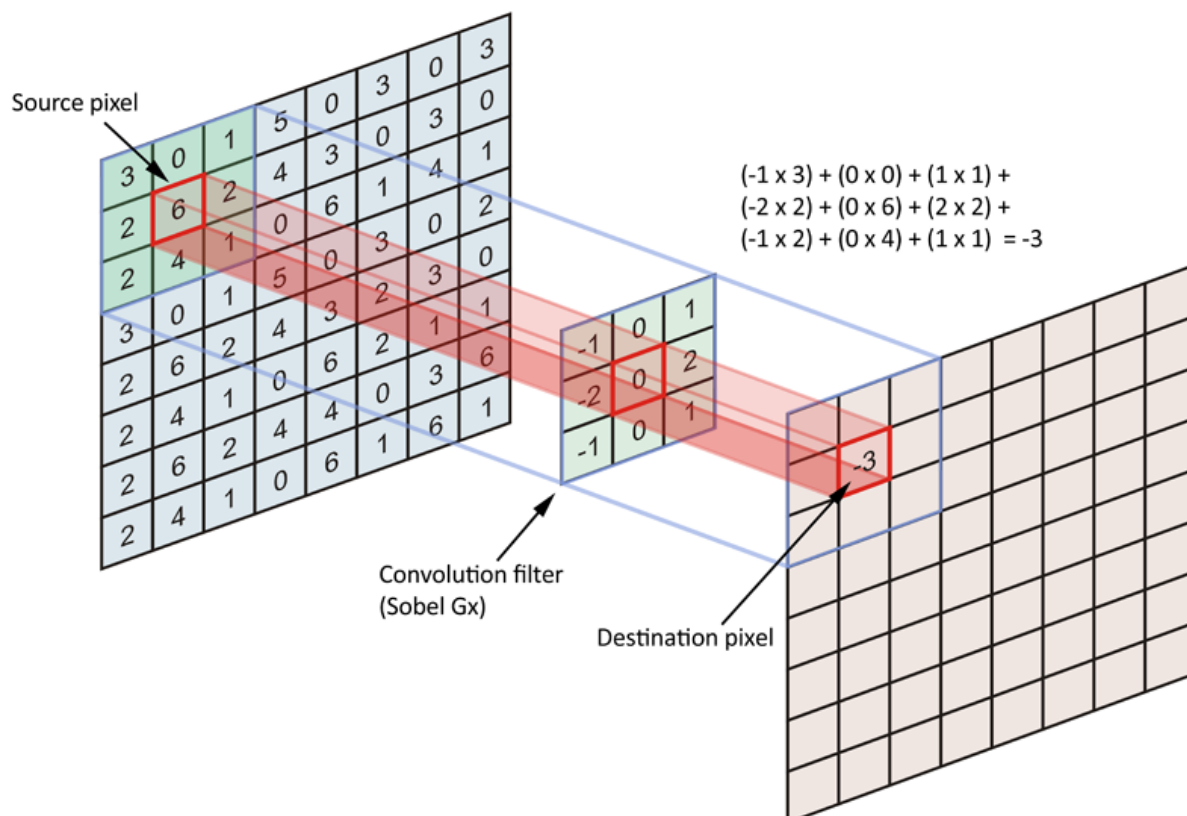
al. (2016). A ideia é que cada camada da rede consiga identificar certas características de uma imagem, que então são combinadas em características mais complexas suficientes para identificar um objeto.

O adjetivo *convolucional* vem do fato de tais redes utilizarem camadas de convolução para gerar mapas de características. Neste contexto, convolução é a operação de aplicar um filtro de janela deslizante sobre uma matriz de entrada (imagem). A Figura 7 mostra o processo de aplicação de uma operação de convolução.

Através da aplicação de várias convoluções, a imagem é transformada em um vetor de características. O filtro de cada camada de convolução é otimizado através do treinamento da rede neural. Outras operações, tais como *pooling* (redução da dimensionalidade através de um filtro que visa reduzir sub-regiões da imagem de entrada e manter as características mais relevantes) e camadas residuais (camadas cuja saída também é adicionada a camadas

posteriores tornando o modelo mais robusto) também são aplicadas visando gerar o conjunto de características mais adequado para a classificação de novas amostras.

Figura 7 – Demonstração da aplicação de convolução sobre um *pixel* de uma imagem. No contexto de CNN, os valores do filtro de convolução serão otimizados através do processo de treinamento da rede.



Fonte: Redmon (2017)

Algoritmos tradicionais baseados em CNNs como AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), VGG16 (SIMONYAN; ZISSERMAN, 2014) e R-CNN (GIRSHICK et al., 2014) procuram primeiramente identificar regiões promissoras na imagem e apenas depois executam um classificador sobre tais regiões.

YOLO utiliza apenas uma CNN para simultaneamente detectar regiões e estimar probabilidades de classe, o que torna o algoritmo extremamente rápido quando comparado com outras abordagens baseadas em CNN e viabiliza sua aplicação em tempo real (REDMON et al., 2016). Isso possibilita seu uso para cenários como identificação em vídeo obtido através de uma *webcam*, por exemplo (REDMON et al., 2016). A arquitetura da rede é apresentada na Figura 8.

Primeiramente, a imagem é redimensionada (a arquitetura padrão YOLOv3 redimensiona qualquer entrada para 608×608 (REDMON; FARHADI, 2018)). A seguir, a imagem é dividida em uma grade de tamanho $S \times S$. Cada célula da grade tenta prever B regiões, um valor de confiança para a existência de um objeto e a precisão da região, e

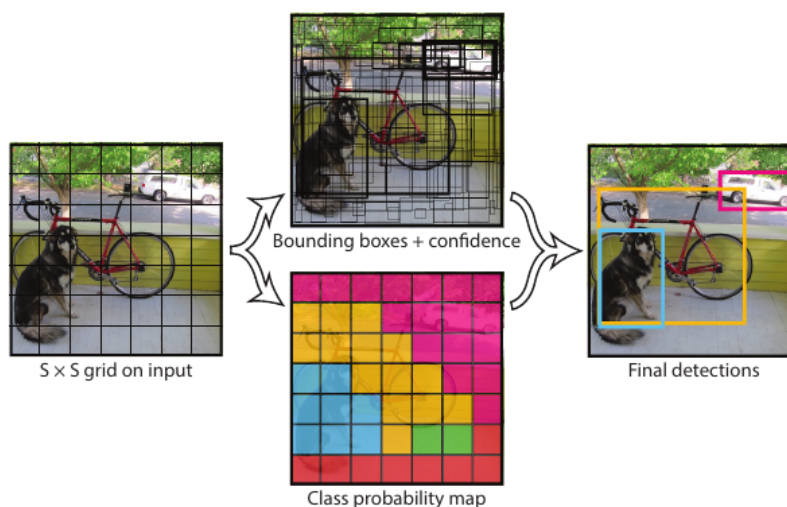
Figura 8 – Arquitetura Darknet-53 (YOLOv3).

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1x	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2x	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8x	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8x	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4x	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Fonte: Redmon e Farhadi (2018)

probabilidades de classe para tais regiões. A classe mais relevante na célula é atribuída a ela. Este processo é ilustrado na Figura 9.

Figura 9 – Modelo YOLO.



Fonte: Redmon et al. (2016)

Porém, em muitos casos, várias previsões de região que se sobrepõem vão detectar o mesmo objeto. Para lidar com isso, é utilizada a técnica de *Non-Maximum Suppression* (NEUBECK; GOOL, 2006). A ideia é suprimir as detecções que possuem muita sobreposição com outras detecções do mesmo objeto através da análise de IOU (*Intersection Over Union*), que é basicamente uma métrica que diz qual o grau de sobreposição entre duas regiões. Ao final, cada instância de objeto é associada a apenas uma região.

O próximo capítulo descreve como estas técnicas foram utilizadas nas abordagens de detecção de espelhos nasais propostas neste trabalho.

2.2 Trabalhos anteriores

Esta seção discute brevemente técnicas utilizadas para extrair a ROI em trabalhos correlatos.

