

Exploratory Study on Mastitis Identification in Bovine Udders Using Thermography

Estudo Exploratório sobre Identificação de Mastite em Úberes Bovinos por Meio de Termografia

Víctor Emmanuel Olmi Louzada^{1*}, Matheus de Freitas Oliveira Baffa², Sandra Helena Gabaldi Wolf³, Isabele Pessoa Ribeiro³, Lucas Grassano Lattari¹

Abstract: Subclinical mastitis is one of the most prevalent and costly diseases in dairy farming, being laborious to detect by traditional methods. This exploratory study proposes a computational system for the early identification of subclinical mastitis in cattle by integrating infrared thermography (IRT) and machine learning (ML) algorithms. From thermal images of udders from Girolando and Jersey cows, 226 shape, texture, and statistical features were extracted. Various ML classifiers were evaluated, with emphasis on SVM models with polynomial and RBF kernels, which achieved AUC values around 97%, demonstrating a high ability to discriminate between healthy and diseased animals. The Extra Trees model reached 100% sensitivity, demonstrating IRT as a promising screening tool. The results validate IRT combined with ML as a non-invasive, low-cost, and automatable approach for monitoring herd health.

Keywords: bovine mastitis — infrared thermography — machine learning — imaging diagnosis — precision animal science

Resumo: A mastite subclínica é uma das doenças mais prevalentes e custosas na pecuária leiteira, sendo de laboriosa detecção pelos métodos tradicionais. Este estudo exploratório propõe um sistema computacional para a identificação precoce da mastite subclínica em bovinos, integrando termografia infravermelha (TIV) e algoritmos de *machine learning* (ML). A partir de imagens térmicas de úberes de vacas das raças Girolando e Jersey, foram extraídas 226 características de forma, textura e estatísticas. Diferentes classificadores de ML foram avaliados, com destaque para os modelos SVM com *kernels* polinomial e RBF, que alcançaram valores de AUC superiores a 97%, demonstrando alta capacidade de discriminar entre animais sadios e doentes. O modelo *Extra Trees* atingiu 100% de sensibilidade, mostrando-se promissor como ferramenta de triagem. Os resultados validam a TIV combinada com ML como uma abordagem não invasiva, de baixo custo e automatizável para o monitoramento da saúde do rebanho.

Palavras-Chave: aprendizado de máquina — diagnóstico por imagem — mastite bovina — termografia infravermelha — zootecnia de precisão

¹ Departamento Acadêmico de Ciência da Computação, Instituto Federal Sudeste de Minas (IF Sudeste MG), Brasil

² Departamento de Computação, Universidade de São Paulo (USP), Brasil

³ Departamento de Ciências Biológicas e da Saúde, Centro Universitário de Adamantina (FAI), Brasil

*Corresponding author: victoremmanuel.louzada@gmail.com

DOI: <http://dx.doi.org/10.22456/2175-2745.150930> • Received: 03/10/2025 • Accepted: 03/11/2025

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introdução

O setor agropecuário brasileiro desempenha um papel fundamental na economia nacional e global, sendo um dos maiores produtores de alimentos do mundo e o quarto maior produtor de leite [1, 2]. Nesse contexto, a pecuária leiteira se destaca por seu constante crescimento e transformação, exigindo cada vez mais inovações para garantir a sustentabilidade e a com-

petitividade [3]. Dada a evidente importância da cadeia do leite no Brasil, é essencial monitorar continuamente as propriedades para identificar riscos e garantir a segurança do alimento, pois o manejo adequado dos animais promove um sistema de produção mais eficiente [4].

Dentre os principais entraves enfrentados por essa cadeia, a mastite bovina é a doença que se destaca por sua alta prevalência, causando perdas significativas na produção e na

qualidade do leite, impactando negativamente a rentabilidade do produtor e da indústria [5]. A doença pode se manifestar de forma clínica, caracterizada por sinais visíveis de inflamação no úbere ou alterações no leite, como grumos, ou subclínica, uma fase assintomática que causa o maior impacto econômico por ser uma infecção silenciosa e de difícil identificação [5]. Contudo, os métodos tradicionais de diagnóstico representam um desafio, sendo frequentemente de alto custo, trabalhosos e inviáveis para uma detecção precoce em larga escala [6].

O processo inflamatório da mastite, mesmo em sua fase subclínica, provoca um aumento da temperatura superficial do úbere devido à resposta do organismo à infecção. Justamente por essa natureza assintomática, a detecção da mastite subclínica demanda ferramentas que auxiliem a inspeção visual, justificando a busca por métodos diagnósticos de alta sensibilidade [7].

Conforme revisado por Schwarz et al. [6], os métodos tradicionais de diagnóstico, como a contagem de células somáticas (CCS) e a cultura microbiana, apesar de serem o “padrão-ouro”, são frequentemente de alto custo, trabalhosos e invasivos, dependendo da intervenção e interpretação humana e, muitas vezes, inviáveis para uma detecção precoce em larga escala. Mesmo testes práticos de triagem como o *california mastitis test* (CMT), embora baratos e rápidos, são subjetivos e podem gerar resultados falso-positivos ou falso-negativos.

Nesse cenário, a TIV surge como uma alternativa promissora, por ser uma técnica não invasiva capaz de detectar essas alterações de temperatura associadas à inflamação mamária. A análise de imagens e a detecção automática da doença tornam-se viáveis com a integração da TIV com algoritmos de *machine learning* (ML). Esses modelos, ao serem treinados, aprendem a identificar padrões térmicos sutis associados à mastite, complementando a subjetividade humana e os custos dos métodos tradicionais [8].

