



UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS AGRÁRIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA TROPICAL

**PREDIÇÃO DO PESO DE GALINHAS CANELA-PRETA COM O USO
DE VISÃO COMPUTACIONAL**

ANA CAROLINE DE ALMEIDA ALVES

Teresina

2023

ANA CAROLINE DE ALMEIDA ALVES

**PREDIÇÃO DO PESO DE GALINHAS CANELA-PRETA COM O USO DE VISÃO
COMPUTACIONAL**

Dissertação de Mestrado apresentada ao programa de Pós-Graduação em Zootecnia Tropical da Universidade Federal do Piauí como requisito necessário para a obtenção do título de Mestre em Zootecnia Tropical.

Orientador: Dr Natanael Pereira da Silva Santos

Coorientador: Dr Romuere Rodrigues Veloso e Silva

Teresina

2023

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Setorial CCA
Serviço de Representação Temática da Informação

A474p Alves, Ana Caroline de Almeida.
Predição do peso de galinhas canela-preta com o uso de visão computacional / Ana Caroline de Almeida Alves. -- 2023.
52 f.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências Agrárias, Programa de Pós-Graduação em Zootecnia Tropical, 2023.
“Orientador: Prof. Dr. Natanael Pereira da Silva Santos.”

1. Galinha caipira. 2. Galinha crioula. 3. Aprendizado de máquina. I. Santos, Natanel Pereira da Silva. II. Título.

CDD 636.5

Bibliotecário: Rafael Gomes de Sousa - CRB3/1163

**PREDIÇÃO DO PESO DE GALINHAS CANELA-PRETA COM O USO DE
VISÃO COMPUTACIONAL**

ANA CAROLINE DE ALMEIDA ALVES

Dissertação aprovada em: 10/04/2023

Banca Examinadora:



Prof. Dr. Natanael Pereira da Silva Santos (Presidente) / DZO /UFPI



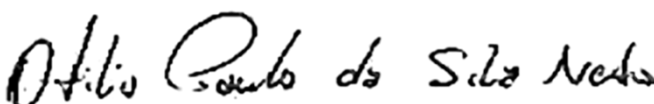
Prof. Dr. José Lindenberg Rocha Sarmento (Interno) / DZO/CCA/UFPI



Prof. Dr. Romuere Rodrigues Veloso e Silva (Interno) / CSHNB/UFPI



Prof. Dr. Flávio Henrique Duarte de Araújo (Interno) / CSHNB/UFPI



Prof. Dr. Otilio Paulo da Silva Neto (Externo) / IFPI

Se eu vi mais longe, foi por estar sobre ombros de gigantes.

Isaac Newton

AGRADECIMENTOS

Agradeço primeiramente a Deus pelo dom da vida e por permitir a realização deste trabalho. A Ele, toda a gratidão. Aos meus pais Adauto Alves Feitosa Filho e Aldenisa de Moura Feitosa e a toda minha família, por ser a base de tudo, por estarem sempre comigo, incentivando e apoiando independente da situação. Agradeço às minhas irmãs pelo apoio e incentivo de sempre.

Agradeço ao meu esposo Mateus Pinto Garcia pelo incentivo e pelo apoio, pela parceria de sempre e durante toda a trajetória deste trabalho.

Gratidão ao Criatório familiar Débora Carvalho e família por ter nos concedido a oportunidade e disponibilidade de realizar a coleta em seu criatório familiar caipira e pela grande ajuda durante a coleta de dados.

Aos meus orientadores Natanael Pereira da Silva Santos e ao meu coorientador Romuere Rodrigues Veloso e Silva por me guiarem, me orientarem e conceder a oportunidade desta pesquisa, pela paciência, pela confiança, ensinamentos transmitidos e pela oportunidade de crescimento acadêmico através dos seus ensinamentos.

Agradeço a todos meus colegas de turma, em especial meu amigo José dos Santos pela parceria, por compartilhar conhecimentos comigo e pela ajuda em todos os momentos necessários.

A todos os professores do Programa de Pós Graduação em Zootecnia Tropical, meus sinceros agradecimentos pelas disciplinas ministradas, pela disponibilidade, minha gratidão por acreditarem e contribuir para a realização deste objetivo.

LISTA DE ILUSTRAÇÕES

REFERENCIAL TEÓRICO

Figura 1 - Galinha da raça Canela-Preta.....	15
Figura 2 - Etapas do processo de visão computacional.....	17

CAPÍTULO I

Figura 1 - Steps of the proposed methodology	49
Figura 2 - Steps of the collection of weight and images of Canela-Preta chickens.....	49

LISTA DE TABELAS

Table 1 - Features based on image regions	50
Table 2 - Results of the K-fold cross-validation	51
Table 3 - Comparison of the approaches used for weight prediction	52

LISTA DE ABREVIATURAS E SIGLAS

ALEPI	Assembleia Legislativa do Piauí
MAE	Erro médio absoluto
MSE	Erro médio quadrático
RMSE	Raiz do erro quadrático médio
MAPE	Erro Percentual Médio Absoluto
R^2	R-Quadrado (Coeficiente de determinação)
ROI	Region of Interest
UFPI	Universidade Federal do Piauí

RESUMO

ALVES, Ana Caroline de Almeida. **Predição do peso de galinhas Canela-Preta com o uso de visão computacional**. 2023. 52p. Dissertação (Mestrado em Zootecnia Tropical) – Universidade Federal do Piauí, Teresina, 2023.

Um fator crucial para produtores comerciais de aves é o acompanhamento do peso corporal dos animais, pois isso possibilita obter o peso ideal para a comercialização. Entretanto, o monitoramento do crescimento animal através de balanças tradicionais pode ser estressante, dispendioso e trabalhoso. Nesse sentido, métodos baseados em visão computacional podem ser úteis para realizar essa tarefa de maneira menos dispendiosa e com maior bem-estar animal. Com o objetivo de desenvolver um algoritmo para prever o peso de galinhas Canela-Preta, foi realizada uma avaliação da capacidade preditiva de cada abordagem, obtendo-se um R^2 de até 45%. Desse modo, este trabalho apresenta uma solução econômica para o acompanhamento do peso de aves comerciais, evitando o estresse causado pelas balanças tradicionais e utilizando métodos de visão computacional para prever o peso das aves.

Palavras-chave: Galinha caipira, Galinha crioula, Aprendizado de máquina, Crescimento.

ALVES, Ana Caroline de Almeida. **Prediction of chicken weight using computer vision**. 2022. 52p. Dissertation (Master in Tropical Animal Science) – Federal University of Piauí, Teresina, 2023.

ABSTRACT

Monitoring broiler bird weight is critical for determining the ideal market weight. However, traditional monitoring methods, such as the use of scales, can be costly, labor-intensive and stressful for animals. Therefore, the application of methods based on computer vision can offer a more efficient and less stressful alternative for monitoring animal growth. The use of computer vision tools allows for the extraction of important information from the images of birds, such as their body weight, without the need for physical contact with the animals. In this study, we aimed to develop an algorithm for predicting the weight of Canela-Preta chickens using methods based on computer vision. The predictive ability of each method was evaluated, and R^2 values of up to 45% were obtained. Our findings showed that computer vision represents a less expensive and more humane solution for monitoring the weight of commercial birds.

Keywords: Free-range chicken, Creole chicken, Machine learning, Image processing, Growth

SUMÁRIO

1 INTRODUÇÃO	12
1.1 Objetivos	13
2 REFERENCIAL TEÓRICO	14
2.1 Galinha da Raça Canela-Preta	14
2.2 Visão Computacional	16
2.3 Segmentação de imagens	18
2.4 Extração de características	19
2.6 Regressão Linear	20
2.7 Validação	21
2.7.1 R^2	21
2.7.2 Erro médio absoluto	22
2.7.3 Erro Percentual Médio Absoluto	23
2.8 Trabalhos Relacionados	24
3. REFERÊNCIAS	27
Capítulo I	33
Introduction	36
Material and methods	37
Ethical certification and experimental characterization	37
Data collection	38
Segmentation	39
Feature extraction	39
Feature selection	40
Prediction tests	41
Results and discussion	42
Conclusion	44
Acknowledgements	45
Funding	45
Author contributions	45
Availability of data and materials	46
Declarations	46
Ethical approval and consent to participate	46
Human and animal rights	46
Competing interests	46
References	46

1 INTRODUÇÃO

A mensuração de tamanho, volume e peso são etapas essenciais na classificação da maioria dos produtos alimentares, agrícolas e pecuária (NYALALA et al., 2021). O acompanhamento do peso é um fator importante para o produtor de aves e outras espécies, pois, por meio deste acompanhamento é possível obter melhores resultados no manejo e na obtenção do peso ideal do animal para o comércio (COMINOTTE., et al, 2020). Monitorar e registrar o peso corporal dos animais permite uma intervenção oportuna e eficaz em dietas e saúde, maior eficiência na seleção genética e na identificação dos momentos ideais para comercializar os animais (WANG et al., 2021).

