

PROJETO DE DISSERTAÇÃO

(Modelo aprovado pelo Colegiado em sua 8ª reunião realizada em 03/09/2015)

**Programa de Pós-graduação em Ciência da Computação
Universidade Federal de Uberlândia**

Aluno: Marcio Salmazo Ramos

Orientador: Maurício Cunha Escarpinati

Coorientador: Daniel Duarte Abdala

**Título do trabalho: Avaliação de Dor em Camundongos com Redes
Neurais e Visão Computacional**

Data de início como aluno regular: 14/03/2024

Previsão de defesa: 01/07/2026

1. Introdução

Um dos principais desafios para as pesquisas em ambientes clínicos veterinários é a avaliação precisa de estímulos dolorosos em animais, uma vez que a ausência da linguagem verbal nos animais exige métodos indiretos de avaliação, como a observação de sinais fisiológicos e comportamentais. A dor não tratada compromete seriamente o bem-estar destes animais, podendo levar a consequências fisiológicas e comportamentais duradouras, como sensibilização do sistema nervoso central, hiperalgesia (sensibilidade exagerada à dor ou estímulos dolorosos), bem como alterações no comportamento social.

Por serem incapazes de verbalizar suas sensações, a dor em camundongos (e outros animais de modo geral) costuma ser avaliada com base em comportamentos observáveis e alterações fisiológicas. Nesse contexto, foi criada a *Grimace Scale*, uma escala de avaliação desenvolvida para mensurar a dor com base em alterações sutis nas expressões faciais dos animais. A *Mouse Grimace Scale* (MGS), por exemplo, observa mudanças em cinco regiões faciais distintas do camundongo (orelhas, olhos, bochechas, nariz e bigodes), as quais sofrem alterações expressivas em resposta a estímulos dolorosos.

A escolha dos camundongos como objetos de experimentação se justifica por sua ampla utilização em pesquisas biomédicas e comportamentais, resultado de uma combinação de fatores biológicos e operacionais. Dentre esses fatores, destacam-se o baixo custo de manutenção, a facilidade de manejo, e o elevado grau de homologia genética com seres humanos — estimado em aproximadamente 95% — o que os torna altamente relevantes para o estudo de respostas fisiológicas e comportamentais.

A metodologia descrita pela MGS foi importante por padronizar os sinais visuais que identificam a presença de dor, permitindo comparações mais rigorosas e sólidas entre estudos. No entanto, os resultados ainda permanecem condicionados à percepção humana, a qual pode comprometer a confiabilidade e precisão das análises em decorrência de vieses. Neste contexto, técnicas automatizadas baseadas em visão computacional e aprendizado de máquina surgem como alternativas promissoras para reduzir o viés humano e a necessidade de mão de obra especializada.

Portanto, este trabalho propõe o desenvolvimento de um software destinado à implementação de técnicas de visão computacional associadas ao aprendizado de máquina, com ênfase em arquiteturas de redes neurais convolucionais (CNNs) e Vision Transformers (ViTs). O objetivo central consiste na identificação automática, em imagens faciais de camundongos, dos marcadores definidos pela MGS, classificando-as quanto à presença e ao grau de dor, reconhecendo os padrões morfológicos sutis de forma rápida, precisa e escalável.

1.1 Motivação

A motivação para este trabalho parte da premissa de que a dependência da interpretação visual por parte do avaliador, mesmo com o uso de uma escala padronizada, constitui um obstáculo à precisão dos diagnósticos e à eficiência no cuidado com os animais. Adicionalmente, a identificação de dor em animais representa um desafio ético e científico. A indução de dor, se não for devidamente controlada, viola princípios fundamentais de bem-estar animal. O problema está na dificuldade em reconhecer os padrões fisiológicos que denotam o sofrimento, o que pode levar à negligência e, conseqüentemente, estender o período de desconforto do animal.

O desenvolvimento de uma ferramenta automatizada, baseada em técnicas de aprendizado de máquina e visão computacional, que seja capaz de detectar em tempo real a presença de dor com base nos critérios estabelecidos pela MGS tem forte potencial para sanar os problemas descritos previamente. Tratamentos preventivos ou paliativos podem ser aplicados de maneira segura e eficiente, levando à redução do sofrimento. Análises clínicas, onde diferentes estímulos precisam ser aplicados ao animal, também podem ser concluídas com maior agilidade, sem demandar a análise de um especialista. Tal imediatismo ajuda a mitigar o período de estresse no animal e torna o processo eficaz e confiável.