Diante do exposto, e considerando o potencial da zootecnia de precisão para solucionar as limitações dos métodos atuais, o objetivo deste trabalho é desenvolver e validar um sistema computacional, baseado na integração de imagens termográficas e algoritmos de ML, para a detecção precoce da mastite subclínica bovina.

As principais contribuições deste trabalho são:

- Elaboração de um *pipeline* para análise de imagens térmicas visando à detecção de mastite bovina.
- Integração de características multimodais (radiômicas, texturais e estatísticas) extraídas das imagens.
- *Benchmark* de múltiplos algoritmos de aprendizado de máquina em dados reais, com validação cruzada estratificada por animal.

2. Trabalhos Relacionados

A eficácia da TIV foi validada por diversos estudos na literatura ao longo dos últimos anos. Trabalhos seminais, como os de Hovinen et al. [9] e Colak et al. [10], demonstraram

a capacidade da técnica em identificar aumentos de temperatura superiores a 1°C em úberes com mastite clínica e sua sensibilidade para detectar variações térmicas de acordo com a gravidade da infecção subclínica. No cenário brasileiro, pesquisadores como Oliveira et al. [11] e Metzker et al. [12] também validaram o uso da TIV para o diagnóstico de mastite subclínica, com foco em raças como a Girolando. O estudo de Ribeiro et al. [13] reforça estes achados, concluindo que a TIV possui potencial no rastreamento de mastite subclínica com capacidade diagnóstica preditiva análoga ao CMT.

Para superar a subjetividade da interpretação humana e as limitações mencionadas, a aplicação de técnicas de inteligência artificial (IA) tem se mostrado um caminho inovador. Pesquisas como a de Zhao et al. [8] exemplificam essa integração, desenvolvendo modelos de ML para o reconhecimento automático de mastite subclínica a partir de imagens de TIV. O potencial dessa abordagem é corroborado por revisões sistemáticas, como a de Borchers, Chang e Wadsworth [14], que destacam a crescente importância do ML na pecuária leiteira. Nessa linha, Silva et al. [15] desenvolveram um método automático baseado em limiarização para diferenciar imagens de animais sadios e doentes, enquanto Zaninelli et al. [16] e Watz et al. [17] criaram softwares para a avaliação automática de imagens térmicas do úbere.

O trabalho de Caracuschanski [18] representa um avanço significativo, utilizando ML para prever a etiologia infecciosa da mastite em sistemas de ordenha robótica. No contexto brasileiro, trabalhos aprofundaram a aplicação de redes neurais convolucionais para a detecção de doenças bovinas, enfrentando desafios comuns como a escassez de imagens e a necessidade de alta precisão. Freitas [19], por exemplo, desenvolveu um sistema para o diagnóstico automático da Ceratoconjuntivite Infecciosa Bovina a partir de imagens termográficas. Para contornar a limitação de dados, o autor empregou técnicas de aumento de dados (*data augmentation*) e *ensemble*, alcançando uma acurácia próxima a 94% e validando o método para a identificação precoce da doença.

Posteriormente, Silva [20] avançou na complexidade metodológica para a detecção da mastite subclínica. Em sua abordagem, o desafio da escassez de dados foi tratado por meio de uma estratégia de transferência de aprendizado sequencial, na qual um modelo (ResNet50) foi adaptado progressivamente para reconhecer imagens no geral (ImageNet), depois em um tema parecido (mamografia humana) e, finalmente, foi focado no problema específico (úberes bovinos). A precisão foi refinada com a otimização bayesiana de hiperparâmetros e a incorporação de um mecanismo de atenção (CBAM), que direciona o foco do modelo para as características mais relevantes da imagem. Essa metodologia resultou em uma acurácia de 97,28%, oferecendo uma ferramenta de alta performance que minimiza custos e a dependência da interpretação humana.

Apesar dos avanços significativos apresentados, a literatura ainda aponta lacunas. A maioria dos estudos de alta precisão depende de bancos de dados limitados e capturados em condições controladas, com carência de validação em

diferentes raças ou sistemas de produção comuns no Brasil. Ademais, a complexidade de modelos pode dificultar sua implementação em *hardware* de baixo custo para uso em tempo real na fazenda. Nesse sentido, o presente trabalho se diferencia ao propor uma arquitetura de ML mais leve e eficiente, especificamente otimizada para dados de gado Girolando e Jersey em clima tropical. A robustez do modelo é então validada em um conjunto de dados que simula condições de campo realistas, notadamente por ser uma amostra mais simples e desbalanceada. Com isso, busca-se não apenas alcançar uma boa acurácia, mas também aumentar a viabilidade e aplicabilidade prática da detecção automatizada de mastite.

3. Materiais e Métodos

O método proposto neste trabalho é organizado em quatro etapas principais, conforme ilustrado na Figura 1. Primeiramente, realiza-se a geração das imagens térmicas a partir das matrizes de temperatura, incluindo o mapeamento dos valores para a escala de intensidade de *pixel* e a remoção dos *pixels* de fundo. Em seguida, na etapa de engenharia de características, são extraídos traços, dados microclimáticos, descritores de forma, textura e estatísticos, seguidos pela normalização e pela seleção das 170 características mais relevantes. O valor de $k = 170$ foi determinado após uma busca iterativa, na qual a quantidade de características foi testada em intervalos de cinco para identificar o subconjunto que resultava no melhor desempenho. Este procedimento de seleção foi realizado com o algoritmo *SelectKBest* e a função *chi2* com o intuito de remover características irrelevantes ou redundantes que pudessem prejudicar o desempenho dos classificadores. Na etapa seguinte, procede-se ao pré-processamento, que inclui a separação dos dados, o balanceamento da base por meio do *synthetic minority over-sampling technique* (SMOTE) e a aplicação de *augmentation* para aumentar a robustez do conjunto frente a variações ambientais. Por fim, diferentes algoritmos de aprendizado de máquina são avaliados utilizando validação cruzada estratificada por grupos. Esta seção está organizada em quatro subseções que detalham cada uma dessas etapas.