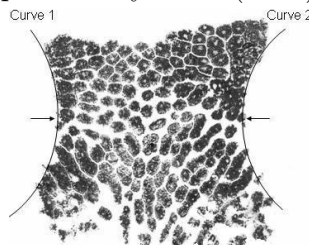
2.2.1 Impressões de tinta do focinho digitalizadas

Em vez de utilizar fotografias digitais, Barry et al. (2007) e Noviyanto e Arymurthy (2013) realizaram pesquisas sobre o uso de impressões de tinta do focinho.

Apesar deste método obter imagens com a ROI já isolada e com boa capacidade de identificação quando bem executado, seu ponto fraco é a dificuldade em obter imagens de qualidade. Isto ocorre devido, principalmente, à dificuldade em controlar os animais e em secar suas narinas, de modo a evitar impressões manchadas (BARRY et al., 2007; NOVIYANTO; ARYMURTHY, 2013). Tais problemas dificultam a aplicação deste método em situações em que centenas de imagens devem ser coletadas.

Barry et al. (2007) utilizaram a derivada zero das funções quadráticas ajustadas aos lados da área das narinas a fim de encontrar os pontos de referência para a extração da ROI, conforme ilustrado na Figura 10.

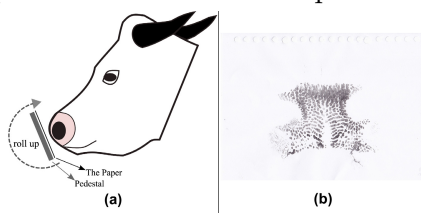
Figura 10 – Pontos de referência para Barry et al. (2007) (encontrados matematicamente).



Fonte: Barry et al. (2007)

Noviyanto e Arymurthy (2013) não entram em detalhes sobre o pré-processamento, mas implicam que este passo não é importante para o procedimento de identificação utilizado por eles. A Figura 11 apresenta o procedimento utilizado para a obtenção das imagens.

Figura 11 – Ilustração do procedimento utilizado por Noviyanto e Arymurthy (2013).

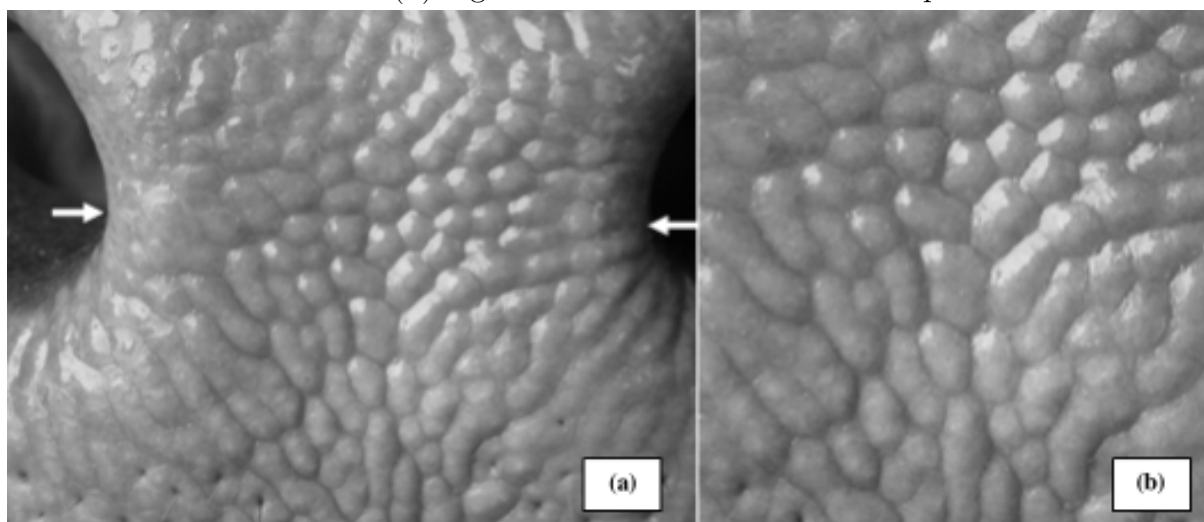


Fonte: Noviyanto e Arymurthy (2013)

2.2.2 Extração manual da ROI

Barry et al. (2007), além de investigar a viabilidade do uso de impressões de tinta do focinho, também realizaram testes com imagens digitais. Para tal, os autores definiram manualmente os pontos de referência para a extração da ROI, conforme exemplificado pela Figura 12.

Figura 12 – Abordagem proposta por Barry et al. (2007): (a) pontos de referência definidos manualmente e (b) região detectada com base em tais pontos.



Fonte: Barry et al. (2007)

Hadad, Mahmoud e Mousa (2015), apesar de discorrerem a respeito do pré-processamento em seu artigo, não tratam da etapa de detecção da ROI. É possível que as ROI foram extraídas manualmente. A Figura 13 apresenta amostras utilizadas pelos pesquisadores.

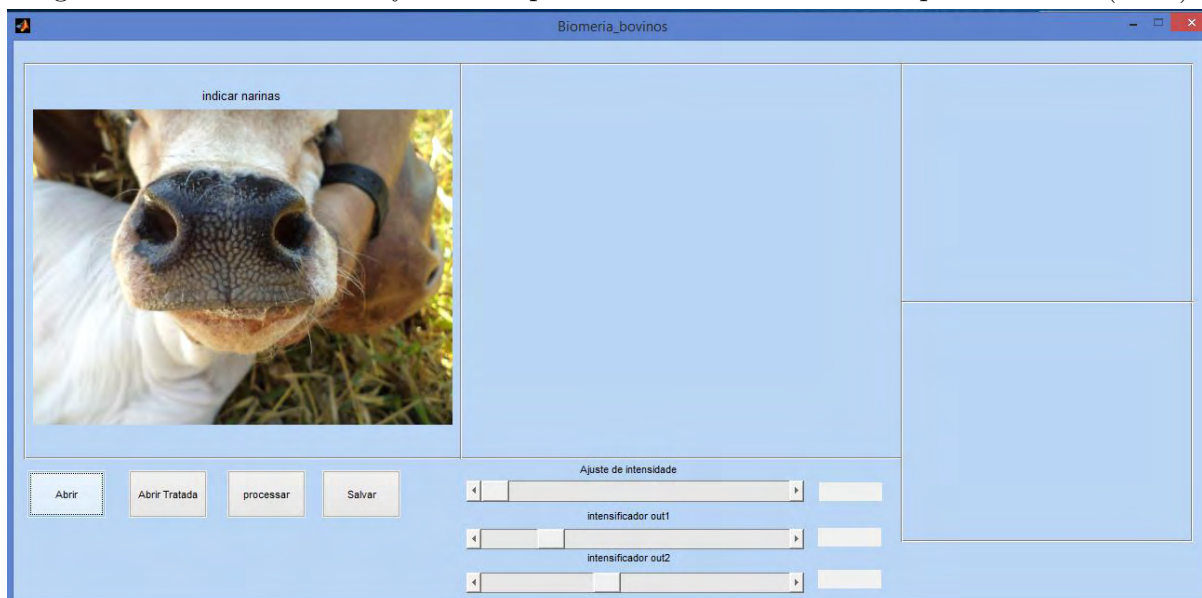
Figura 13 – Amostra de imagens de dois focinhos bovinos utilizados por Hadad, Mahmoud e Mousa (2015).



Fonte: Hadad, Mahmoud e Mousa (2015)

Gimenez (2015) pede ao usuário para que identifique as narinas para a detecção, conforme ilustrado na Figura 14. Apesar da abordagem de classificação ser automática, essa etapa dependente do usuário limita o processamento de grandes quantidades de dados.

Figura 14 – Interface do *software* de processamento desenvolvido por Gimenez (2015).