1.2 Objetivos e Desafios da Pesquisa

Este trabalho tem como principal objetivo a contribuição para o setor clínico veterinário por meio da construção de uma ferramenta automatizada, voltada para a detecção e classificação de padrões faciais que expressam a presença de dor em animais. A proposta se baseia na MGS associada a técnicas de visão computacional e aprendizado de máquina, visando superar as limitações dos métodos convencionais, que ainda dependem predominantemente da observação humana manual.

O estudo também se propõe a avaliar, de forma abrangente, a aplicabilidade da ferramenta em contextos de produção. Espera-se, portanto, agregar maior confiabilidade aos resultados das análises, tornando-a atrativa para instituições científicas e clínicas. A partir dessa premissa, o estudo se desdobra nos seguintes objetivos específicos:

1. Levantamento do estado da arte, identificar os métodos de classificação de imagens mais adequados que estão disponíveis na atualidade;
2. Especificar detalhadamente os equipamentos e os protocolos empregados para a captura das imagens;
3. Desenvolvimento de um software capaz de extrair e organizar as imagens, crucial para a geração da base de dados;
4. Realizar a organização e o pré-processamento das imagens, a fim de adequá-las à entrada dos classificadores;
5. Implementar os diferentes modelos de classificadores, essenciais para o processamento dos dados coletados previamente;
6. Treinar os modelos implementados e armazenar seus resultados;
7. Realizar uma análise comparativa entre diferentes modelos aplicados, buscando determinar quais abordagens são mais eficazes para o projeto;
8. Refinar os modelos, ajustando seus hiperparâmetros para extrair o melhor resultado possível;

9. Investigar as implicações práticas da implementação dessa tecnologia no setor veterinário, o que envolve uma análise detalhada de custo-benefício, impactos positivos, potenciais desafios técnicos e operacionais.

1.3 Hipótese

A hipótese deste estudo parte da premissa de que classificadores inteligentes possam identificar diferentes regiões de interesse em imagens de camundongos que, ao serem analisadas seguindo os padrões estabelecidos pela MGS, permitem apontar a presença de dor e quantificá-la. Acredita-se que técnicas de visão computacional associadas ao treinamento de máquina, sejam capazes de detectar nuances sutis nas imagens que não são facilmente perceptíveis ao observador não especialista ou que não são capturadas por métodos convencionais.

1.4 Contribuições

Liste as contribuições do seu trabalho.

2. Revisão da Literatura Correlata

O trabalho em questão lida essencialmente com um problema classificatório baseado em imagens, mais especificamente, a classificação de dor em camundongos com base em seus retratos faciais, utilizando como parâmetro a MGS. Neste contexto, diferentes estudos foram conduzidos na literatura científica a para a implementação de técnicas inovadoras na tarefa de classificação baseada em visão computacional.

O estudo desenvolvido por (TUTTLE et al., 2018) propõe uma abordagem automatizada da MGS para identificar dor em camundongos com base em imagens de suas expressões faciais após laparotomia (procedimento cirúrgico que envolve a abertura da cavidade abdominal por meio de uma incisão na parede abdominal), dessa forma, é possível verificar a eficácia dos medicamentos analgésicos administrados e reduzir a necessidade de observação humana intensiva e altamente especializada para pontuar imagens.

O modelo demonstrou não apenas alta precisão interna, mas também boa capacidade de generalização, identificando corretamente a dor nas imagens avaliadas. Contudo, os autores chamaram a atenção para a dificuldade do modelo, e também dos avaliadores humanos, em diferenciar imagens de camundongos adormecidos daquelas com expressões de dor sutil. Adicionalmente, reforçaram a necessidade de novos conjuntos de treinamento para calibrar o sistema a fim de abranger diferentes raças, como as de pelagem preta ou aguti. O trabalho apresentado nesta dissertação apresenta uma abordagem mais generalizada, com o intuito de classificar dor em diferentes contextos, não se limitando apenas ao cenário pós-cirúrgico.