3.1 Descrição dos Dados e Geração das Imagens

Os dados utilizados neste estudo são provenientes da pesquisa [13], cujos autores cederam o material para a presente investigação. O conjunto de dados é composto por 45 quartos mamários de vacas da raça Girolando e 84 quartos de vacas da raça Jersey, oriundos de propriedades rurais da região de (removido para *blind review*). A coleta original foi realizada com uma câmera termográfica FLIR C5sc, posicionada a uma distância de 50 a 100 cm do úbere, no momento da ordenha matinal.

Para este trabalho, as imagens foram rotuladas com base no diagnóstico de mastite subclínica obtido no estudo original, que utilizou o *california mastitis test* (CMT) como padrão-ouro. Dessa forma, os quartos mamários foram classificados em duas categorias para fins de treinamento e avaliação dos algoritmos de aprendizado de máquina: “Sadio”, correspon-

dente a resultados CMT negativos ou com traços, e “Mastite Subclínica”, correspondente a resultados positivos (+, ++ ou +++).

A geração das imagens térmicas a partir das matrizes de temperatura seguiu um processo em três etapas principais. Primeiramente, os valores das matrizes foram normalizados para a escala [0,1] por meio do método Min-Max. Em seguida, os valores normalizados foram convertidos em intensidades de *pixel* variando entre 0 e 255, de modo a permitir a representação em formato de imagem digital. Por fim, aplicou-se uma etapa de remoção de fundo, em que *pixels* com intensidade inferior a 107 foram eliminados. Esse procedimento assegura a construção de imagens que representam de forma mais fiel a distribuição térmica global associada à presença ou ausência de mastite. A Figura 2 ilustra alguns exemplos das imagens resultantes do processo de geração.

3.2 Extração de Características

A etapa de extração de características foi realizada a partir da região de interesse (*region of interest* - ROI) nas imagens térmicas. Essa ROI foi obtida por meio de um processo de limiarização. Utilizou-se um limiar de intensidade de 107, valor definido empiricamente por inspeção visual para melhor separar o úbere do fundo. Os *pixels* com intensidade superior a 107 foram preservados e considerados como área de análise, enquanto o fundo foi descartado. Dessa forma, as características extraídas refletem diretamente a distribuição térmica do tecido mamário, minimizando a influência de artefatos externos.

Para a caracterização quantitativa, foram combinadas três abordagens complementares. O *PyRadiomics* foi empregado para a extração de 102 descritores, incluindo 9 atributos de forma bidimensional (*shape2D*), 18 de primeira ordem, 24 da matriz de coocorrência de níveis de cinza (GLCM), 16 da matriz de comprimentos de *run* (GLRLM), 16 da matriz de zonas (GLSZM), 14 da matriz de dependência de níveis de cinza (GLDM) e 5 da matriz de diferenças de tons vizinhos (NGTDM). Complementarmente, a biblioteca *Mahotas* [21] foi utilizada para calcular 25 momentos de Zernike, 36 histogramas de padrões locais binários (LBP) e 54 estatísticas de adjacência de limiarização (*thresholding adjacency statistics* - TAS), totalizando 115 atributos adicionais. Por fim, a partir das matrizes de temperatura originais, foram derivadas 9 estatísticas globais, como média, desvio-padrão, máximo, mínimo, mediana, percentis 25 e 75, assimetria e curtose. No total, cada amostra foi representada por 226 características, integrando aspectos morfológicos, texturais e estatísticos da distribuição térmica dos úberes. Um resumo dos grupos de características e respectivas quantidades é apresentado na Tabela 1.

3.3 Pré-processamento e *Augmentation*

O pré-processamento dos dados envolveu inicialmente a separação estratificada em grupos, considerando cada animal como uma unidade única durante a divisão dos conjuntos de treino e teste. Essa estratégia, conhecida como *Stratified Group*

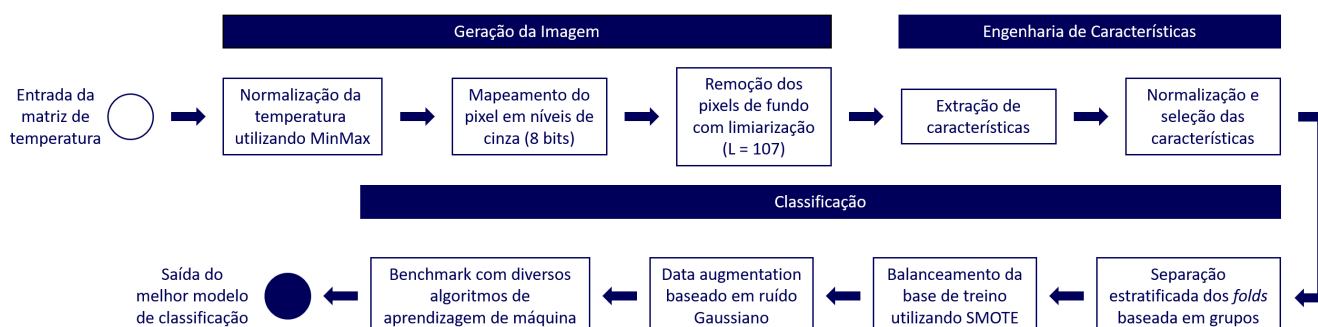


Figure 1. Fluxograma do método proposto para identificação computacional de mastite por termografia, incluindo geração das imagens térmicas, engenharia de características e classificação.