Existem métodos para determinar a área corporal das aves, como o escalpelamento e o uso de moldes à base de gases que aderem à superfície do animal, além da pesagem através de balanças. No entanto, esses métodos causam estresse e desgaste ao animal, além de exigir tempo e recursos humanos (ABREU et al., 2012). Isso ocorre porque a pesagem manual de animais é demorada, trabalhosa e cara, pois requer que o animal esteja imóvel. Além disso, é um método arriscado tanto para humanos quanto para animais, podendo gerar lesões e traumas (DOHMEN; CATAL; LIU, 2021). Portanto, é importante desenvolver métodos mais seguros para determinar o peso corporal das aves, a fim de minimizar o estresse e o desconforto para os animais e evitar riscos para os humanos envolvidos no processo.

Pesquisas conduzidas em condições experimentais e em fazendas comerciais mostraram que o manejo dos animais de forma estressante e que estimule o medo no animal, pode afetar a fisiologia do estresse e a produtividade, e a qualidade da carne (HEMSWORTH et al., 2011). Dessa forma, muitos estudos têm sido feitos para reduzir o estresse e, assim, melhorar o bem-estar animal durante o período de criação. A ciência do bem-estar animal fez avanços significativos na compreensão dos sentimentos dos animais na produção e na proposição de uma variedade de novas práticas para criar um equilíbrio justo entre as necessidades dos animais, os custos de produção e as exigências dos consumidores (FAUCITANO. et al., 2022).

Uma ferramenta que possa prever o peso corporal de animais seria de grande utilidade para os produtores, incluindo os produtores primários. Com essa ferramenta, seria possível evitar o desgaste de pesar as aves manualmente e obter uma estimativa

precisa do peso dos animais. Além disso, a predição automática do peso corporal pode ser realizada de forma mais rápida e eficiente, permitindo que o produtor tome decisões mais precisas sobre a nutrição e o manejo dos animais. Essa ferramenta pode ser uma vantagem competitiva para os produtores do setor primário, contribuindo para a melhoria da qualidade e produtividade da produção animal, e consequentemente, para o aumento da rentabilidade. Além disso, a ferramenta de predição de peso corporal também pode ajudar na identificação precoce de animais com problemas de saúde ou nutrição, permitindo que o produtor intervenha antes que a situação se agrave. Isso pode resultar em menor mortalidade animal e redução de perdas financeiras.

Pesquisadores têm mostrado que a análise digital de imagens pode ser utilizada para determinar medidas e peso corporal de animais. Na literatura foram encontrados diversos trabalhos com animais voltados ao processamento de imagens com medição, contagem e predição de peso (SANT'ANA et al., 2021). Assim, uma ferramenta capaz de prever o peso corporal de animais seria de grande auxílio ao produtor, incluindo o produtor primário, visto que não seria necessário o desgaste de pesar as aves manualmente.

Desse modo, os métodos de visão computacional podem ser utilizados na determinação do peso de animais de forma não-invasiva. Como há uma correlação importante entre as medidas corporais e o peso corporal, de modo que a massa corporal pode ser estimada por meio de índices biométricos obtidos com o uso de imagens, é possível estimar a área e o peso do animal através de um método que reduza o estresse animal. Desse modo, a criação de um método que possibilita o acompanhamento tornar o acompanhamento do peso corporal de aves mais simples e com menor estresse para a ave.

1.1 Objetivos

O objetivo geral desta ferramenta é monitorar e prever o peso de galinhas. Os objetivos específicos deste projeto são:

1. Construir um conjunto de dados de imagens das galinhas de raça Canela-Preta;

2. Desenvolver um sistema de visão computacional para realizar a predição do peso das aves;
3. Avaliar o desempenho dos modelos de predição desenvolvidos para validar os resultados preditivos obtidos.

Esta dissertação está estruturada conforme normas do Manual de Normalização de Monografia, Dissertação e Tese da Universidade Federal do Piauí. Dessa forma, o Capítulo I segue normas de revista e o Capítulo seguinte segue as normas de publicação da *BMC Bioinformatics* (ISSN 1471-2105), o qual será submetido à análise para posterior publicação. O estudo foi segmentado a fim de melhorar a acomodação dos diferentes tópicos. Inicialmente, tem-se a revisão de literatura que engloba os temas abordados no Capítulo I, o qual servirá de referencial teórico, bem como as respectivas referências bibliográficas utilizadas. A finalização desta pesquisa é alcançada com as Considerações Finais e Perspectivas Futuras.

2 REFERENCIAL TEÓRICO

Esta seção aborda tópicos relevantes ao contexto deste trabalho: galinha da Raça Canela-Preta, Visão Computacional, Segmentação de imagens, extração de características, descritores de forma, Regressão Linear, Redes Neurais e Validação.

2.1 Galinha da Raça Canela-Preta

A criação de aves caipiras tem grande aceitação entre os pequenos produtores, sobretudo, devido à resistência das aves, menores índices de mortalidade e boa produtividade (BENTO; GALVÃO; SOUZA, 2009). Ao longo do tempo, as galinhas caipiras passaram por um processo de adaptação e seleção natural, dessa maneira, surgiram as raças de galinhas crioulas brasileiras (CARVALHO et al., 2018).

Dentre essas raças, existem as galinhas caipiras da raça Canela-Preta, que são descendentes de galinhas introduzidas no Brasil pelos portugueses ao longo da colonização. Essas aves foram criadas soltas em quintais e fazendas, ocasionando em cruzamentos aleatórios (CARVALHO et al., 2018).

Em 2021, o Governo do Estado do Piauí sancionou a Lei nº 7615/21, declarando a galinha caipira da raça Canela-Preta como patrimônio histórico, cultural

e genético do estado. Essa raça é considerada genuinamente nativa, agregando valor econômico e social ao Piauí, uma vez que é uma importante fonte de proteína e renda para a população local. A galinha da raça Canela-Preta (Figura 1) é uma raça nativa brasileira, geneticamente estruturada e encontrada principalmente no Nordeste do país, especialmente no Piauí, conforme afirmam Carvalho et al. (2016). Essa raça é conhecida por sua rusticidade e adaptação às condições climáticas e ambientais da região, além de apresentar um sabor característico que a torna valorizada pelos consumidores locais e de outras regiões do país. A valorização e preservação da galinha da raça Canela-Preta é importante não apenas do ponto de vista econômico e cultural, mas também do ponto de vista genético, uma vez que essa raça é uma importante fonte de diversidade genética para o melhoramento das aves comerciais. Portanto, é fundamental a adoção de medidas que visem a proteção e a conservação dessa raça tão importante para o Piauí e para o todo o país.

Figura 1 - Galinha da raça Canela-Preta



Fonte: AGERP

As galinhas da raça Canela-Preta possuem tarso e falanges de coloração preta, corpo predominantemente preto com pigmentações nas penas do pescoço nas cores branca, dourada ou vermelha e por possuírem carne de coloração mais escura, o que as caracterizam por possuírem plumagens, ovos e sabor diferenciados em relação às outras raças crioulas (CARVALHO et al., 2016). As galinhas Canela-Preta possuem dimorfismo sexual, ou seja, os machos são diferentes das fêmeas. As fêmeas possuem coloração de plumagem preta com dourada, prateada e toda preta. Machos possuem coloração nas extremidades preta e região dorsal prateada branca, prateada amarelada ou avermelhada. Ambos sexos possuem falanges predominantemente preta.

Elas costumam ser criadas por agricultores familiares com pouco ou nenhum investimento tecnológico (CARVALHO et al., 2020). A partir de 1980, houve uma crescente valorização de produtos naturais, desse modo, essas galinhas crioulas passaram a ser mais valorizadas, devido a rusticidade, além de serem apropriadas para a criação em sistemas tradicionais (CARVALHO et al., 2016).

É possível diferenciar três sistemas de exploração para criação de aves: extensivo, semi-intensivo e intensivo (BENTO; JÚNIOR; SOUZA, 2009). As galinhas da raça Canela-Preta são criadas predominantemente da forma extensiva, sendo manejadas soltas e são alimentadas em regime de pastejo (CARVALHO et al., 2017). Essas aves adquirem peso médio para comercialização a partir dos seis meses, se criadas no sistema semi-intensivo e extensivo após o oitavo mês (CARVALHO et al., 2017).

2.2 Visão Computacional

Diversas técnicas têm sido propostas e validadas para predição da massa corporal animal com base em visão 2D e 3D na literatura (DOHMEN; CATAL; LIU, 2021). Isso ocorre porque é de grande interesse para o produtor ter como alternativa um método que permita determinar medidas e peso corporal de animais de forma menos invasiva, que diminua o estresse animal, bem como, diminui a necessidade de recursos humanos. Contudo, não existem estudos neste viés visando a obtenção do peso de galinhas caipiras da raça Canela-Preta. Uma vez que este método é baseado em visão computacional, ele garante que as aves sejam pesadas regularmente e não causem estresse ou a necessidade de trabalho humano considerável (NYALALA et al., 2021).

A visão computacional é uma área da ciência da computação cujo objetivo é reproduzir a capacidade humana de ver, analisar e compreender imagens e, para isso, utiliza técnicas computacionais (WEBER et al., 2020). Esta área de estudo usa de tecnologias de obtenção de informações sobre o ambiente físico a partir de imagens. Desse modo, suas origens estão na área do que se denomina inteligência artificial, que visa imitar o comportamento humano inteligente (FÖRSTNER; WROBEL, 2016). Tem-se buscado desenvolver métodos que sejam capazes de replicar uma das capacidades mais incríveis do sistema visual humano, ou seja, inferir características

do mundo real puramente usando a luz refletida para os olhos de vários objetos (KHAN et al., 2018).