Outro estudo de destaque nesta área foi desenvolvido por (WOTTON et al., 2020), que propôs um sistema automatizado de fenotipagem comportamental para camundongos submetidos ao teste da formalina. O trabalho busca atender a uma demanda de métodos objetivos, reprodutíveis e escaláveis para avaliação de comportamentos nocifensivos (indicam resposta à percepção de dor), como lamber, morder ou levantar a pata. Neste estudo os padrões estipulados pela MGS não foram utilizados para a análise, uma vez que tais comportamentos não se restringem à estímulos faciais. Como resultado, o sistema foi capaz de alcançar 98% de concordância com a avaliação humana.

Trabalhos similares de classificação de dor também foram desenvolvidos levando em consideração outras espécies de animais. Ainda no âmbito veterinário, o estudo conduzido por (LENCIONI et al., 2021). buscou monitorar a dor em cavalos por meio da classificação automática de suas expressões faciais, seguindo os padrões da *Grimace Scale*. A construção do sistema envolveu o treinamento de CNNs específicas para cada ponto chave da face (orelhas, olhos e boca/narinas), as quais foram posteriormente combinadas em um classificador mais abrangente para integrar as decisões regionais e classificar a dor em imagens completas dos cavalos.

Os resultados demonstraram que o modelo treinado para analisar a posição das orelhas apresentou a maior acurácia individual (90.3%), enquanto os modelos de olhos e boca/narinas obtiveram acurácias de 65.5% e 74.5%, respectivamente. Ao combinar essas informações para avaliar imagens completas, o sistema atingiu uma acurácia geral de 75.8% na classificação em três níveis de dor. Quando a tarefa foi simplificada para distinguir apenas entre 'presença' ou 'ausência' de dor, a acurácia aumentou significativamente para 88.3%, revelando um comparativo importante entre a relevância particular de cada elemento chave para o treinamento.

Diferentes estudos de classificação de dor também foram desenvolvidos para um contexto humano, como é o caso do trabalho descrito por (KARAMITSOS; SLADJI; MODAK, 2021), que propõe uma abordagem baseada em CNNs para detecção automática de dor em pacientes clínicos, utilizando expressões faciais como principal fonte de informação. O estudo busca oferecer uma solução que possa ser integrada a sistemas hospitalares para monitoramento contínuo e automático da dor em pacientes com limitações verbais.

A metodologia emprega a base de dados *UNBC-McMaster Shoulder Pain Expression Archive*, um dos conjuntos mais amplamente utilizados na literatura, contendo vídeos de 100 pacientes com dor genuína induzida por movimentos articulares. Após o pré-processamento, as imagens foram utilizadas para treinar uma arquitetura modificada da VGG16, adaptada especificamente para a tarefa de classificação binária. A avaliação do modelo revelou acurácia de 92,5% e recall de 86,96% (mede a capacidade do modelo de identificar corretamente os casos positivos) em um conjunto de teste de 400 imagens. Esses resultados superam os de outras abordagens mencionadas na literatura, incluindo modelos híbridos e redes recorrentes.

Apesar dos avanços alcançados, os autores reconhecem limitações importantes: a necessidade de ampliar e balancear ainda mais a base de dados, melhorar a qualidade das imagens utilizadas. Também destacam que o modelo atual ainda não contempla comportamentos faciais que possam ser confundidos com dor, sugerindo que etapas futuras de aprimoramento envolvam a detecção e correção de falsos positivos e a inclusão de mais variabilidade nos dados de treinamento.

Apesar dos avanços proporcionados, o foco em padrões locais dificulta a integração de informações mais amplas dentro da imagem, o que pode comprometer o desempenho em tarefas que exigem maior contextualização espacial. Além disso, operam sobre grids fixos e requerem pré-processamentos específicos, o que pode limitar sua flexibilidade frente a diferentes tipos de entrada. Nesse cenário, surgem alternativas, como as ViTs, que propõem uma abordagem distinta proveniente da arquitetura usada para o processamento de linguagem natural, capaz de capturar relações globais entre as entradas desde o início do treinamento.