Fonte: Gimenez (2015)

2.2.3 Detecção automática usando *MB-LBP Cascade*

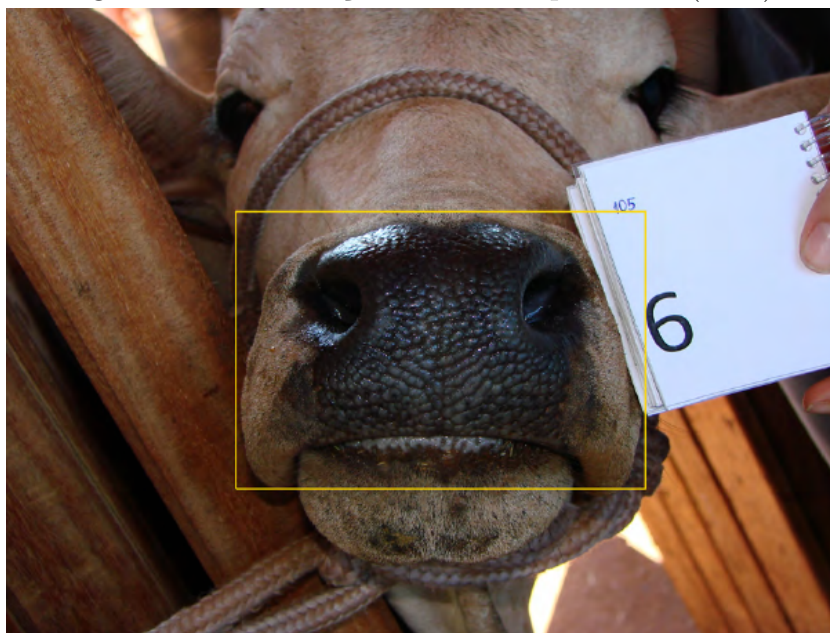
Leick (2016) fez uso de *MB-LBP Cascade* para identificar as regiões de interesse. Foram treinados três classificadores para identificar *Focinho*, *Narina Direita* e *Narina Esquerda*. Primeiramente, é extraída a região do focinho para, segundo Leick (2016), reduzir o espaço de imagem a ser tratado. A Figura 15 contém um exemplo de detecção de focinho.

A seguir, são identificadas as narinas. Segundo Leick (2016), este passo é importante para determinar a inclinação do focinho e para definir a dimensão do espelho nasal, que é dada pela distância entre as narinas. É importante que a imagem esteja paralela ao eixo x para que a determinação da área correspondente ao espelho nasal fique correta.

Após o ajuste da inclinação da imagem, é obtida a área entre as narinas, que é o espelho nasal. A Figura 16 contém um exemplo de detecção de narinas e espelho nasal com base na região do focinho da Figura 15.

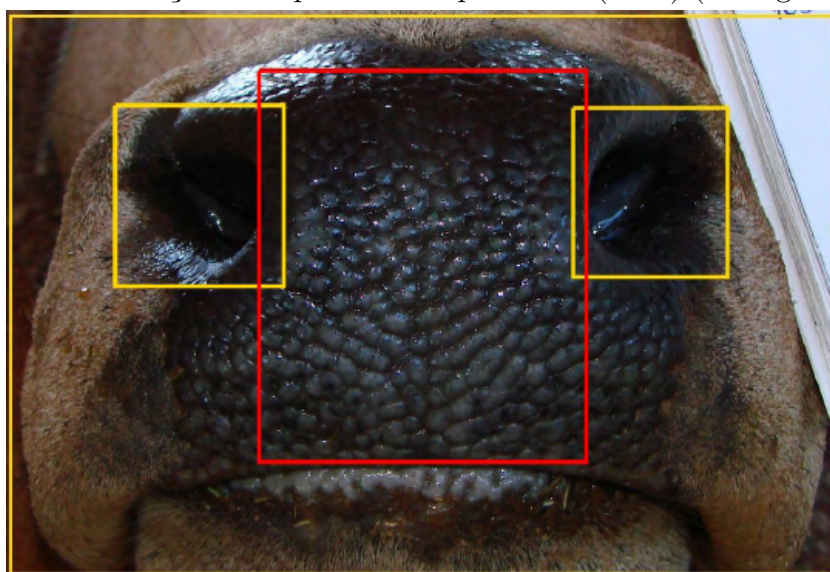
Para o treinamento, Leick (2016) utilizou 101 imagens para cada característica (*Focinho*, *Narina Direita* e *Narina Esquerda*) e uma igual quantidade para os contra-exemplos, totalizando 606 arquivos de imagens base. Para cada característica, foram

Figura 15 – Localização do focinho por Leick (2016).



Fonte: Leick (2016)

Figura 16 – Localização do espelho nasal por Leick (2016) (retângulo central).



Fonte: Leick (2016)

artificialmente geradas mais 4000 imagens rotacionadas e distorcidas para aumentar o tamanho da amostra de treinamento (*data augmentation*).

A partir dessas amostras, foi treinado um classificador *MB-LBP Cascade* com 20 estágios para cada característica. Não são mencionadas métricas que permitam avaliar a capacidade de detecção do detector treinado.

3 MATERIAIS E MÉTODOS

Neste capítulo, a Seção 3.1 apresenta informações sobre as bases de imagens utilizadas, enquanto a Seção 3.2 discute detalhes sobre a implementação dos algoritmos e a Seção 3.3 traz informações sobre o experimento e as condições sobre as quais foi executado.

3.1 Bases de imagens utilizadas

Foi utilizado um total de 14324 fotos coletadas pelo Instituto Agrônomo do Paraná (IAPAR) e 1008 fotos coletadas pelo Laboratório de Física Aplicada e Computacional da Universidade de São Paulo (LAFAC - USP). O Quadro 1 mostra detalhes sobre cada uma das bases. Todas as bases, com exceção da Nelore_USP, foram obtidas como resultado de uma parceria entre o IAPAR e a UTFPR durante os anos de 2018 e 2019.

Quadro 1 – Bases de imagens.

Base	Imagens	Animais	Equipamento	Resolução
Jersey_1	668	33	celular Motorola Moto X4	4032x3024
Jersey_2	319	18	celular Samsung Galaxy S4 SGH-M919	4128x3096
Jersey_3	507	51	câmera Canon EOS Rebel T5	5184x3456
Puruna_S1	2723	112	celular Motorola Moto X4	4032x3024
Puruna_S2	4085	169	celular Motorola Moto X4	4032x3024
Puruna_S3	5645	181	celular Motorola Moto X4	4032x3024
Nelore_USP	1008	51	câmera Sony DSC-H5	2048x1536 ou 3072x2304

Fonte: Autoria própria

Na sequência, cada base é descrita individualmente e são mostrados exemplos de imagens.

3.1.1 Bases Jersey

Estas bases são compostas de fotos de bovinos da raça Jersey. A identificação do animal está incluída na própria foto (uma placa com um número identificador) e, em geral, o bovino ocupa uma área maior da imagem.

As fotos das bases Jersey_1 e Jersey_2 foram obtidas em setembro de 2018 e contém animais diferentes. A base Jersey_3 contém fotos de todos os animais da Jersey_1 e Jersey_2, mas obtidas com outro equipamento em fevereiro de 2019.

As Figuras 17, 18 e 19 mostram alguns exemplos de imagens destas bases. Observe que, além da variação da orientação do espelho nasal e da iluminação, em alguns casos pode ocorrer oclusão de parte da ROI. Além disso, a aquisição de imagens do mesmo animal

em dias diferentes e com equipamentos distintos aumenta a variabilidade das imagens de treinamento.

Figura 17 – Amostras da base Jersey_1.



Fonte: IAPAR

Figura 18 – Amostras da base Jersey_2.



Fonte: IAPAR

Figura 19 – Amostras da base Jersey_3.