O processo de visão computacional normalmente contém as seguintes etapas: Se inicia pela aquisição de imagens e dados e, em seguida, etapas de pré-processamento e reconhecimento de padrões para executar determinada tarefa (SAMPATH et al., 2021). A metodologia usando visão computacional está representada conforme visto na Figura 2.

Figura 2 - Etapas do processo de visão computacional



Fonte: Do autor

Os sistemas de visão computacional têm se mostrado uma ferramenta poderosa para a medição do peso de animais vivos sem causar estresse animal. Medidas automáticas de animais podem ser uma ferramenta extremamente útil para monitorar o crescimento individual de cada animal (FERNANDES et al., 2020). Dessa forma, com esses avanços, é possível avaliar o desempenho de crescimento do gado individualmente com mais precisão, permitindo a possibilidade de medições múltiplas de características de interesse com o mínimo de estresse animal (FERNANDES et al., 2019). O uso de imagens digitais para obter estimativa do peso corporal de aves é possível através da visão computacional, visto que, esses métodos possuem a vantagem de não serem invasivos (ABREU et al., 2012). Há vários benefícios ao se aplicar a visão computacional neste tipo de problema, tais como precisão e confiabilidade na classificação e velocidade de classificação em comparação com os processos manuais (NYALALA et al., 2021).

A aplicação de técnicas usando a visão computacional tem se tornado cada vez mais importantes e eficazes devido às suas aplicações abrangentes em áreas tão diversas como vigilância e monitoramento, saúde e medicina, esportes e recreação, robótica, drones e carros autônomos (KHAN et al., 2018). O uso da Visão Computacional no manejo de animais, pode ser uma alternativa para avaliar curvas de crescimento em tempo real com menor impacto negativo na saúde do animal e garantindo o bem estar animal (FERNANDES et al., 2019). O sistema automatizado

de visão computadorizada possui o potencial de prever o peso corporal de animais em rebanho (COMINOTTE et al., 2020). Contudo, é possível aprimorar os métodos de avaliação de imagens visando reduzir o erro de predição (ABREU et al., 2012).

2.3 Segmentação de imagens

A segmentação de imagens é considerada uma tarefa importante em visão computacional e processamento de imagens (MINAEE et al., 2021). Esta etapa desempenha papel fundamental em ferramentas que lidam com processamento de imagem e visão computacional, porque o resultado da segmentação afetará diretamente o desempenho das etapas de processamento subsequentes (ZHOU et al., 2016).

A segmentação de imagem é a divisão de uma imagem específica em um número de segmentos homogêneos, portanto, a representação de uma imagem em formas simples e fácil aumenta a eficácia do reconhecimento de padrões (ABDULATEEF; SALMAN 2021). O nível a que esta subdivisão é realizada depende do problema a ser resolvido, ou seja, a segmentação deve parar quando os objetos de interesse em uma aplicação estiverem isolados (ZAITOUN; AQEL, 2015).

Este processo é usado para distinguir diferentes objetos na imagem, realizado por meio de técnicas desenvolvidas pelos pesquisadores, a fim de tornar as imagens fáceis de avaliar, pois, tais técnicas trabalham dividindo a imagem digital em várias regiões para analisá-las (KHAN, 2014). Os objetos nos conjuntos de dados reais ocupam apenas uma pequena parte da imagem, enquanto o restante da imagem é o fundo (SAMPATH et al., 2021). No caso da aplicação da segmentação no corpo de animais, as bordas são calculadas a fim de identificar a possível região do animal.

2.4 Extração de características

Uma imagem possui propriedades que podem ser medidas sendo chamadas de características do objeto. Tais características são normalmente agrupados em uma matriz escalar chamada "vetor de características" ou "descriptor de recursos" (SILVA et al., 2013). Para realizar a extração de características de uma imagem, é comum aplicar um descriptor que possa armazenar informações importantes que representam a imagem em um vetor. Esse processo tem como objetivo encontrar características

relevantes da imagem que possam ser usadas em tarefas como classificação, reconhecimento ou detecção de objetos.

As novas técnicas baseadas em dados usam algoritmos de aprendizado de máquina para construir descritores de forma precisa e confiável (ROSTAMI et al., 2019).

As características podem incluir bordas, contornos, texturas, formas, padrões, cores e outras propriedades que possam ser extraídas das imagens. Essas características são selecionadas com base em sua capacidade de discriminar objetos ou padrões de interesse em relação ao fundo, ou outras classes de objetos. Para aplicar tal abordagem, é necessário frisar que uma das características básicas de um objeto é sua forma (LÓPEZ et al., 2017).

2.5 Descritores de forma

A característica de forma se refere a uma descrição informativa que fornece a um objeto uma identificação como membro de alguma categoria. A característica de forma consiste na área da superfície do objeto com interseção em cada caixa (ZHONG, 2009). De modo geral, as características de forma consistem em uma representação simplificada de uma forma 2D ou 3D, no qual é convertido em um vetor contendo um conjunto de valores numéricos ou uma estrutura semelhante a um gráfico, onde, será usado para descrever a forma da imagem geometricamente ou topologicamente possibilitando o uso desses valores para as etapas subsequentes, tais como a predição de determinados valores (KAZMI; YOU; ZHANG, 2013). Este mecanismo permite realizar processos de classificação, recuperação e comparação para correspondência de objetos (LÓPEZ et al., 2017).

Os descritores baseados em contorno expressam as propriedades das formas pelo seu contorno, enquanto que os descritores baseados em região expressam distribuições de pixels como um objeto 2D, podendo descrever objetos complexos que consistem de múltiplas regiões desconexas, bem como, objetos simples (ERPEN, 2004). A escolha de um descritor não é uma tarefa trivial, para isto, é necessária uma série de pesquisas e testes para a avaliação do desempenho de cada abordagem, a fim de descobrir qual a melhor abordagem para determinados tipos de problema (ERPEN, 2004).

2.6 Regressão Linear

A regressão linear desempenha um papel fundamental na modelagem estatística (SU; YAN; TSAI, 2012). A regressão consiste em uma tarefa de aprendizado de máquina para prever valores reais, descrevendo a relação estatística entre uma ou mais variáveis preditoras e a variável resposta. De forma geral, a análise de regressão linear é uma forma de modelar uma saída contínua, por meio de uma variável chamada 'Y', no qual será o valor previsto como uma função linear de um ou vários preditores 'X' que consistem em valores reais (ANTTALAINEN et al., 2021).

A fórmula da regressão linear simples é dada pela Equação 1:

$$y = a + bx$$

(1)

Em que:

y é a variável dependente (a que se deseja prever);

x é a variável independente (a que se utiliza para prever);

a é o intercepto da reta (o valor de y quando x é igual a zero);

b é o coeficiente angular (a inclinação da reta).

Uma vez que os valores de a e b são estimados, é possível utilizar a equação da reta de regressão para fazer previsões de y para novos valores de x.

Existem vários métodos de estimativa disponíveis para modelos lineares, incluindo mínimos quadrados, máxima verossimilhança, abordagem bayesiana, estimativa robusta, dentre outros (SU; YAN; TSAI, 2012). Primeiramente, as análises de regressão são geralmente usadas para previsão. Em segundo lugar, a análise de regressão pode ser usada em alguns casos para determinar relações causais entre as variáveis independentes e dependentes. A regressão mostra apenas uma relação entre as variáveis e não estabelece causalidade. Isso significa que, embora uma relação possa ser observada entre as variáveis, isso não significa necessariamente que uma causa a outra.(MAULUD; ABDULAZEEZ, 2020).

2.7 Validação

Uma etapa essencial em qualquer modelo de aprendizado de máquina é avaliar a acurácia do modelo. Para realizar a avaliação é necessário estudar e aplicar algumas métricas descritas na literatura. Neste estudo foram exploradas as seguintes métricas: R^2 , MAE, MSE, RMSE.

2.7.1 R^2

O R^2 corresponde à precisão obtida e também é conhecido por coeficiente de determinação. Esta métrica visa expressar a quantidade da variância dos dados. Essa medida calcula qual a porcentagem da variância que pôde ser prevista pelo modelo de regressão e, assim, é possível avaliar o quanto as medidas reais estão próximas do nosso modelo. O valor do R^2 varia de 0 a 1 e geralmente é representado em porcentagem. A interpretação dos resultados do R^2 funcionam da seguinte maneira:

$R^2 \geq 0$: Com regressão linear sem restrições, é não negativo e corresponde ao quadrado do coeficiente de correlação múltipla.

$R^2 = 0$: A linha ajustada é horizontal. Com duas variáveis numéricas este é o caso se as variáveis forem independentes, ou seja, não correlacionadas (CHICCO; WARRENS; JURMAN, 2021).