O estudo desenvolvido por (BARMAN et al., 2024) propõe uma solução para a detecção de doenças em folhas de tomate, trazendo comparativos entre a ViT e a Inception v3 – um modelo baseado em CNNs. Adicionalmente, é proposto o desenvolvimento de um aplicativo para dispositivos Android capaz de realizar a detecção em tempo real a partir de imagens capturadas por smartphones, com potencial aplicação na agricultura de precisão. Como resultado, a arquitetura baseada na ViT demonstrou maior precisão, recall e F1-score em quase todas as classes de doença, obtendo acurácia de 97,34% no treino e 95,76% na validação. A Inception V3 obteve acurácia de 94,8% no treino e 94,02% na validação. O que demonstra um forte potencial das ViTs para o processo classificatório.

No contexto de detecção de dor, o trabalho conduzido por (FIORENTINI; ERTUGRUL; SALAH, 2022) propõe a primeira pipeline totalmente baseada em *Vision Transformer* e *Video Vision Transformers* (ViViT) na tarefa de classificação binária de dor facial. Para o treinamento, foi utilizado o *dataset UNBC-McMaster Shoulder Pain*, que contém vídeos de pacientes com dor induzida clinicamente e rotulados com base na escala PSPI (*Prkachin and Solomon Pain Intensity*), a qual codifica ações faciais relacionadas à dor. Importante salientar que o ViViT é uma extensão do *Vision Transformer*, adaptada especificamente para dados de vídeo.

Enquanto o ViT processa imagens estáticas, o ViViT processa sequências de quadros, capturando informações espaciais e temporais simultaneamente.

Compreender as características e diferenças entre as arquiteturas ViTs e CNNs é essencial para contextualizar os avanços recentes em classificação de imagens no campo da visão computacional. O estudo descrito por (MAURICIO; DOMINGUES; BERNARDINO, 2023) apresenta uma revisão sistemática da literatura, com foco na comparação entre ambas as arquiteturas em tarefas de classificação. Considerando a consolidação das CNNs como abordagem predominante e o surgimento recente dos ViTs, os autores analisam as condições em que cada arquitetura apresenta melhor desempenho, além de discutir suas vantagens, limitações e aplicações práticas.

Os resultados mostraram que os ViTs tendem a apresentar melhor desempenho quando aplicados a datasets menores, por conseguirem capturar relações globais via mecanismos de autoatenção. Em contrapartida, as CNNs foram mais robustas em conjuntos de dados maiores e mais diversos, além de manterem melhor generalização em tarefas com menos ruído. Arquiteturas híbridas — que combinam convoluções com atenção — também se destacaram por unir o melhor dos dois mundos. A revisão conclui que não há um “vencedor absoluto” entre ViTs e CNNs, mas sim cenários mais adequados para cada arquitetura.

3. Método de Pesquisa

As estratégias metodológicas adotadas para o desenvolvimento deste estudo foram organizadas em cinco etapas principais. A primeira refere-se ao processo de aquisição da base de dados, cujo objetivo foi garantir a captura de vídeos dos camundongos ao longo do protocolo de indução à dor, buscando registrar as expressões de desconforto de acordo com os critérios estabelecidos pela *Mouse Grimace Scale* (MGS). A adoção da MGS foi fundamental para padronizar o registro das alterações morfológicas associadas à manifestação de dor, assegurando consistência na classificação.

Para a execução desta etapa, inicialmente serão definidos os protocolos de coleta de dados, incluindo a padronização do ambiente experimental, o posicionamento e configuração dos equipamentos de captura, bem como os procedimentos de manipulação e contenção dos animais. Em seguida, o protocolo de indução controlada de dor, será definido e aplicado de acordo com a metodologia validada na no estudo de Langford et al. (2010) sobre codificação de expressões faciais de dor em roedores. Tal abordagem possibilita a observação de respostas consistentes com o estímulo nociceptivo, além de assegurar a reprodutibilidade experimental e rigor ético no manejo dos animais.

A segunda etapa desta pesquisa consiste na revisão bibliográfica exploratória, com foco em modelos de aprendizado profundo empregados em tarefas de classificação de imagens. Considerando que o classificador representa a principal ferramenta computacional responsável pelo processo de aprendizado de máquina, a seleção das arquiteturas demandou uma análise criteriosa de critérios como desempenho, capacidade de generalização, robustez e adequação ao contexto da visão computacional aplicada à detecção de dor em animais.