Fonte: IAPAR

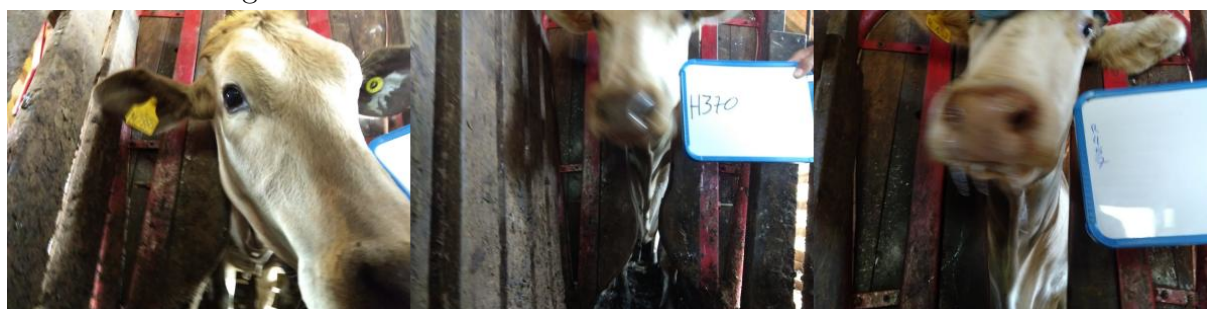
3.1.2 Bases Purunã

Estas bases são compostas de fotos de bovinos da raça Purunã. A raça é o resultado de décadas de pesquisa com acasalamentos entre várias raças, controlados por pesquisadores do IAPAR (PEROTTO et al., 2000).

As fotos foram obtidas ao longo de três semanas entre março e abril de 2019. Em cada semana, foram obtidas fotos do mesmo bovino em dois dias diferentes, e os resultados foram compilados em uma base. Tal procedimento resultou nas bases Puruna_S1, Puruna_S2 e Puruna_S3.

Não houve nenhum tipo de controle sobre as condições para a obtenção das fotos, o que resultou em uma grande proporção de imagens inviáveis para fins de identificação, conforme exemplificado na Figura 20.

Figura 20 – Amostras de fotos inviáveis nas bases Purunã.



Fonte: IAPAR

Para resolver este problema, três bolsistas do IAPAR (incluindo o autor) foram atribuídos a cada imagem com o fim de determinar se ela é viável ou não para a identificação do bovino. Cada trabalhador votou “viável” ou “não viável” para cada imagem. Com base nos resultados, foi gerada uma lista das imagens que foram aprovadas por maioria (dois votos favoráveis). No total, foram aprovadas 3251 de 12453 imagens (26,11%), conforme mostra a Tabela 1.

Tabela 1 – Imagens aprovadas nas Bases Purunã.

	Puruna_S1	Puruna_S2	Puruna_S3
Dia 1	211/1250 (16,88%)	480/1931 (24,86%)	685/2876 (23,82%)
Dia 2	511/1473 (34,69%)	646/2154 (29,99%)	718/2769 (25,93%)
Total	722/2723 (26,51%)	1126/4085 (27,56%)	1403/5645 (24,85%)

Fonte: Autoria própria

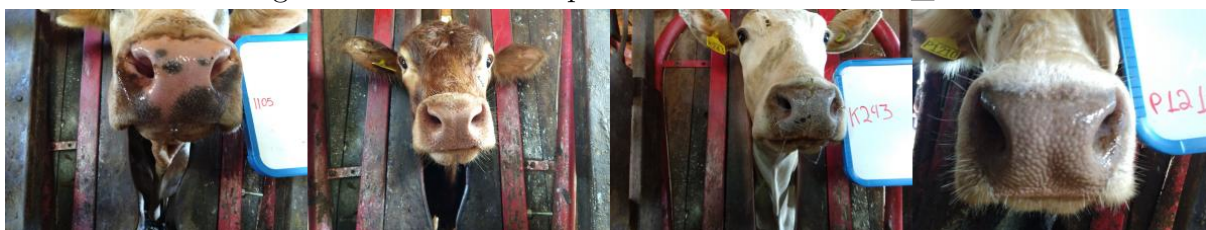
As Figuras 21, 22 e 23 apresentam amostras de imagens aprovadas nas bases em questão.

Figura 21 – Amostras aprovadas da base Puruna_S1.



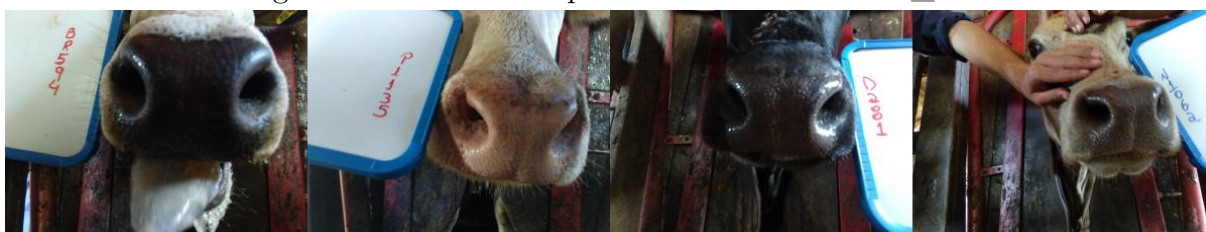
Fonte: IAPAR

Figura 22 – Amostras aprovadas da base Puruna_S2.



Fonte: IAPAR

Figura 23 – Amostras aprovadas da base Puruna_S3.

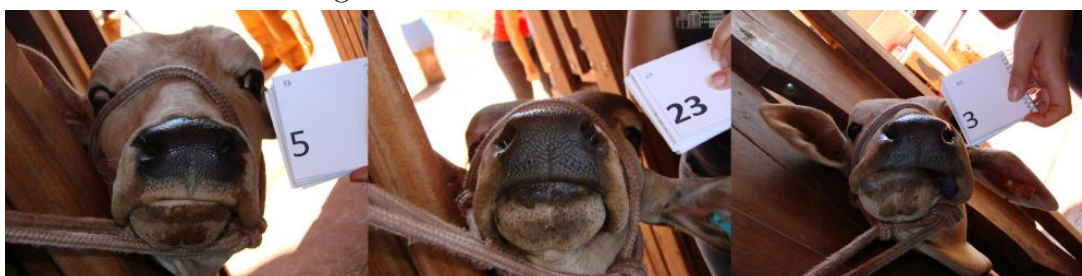


Fonte: IAPAR

3.1.3 Base Nelore_USP

A base Nelore_USP é a mesma que foi utilizada por Leick (2016). Os animais são da raça Nelore e as fotos foram obtidas em outubro de 2008, em condições controladas. A Figura 24 apresenta algumas imagens da base.

Figura 24 – Amostras da base USP.