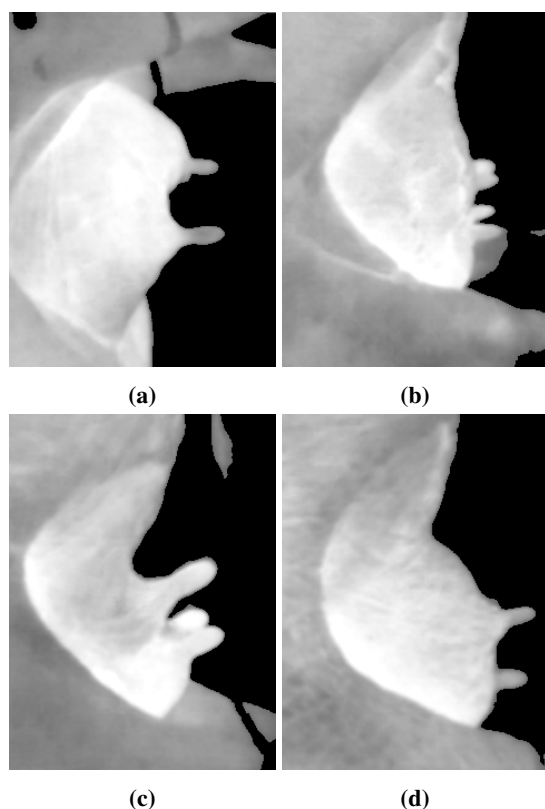


Figure 2. Exemplos de imagens térmicas geradas a partir das matrizes de temperatura, após normalização e remoção do fundo. São apresentadas amostras de úberes classificados como saudáveis (Figuras (a) e (b)) e com mastite subclínica (Figuras (c) e (d)), em diferentes raças.

K-Fold, tem como objetivo evitar vazamento de informação entre os *folds*, assegurando que todas as imagens de um mesmo animal estejam sempre no mesmo conjunto. Além disso, a estratificação garante a preservação da proporção entre as classes (sadio e mastite subclínica) em cada partição, reduzindo o risco de vieses na avaliação.

Outro passo fundamental foi o balanceamento da base de treino utilizando o algoritmo *Synthetic Minority Over-sampling Technique* (SMOTE). Esse algoritmo gera amostras

Table 1. Grupos de características extraídas e suas respectivas quantidades.

Características	Quantidade
Forma	9
Estatísticas de primeira ordem	18
GLCM	24
GLRLM	16
GLSZM	16
GLDM	14
NGTDM	5
Momentos de Zernike	25
LBP	36
TAS	54
Estatísticas da temperatura	9
Total	226

sintéticas da classe minoritária ao interpolar instâncias existentes em seu espaço de características, promovendo maior equilíbrio entre as classes e melhorando a capacidade dos modelos em reconhecer padrões relacionados aos casos positivos de mastite.

Por fim, foi aplicada uma etapa de *data augmentation* nas características de treino. O método consistiu na adição de ruído gaussiano a cada amostra, cuja distribuição normal é definida por:

$$\varepsilon \sim \mathcal{N}(0, \sigma^2) \quad (1)$$

em que ε representa o ruído adicionado a cada atributo, com média zero e variância σ^2 proporcional à variabilidade da própria característica. Essa estratégia contribui para aumentar a diversidade dos dados de treino, tornando os modelos mais robustos a pequenas variações na distribuição térmica e menos suscetíveis a sobreajuste. É importante salientar que as técnicas de balanceamento e *augmentation* foram aplicadas somente na base de treino.

3.4 Metodologia de Classificação

A etapa de classificação foi estruturada como um *benchmark* envolvendo múltiplos algoritmos de aprendizagem de

máquina, com o objetivo de avaliar diferentes abordagens para a identificação de mastite subclínica a partir das características extraídas. Foram incluídos métodos lineares, baseados em árvores de decisão, modelos *ensemble*, algoritmos de *gradient boosting* e redes neurais. A maior parte dos classificadores foi utilizada com seus parâmetros padrões da *framework Scikit-Learn*, com exceção de alguns casos específicos em que ajustes se mostraram necessários. O classificador *K-Nearest Neighbors*, por exemplo, foi configurado com $k = 46$ vizinhos. Este valor foi escolhido como um ponto intermediário para mitigar os riscos tanto de sobreajuste (*overfitting*), associado a um k baixo, quanto de subajuste (*underfitting*), associado a um k alto, buscando assim um modelo com boa capacidade de generalização. Todos os experimentos foram conduzidos com a definição do valor da semente (*random seed*) igual a 42, garantindo a reprodutibilidade dos resultados.