A terceira etapa do projeto consiste no tratamento dos vídeos coletados para a construção da base de imagens utilizada no estudo. Inicialmente, os vídeos serão submetidos a um processo de extração de *frames*, conduzido por profissionais da equipe veterinária. O foco será a identificação e seleção de quadros que evidenciassem alterações faciais relevantes para a avaliação da dor segundo os critérios da *Mouse Grimace Scale* (MGS), com ênfase em regiões como orelhas, olhos, focinho e boca.

Para viabilizar o processo de forma eficiente e padronizada, será desenvolvido um software dedicado à manipulação e análise dos vídeos. Entre as funcionalidades implementadas, destaca-se um módulo específico para extração de imagens, o qual possibilita o recorte e armazenamento individual das regiões de interesse em um mesmo frame, associando rótulos pré-definidos referentes à presença e intensidade dor.

Espera-se também um feedback ao longo de sua utilização, a fim de incorporar melhorias iterativamente ao software, resultando em melhorias na usabilidade e adequação às necessidades práticas da atividade. Além de otimizar o trabalho da equipe, o sistema desenvolvido contribuiu para a construção de uma base de dados consistente e bem estruturada, assegurando a qualidade dos experimentos subsequentes.

A quarta etapa deste trabalho tem como foco a construção e o treinamento dos modelos classificadores selecionados previamente por meio da pesquisa exploratória. Essa etapa compreende a definição e configuração da estrutura matemática e computacional necessária para que as redes fossem capazes de processar os dados e aprender os padrões relevantes para o problema proposto. Em termos gerais, ela estabelece como cada modelo é organizado, como os dados são transformados ao longo de suas camadas e de que forma seus parâmetros são ajustados durante o processo de aprendizado.

Com o intuito de automatizar e padronizar o processo experimental, serão desenvolvidas aplicações independentes para a construção e treinamento dos classificadores. Entre as funcionalidades implementadas destacam-se: a implementação de interfaces gráficas para a melhor experiência do usuário, a seleção da base de dados e definição dos parâmetros de entrada, a construção automatizada das arquiteturas com possibilidade de ajuste de hiperparâmetros e o monitoramento das métricas resultantes do treinamento.

Por fim, a etapa final do desenvolvimento deste projeto consiste na definição das técnicas de avaliação dos resultados, uma atividade crucial para mensurar a eficácia dos modelos e assegurar a validade das conclusões obtidas. Para isso, será elaborado um protocolo de avaliação com os seguintes objetivos: medir a capacidade dos modelos em detectar e quantificar a dor, comparar o desempenho entre as arquiteturas testadas e avaliar a robustez e a capacidade de generalização das abordagens. Além disso, a análise será complementada por técnicas estatísticas, como a construção de matrizes de confusão e o cálculo de precisão (*precision*), revocação (*recall*) e *F1-Score*, permitindo uma avaliação abrangente do desempenho dos modelos.

4. Resultados Esperados

Espera-se que o desenvolvimento deste trabalho resulte na construção de um software robusto, capaz de realizar a detecção automática de dor em camundongos, bem como quantificar precisamente sua intensidade, seguindo critérios estabelecidos pela MGS. Contribuindo para a eficácia de equipes veterinárias no contexto de experimentos clínicos, além de conferir maior conforto aos animais mediante intervenções mais rápidas (conforme necessário).

No âmbito dos resultados quantitativos, espera-se que os modelos de classificação desenvolvidos alcancem índices satisfatórios de acurácia, precisão, revocação e *F1-Score*, demonstrando capacidade de identificar padrões faciais sutis relacionados à manifestação de dor, superando a variabilidade humana típica desse tipo de avaliação. Também se espera observar diferenças de desempenho entre as arquiteturas adotadas, contribuindo para a compreensão das condições nas quais cada modelo apresenta maior eficiência.

Além dos resultados métricos, espera-se que a ferramenta desenvolvida apresente aplicabilidade prática em cenários reais de pesquisa e manejo veterinário, permitindo a automatização parcial ou total da avaliação da dor, redução de vieses subjetivos, minimização do tempo de análise.

Por fim, também busca-se alcançar a geração de um repositório relevante para a comunidade científica, servindo como uma base padronizada para estudos futuros, o que estimula a replicabilidade, comparabilidade e extensão dos resultados aqui obtidos. Contribuindo de forma significativa para o avanço em pesquisas no setor veterinário e no bem-estar animal.