Na Equação 2, a variância Residual é dividida pela variância total e subtraída por 1. O numerador $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ é a soma dos quadrados total (variância Residual) e o denominador $\sum_{i=1}^n (y_i - \underline{y})^2$ é a soma dos quadrados da regressão que resulta na variância total. Conforme visto na Equação (2).

$$R^2 = 1 - \frac{\text{Variação Residual}}{\text{Variança Total}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \underline{y})^2}$$

(2)

Em que:

y_i representa o valor real;

\hat{y} representa o valor predito;

\underline{y} representa o valor médio das amostras.

Em resumo, essa fórmula mede o grau de ajuste do modelo de regressão aos dados observados. Quanto mais próximo de 1 for o valor do R^2 , melhor será o ajuste do modelo aos dados, indicando que a maioria da variabilidade dos dados foi explicada pelo modelo. Por outro lado, se o valor do R^2 for próximo de 0, isso indica que o modelo não explica bem a variabilidade dos dados e não é um bom ajuste para os dados observados.

2.7.2 Erro médio absoluto

O Erro médio absoluto (MAE) corresponde à média das distâncias entre valores preditos e os valores reais. Seu cálculo corresponde ao somatório de cada um dos dados subtraídos pela sua referida média, resultando na diferença absoluta. O cálculo do MAE é relativamente simples, onde envolve somar as magnitudes (valores absolutos) dos erros para obter o 'erro total' e, em seguida, divide o erro total pelo número de amostras (WILLMOTT; MATSUURA, 2005). Na fórmula do MAE, o n é o número de amostras, y_i é o valor previsto subtraído por \hat{y}_i que corresponde ao valor real. Com esta fórmula é possível obter a média dos resíduos no conjunto de dados. O resultado desta subtração retorna o valor absoluto residual. Podemos ver a fórmula dessa métrica na Equação (3).

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (3)$$

Em que:

n é o número de amostras;

y_j é o valor previsto subtraído por \hat{y}_j que corresponde ao valor real.

O resultado desta subtração retorna o valor absoluto residual (SARMENTO et al., 2006).

2.7.3 Erro Percentual Médio Absoluto

O erro percentual médio absoluto (MAPE) é uma medida de erro comumente utilizada para avaliar a precisão de modelos de previsão. O MAPE é calculado como

a média do valor absoluto das diferenças percentuais entre os valores previstos e os valores reais. Ele é expresso como uma porcentagem e indica o quão longe, em média, as previsões estão dos valores reais. Em aplicações reais, o MAPE é frequentemente usado quando se sabe que a quantidade a prever permanece muito acima de zero. Foi usado, por exemplo, como medida de qualidade em um concurso de previsão de consumo de eletricidade (MYTTENAERE et al., 2016). Esta métrica é calculada como a média do valor absoluto das diferenças percentuais entre os valores previstos e os valores reais. Ele é expresso como uma porcentagem e indica o quão longe, em média, as previsões estão dos valores reais, sendo dada pela Equação (4):

$$\text{MAPE} = \left(\frac{100\%}{n} \right) * \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (4)$$

Em que:

n é o número de amostras;

y_i é o valor real;

\hat{y}_i é o valor previsto

2.8 Trabalhos Relacionados

No trabalho de Fernandes et al. (2019) foi desenvolvido um sistema de visão computacional para extração das características corporais relacionadas à forma e posterior predição do peso corporal em porcos a partir de imagens. As imagens dos suínos que tiveram dados coletados foram usadas para o desenvolvimento dos modelos preditivos. Foi aplicada a validação cruzada para os modelos desenvolvidos com dados de suínos, onde, apresentou um R^2 variando de 0,86 a 0,94. Das medidas corporais extraídas da imagem, volume corporal, área e comprimento foram os mais úteis para a predição do peso corporal. Houve a inclusão das demais medidas corporais como: largura e altura ou descritores de forma ao modelo, o que promoveu uma melhora significativa das previsões.

Na pesquisa de Adamczak et al. (2018) o objetivo do estudo foi explorar a possibilidade de aplicação de digitalização 3D para determinar o peso de todo o peito e o peso de peitoral maior e menor em carcaças de frango. O material para o estudo

foram carcaças de frango significativamente diferenciadas em termos de peso. As imagens 3D foram obtidas por meio de um conjunto composto por um scanner e uma mesa giratória, sobre a qual foi colocado um tripé com estribo industrial padrão e penduradas as carcaças de frango. Após a análise das carcaças, foi determinado o peso do peito inteiro e seus músculos, que foram visto por meio de coeficientes de correlação e equações de regressão

O objetivo do estudo de Dohmen, Catal, Liu (2021), foi projetar um novo modelo de predição de massa usando algoritmos *deep learning* para rebanhos em fazendas leiteiras. O algoritmo de segmentação *Mask-RCNN* foi utilizado para segmentar as imagens das novilhas e remover o fundo. Um modelo CNN foi desenvolvido para prever a massa corporal de novilhas. Para o estudo de caso, foi construído um novo conjunto de dados baseado em imagens de 63 novilhas. O modelo de vista lateral teve um R^2 de 0,91 e um RMSE de 27 kg, o modelo de vista superior teve um R^2 de 0,96 e um RMSE de 20 kg.

O objetivo do estudo de Weber et al. (2020) foi extrair automaticamente medidas de imagens 2D da área dorsal de bovinos da raça Nelore para estimar o peso desses bovinos por meio de algoritmos de regressão. Para tanto, foram selecionadas as distâncias euclidianas dos pontos gerados pelo *Active Contour Model*, juntamente com as características obtidas no dorsal *Convex Hull*. Estes foram submetidos aos algoritmos *Bagging*, *Regression by Discretization* e *Random Forest* para análise das métricas de erro predito. Assim, o algoritmo de Bagging apresentou os melhores resultados.

Na pesquisa de Gümüs et al. (2021) foram avaliadas a estimativa de peso e cores por análise de imagem de duas espécies de bagres cultivados (peixe europeu *Silurus glanis* e bagre africano *Clarias gariepinus*). Um total de 60 peixes inteiros de cada espécie foram primeiramente pesados e, em seguida, suas fotos tiradas por uma câmera digital em uma caixa de luz. Comprimento, peso e cor de todos os peixes foram determinados por análise de imagem e os resultados foram comparados entre os dois peixes. A análise de imagens pode determinar o comprimento, a área de visualização e a cor dos peixes e ajudar no desenvolvimento das relações peso/ altura.

O objetivo do estudo de Yu; Lee; Morota, (2021) foi estabelecer um sistema de visão computacional por meio da coleta de vídeos para capturar imagens de visão de

cima e de profundidade de porcos em crescimento sem restrição para prever seu peso corporal ao longo do tempo. Os porcos foram pesados manualmente usando uma balança eletrônica. Um pipeline de processamento de imagens foi desenvolvido para processar as imagens. Cada porco dentro do quadro de imagens foi segmentado por um algoritmo de limiar, e o contorno do porco foi identificado para extrair seu comprimento e largura. A altura do porco foi estimada a partir das imagens de profundidade capturadas pelo sensor infravermelho de profundidade. O controle de qualidade incluiu a remoção de porcos que estavam tocando a cerca e sentados, bem como aqueles que apresentavam formas extremamente distorcidas ou borrões de movimento devido ao seu movimento frequente. O sistema pôde fornecer uma abordagem eficaz para prever o peso corporal vivo de porcos a partir de imagens.

No estudo de Cominotte et al. (2020) os objetivos foram avaliar a qualidade preditiva de um sistema automatizado de visão computacional usado para prever o peso corporal e o ganho médio diário em bovinos de corte; e comparar diferentes abordagens preditivas. Para isso, foram coletadas 234 imagens de bovinos de corte durante as fases de desmame. Primeiramente, foram realizadas medidas biométricas do corpo de cada animal, como volume corporal, área, comprimento e outros, por meio de imagens tridimensionais captadas com o sensor Kinect, e seus respectivos pesos corporais foram adquiridos por meio de balança eletrônica. Em seguida, as medidas biométricas foram utilizadas como variáveis nas quatro abordagens preditivas. Para avaliar a qualidade da previsão, uma validação cruzada *leave-oneout* foi adotada. Foi concluído que a CNN foi a melhor abordagem de predição.

No estudo de Zhang (2021) o objetivo foi desenvolver em uma CNN de regressão múltipla de saída para estimar o peso e o tamanho corporal dos suínos. Para treinar e avaliar o peso do porco e o modelo de estimativa do tamanho corporal, foram necessários dados de peso dos suínos, dados do tamanho do corpo e imagens 3D dos lombos dos suínos. Os dados de modelagem foram usados para treinar os modelos e os dados de teste foram usados como dados desconhecidos para testar a capacidade de generalização dos modelos treinados. O valor do R^2 entre os resultados estimados e medidos estavam na faixa de 0,98 a 0,99. Combinado com a plataforma de desenvolvimento de software LabVIEW. Com este método foi possível estimar o peso do porco e o tamanho do corpo de forma precisa, rápida e automática.