Os algoritmos avaliados foram: Regressão Logística (*Logistic Regression*), Floresta Aleatória (*Random Forest*), Máquinas de Vetores de Suporte (*Support Vector Machines - SVM*) com funções de *kernel* linear, polinomial e radial (*radial basis function - RBF*), KNN (*K-Nearest Neighbors*), *XGBoost*, *AdaBoost*, Árvore de Decisão (*Decision Tree*), *Gradient Boosting*, *Extra Trees*, *Naive Bayes* Gaussiano, *Bagging*, *Stochastic Gradient Descent* (SGD) e um *Multilayer Perceptron* (MLP). No caso do *XGBoost*, foi definido explicitamente o uso da métrica de avaliação *log-loss* e desativado o codificador interno de rótulos, enquanto a Regressão Logística, o SGD e o MLP foram configurados para até 1000 iterações, a fim de prevenir a interrupção prematura do treinamento.

4. Experimentos e Resultados

O desenvolvimento do trabalho foi realizado em um *workstation* equipado com *Windows 11*, processador *Intel Core i9*, 128 GB de memória RAM, 8 TB de armazenamento em SSD NVMe e uma GPU NVIDIA RTX 4090. A implementação foi feita em *Python 3.11*, utilizando as bibliotecas *OpenCV 4.12*, *Scikit-learn 1.6*, *PyRadiomics 3.1* para extração de atributos radiômicos e *Mahotas 1.4* para extração de atributos de textura.

Os experimentos seguiram um protocolo de validação cruzada estratificada com agrupamento, de forma que todas as imagens pertencentes a um mesmo animal fossem mantidas no mesmo conjunto de treino ou de teste, evitando o vazamento de informação entre os *folds*. O processo foi estratificado para preservar a proporção entre as classes e utilizou-se $k = 10$ *folds*.

A avaliação dos experimentos foi conduzida com base em seis métricas: acurácia, sensibilidade, especificidade, precisão, *F1-Score* e AUC. A acurácia indica a proporção total de diagnósticos corretos. A sensibilidade mede a capacidade do modelo em detectar vacas realmente com mastite, sendo importante para evitar falsos negativos que poderiam atrasar o tratamento. A especificidade avalia a capacidade de identificar corretamente os animais saudáveis, reduzindo falsos positivos que poderiam levar a exames desnecessários. A precisão rep-

resenta a proporção de casos classificados como mastite que de fato correspondem à doença. O *F1-Score* integra precisão e sensibilidade em uma medida única, importante no contexto de base desbalanceada entre vacas saudáveis e doentes. Já a AUC quantifica a capacidade global do modelo em distinguir entre mastite e normalidade em diferentes limiares de decisão.

A escolha dessas métricas permitiu uma análise abrangente do desempenho no contexto veterinário. Acurácia fornece uma visão geral, mas sensibilidade e especificidade são fundamentais em diagnósticos biomédicos, pois lidam diretamente com os riscos de erros clínicos. O *F1-Score* ajuda a equilibrar os impactos do desbalanceamento de dados, e o AUC oferece uma avaliação robusta da capacidade discriminativa do modelo, independentemente do ponto de corte usado para decidir se um animal tem ou não mastite.

4.1 Resultados

A Tabela 2 apresenta o desempenho dos diferentes modelos testados. Observa-se que classificadores mais sofisticados, como SVM com *kernel* polinomial e RBF, e modelos de *ensemble* como *Random Forest* e *Extra Trees*, alcançaram as maiores acurácias médias, consistentemente próximas de 90%. Por outro lado, modelos mais simples, como KNN, tiveram desempenho bastante inferior, revelando que a separação entre classes exige abordagens capazes de lidar com relações não lineares entre as características extraídas.

Um aspecto importante é a métrica AUC, que se manteve elevada em vários modelos, atingindo valores acima de 0.95, com destaque para o SVM com *kernel* polinomial (97,83%) e RBF (97,13%). Diferentemente da acurácia, a AUC é invariante ao desbalanceamento entre classes no conjunto de teste, o que a torna mais adequada ao contexto do estudo. Esse resultado sugere que, apesar das limitações do conjunto de dados, alguns modelos conseguem capturar padrões relevantes na distribuição térmica global, apresentando boa capacidade discriminativa em diferentes limiares de decisão.

Ainda assim, os valores de sensibilidade e especificidade mostraram grande variação entre os modelos. Enquanto alguns classificadores conseguiram identificar bem os casos positivos de mastite, outros apresentaram maior acerto em animais saudáveis, evidenciando dificuldades de generalização. Esses resultados reforçam que os achados devem ser vistos como indicativos preliminares de viabilidade da abordagem, mas que só poderão se consolidar com protocolos de coleta mais robustos e com o aumento significativo da base de dados disponível.

4.2 Discussão

Os resultados deste estudo indicam que a distribuição térmica global do úbere, capturada por termografia, contém padrões informativos que possibilitam a detecção de mastite subclínica. O desempenho dos melhores modelos de IA avaliados superou as expectativas para um estudo exploratório, com acurácias próximas de 90% e valores de AUC superiores a 97%, os achados estão alinhados com a literatura, que aponta a TIV como uma ferramenta promissora, mas reforçam a necessidade de