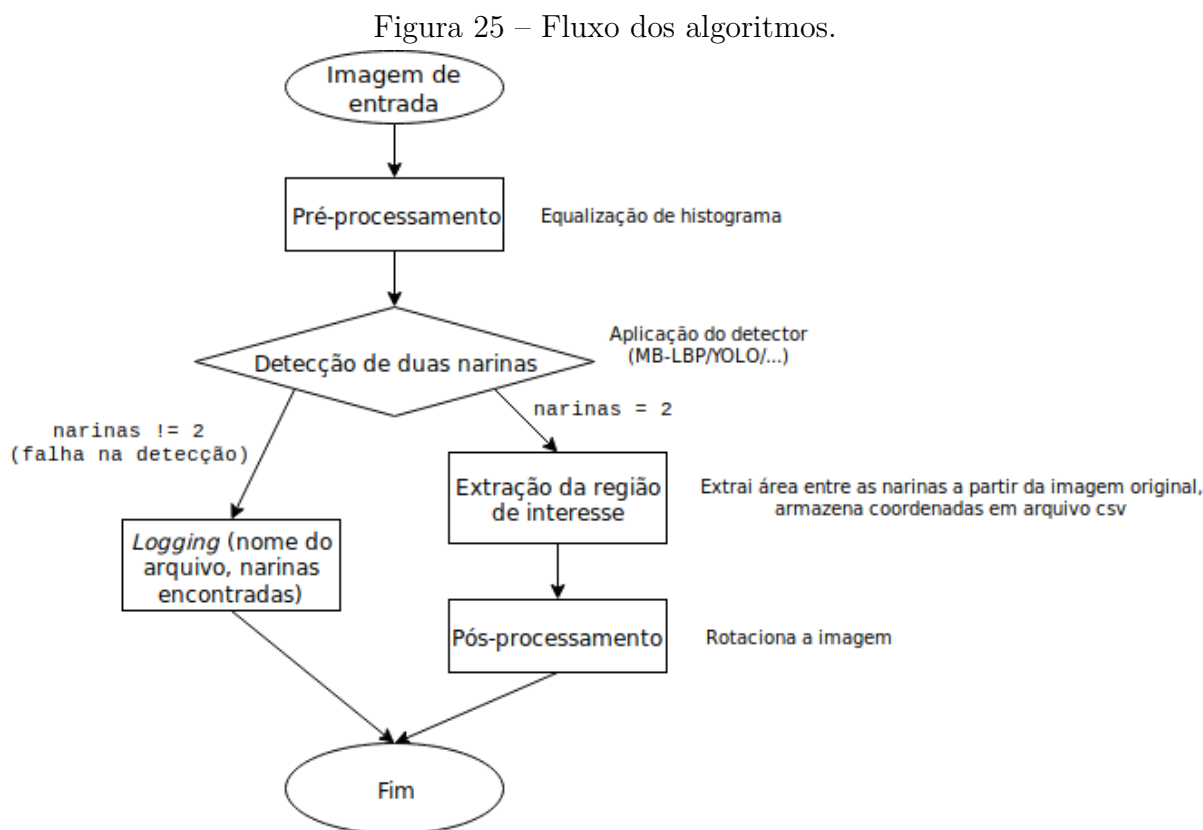
Fonte: LAFAC - USP

3.2 Implementação dos algoritmos

Em trabalhos anteriores sob a supervisão do IAPAR, foram implementados (em linguagem *Python* e com uso da biblioteca *OpenCV*) um algoritmo utilizando *MB-LBP Cascade* (ROSA, 2019) e um algoritmo utilizando YOLO (NOLASCO, 2019). A implementação de Rosa (2019) foi usada como base para a do autor.

Ambas as implementações seguem a ideia base de Leick (2016): detectar duas narinas e então dizer que a região entre as narinas é a ROI. As narinas são características bem definidas e esta abordagem permite corrigir a inclinação da imagem, o que, segundo Gimenez (2015), melhora os resultados finais da identificação de bovinos.

A Figura 25 apresenta o fluxo dos algoritmos.



Fonte: Autoria própria

Os algoritmos foram treinados usando 667 das 668 fotos da base Jersey_1 (uma imagem foi removida por ser inviável para fins de identificação). Nolasco (2019) anotou a posição das narinas para 637 fotos, sendo as restantes anotadas pelo autor. Idealmente, a amostra de treinamento conteria amostras de todas as bases, mas isto não foi possível por causa da indisponibilidade de tempo para realizar novos treinamentos da rede. Por este motivo, optou-se por utilizar a rede treinada por Nolasco (2019). A Figura 26 apresenta alguns exemplos de anotação de narinas.

Figura 26 – Exemplos de anotação de narinas.



Fonte: Nolasco (2019)

3.2.1 MB-LBP Cascade

Seguindo a ideia de Leick (2016), foram treinados três classificadores *MB-LBP Cascade*: um com 50 estágios para identificar Focinho, e dois com 15 estágios, para identificar Narina Direita e Narina Esquerda. O detector de Focinho é útil para evitar detecções de Narinas fora da região de interesse, como sugere Leick (2016).

Para preparar o treinamento, as imagens são convertidas para escala de cinza. A seguir, são extraídas amostras positivas e negativas de Focinho, Narina Direita e Narina Esquerda. As amostras positivas de Narina Direita e Narina Esquerda simplesmente são as regiões anotadas conforme exemplificado na Figura 26.

Para as amostras positivas de Focinho, considerando as anotações de narina como ponto de referência, foi considerado um retângulo com bordas 750 *pixels* além da borda inferior da narina posicionada mais abaixo e 250 *pixels* além das bordas dos outros lados, conforme exemplificado na Figura 27.

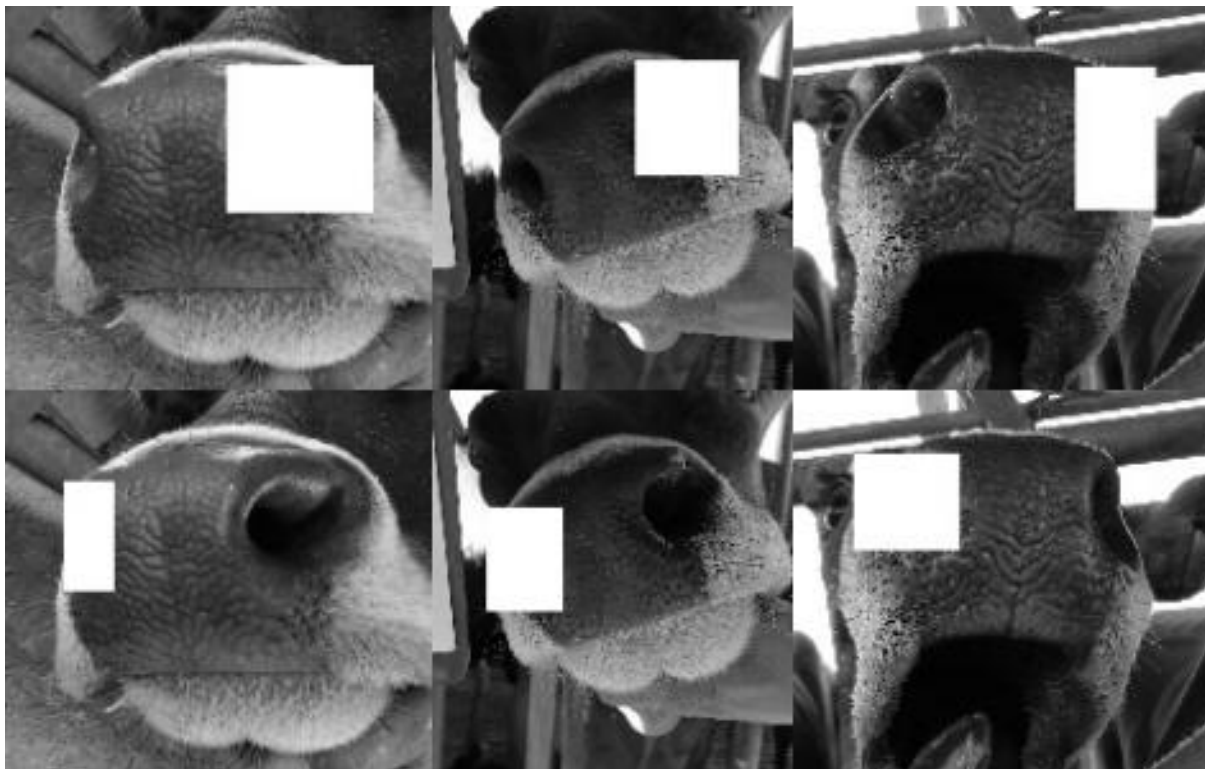
Figura 27 – Exemplos de amostras positivas de Focinho (retângulo maior).



Fonte: Autoria própria.

Para as amostras negativas, primeiramente foram geradas versões das imagens originais com as regiões positivas de narinas e focinho removidas, conforme exemplificado nas Figuras 28 e 29.

Figura 28 – Exemplos de amostras negativas de Narina Direita (acima) e Narina Esquerda (abaixo)



Fonte: Autoria própria.

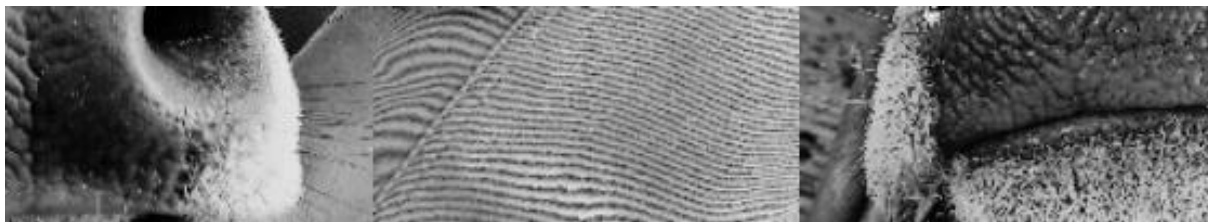
Figura 29 – Exemplos de amostras negativas de Focinho.