O trabalho de Nyalala et al. (2021) apresenta uma previsão do peso da carcaça de frango com base em imagens de profundidade. Um *Active Shape Model* foi desenvolvido para segmentar a carcaça em quatro cortes. Cinco modelos de regressão foram desenvolvidos com base nas características da imagem para cada estimativa de peso (carcaça e seus cortes). O desempenho dos modelos de estimativa de peso foi avaliado com base em uma comparação do peso medido com o peso estimado. O algoritmo de segmentação das partes da carcaça foi avaliado com base no número de objetos após a segmentação (contagem) em relação ao número de cortes. O modelo Bayesian-ANN superou todos os outros modelos de regressão em R^2 de 0,99 e R^2 0,98 na estimativa do peso da carcaça inteira e da cabeça e pescoço. Esta técnica proposta pode ser aplicada como um método não intrusivo e preciso.

No estudo de Jia et al. (2021) é apresentado um método de visão computacional que identifica automaticamente nove partes do corpo da vaca, ou seja, cabeça, tronco, úbere, barriga (ou traseira), perna dianteira esquerda, perna dianteira direita, perna esquerda, posterior direito posterior e cauda. Usamos um método para detecção de partes do corpo usando uma versão aprimorada de correspondência de modelos. Todo o processo de imagem inclui gravação de vídeo, restauração e pré-processamento de imagem, extração de esqueleto, correspondência de modelos, reconhecimento de perna direita ou esquerda, calibração e identificação de partes do corpo. 1421 imagens de vista lateral de 113 vacas e 859 imagens de vista traseira de 75 vacas foram coletadas usando uma câmera de profundidade RGB instalada no celeiro para vacas holandesas em lactação. A precisão média de segmentação de partes do corpo para visão lateral e visão traseira foi de 96% e 91%, respectivamente. Os resultados indicam que é possível detectar e extrair automaticamente partes do corpo de imagens de profundidade RGB sem qualquer interferência manual.

Na literatura não foram encontrados relatos de trabalhos relacionados à predição do peso de galinhas caipiras da raça Canela-Preta. Foram encontrados na literatura trabalhos de predição de peso apenas com galinhas de granja.

3. REFERÊNCIAS

- ADAMCZAK, L. et al. **The use of 3d scanning to determine the weight of the Chicken breast.** Computers and Electronics in Agriculture, Elsevier, v. 155, p. 394–399, 2018.
- ABDULATEEF, Salwa Khalid; SALMAN, Mohanad Dawood. A Comprehensive Review of Image Segmentation Techniques. **Iraqi Journal for Electrical & Electronic Engineering**, v. 17, n. 2, 2021.
- ABREU, P. G. d. et al. **Medidas morfológicas em função do peso e da idade da ave, por meio de imagens.** Revista Brasileira de Engenharia Agrícola e Ambiental, SciELO Brasil, v. 16, n. 7, p. 795–801, 2012.
- ALEPI. **Galinha Canela-Preta é patrimônio histórico, cultural e genético do Piauí.** Parlamento Piauí, 2021. Disponível em: <https://www.parlamentopiaui.com.br/noticias/alepi/galinha-canela-preta-e-patrimonio-historico-cultural-e-genetico-do-piaui-193178.html>. Acesso em: 02 de agosto de 2022.
- ANTTALAINEN, A. et al. **Predicting lecithin concentration from differential mobility spectrometry measurements with linear regression models and neural networks.** Talanta, Elsevier, v. 225, p. 121926, 2021.
- BASHIRI, F. S. et al. **An application of manifold learning in global shape descriptors.** Algorithms, Multidisciplinary Digital Publishing Institute, v. 12, n. 8, p. 171, 2019.
- BENTO, E. F. GALVÃO, J. G. B. G. SOUZA, A. F. d. **Sistema alternativo de produção de aves.** Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte, 2009.
- CARVALHO, D. A et al. **Padrão racial fenotípico de galinhas brasileiras da raça Canela-Preta.** Archivos de zootecnia, Universidad de Córdoba, v. 66, n. 254, p. 195–202, 2017.
- CARVALHO, D. A et al. **Variabilidade genética de 12 loci de microssatélites em galinhas crioulas Canela-Preta.** Arquivo Brasileiro de Medicina Veterinária e Zootecnia, SciELO Brasil, v. 70, n. 4, p. 1275–1281, 2018.

CARVALHO, D. A. et al. **Caracterização genética e estrutura populacional de galinhas crioulas Canela-Preta.** Pesquisa Agropecuária Brasileira, SciELO Brasil, v. 51, n. 11, p.1899–1906, 2016.

CARVALHO, D. A. et al. **Caracterização genética e estrutura populacional de galinhas crioulas Canela-Preta.** Pesquisa Agropecuária Brasileira, SciELO Brasil, v. 51, n. 11, p.1899–1906, 2016.

CARVALHO, D. A. et al. **Diversity and genetic relationship of free-range chickens from the northeast region of Brazil.** Animals, Multidisciplinary Digital Publishing Institute, v. 10, n. 10, p. 1857, 2020.

CHICCO, D.; WARRENS, M. J.; JURMAN, G. **The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation.** PeerJ Computer Science, PeerJ Inc., v. 7, p. e623, 2021.

CLEOPHAS, T. J.; ZWINDERMAN, A. H. **Linear regression (regression coefficient, correlation coefficient and their standard errors).** In: Clinical Data Analysis on a Pocket Calculator. [S.l.]: Springer, 2016.

COMINOTTE, A. et al. **Automated computer vision system to predict body weight and average daily gain in beef cattle during growing and finishing phases.** Livestock Science, Elsevier, v. 232, p. 103904, 2020.

DOHMEN, R.; CATAL, C.; LIU, Q. **Image-based body mass prediction of heifers using deep neural networks.** Biosystems Engineering, Elsevier, v. 204, p. 283–293, 2021.

ERPEN, L. R. C. **Reconhecimento de padrões em imagens por descritores de forma.** 2004.

FAUCITANO, L. et al. **Fundamentals of animal welfare in meat animals and consumer attitudes to animal welfare.** In: New aspects of meat quality. Woodhead Publishing, 2022. p. 667-703.

FANG, Y. et al. **3d deep shape descriptor.** In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.: s.n.], 2015.

FERNANDES, A. F. et al. **A novel automated system to acquire biometric and morphological measurements and predict body weight of pigs via 3d computer vision.** Journal of animal science, Oxford University Press US, v. 97, n. 1, p. 496–508, 2019.

FÖRSTNER, W.; WROBEL, B. P. **Photogrammetric computer vision.** [S.l.]: Springer, 2016.

GÜMÜS, E. et al. **Evaluation of body weight and color of cultured european catfish (silurus glanis) and african catfish (clarias gariepinus) using image analysis.** Aquacultural Engineering, Elsevier, v. 93, p. 102147, 2021.

HEMSWORTH, P. H. et al. **Human–animal interactions at abattoirs: Relationships between handling and animal stress in sheep and cattle.** Applied animal behaviour science, Elsevier, v. 135, n. 1-2, p. 24–33, 2011.

HE, Kaiming et al. Mask r-cnn. In: **Proceedings of the IEEE international conference on computer vision.** 2017. p. 2961-2969.

JIA, N. et al. **Segmentation of body parts of cows in rgb-depth images based on template matching.** Computers and Electronics in Agriculture, Elsevier, v. 180, p. 105897, 2021.

KHAN, S. et al. **A guide to convolutional neural networks for computer vision.** Synthesis Lectures on Computer Vision, Morgan & Claypool Publishers, v. 8, n. 1, p. 1–207, 2018.

KAZMI, I. K.; YOU, L.; ZHANG, J. J. **A survey of 2d and 3d shape descriptors.** In: 2013 10th International Conference Computer Graphics, Imaging and Visualization. [S.l.: s.n.], 2013.

LÓPEZ, G. L. et al. **Comparative analysis of shape descriptors for 3d objects.** Multimedia Tools and Applications, Springer, v. 76, n. 5, p. 6993–7040, 2017.

MAULUD, D.; ABDULAZEEZ, A. M. **A review on linear regression comprehensive in machine learning.** Journal of Applied Science and Technology Trends, v. 1, n. 4, p. 140–147, 2020.

MINAEE, S. et al. **Image segmentation using deep learning: A survey.** *IEEE transactions on pattern analysis and machine intelligence*, IEEE, 2021.

MYTTENAERE, A. D. et al. **Mean absolute percentage error for regression models.** *Neurocomputing*, Elsevier, v. 192, p. 38–48, 2016.

ROSTAMI, R. et al. **A survey on data-driven 3d shape descriptors.** In: WILEY ONLINE LIBRARY. *Computer Graphics Forum*. [S.l.], 2019.

SARMENTO, J. L.R et al. **Estudo da curva de crescimento de ovinos Santa Inês.** *Revista Brasileira de Zootecnia*, v. 35, p. 435-442, 2006.

SAMPATH, V. et al. **A survey on generative adversarial networks for imbalance problems in computer vision tasks.** *Journal of big Data*, Springer, v. 8, n. 1, p. 1–59, 2021.

SILVA, R. et al. **Automatic motorcycle detection on public roads.** *CLEI Electronic Journal*, Centro Latinoamericano de Estudios en Informática, v. 16, n. 3, p. 4–4, 2013.

SU, X.; YAN, X.; TSAI, C.-L. **Linear regression.** *Wiley Interdisciplinary Reviews: Computational Statistics*, Wiley Online Library, v. 4, n. 3, p. 275–294, 2012.