Table 2. Resultados médios dos modelos na identificação de mastite bovina

Modelo	Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score	AUC
Logistic Regression	77,28% ± 13,52	82,28% ± 16,47	57,35% ± 41,20	87,32% ± 18,55	81,57% ± 12,98	87,28% ± 7,49
Random Forest	86,61% ± 10,56	97,50% ± 5,00	42,81% ± 37,69	83,60% ± 15,13	89,16% ± 10,03	92,80% ± 8,29
SVM Linear	80,89% ± 13,88	88,51% ± 14,30	60,35% ± 36,59	85,75% ± 19,27	84,39% ± 12,95	91,99% ± 8,31
SVM Poly	89,06% ± 7,53	93,88% ± 6,54	67,02% ± 38,12	90,78% ± 13,77	91,35% ± 7,37	97,83% ± 3,95
SVM RBF	89,33% ± 10,62	95,72% ± 5,44	64,24% ± 36,76	88,02% ± 16,20	90,79% ± 10,71	97,13% ± 4,12
KNN (k=46)	52,39% ± 21,63	35,66% ± 22,33	68,57% ± 45,08	85,28% ± 29,92	47,68% ± 24,34	74,05% ± 19,66
XGBoost	80,28% ± 14,46	90,09% ± 11,21	50,59% ± 33,38	81,73% ± 20,67	83,77% ± 14,61	88,42% ± 10,95
AdaBoost	82,61% ± 8,39	86,34% ± 12,04	63,68% ± 36,51	87,39% ± 14,65	85,46% ± 9,03	89,00% ± 12,14
Decision Tree	77,44% ± 12,80	87,21% ± 9,71	54,14% ± 39,18	82,89% ± 20,11	82,87% ± 11,08	77,93% ± 15,46
Gradient Boosting	79,94% ± 9,84	88,43% ± 12,06	46,78% ± 37,18	84,34% ± 16,31	84,41% ± 8,43	95,38% ± 5,37
Extra Trees	88,28% ± 9,65	100,00% ± 0,00	45,83% ± 33,62	83,68% ± 14,02	90,47% ± 8,50	95,71% ± 4,32
Gaussian NB	59,28% ± 21,44	57,96% ± 28,86	52,94% ± 40,48	83,53% ± 23,78	62,34% ± 26,06	65,39% ± 26,19
Bagging	85,78% ± 7,01	92,01% ± 11,31	58,44% ± 36,73	87,57% ± 14,26	88,22% ± 7,88	88,99% ± 8,98
SGD	79,00% ± 14,13	84,97% ± 19,11	54,27% ± 38,33	86,18% ± 17,06	82,45% ± 13,69	91,77% ± 5,89
MLP	81,89% ± 9,24	87,83% ± 9,54	63,37% ± 37,53	87,81% ± 18,25	85,87% ± 10,43	91,73% ± 9,29

ampliação da base de dados e do refinamento dos protocolos de aquisição de imagem para alcançar maior acurácia e robustez.

A implantação de um sistema como o proposto na cadeia produtiva é relativamente simples, dependendo, em um primeiro momento, apenas da aquisição de uma câmera térmica adequada. A análise das imagens pode ser adaptada à infraestrutura de cada fazenda: pode ser realizada localmente, em um computador na própria propriedade (*edge computing*), oferecendo diagnósticos rápidos para decisões imediatas; ou de forma remota (*cloud computing*), onde as imagens são enviadas para um servidor com maior poder de processamento, permitindo o uso de modelos mais complexos e a integração com dados de outras fontes.

Contudo, a implementação de um sistema térmico em campo enfrenta desafios significativos, principalmente a necessidade de que a análise considere as condições de temperatura e umidade do ambiente. Conforme apontado pelos autores [11, 13, 18], fatores externos influenciam a temperatura superficial da pele. Estabelecer especificações de coleta muito restritivas, como horários fixos e ambientes climatizados, pode “melhorar os resultados” em um contexto experimental, mas torna o sistema impraticável no uso real e no cotidiano de uma fazenda. O avanço metodológico neste trabalho consiste em modelos de IA que incorporam explicitamente a temperatura e a umidade como características no processo de classificação. A inclusão dessas características se mostrou um fator fundamental para o desempenho e a capacidade de generalização dos algoritmos, assim como no trabalho de Caracuschanski [18], cujos algoritmos para predição da etiologia da mastite também dependeram de características ambientais para alcançar alta performance.

Adicionalmente, o fator racial demonstrou ser de grande importância. O trabalho de Ribeiro et al. [13] aborda diretamente essa lacuna ao comparar as raças Girolando e Jersey, indicando que, embora a raça Girolando apresente temperaturas superficiais mais baixas, a TIV é eficaz para o diagnóstico em ambas, contanto que as particularidades térmicas de cada raça sejam consideradas. Nosso estudo reforça essa conclusão, ao tratar essas particularidades como características essenciais que foram incorporadas ao modelo de *machine learning* para garantir sua robustez e aplicabilidade entre diferentes raças.

As vantagens da abordagem proposta, que une TIV e ML, são notáveis quando comparadas aos métodos tradicionais. Trata-se de um método não invasivo, rápido e com potencial para análise em tempo real, contrastando com os diagnósticos laboratoriais, que são caros, trabalhosos e demorados [20]. Após a aquisição inicial do equipamento, o custo por animal torna-se muito baixo, e a automação do processo elimina a subjetividade da interpretação humana, uma das principais limitações de testes como o CMT.

A análise do desempenho preditivo revela uma nuance importante. Modelos como o SVM (com *kernels* polinomial e RBF) apresentaram um excelente equilíbrio entre as métricas, com AUCs próximos de 98%, indicando uma capacidade discriminativa quase perfeita. Por outro lado, o modelo *Extra Trees*, embora com acurácia levemente inferior, atingiu 100% de sensibilidade, mostrando-se promissor como ferramenta de triagem. Um animal sinalizado por ele poderia ser encaminhado para um teste confirmatório, otimizando recursos e garantindo que nenhum caso passe despercebido.