Fonte: Autoria própria.

A partir destas amostras, foram treinados vários modelos simples permitindo até 80% de falsos positivos. Os falsos positivos detectados por estes modelos foram repetidamente adicionados às amostras negativas e os modelos foram retreinados, até existir um total de pelo menos 2000 amostras negativas para o treinamento de cada detector. As Figuras 30 e 31 apresentam exemplos de falsos positivos detectados pelos modelos simples.

Figura 30 – Exemplos de amostras negativas de Focinho geradas pelos modelos simples.



Fonte: Autoria própria.

Figura 31 – Exemplos de amostras negativas de Narina Direita (acima) e Narina Esquerda (abaixo) geradas pelos modelos simples.



Fonte: Autoria própria.

Por fim, foram treinados os detectores definitivos utilizando mais 5000 amostras rotacionadas e distorcidas geradas artificialmente. Os detectores de Narinas foram treinados com 15 estágios, capacidade mínima de detecção de 99.5% e máximo de falsos positivos em 10%, enquanto o detector de Focinho foi treinado com 50 estágios, capacidade mínima de detecção de 99.2% e máximo de falsos positivos em 15% (valores determinados empiricamente).

3.2.2 YOLO (*You Only Look Once*)

Para o treinamento do detector utilizando YOLO, simplesmente foram utilizadas as regiões de Narina Direita e Narina Esquerda anotadas conforme exemplificado na Figura 26. Como o detector se mostrou robusto com relação a ruídos na imagem, não se julgou necessário detectar a região do Focinho antes.

Foram realizadas 30000 iterações, sobre subconjuntos de 24 imagens a cada iteração.

A arquitetura da rede é baseada na *Darknet-53*, como apresentada por Redmon e Farhadi (2018).

3.3 Experimento

Esta seção apresenta as métricas utilizadas nos cenários experimentais e o fluxo seguido por estes.

3.3.1 Métricas utilizadas

São utilizadas três métricas nos cenários experimentais: taxa de detecção, taxa de falsos positivos e taxa de identificação.

Taxa de detecção é a quantidade de imagens em que o detector encontrou algo que considera ser duas narinas, dividida pela quantidade total de imagens. Conforme explicado na Seção 3.2, considerando que nas implementações realizadas a ROI é definida como uma área entre as duas narinas, o fato de duas regiões serem identificadas como narinas significa que é possível extrair algo que o detector considera ser o resultado desejado.

Falsos positivos são definidos como erros em que um teste incorretamente indica a presença de um resultado, quando na realidade ele não está presente. No caso específico deste trabalho, a taxa de falsos positivos consiste na quantidade de imagens em que o detector considerou regiões incorretas da imagem como sendo as narinas, dividida pela quantidade total de imagens detectadas. A detecção incorreta das narinas implica em uma detecção incorreta da ROI e, portanto, tais detecções não podem ser usadas para o teste de identificação.

Taxa de identificação é a quantidade de ROIs válidas que o classificador associou ao animal correto, dividida pela quantidade total de ROIs válidas. É de suma importância que as regiões detectadas possuam, de fato, características úteis para a identificação única dos animais. Portanto, se faz lógico testar a capacidade de identificação das ROIs detectadas em cenários de classificação.

3.3.2 Fluxo experimental

O experimento consiste em comparar os detectores sobre diversas bases de imagens. Para cada base, primeiramente, os detectores são executados sobre cada imagem da base, sendo armazenadas a taxa de detecção e as ROIs detectadas. A seguir, sobre as ROIs detectadas, é realizada uma avaliação qualitativa binária sobre a qualidade de cada detecção (aprovada/reprovada), sendo armazenadas a taxa de falsos positivos e as ROIs aprovadas. Por fim, sobre as ROIs aprovadas, a taxa de identificação é testada utilizando um classificador VGG16 com *5-Fold Cross-Validation*.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Este capítulo apresenta os resultados de cada detector em diferentes combinações de bases de imagens. Conforme mencionado na Seção 3.3, as métricas utilizadas são:

- **Taxa de detecção:** quantidade de imagens em que o detector encontrou algo que considera ser duas narinas, dividida pela quantidade total de imagens.
- **Taxa de falsos positivos:** quantidade de imagens em que o detector considerou regiões incorretas da imagem como sendo as narinas, dividida pela quantidade total de imagens detectadas.
- **Taxa de identificação:** quantidade de ROIs válidas que o classificador associou ao animal correto, dividida pela quantidade total de ROIs válidas.

Para a avaliação da viabilidade dos métodos apresentados, foram definidos sete cenários através da combinação de diferentes bases de imagens e métodos de extração da ROI. É importante reiterar que a base Jersey_1 não foi utilizada para os testes por ter sido utilizada para o treinamento dos modelos.

4.1 Cenários 1-3: Bases Jersey_2 + Jersey_3

Estes cenários tratam de testes com bases de imagens contendo animais que foram utilizados no treinamento dos modelos.

- **Cenário 1:** imagens bases Jersey_2 + Jersey_3, ROI YOLO
- **Cenário 2:** imagens bases Jersey_2 + Jersey_3, ROI MB-LBP
- **Cenário 3:** imagens bases Jersey_2 + Jersey_3, ROI extraída manualmente

Além disso, Santos e Siqueira (2019) disponibilizaram anotações manuais da ROI para as bases Jersey_1, Jersey_2 e Jersey_3, tais anotações são utilizadas no Cenário 3. Como elas tratam apenas diretamente da região do espelho nasal e não das narinas, não puderam ser utilizadas para o treinamento dos modelos.

Estas anotações de ROI consideram uma região um pouco mais extensa do que a região detectada pelos algoritmos, conforme ilustrado na Figura 32.

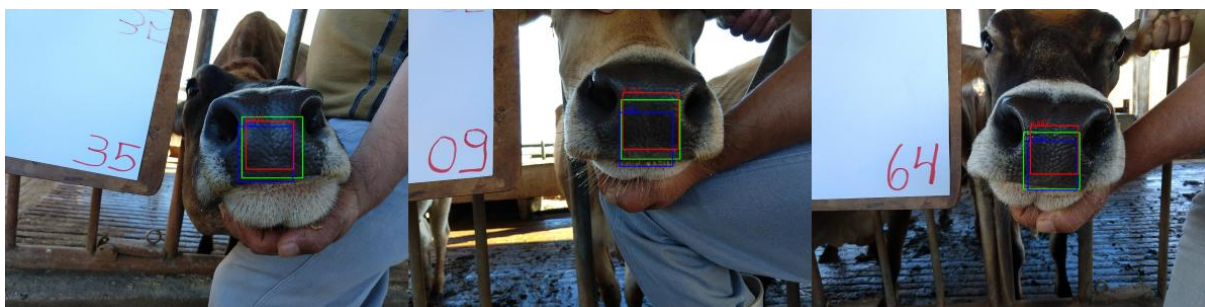
A Tabela 2 apresenta os resultados para os cenários em questão.

Tabela 2 – Resultados para os Cenários Jersey (1, 2 e 3).

Detector	Taxa de Detecção	Falsos Positivos	Taxa de Identificação
YOLO	792/826 (96,00%)	0/792 (0,00%)	789/792 (99,62%)
MB-LBP	570/826 (66,46%)	59/570 (10,35%)	477/511 (93,35%)
Manual			812/817 (99,39%)

Fonte: Autoria própria.

Figura 32 – Exemplos de ROIs detectadas pelos algoritmos e anotação manual. Detecção manual em verde, detecção pelo YOLO em vermelho e detecção pelo MB-LBP em azul



Fonte: Autoria própria.