TERLOUW, E. et al. **Pre-slaughter conditions, animal stress and welfare: current status and possible future research.** *animal*, Elsevier, v. 2, n. 10, p. 1501–1517, 2008.

NYALALA, I. et al. **Weight and volume estimation of poultry and products based on computer vision systems: a review.** *Poultry Science*, Elsevier, p. 101072, 2021.

NYALALA, I. et al. **On-line weight estimation of broiler carcass and cuts by a computer vision system.** *Poultry Science*, Elsevier, v. 100, n. 12, p. 101474, 2021.

SANT'ANA, D. A. et al. **Weighing live sheep using computer vision techniques and regression machine learning.** *Machine Learning with Applications*, Elsevier, p. 100076, 2021.

TERLOUW, E. et al. **Pre-slaughter conditions, animal stress and welfare: current status and possible future research.** *animal*, Elsevier, v. 2, n. 10, p. 1501–1517, 2008.

WANG, Z. et al. **Asas-nanp symposium: Applications of machine learning for livestock body weight prediction from digital images.** *Journal of Animal Science*, Oxford University Press US, v. 99, n. 2, p. skab022, 2021.

WEBER, V. A. M. et al. **Cattle weight estimation using active contour models and regression trees bagging.** *Computers and Electronics in Agriculture*, Elsevier, v. 179, p. 105804, 2020.

WILLMOTT, C. J.; MATSUURA, K. **Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance.** *Climate research*, v. 30, n. 1, p. 79–82, 2005.

YANG, L.; MEER, P.; FORAN, D. J. **Unsupervised segmentation based on robust estimation and color active contour models.** *IEEE Transactions on Information Technology in Biomedicine*, IEEE, v. 9, n. 3, p. 475–486, 2005.

YU, H.; LEE, K.; MOROTA, G. **Forecasting dynamic body weight of nonrestrained pigs from images using an rgb-d sensor camera.** *Translational Animal Science*, Oxford University Press US, v. 5, n. 1, 2021.

ZHANG, J. et al. **Pig weight and Body Size Estimation Using a Multiple Output Regression Convolutional Neural Network: A Fast and Fully Automatic Method.** *Sensors Multidisciplinary Digital Publishing Institute*, v. 21, n. 9, p. 3218, 2021.

ZAITOUN, N. M.; AQEL, M. J. **Survey on image segmentation techniques.** *Procedia Computer Science*, Elsevier, v. 65, p. 797–806, 2015.

ZHONG, Y. **Intrinsic shape signatures: A shape descriptor for 3d object recognition.** In: IEEE. 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops . [S.l.], 2009.

ZHOU, S. et al. **Active contour model based on local and global intensity information for medical image segmentation.** Neurocomputing, Elsevier, v. 186, p. 107–118, 2016.

Capítulo I

Weight prediction of Canela-Preta chicken using computer vision

Weight prediction of Canela-Preta chicken using computer vision

Ana Caroline de Almeida Alves^{1*}, Mateus Pinto Garcia², José dos Santos de Moura^{1,3},
Natanael Pereira da Silva Santos^{1,4}, Romuere Rodrigues Veloso e Silva²

¹ Graduate Program in Tropical Animal Science, Agrarian Sciences Center (CCA),
Federal University of Piauí (UFPI), Campus Universitário Ministro Petrônio Portella,
Teresina, PI 64049-550, Brazil

² Information Systems, UFPI, Campus Senador Helvídio Nunes de Barros, Picos, PI
64607-670, Brazil

³ Federal Institute of Education, Science and Technology of Piauí (IFPI), Rua da
Barragem do Bezerra, s/n, José de Freitas, PI 64110-000, Brazil

⁴ Department of Animal Science, CCA, UFPI, Campus Universitário Ministro Petrônio
Portella, Teresina, PI 64049-550, Brazil

*Correspondence: carol_alves@ufpi.edu.br

Highlights

- Computer vision has proven to be useful in predicting animal weight.
- The scores obtained in the evaluation of the segmentation metrics reached 45% of R^2 .
- The best performance for the prediction test was obtained using the Random Forest regressor.

Abstract

Monitoring broiler bird weight is critical for determining the ideal market weight. However, traditional monitoring methods, such as the use of scales, can be costly, labor-intensive and stressful for animals. Therefore, the application of methods based on computer vision can offer a more efficient and less stressful alternative for monitoring animal growth. The use of computer vision tools allows for the extraction of important information from the images of birds, such as their body weight, without the need for physical contact with the animals. In this study, we aimed to develop an algorithm for predicting the weight of Canela-Preta chickens using methods based on computer vision. The predictive ability of each method was evaluated, and R^2 values of up to 45% were obtained. Our findings showed that computer vision represents a less expensive and more humane solution for monitoring the weight of commercial birds.

Keywords: Free-range chicken, Creole chicken, Machine learning, Image processing, Growth

Introduction

Monitoring body weight is an important factor in livestock systems because frequent measurements allow farmers to obtain better results in management and determine the ideal market weight of the animal [1]. Previous monitoring enables decision-making in relation to the diet and management of the flock.

Research conducted under experimental conditions and on commercial farms has shown that rough handling affects the animal's fear of humans and can affect the stress physiology, productivity, and meat quality of livestock [2]. Thus, different studies have been conducted aiming to reduce stress and improve animal welfare in farming conditions. Over the last two decades, an increasing number of consumers have demanded animal products that reflect their animal welfare concerns. In this context, the Brazilian poultry industry is undergoing a process of increasing rigor in relation to production practices related to breeding and pre-slaughter operations since welfare is an extremely important subject that has received great attention in developed countries [3]. Therefore, it is important that livestock producers follow good animal management practices, ensuring that their animals receive proper treatment and are raised in environments that meet their natural needs.

Researchers have shown that digital image analysis can be used to determine measurements and body weight of animals. There are several studies with animals aimed at image processing with measurement, counting, and weight prediction [4].

A tool for animal body weight prediction would be very useful for chicken producers, since it would allow for the accurate estimation of the weight of birds without manual measurements. In addition, automatic prediction of body weight can be performed more quickly and efficiently, allowing the producer to make more accurate decisions about animal nutrition and management. An automated system to predict

body weight can be a competitive advantage for livestock producers, which would improve quality and productivity and consequently increase profitability. In addition, a body weight prediction tool can help to identify animals with health or nutrition problems, which can reduce animal mortality and financial losses.

Previous studies have reported the development of algorithms using computer vision to extract features such as body measurements and shape descriptors to predict the body weight of animals (e.g., pigs, cattle, and chickens) from images. In those studies, images of animals whose data were collected were used for the development of predictive models. However, there is a lack of published information about the use of those approaches on free-range chickens.

Computer vision methods can be used to determine the live weight of farm animals without human intervention. As there is an important correlation between body measurements and body weight, body mass can be estimated based on biometric indices obtained using images. Therefore, it is possible to estimate the animal weight using a method that reduces stress.

Materials and methods

Ethical certification and experimental characterization

The experimental procedures carried out in this study were approved by the Committee on Ethics in the Use of Animals (CEUA) of the Federal University of Piauí (UFPI), Brazil (protocol number 692/2021).

For this experiment, we collected images of healthy free-roaming Canela-Preta hens raised on the Débora Carvalho Farm, located in the rural settlement Resistência Camponesa, in the municipality of Demerval Lobão, PI, Brazil (latitude 5°21'29"S, longitude 42°40'32"W). The experimental birds were divided into two groups of

individuals with predominant ages of 9 and 16 months. The birds were dewormed and vaccinated against Newcastle disease, Gumboro disease, infectious bronchitis, fowl pox, and avian infectious coryza. Each vaccine was administered according to the protocol used on the farm.

Data collection

Images were captured using an Xbox One camera equipped with the Microsoft Kinect sensor. A Ryzen 7 5700G computer, 16 Core, 16 GB RAM was used for the development of algorithms and tests with computer vision. The essential steps for the development of the study are shown in Fig. 1. The first step consisted of the collection of images and weights of animals. In the second step, the Mask R-CNN approach was applied. Subsequently, the features were extracted from images, followed by the selection of features for prediction tests. Finally, the results were evaluated.

Our database consisted of images of 161 Canela-Preta chickens, which were arranged within different folders containing several images (frames with resolution 512 × 424) of each bird. The animals used in the current study were kept in a wide surrounded coop. The weight and images of each hen were individually collected outside the coop. The actual weights obtained were used later, both in the training phase and in the evaluation, to measure the effectiveness of the prediction. The birds were placed in a corridor with a camera positioned above the target to capture the image of each chicken as it walked (Fig. 2).

When the acquired images are not adequate, the desired task may not be achievable [5]. Therefore, we manually selected the best frames to obtain the standard angle and quality of the images. This step is necessary, as there are many similar images of the same animal, as well as inadequate images. Thus, images without any

variation were discarded, and images in which the chickens were predominantly dorsal and with the background image at the same angle or with the most similar angle possible were selected. As a result, 104 chickens with satisfactory images remained for the tests.