No entanto, esse resultado deve ser visto como preliminar, pois pode refletir limitações do conjunto de dados e não necessariamente se reproduzir em cenários de campo. Contudo, a alta taxa de falsos positivos, indicada pela baixa especificidade, é uma consequência direta da sobreposição das assinaturas térmicas entre úberes sadios e aqueles com inflamação subclínica, um desafio inerente ao problema que motiva o desenvolvimento de protocolos de coleta de dados ainda mais robustos.

Em comparação da nossa proposta com as demais, nosso método apresenta uma troca entre performance e viabilidade de implementação. A principal desvantagem da nossa abordagem é que a acurácia dos melhores modelos, embora alta (próxima de 90%), não alcança os picos de desempenho reportados por modelos de aprendizado profundo como o de Silva [20], que superou 97%. Contudo, essa diferença é justificada pelo cenário de recursos limitados. As abordagens de deep learning dependem de arquiteturas complexas, grande poder computacional e técnicas avançadas para funcionarem, o que é um obstáculo para a implantação em tempo real na fazenda. Nossa solução, por outro lado, é mais leve, utilizando de abordagens consolidadas de extração de características e classificadores clássicos, que demandam menos recursos e são

mais estáveis com um conjunto de dados reduzido. Portanto, o método proposto tende a obter uma acurácia um pouco inferior a literatura, mas se destaca em viabilidade, estabilidade e simplicidade de implementação em condições de campo realistas.

Ainda no âmbito dos métodos baseados em aprendizagem profunda, conforme amplamente consolidado na literatura, tais abordagens são altamente dependentes de grandes volumes de dados rotulados para alcançar desempenho e generalização adequados [22]. No contexto deste estudo, o número limitado de amostras disponíveis inviabilizou o treinamento eficaz de modelos profundos, tornando-os suscetíveis a *overfitting* e instabilidade nos resultados. Além disso, o elevado custo computacional e a menor interpretabilidade desses modelos representam obstáculos adicionais em uma investigação de caráter exploratório. Por essas razões, optou-se por uma abordagem baseada em engenharia de características, que permite a extração explícita e interpretável de descritores robustos de forma, textura e estatística, garantindo maior estabilidade, transparência e validade dos resultados obtidos.

Dentre as limitações do presente estudo, duas principais devem ser consideradas. Primeiramente, o tamanho reduzido da amostra, que compromete a capacidade de generalização dos modelos. Adicionalmente, a utilização de imagens laterais do úbere, que limita o campo de visão. Como cada úbere possui quatro quartos mamários que podem apresentar mastite de forma independente ou combinada, uma visão lateral corre o risco de não capturar uma inflamação localizada ou de ter seu sinal térmico atenuado pela média da temperatura de quartos sadios adjacentes. É fundamental ter em mente que os dados foram originalmente coletados para uma pesquisa com foco mais estatístico, e não especificamente para análise de imagens por IA, o que resultou em um protocolo de coleta que não é o mais apropriado para esta finalidade.

As implicações futuras desta pesquisa apontam para diversas linhas de investigação, sendo a principal meta a expansão da base de dados. Planeja-se a coleta de mais imagens, com um protocolo já desenhado para aplicações em visão computacional e IA, incluindo maior diversidade de animais, raças e condições ambientais para treinar modelos mais robustos. Sugere-se o estabelecimento de um novo protocolo de aquisição de imagem que capture a região inferior do úbere, representando os quatro quartos em uma única imagem, além de imagens individuais por lado para uma análise mais detalhada e localizada. Futuramente, a exploração de modelos multimodais, que combinem dados termográficos com informações laboratoriais (CCS) ou de ordenhas robotizadas, representa um caminho promissor. O objetivo final é a criação de uma ferramenta de monitoramento da saúde animal em larga escala, de baixo custo e alta eficiência, que possa ser integrada aos sistemas de gestão das propriedades leiteiras modernas.

5. Conclusão

Os resultados obtidos neste estudo confirmam o potencial da integração entre TIV e modelos de ML como alternativa viável para a identificação precoce da mastite subclínica bovina. Apesar das limitações relacionadas ao tamanho da amostra e às condições de coleta, os modelos mais avançados, em especial o SVM com *kernel*s polinomial e RBF, alcançaram valores de AUC superiores a 97%, o que reforça o forte potencial das imagens térmicas como fonte de informação para diferenciar animais sadios de infectados.

A principal contribuição deste trabalho foi validar a abordagem que combina a engenharia de características extraídas de imagens térmicas com um *benchmark* de classificadores, mas também identificar o perfil de desempenho de cada modelo. A análise revelou que, enquanto os SVMs oferecem um diagnóstico equilibrado, modelos como o *Extra Trees* podem atingir 100% de sensibilidade, qualificando-os como uma ferramenta de triagem ideal. Essa distinção é fundamental para a aplicação prática da tecnologia no campo, demonstrando que a análise computacional pode superar a subjetividade da inspeção visual e as limitações dos métodos de diagnóstico tradicionais.

A abordagem proposta apresenta vantagens claras em relação aos métodos convencionais, por ser não invasiva, de baixo custo e passível de automatização, o que favorece sua aplicação em sistemas de produção leiteira modernos. Entretanto, para consolidar sua adoção prática, torna-se essencial ampliar a base de dados com protocolos de aquisição de imagens otimizados para IA, visando principalmente melhorar a especificidade dos modelos de alta sensibilidade. Dessa forma, este trabalho contribui para o avanço da zootecnia de precisão ao demonstrar a viabilidade do uso de TIV associada ao ML como ferramenta de apoio ao diagnóstico veterinário, apontando um caminho promissor para soluções robustas e acessíveis, com impacto direto na saúde animal, na produtividade do rebanho e na sustentabilidade da cadeia leiteira.