O detector YOLO apresentou excelentes resultados nestes cenários. A maioria das imagens de falha não são detectáveis através da abordagem proposta, devido a ela ser baseada na detecção de duas narinas, porém uma delas não ser completamente visível nas imagens em questão. Alguns exemplos são apresentados na Figura 33.

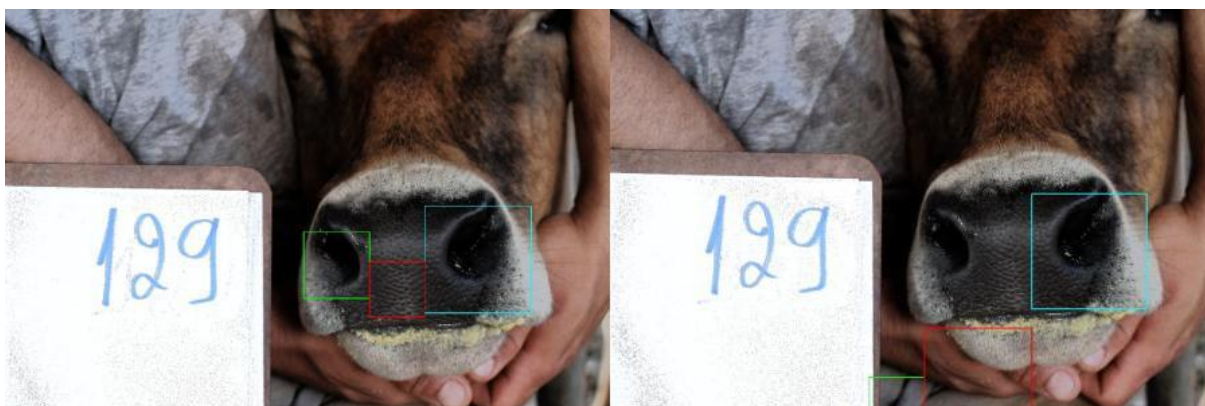
Figura 33 – Amostras de falha do detector YOLO nas imagens Jersey_2 + Jersey_3. Duas são indetectáveis pelo fato de a narina direita não ser completamente visível na imagem. A imagem à direita apresenta uma falha em uma imagem normal, provavelmente devido à combinação de ângulo e iluminação da narina direita.



Fonte: Autoria própria.

O detector MB-LBP mostrou-se inferior em todas as métricas, o que também ocorreu em todos os outros cenários. Aparentemente, o detector MB-LBP é sensível à pequenas variações na imagem em alguns casos. A Figura 34 apresenta um caso de uma detecção aceitável e um falso positivo em fotos do mesmo animal obtidas na mesma sessão.

Figura 34 – Uma detecção aceitável e um falso positivo do detector MB-LBP em fotos do mesmo animal obtidas na mesma sessão.



Fonte: Autoria própria.

4.2 Cenários 4-5: Bases Puruna_S1 + Puruna_S2 + Puruna_S3

Estes cenários tratam de testes com as bases de imagens mais desafiadoras, devido à grande variedade de condições de obtenção das fotos e variedade de texturas entre os bovinos.

- **Cenário 4:** imagens bases Puruna_S1 + Puruna_S2 + Puruna_S3, ROI YOLO
- **Cenário 5:** imagens bases Puruna_S1 + Puruna_S2 + Puruna_S3, ROI MB-LBP

A Tabela 3 apresenta os resultados para os cenários em questão.

Tabela 3 – Resultados para os Cenários Purunã (4 e 5).

Detector	Taxa de Detecção	Falsos Positivos	Taxa de Identificação
YOLO	1876/3251 (57,71%)	0/1876 (0,00%)	1617/1876 (86,19%)
MB-LBP	290/3251 (8,92%)	42/290 (14,48%)	176/248 (70,97%)

Fonte: Autoria própria.

Muitos dos casos de falha do detector YOLO provavelmente podem ser explicados por narinas em menor proporção com relação ao tamanho da imagem ou texturas incomuns. A Figura 35 apresenta exemplos de falhas de detecção.

Apesar da taxa de detecção relativamente baixa do detector MB-LBP, ele conseguiu detectar corretamente algumas narinas que o detector YOLO não conseguiu. A Figura 36 apresenta alguns exemplos de tais casos, incluindo a detecção de ambos MB-LBP e YOLO.

A Figura 37 apresenta exemplos aleatórios de falsos positivos do MB-LBP.

Figura 35 – Falhas de detecção do detector YOLO em imagens das Bases Purunã. As narinas não foram detectadas em nenhuma das cinco imagens do bovino mais à esquerda.



Fonte: Autoria própria.

Figura 36 – Exemplos onde MB-LBP detectou narinas e YOLO não detectou. Detecções pelo MB-LBP na linha de cima e detecções pelo YOLO na linha de baixo.



Fonte: Autoria própria.

Figura 37 – Exemplos aleatórios de falsos positivos do detector MB-LBP nas Bases Purunã.



Fonte: Autoria própria.

4.3 Cenários 6-7: Base Nelore_USP

Estes cenários tratam de testes com uma base de imagens com imagens obtidas em condições controladas, mas com bovinos de uma raça diferente.

- **Cenário 6:** imagens base Nelore_USP, ROI YOLO
- **Cenário 7:** imagens base Nelore_USP, ROI MB-LBP

A Tabela 4 apresenta os resultados para os cenários em questão.

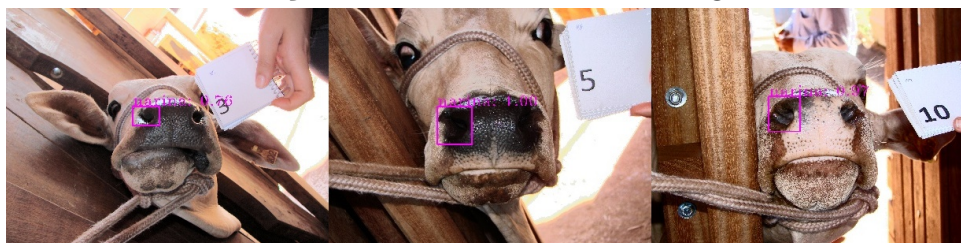
Tabela 4 – Resultados para os Cenários Nelore (6 e 7).

Detector	Taxa de Detecção	Falsos Positivos	Taxa de Identificação
YOLO	987/1008 (97,82%)	0/987 (0,00%)	976/987 (98,89%)
MB-LBP	727/1008 (72,05%)	45/727 (6,19%)	661/682 (96,92%)

Fonte: Autoria própria.

A Figura 38 apresenta exemplos de falhas do detector YOLO. Assim como nas Bases Purunã, muitos dos casos de falha provavelmente podem ser explicados por narinas em menor proporção com relação ao tamanho da imagem ou texturas incomuns.

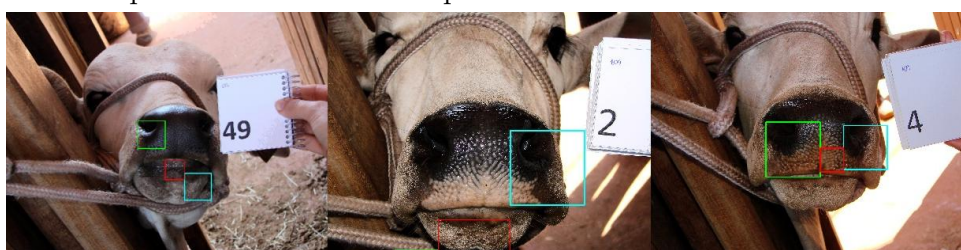
Figura 38 – Falhas de detecção do detector YOLO em imagens da base Nelore_USP.



Fonte: Autoria própria.

A Figura 39 apresenta exemplos aleatórios de falsos positivos do detector MB-LBP. O desempenho apresentado com relação às métricas utilizadas foi estritamente inferior ao do detector YOLO.

Figura 39 – Exemplos aleatórios de falsos positivos do detector MB-LBP nas Bases Purunã.