Segmentation

After preparing the database and selecting the best frames, segmentation was conducted. In our database, the chicken's body is predominantly in the center of the image with some noise around it, such as objects, leaves or image noise that could be identified by the algorithm as a region of interest (ROI) and cause algorithm errors. Thus, all objects connected to the edge of the ROI were removed from the image to isolate only the animal's body in the image.

For segmentation, we used Mask R-CNN, which is a Convolutional Neural Network (CNN) in terms of image segmentation [6]. In the current study, this approach was used to detect objects and extract their segmented masks from the images. The resulting binary mask was used to extract the region of interest from the original image.

Feature extraction

Relevant information was extracted from the images using algorithms based on feature descriptors, which allowed obtaining numerical values representative of specific image features. The resulting values were further processed to provide useful information about the image.

The extracted features were grouped into a feature vector, which was used in subsequent steps of this study. The image characteristics were exported and saved in a document, in which they were used as variables in the prediction tests.

The extraction of characteristics from the images was performed using the scikit-image library available for the Python programming language, which is focused on image processing and allowed us to obtain several measures and characteristics of each object present in a binary image. The extraction of characteristics related to the shape of chickens was performed using the regionprops module.

The extracted features are essential for the regression process, as they provide relevant information for predicting chicken weight. The extraction approach allows obtaining detailed information about the analyzed images and creating a consistent dataset for training the regression algorithm.

A total of 54 features were extracted from the images, which was followed by the selection of characteristics to remove the features with the lowest correlation with the response variable, i.e., the weight to be predicted.

Feature selection

After extracting features, the Pearson correlation coefficient was estimated. It is important to mention that this coefficient only measures the linear relationship between two variables and is not suitable for measuring nonlinear correlations. Pearson's correlation ranges from -1 to +1, where -1 represents a perfect negative correlation (when one variable increases, the other decreases in the same proportion), +1 represents a perfect positive correlation (two variables increase simultaneously) and 0 represents the absence of a linear correlation [7]. When there are various features in a regression model, Pearson's correlation can be used to identify which variables are the most associated with the response variable. This assists in the selection of the most relevant features for the regression model.

In the current study, a total of 54 features were extracted. The Pearson correlation analysis showed which characteristics had the highest correlation with the response variable (weight to be predicted). Finally, the characteristics that had the lowest correlation with the response variable were excluded. The features described in Table 1 remained and were used for prediction.

Prediction tests

Cross-validation and division into training and testing data were used for evaluation of the regression models.

The K-fold cross-validation procedure was used to validate the generalization of the base, where $K = 5$. This analysis allowed us to evaluate the effectiveness of the prediction models in relation to the body weights of birds. Cross-validation is widely used for the estimation of prediction error [8]. This approach consists of dividing the dataset into K roughly equal-sized subsets (folds). Subsequently, the model is trained on $K-1$ of these subsets and tested on the remaining subset. This process is repeated K times, using each subset as the test set once and using the other $K-1$ subsets as the training set. At the end of the process, the average of the evaluation metrics is calculated. K-fold is useful to assess the generalizability of a prediction model on unseen data and reduce the risk of overfitting (too much fit to training data).

The next step consisted of using predictive approaches to estimate the body weight of the chickens. For this, 70% of the data were used to train the algorithm and 30% for testing. During the training phase, we used both the real weights of the animals and the features extracted from the images to fit the prediction models.

In this step, four regressors were applied to create the predictive models: linear regression, lasso, ridge, and random forest. These algorithms were applied using

several individual images (frames) of 104 chickens, considering the division of training and testing mentioned above. In order to evaluate the quality of the predictions generated by each model, we used different performance metrics, such as R-squared (R^2), mean absolute error (MAE), and mean absolute percentage error (MAPE).

Results and Discussion

In the current study, we evaluated the results of the predictions according to the performance of the algorithms used. Therefore, in this section we discuss the results, as well as the metrics used for evaluation.

Cross-validation was applied to the complete database to evaluate the performance of the regression model and for selection of the best model with the results obtained. For cross-validation, we considered $K = 5$.

Table 2 shows the results of the k-fold cross-validation with $K = 5$ for linear regression, lasso, ridge, and random forest. Adamczak et al. [9] and Cominotte et al. [1] mentioned that it is important to perform cross-validation to determine the performance of the regression model. Cominotte et al. [1] adopted leave-one-out cross-validation to assess the predictive quality of four approaches. Cross-validation is widely used for the estimation of prediction error [8]. In the current study, the means and standard deviations of the R^2 , MAE and MAPE metrics were calculated for each algorithm. The results indicate that the random forest algorithm had the best performance in all metrics, with average values of 0.37, 66.71, and 0.03 for R^2 , MAE, and MAPE, respectively. On the other hand, linear regression had the poorest performance, with an average negative R^2 of -0.28 and an average MAE of 108. The Lasso and Ridge algorithms performed intermediately regarding R^2 , with averages of 0.27 and 0.22, respectively.

The standard deviation of evaluation metrics is also important to assess the variability of the results and stability of the model. In the current study, we observed that the standard deviation of the metrics used for Random Forest was lower compared to those obtained using the other models (Table 2). This indicated that Random Forest yielded more stable and consistent predictions, which is why this algorithm was considered the most suitable for the proposed task.

In the second moment, 70% of the data were used for training and 30% for testing to evaluate the performance of the algorithms for the testing set using R^2 , MAE, and MAPE (Table 3).

R^2 is a measure of goodness of fit of a model that ranges from 0 to 1 and indicates the proportion of the total variance in the data explained by the model. Hence, a model with a higher value of R^2 is considered to be the best. In this study, the R^2 values ranged from 0.21 to 0.45, which indicates a moderate goodness of fit.

MAE measures the average absolute error (difference) between actual and predicted values. We obtained MAE values between 74.96 and 104.00 (Table 3). These findings indicate that the model predictions may have average errors ranging from 74.96 to 104 g (grams).

MAPE is the mean of absolute percentage errors between model predictions and actual values. In the current study, the MAPE values ranged from 4.08 to 5.77%, which indicates that the model predictions are, on average, 4.08 to 5.77% off from the actual values.

In summary, the results indicate that the models in this study have reasonable fit, with predictions that are probably, on average, 74.96 to 104 units of measurement or 4.08 to 5.77% off from the actual values. It is important to highlight that the choice of the best model to predict the live body weight of chickens depends on the specific

context and characteristics of the data. Furthermore, model selection must be evaluated considering other metrics and performance criteria.

Our findings indicate that Random Forest is the model that best fits the data used in this work and shows the best predictions for new data. This can be explained by the ability of Random Forest to handle complex and correlated data, as well as to its ability to handle overfitting and generalize well to new data.

We emphasize that the performance of different techniques can vary depending on the dataset and the forecast objective. Therefore, it is important to evaluate multiple methods before selecting the most appropriate for each study. For example, Weber et al. [10] submitted image features to Bagging, Regression by Discretization, and Random Forest algorithms for cattle weight prediction. They observed that the Bagging algorithm showed the best results, with MAE of 13.44 kg (± 2.76), RMSE of 15.88 kg (± 2.86), MAPE of 2.27%, and R^2 at 0.75. On the other hand, they reported that Random Forest presented inferior results in relation to other metrics, for example, an R^2 of 0.72.

Conclusion

This work presents an algorithm that represents the first step for the development of a computational tool to monitor and measure the weight of birds in real time, enabling poultry producers to make more accurate decisions regarding management, zootechnical control, welfare, and animal feeding.

Our findings demonstrated that it is possible to use images and image processing techniques to predict the body weight of chickens. Nevertheless, there were some limitations to obtaining significant weight predictions in this study (e.g., errors in image capture because of feather coverage or chicken movement). Probably the proposed

algorithm would achieve more satisfactory results using chicken carcasses, as it would exclude the limitations mentioned above.

We believe that the application of the proposed algorithm with its necessary improvements can be a breakthrough in the poultry industry, which could benefit small- and medium-scale producers of Canela-Preta chickens.

Acknowledgements

We thank the National Council for Scientific and Technological Development (CNPq) and the Coordination for the Training of High Qualified Education Personnel (CAPES).

Funding

The National Council for Scientific and Technological Development (CNPq) granted a scholarship to the first author (Process: 138168/2021-0). This study was partially financed by the Coordination for the Training of High Qualified Education Personnel–Brazil (CAPES; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) Finance Code 001.

Author contributions

ACAA for Ana Caroline de Almeida Alves, MPG for Mateus Pinto Garcia, JSM for José dos Santos de Moura, RRVS for Romuere Rodrigues Veloso e Silva, NPSS for Natanael Pereira da Silva Santos. Designed the study: ACAA and RRVS. Performed data collection: ACAA, MPG, JSM, RRVS, and NPSS. Formulated and planned the experiments: ACAA, MPG, RRVS, and NPSS. Conducted the experiments: ACAA and MPG. Interpreted the outcomes: ACAA and MPG. Drafted the article: ACAA. Reviewed the article: ACAA and NPSS. All authors read and approved the final manuscript.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations**Ethical approval and consent to participate**

The experimental procedures carried out in this study were approved by the Committee on Ethics in the Use of Animals (CEUA) of the Federal University of Piauí (UFPI), Brazil (protocol number 692/2021).