Agradecimentos

Este trabalho é apoiado parcialmente pelo Fundo Nacional de Desenvolvimento da Educação (FNDE) através do Programa de Educação Tutorial (PET - Conexões Ciência da Computação).

Contribuição dos Autores

Víctor Emmanuel Olmi Louzada atuou na conceituação, metodologia, software, investigação e redação do rascunho original; Lucas Grassano Lattari foi responsável pela supervisão e administração do projeto; Matheus de Freitas Oliveira Baffa contribuiu com a supervisão (co-supervisão) e redação (revisão e edição); Sandra Helena Gabaldi Wolf e Isabele Pessoa Ribeiro contribuíram com recursos (dados) e validação (especialidade veterinária).

Referências

- [1] USDA. *Dairy: World Markets and Trade*. Washington, D.C.: [s.n.], 2023. Relatório técnico. Acesso em: 25 out. 2025. Disponível em: <https://apps.fas.usda.gov/psdonline/circulars/dairy.pdf>.
- [2] EMBRAPA. *Relatório anual da produção de leite no Brasil*. Juiz de Fora, MG: [s.n.], 2024. Relatório institucional. Acesso em: 25 out. 2025. Disponível em: <https://www.embrapa.br/gado-de-leite>.
- [3] CARVALHO, G. R. 10 reflexões sobre competitividade e desafios do leite brasileiro. In: EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA GADO DE LEITE (Ed.). *Anuário Leite 2024*. São Paulo: Texto Comunicação Corporativa, 2024. p. 16–19.
- [4] PEGORARO, L. M. C. (Ed.). *Biossegurança na bovinocultura leiteira*. Brasília, DF: Embrapa, 2018.
- [5] FRANCO, A. B. et al. Mastite bovina e as suas consequências na saúde pública. *PUBVET*, v. 16, n. 10, p. 1–10, oct 2022.
- [6] SCHWARZ, D. et al. A review of the most common and emerging technologies for the diagnosis of subclinical mastitis in dairy cows. *Animals*, v. 11, n. 6, p. 1761, jun 2021.
- [7] RUEGG, P. L. A 100-year review: Mastitis detection, management, and prevention. *Journal of Dairy Science*, v. 100, n. 12, p. 10381–10397, 2017.
- [8] ZHAO, K. et al. An automatic recognition of subclinical mastitis in dairy cattle based on infrared thermography and machine learning. *Computers and Electronics in Agriculture*, v. 182, p. 106001, mar 2021.
- [9] HOVINEN, M. et al. Detection of clinical mastitis with the help of a thermal camera. *Journal of Dairy Science*, v. 91, n. 12, p. 4592–4598, dec 2008.
- [10] COLAK, A. et al. Short Communication: Early detection of mastitis using infrared thermography in dairy cows. *Journal of Dairy Science*, v. 91, n. 11, p. 4244–4248, 2008.
- [11] OLIVEIRA, A. V. D. et al. Termografia infravermelha como técnica para detecção de mastite bovina subclínica. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, v. 74, n. 6, p. 992–998, 2022.
- [12] METZKER, C. A. et al. Uso da termografia infravermelha para diagnóstico de mastite clínica e subclínica em vacas girolando. *Pesquisa Veterinária Brasileira*, v. 40, n. 8, p. 752–758, ago 2020.
- [13] RIBEIRO, I. P. et al. Termografia infravermelha na detecção de mastite clínica e subclínica em bovinos de leite: comparação entre as raças Girolando e Jersey. *Ciência Animal Brasileira*, v. 24, p. e–76726P, 2023.
- [14] BORCHERS, M. R.; CHANG, Y. M.; WADSWORTH, B. A. A systematic review of machine learning and computer vision in dairy cattle. *Journal of Dairy Science*, v. 104, n. 12, p. 13005–13026, dec 2021.
- [15] SILVA, A. S. et al. Ferramenta para segmentação e classificação de mastite bovina em imagens termográficas. In: *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 336–347. Evento Online.
- [16] ZANINELLI, M. et al. First evaluation of infrared thermography as a tool for the monitoring of udder health status in farms of dairy cows. *Sensors*, v. 18, n. 3, p. 862, 2018.
- [17] WATZ, S. et al. Technical note: Automatic evaluation of infrared thermal images by computerized active shape modeling of bovine udders challenged with *Escherichia coli*. *Journal of Dairy Science*, v. 102, n. 5, p. 4541–4545, 2019.
- [18] CARACUSCHANSKI, M. *Uso de aprendizado de máquina para predição da etiologia infecciosa da mastite em sistemas de ordenha robótica*. Tese (Tese de Doutorado) — Universidade Federal de Minas Gerais, Belo Horizonte, 2023.
- [19] FREITAS, P. P. *Diagnóstico automático da ceratoconjuntivite infecciosa bovina usando imagens termográficas e redes neurais convolucionais*. Dissertação (Dissertação de Mestrado) — Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto, 2019.
- [20] SILVA, L. H. G. *Detecção da mastite subclínica bovina em imagens termográficas utilizando transferência de aprendizado sequencial e otimização bayesiana*. Dissertação (Dissertação de Mestrado) — Universidade Federal de Campina Grande, Campina Grande, 2023.
- [21] COELHO, L. P. Mahotas: Open source software for scriptable computer vision. *Journal of Open Research Software*, v. 1, n. 1, p. e3, 2013.
- [22] ALZUBAIDI, L. et al. Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *Journal of Big Data*, Springer, v. 8, n. 1, p. 53, 2021.