Fonte: Autoria própria.

5 CONCLUSÃO

Este trabalho apresenta uma revisão da literatura sobre técnicas de detecção de objetos aplicadas no cenário de identificação de bovinos. A aplicação de uma abordagem utilizando uma rede neural convolucional (YOLO) mostrou-se superior a uma abordagem clássica explorada em um trabalho anterior (MB-LBP *Cascade*, Leick (2016)), e abre o caminho para explorações mais profundas.

Este trabalho limitou-se à simples detecção da ROI. Outros aspectos importantes são a validação da qualidade da ROI e velocidade de execução. Se o focinho está desfocado (e, portanto, não há minúcias suficientes), a detecção da ROI ocorre normalmente, mas a ROI não é útil para fins de identificação e deve ser rejeitada. Para a viabilidade de um sistema de identificação de bovinos em um sistema embarcado, é necessário que todos os algoritmos executem sem grande demandas de recursos.

Os resultados nas bases Purunã mostram que é necessária uma maior variedade nas imagens utilizadas para treinar os detectores. O fato de apenas a base Jersey_1 ser utilizada para treinamento provavelmente enfraquece os modelos.

Referências

- ALLEN, A. et al. Evaluation of retinal imaging technology for the biometric identification of bovine animals in northern ireland. **Livestock science**, Elsevier, v. 116, n. 1-3, p. 42–52, 2008. Citado na página 11.
- AWAD, A. I. From classical methods to animal biometrics: A review on cattle identification and tracking. **Computers and Electronics in Agriculture**, Elsevier, v. 123, p. 423–435, 2016. Citado na página 11.
- BARRY, B. et al. Using muzzle pattern recognition as a biometric approach for cattle identification. **Transactions of the ASABE**, American Society of Agricultural and Biological Engineers, v. 50, n. 3, p. 1073–1080, 2007. Citado 4 vezes nas páginas , 11, 20 e 21.
- BOWLING, M. et al. Identification and traceability of cattle in selected countries outside of north america. **The Professional Animal Scientist**, Elsevier, v. 24, n. 4, p. 287–294, 2008. Citado na página 11.
- CROW, F. C. Summed-area tables for texture mapping. In: ACM. **ACM SIGGRAPH computer graphics**. [S.l.], 1984. v. 18, n. 3, p. 207–212. Citado na página 14.
- FREUND, Y.; SCHAPIRE, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. **Journal of computer and system sciences**, Elsevier, v. 55, n. 1, p. 119–139, 1997. Citado na página 14.
- GIMENEZ, C. M. **Identificação biométrica de bovinos utilizando imagens do espelho nasal**. Tese (Doutorado em Qualidade e Produtividade Animal) — Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, 2015. Citado 4 vezes nas páginas , 11, 22 e 28.
- GIRSHICK, R. et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2014. p. 580–587. Citado na página 17.
- HADAD, H. M. E.; MAHMOUD, H. A.; MOUSA, F. A. Bovines muzzle classification based on machine learning techniques. **Procedia Computer Science**, Elsevier, v. 65, p. 864–871, 2015. Citado 3 vezes nas páginas , 11 e 21.
- HEATON, M. P. et al. Selection and use of snp markers for animal identification and paternity analysis in us beef cattle. **Mammalian genome**, Springer, v. 13, n. 5, p. 272–281, 2002. Citado na página 11.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: **Advances in neural information processing systems**. [S.l.: s.n.], 2012. p. 1097–1105. Citado na página 17.
- KUMAR, S.; SINGH, S. K. Automatic identification of cattle using muzzle point pattern: a hybrid feature extraction and classification paradigm. **Multimedia Tools and Applications**, Springer, v. 76, n. 24, p. 26551–26580, 2017. Citado na página 11.

- LEICK, W. d. S. **Tecnologia computacional de apoio a rastreabilidade biométrica de bovinos**. Dissertação (Mestrado em Gestão e Inovação na Indústria Animal) — Faculdade de Zootecnia e Engenharia de Alimentos, Universidade de São Paulo, Pirassununga, 2016. Citado 8 vezes nas páginas , 11, 22, 23, 27, 28, 29 e 38.
- LIAO, S. et al. Learning multi-scale block local binary patterns for face recognition. In: SPRINGER. **International Conference on Biometrics**. [S.l.], 2007. p. 828–837. Citado na página 15.
- LIENHART, R.; MAYDT, J. An extended set of haar-like features for rapid object detection. In: IEEE. **Proceedings. International Conference on Image Processing**. [S.l.], 2002. v. 1, p. I–I. Citado na página 14.
- NEUBECK, A.; GOOL, L. V. Efficient non-maximum suppression. In: IEEE. **18th International Conference on Pattern Recognition (ICPR'06)**. [S.l.], 2006. v. 3, p. 850–855. Citado na página 19.
- NOLASCO, L. **Pesquisa IAPAR - Referência a determinar**. 2019. Citado 2 vezes nas páginas 28 e 29.
- NOVIYANTO, A.; ARYMURTHY, A. M. Beef cattle identification based on muzzle pattern using a matching refinement technique in the sift method. **Computers and Electronics in Agriculture**, Elsevier, v. 99, p. 77–84, 2013. Citado 3 vezes nas páginas , 11 e 20.
- OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, IEEE, n. 7, p. 971–987, 2002. Citado na página 15.
- PEROTTO, D. et al. Características quantitativas de carcaça de bovinos zebu e de cruzamentos bos taurus x zebu. **Revista Brasileira de Zootecnia**, v. 29, n. 6, p. 2019–2029, 2000. Citado na página 25.
- REDMON, J. **Lesson 1 cnn-primer**. 2017. <<https://github.com/DarknetForever/cnn-primer/tree/master/1>>. Accessed: 2019-06-22. Citado na página 17.
- REDMON, J. et al. You only look once: Unified, real-time object detection. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 779–788. Citado 3 vezes nas páginas 16, 17 e 18.
- REDMON, J.; FARHADI, A. Yolov3: An incremental improvement. **arXiv preprint arXiv:1804.02767**, 2018. Citado 3 vezes nas páginas 17, 18 e 31.
- ROSA, G. A. **Pesquisa IAPAR - Referência a determinar**. 2019. Citado na página 28.
- SANTOS, I. V.; SIQUEIRA, M. F. R. **Pesquisa IAPAR - Anotações de IOU bases Jersey_1, Jersey_2 e Jersey_3**. 2019. Citado na página 33.
- SILVA, S. Z. d.; TRICHES, D.; MALAFAIA, G. Análise das barreiras não tarifárias à exportação na cadeia da carne bovina brasileira. **Revista de Política Agrícola**, v. 20, n. 2, p. 23–39, 2011. Citado na página 11.

- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014. Citado na página 17.
- SUN, S.; YANG, S.; ZHAO, L. Noncooperative bovine iris recognition via sift. **Neuro-computing**, Elsevier, v. 120, p. 310–317, 2013. Citado na página 11.
- SZELISKI, R. **Computer Vision: Algorithms and Applications**. [S.l.]: Springer, 2010. ISBN 9781848829343. Citado na página 11.
- VALE, A. R. V. **Barreiras às exportações de carne bovina brasileira com a União Europeia, Rússia e Irã**. Dissertação (Ciências Econômicas) — Universidade Federal do Paraná, Curitiba, 2016. Citado na página 11.
- VIOLA, P.; JONES, M. et al. Rapid object detection using a boosted cascade of simple features. **CVPR (1)**, v. 1, p. 511–518, 2001. Citado na página 13.
- ZENO, B.; YUDIN, D.; ALKHATIB, B. Event recognition on images using support vector machine and multi-level histograms of local patterns. **ARN J. Eng. Appl. Sci**, v. 11, n. 20, p. 12282–12287, 2016. Citado na página 16.
- ZHANG, L. et al. Face detection based on multi-block lbp representation. In: SPRINGER. **International Conference on Biometrics**. [S.l.], 2007. p. 11–18. Citado 2 vezes nas páginas 15 e 16.