Human and animal rights

This study complies with current laws in Brazil and was approved by the CEUA/UFPI, Brazil (protocol number 692/2021).

Consent for publication

Not applicable

Competing interests

The authors have no competing interests to disclose.

References

1. Cominotte A, Fernandes AFA, Dorea JRR, Rosa GJM, Ladeira MM, van Cleef EHCB, Pereira GL, Baldassini WA, Machado Neto OR. Automated computer vision system to predict body weight and average daily gain in beef cattle during growing

- and finishing phases. *Livestock Science*. 2020; 232: 103904. <https://doi.org/10.1016/j.livsci.2019.103904>.
2. Hemsworth PH, Rice M, Karlen MG, Calleja L, Barnett JL, Nash J, Coleman GJ. Human–animal interactions at abattoirs: Relationships between handling and animal stress in sheep and cattle. *Applied animal behavior science*. 2011; 135: 24-33. <https://doi.org/10.1016/j.applanim.2011.09.007>.
 3. Sakamoto KS, Benincasa NC, Silva IJO, Lobos CMV. The challenges of animal welfare in modern Brazilian poultry farming. *Journal of Animal Behaviour and Biometeorology*. 2020; 8: 131-135. <http://dx.doi.org/10.31893/jabb.20018>.
 4. Sant’Ana DA, Pache MCB, Martins J, Soares WP, Melo SLN, Garcia V, Weber VAM, Heimbach NS, Mateus RG, Pistori H. Weighing live sheep using computer vision techniques and regression machine learning. *Machine Learning with Applications*. 2021; 5: 100076. <https://doi.org/10.1016/j.mlwa.2021.100076>.
 5. Sampath V, Maurtua I, Aguilar Martín JJ, Gutierrez A. A survey on generative adversarial networks for imbalance problems in computer vision tasks. *Journal of big Data*. 2021; 8: 27. <https://doi.org/10.1186/s40537-021-00414-0>.
 6. Abdulla W. Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow. [S.l.]: Github, 2017. https://github.com/matterport/Mask_RCNN.
 7. Baak M, Koopman R, Snoek H, Klous S. A new correlation coefficient between categorical, ordinal and interval variables with Pearson characteristics. *Computational Statistics & Data Analysis*. 2020; 152: 107043. <https://doi.org/10.1016/j.csda.2020.107043>.
 8. Fushiki T. Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*. 2011; 21: 137-146. <https://doi.org/10.1007/s11222-009-9153-8>.

9. Adamczak L, Chmiel M, Florowski T, Pietrzak D, Witkowski M, Barczak T. The use of 3D scanning to determine the weight of the chicken breast. *Computers and Electronics in Agriculture*. 2018; 155: 394-399. <https://doi.org/10.1016/j.compag.2018.10.039>.
10. Weber VAM, Weber FL, Oliveira AS, Astolfi G, Menezes GV, Porto JVA, Rezende FPC, Moraes PH, Matsubara ET, Mateus RG et al. Cattle weight estimation using active contour models and regression trees Bagging. *Computers and Electronics in Agriculture*, 2020; 179: 105804. <https://doi.org/10.1016/j.compag.2020.105804>.

Figure captions

Fig. 1 Steps of the proposed methodology

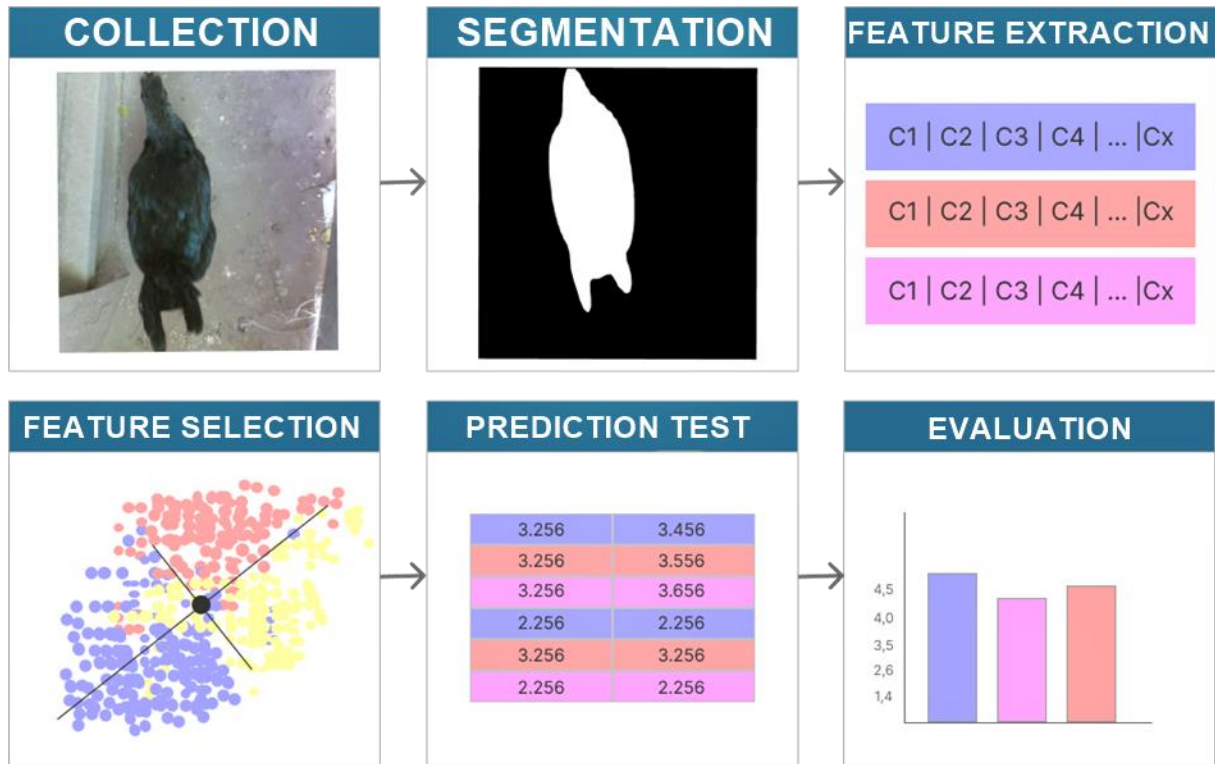


Fig. 2 Steps of the collection of weight and images of Canela-Preta chickens

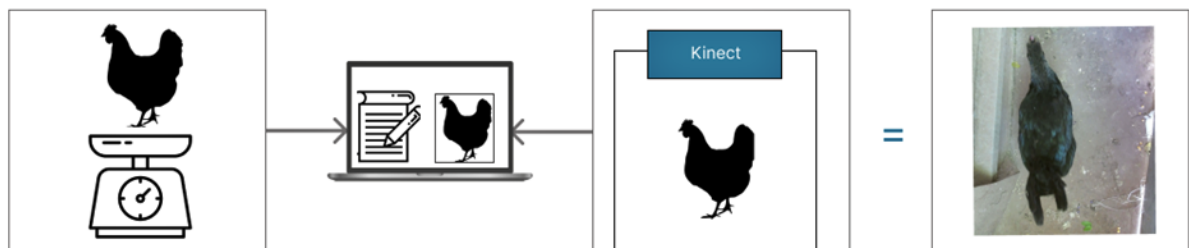


Table 1 Features based on image regions

Extracted feature	Description
<i>area</i>	Area of the region delimited by the contour of the image
<i>bbox_area</i>	Area of the smallest axis-aligned rectangle containing the region
<i>convex_area</i>	Area of the region delimited by the convex contour of the region
<i>equivalent_diameter</i>	Diameter of the circle with the same area as the region
<i>filled_area</i>	Area of the region with all holes filled
<i>major_axis_length</i>	Length of the long axis of the ellipse that best fits the region
<i>minor_axis_length</i>	Length of the minor axis of the ellipse that best fits the region
<i>perimeter</i>	Perimeter of the region delimited by the contour of the image
<i>feret_diameter_max</i>	Maximum distance between two tangent parallel points on opposite sides of the particle or object
<i>moments</i>	Moments 1 of order 1 to 3 of the region
<i>moments_central</i>	Central moments of order 1 to 3 of the region

Table 2 Results of the K-fold cross-validation

Average	Standard deviation
----------------	---------------------------

Algorithm	R²	MAE	MAPE	R²	MAE	MAPE
Linear regression	-0.28	108.00	0.06	0.55	31.74	0.01
Lasso	0.27	80.62	0.04	0.08	7.89	0.0043
Ridge	0.22	82.41	0.04	0.08	9.51	0.0051
Random Forest	0.37	66.71	0.03	0.11	7.30	0.0045

R²: coefficient of determination; MAE: mean absolute error; MAPE: mean absolute percentage error; Lasso: least absolute shrinkage and selection operator.

Table 3 Comparison of the approaches used for weight prediction

Algorithm	Evaluation of the testing set		
	R ²	MAE	MAPE (%)
Linear regression	0.21	102.00	5.66
Lasso	0.21	99.28	5.45
Ridge	0.22	104.00	5.77
Random Forest	0.45	74.96	4.08

R²: coefficient of determination; MAE: mean absolute error; MAPE: mean absolute percentage error; Lasso: least absolute shrinkage and selection operator.