5. Cronograma de Execução

Detalhe o cronograma das principais etapas de seu trabalho finalizando pela data da defesa.

6. Justificativa pelo atraso na entrega do projeto (tópico obrigatório somente no caso de entrega do projeto foram do prazo regulamentar*)

O manejo dos camundongos e a fase inicial do estudo — referente à extração de quadros relevantes dos vídeos e à posterior classificação destinada à construção do *ground truth* — precisaram ser conduzidas por uma equipe especializada da área veterinária, uma vez que tais profissionais detêm a expertise técnica e o treinamento adequados tanto para a manipulação ética dos animais quanto para a avaliação precisa das expressões associadas à dor, garantindo a confiabilidade da base gerada.

Em virtude dessa dependência técnica, o desenvolvimento do trabalho ficou sujeito à colaboração direta com a equipe veterinária para a obtenção dos dados necessária ao treinamento dos modelos de aprendizado de máquina, os quais, demandam um volume significativo de amostras a fim de alcançar níveis satisfatórios de desempenho, confiabilidade e capacidade de generalização. O início da coleta das imagens sofreu atrasos substanciais em decorrência dos trâmites relacionados ao Comitê de Ética para o uso de animais, cuja análise e aprovação prévias são obrigatórias para a realização de experimentos propostos nesse projeto.

Infelizmente, os contratempos associados à construção da base de dados inviabilizaram a conclusão das etapas finais do projeto dentro do prazo inicialmente estabelecido pelo cronograma. Ainda assim, destaca-se que houveram avanços significativos na execução das atividades previstas, demonstrando o comprometimento com o desenvolvimento da pesquisa e o progresso efetivamente alcançado até o presente momento.

7. Bibliografia

BARMAN, U. et al. Vit-smartagri: Vision transformer and smartphone-based plant disease detection for smart agriculture. *Agronomy*, MDPI, v. 14, 2024. Disponível em: <<https://doi.org/10.3390/agronomy14020327>>.

FIORENTINI, G.; ERTUGRUL, I. O.; SALAH, A. A. Fully-attentive and interpretable: vision and video vision transformers for pain detection. *Research Gate*, 2022. Disponível em: <<https://doi.org/10.48550/arXiv.2210.15769>>.

KARAMITSOS, I.; SLADJI, I.; MODAK, S. A modified cnn network for automatic pain identification using facial expressions. *Journal of Software Engineering and Applications*, Scientific Research Publishing, v. 14, p. 400–417, 2021. Disponível em: <<https://doi.org/10.1371/journal.pone.0258672>>.

LENCIONI, G. C. et al. Pain assessment in horses using automatic facial expression recognition through deep learning-based modeling. *Plos one*, 2021. Disponível em: <<https://doi.org/10.1371/journal.pone.0258672>>.

MAURICIO, J.; DOMINGUES, I.; BERNARDINO, J. Comparing vision transformers and convolutional neural networks for image classification: A literature review. *Applied Sciences*, MDPI, v. 13, 2023. Disponível em: <<https://doi.org/10.3390/app13095521>>.

TUTTLE, A. H. et al. Deep neural networks to assess pain from mouse facial expressions. *Molecular Pain*, SAGE Publications, v. 14, p. 1–9, 2018. Disponível em: <<https://journals.sagepub.com/doi/full/10.1177/1744806918763658>>.

WOTTON, J. et al. Machine learning-based automated phenotyping of inflammatory nocifensive behavior in mice. *Molecular Pain*, SAGE Publications, v. 16, p. 1–16, 2020. Disponível em: <<https://journals.sagepub.com/doi/full/10.1177/1744806920958596>>.

***RI-PPGCO/UFU, art. 18, parágrafo único:**

O projeto de Dissertação de Mestrado deverá ser apresentado pelo estudante até o final do segundo semestre letivo, contado a partir da matrícula de ingresso como aluno regular.

Uberlândia, ____ de _____ de 20__.

O projeto deve ser encaminhado para a coordenação pelo orientador, conforme instruções disponíveis no link:
<http://www.ppgco.facom.ufu.br/servicos/projeto-de-dissertacao>