

**MICAEL BARREIROS RICATO**

**COMPARAÇÃO DE ALGORITMOS DE MACHINE LEARNING PARA PREDIÇÃO  
DO PESO CORPORAL DE CABRAS LEITEIRAS USANDO DADOS DE  
PROCESSAMENTO DE IMAGENS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Zootecnia, para obtenção do título de *Magister Scientiae*.

Orientadora: Érica Beatriz Schultz

Coorientadora: Karina Costa Busato

**VIÇOSA - MINAS GERAIS  
2024**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

R488c  
2024

Ricato, Micael Barreiros, 1998-

Comparação de algoritmos de *machine learning* para  
predição do peso corporal de cabras leiteiras usando dados de  
processamento de imagens / Micael Barreiros Ricato. – Viçosa,  
MG, 2024.

1 dissertação eletrônica (34 f.): il. (algumas color.).

Orientador: Erica Beatriz Schultz.

Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Zootecnia, 2024.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2024.256>

Modo de acesso: World Wide Web.

1. Caprinos - Pesos e medidas. 2. Aprendizado do  
computador. 3. Algoritmos. 4. Morfometria. 5. Imagens digitais.  
I. Schultz, Erica Beatriz, 1992-. II. Universidade Federal de  
Viçosa. Departamento de Zootecnia. Programa de  
Pós-Graduação em Zootecnia. III. Título.

CDD 22. ed. 636.39


**MICAEL BARREIROS RICATO**

**COMPARAÇÃO DE ALGORITMOS DE MACHINE LEARNING PARA PREDIÇÃO  
DO PESO CORPORAL DE CABRAS LEITEIRAS USANDO DADOS DE  
PROCESSAMENTO DE IMAGENS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Zootecnia, para obtenção do título de *Magister Scientiae*.


APROVADA: 23 de fevereiro de 2024.

Assentimento:

Documento assinado digitalmente  
 **MICAEL BARREIROS RICATO**  
Data: 18/07/2024 08:32:16-0300  
Verifique em <https://validar.iti.gov.br>

---

Micael Barreiros Ricato  
Autor

Documento assinado digitalmente  
 **ERICA BEATRIZ SCHULTZ**  
Data: 19/07/2024 15:41:58-0300  
Verifique em <https://validar.iti.gov.br>

---

Érica Beatriz Schultz  
Orientadora

*Ao meu Deus, meus pais, minha irmã e a todos que contribuíram na minha trajetória.*

## AGRADECIMENTOS

Ao meu amado Deus, por seu amor que constrange e sua graça infinita.

Aos meus pais e minha irmã, pelo apoio de sempre e orações constantes. Foram fundamentais para que eu chegasse até aqui. Realmente um alicerce na minha vida.

A minha orientadora, pelo acompanhamento durante todo o percurso. Com calma, atenção, dedicação e carinho me conduziu neste tempo, ainda que curto, com muita maestria. Mesmo com muitas atividades e trabalhos que possuía, de pronto me atendeu e concedeu a oportunidade de trabalhar contigo. Muito obrigado pelos ensinamentos, momentos de confraternização e pela paciência comigo, Professora Érica!!

A minha coorientadora, Professora Karina, pela oportunidade de ter sido seu orientado. Agradeço pelos muitos aprendizados que recebi e a oportunidade que me concedeu de vivenciar a rotina do setor de ovinos de maneira singular. Também sou grato pela parceria de sempre e pelos incontáveis almoços, programações gastronômicas e pela vida dos seus pais.

Aos funcionários da antiga granja de melhoramento. Quanto aprendizado, quantas alegrias e momentos de descontração!

Aos funcionários do Capril, pela amizade e auxílio em diversos momentos.

Ao José Geraldo, pela parceria. Aprendi muito, não só de ovinocultura, mas de diversas outras coisas. Valeu pelas opiniões e conversas sempre agradáveis Zé!

Aos meus irmãos e Pastores da Igreja Assembleia de Deus, ministério de Madureira, que sem sombra de dúvidas foram instrumentos de Deus na minha vida. Só a eternidade pode retribuir a vocês o que fizeram por mim.

Ao Mateus, pela ajuda nesta reta final do mestrado. Sua participação foi fundamental para que eu conseguisse chegar até aqui.

À Daniela pela força e apoio durante a coleta.

À Eula pela ajuda com os dados.

À Taynah e sua família pela parceria e amizade.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos.

*“ Melhor é o fim das coisas do que o princípio delas”.*  
(Eclesiastes 7.8)

## RESUMO

RICATO, Micael Barreiros, M.Sc., Universidade Federal de Viçosa, fevereiro de 2024. **Comparação de algoritmos de *machine learning* para predição do peso corporal de cabras leiteiras usando dados do processamento de imagens.** Orientadora: Érica Beatriz Shultz. Coorientadora: Karina Costa Busato

Objetivou-se verificar qual o melhor algoritmo de aprendizado de máquina associado a atributos de imagens digitais para predição de peso corporal de caprinos leiteiros. Foram utilizadas 186 fêmeas das raças Saanen (n = 98) e Alpina (n = 88) das categorias de cria, recria e lactação. As gravações foram realizadas com câmera Intel® RealSense™ (D435) para a extração das imagens 2D e processadas no software ImageJ para a obtenção das medidas lineares de altura de cernelha, altura de garupa, comprimento de garupa, largura de garupa, comprimento corporal, largura do peito, profundidade e altura da pata. Os animais foram pesados, sem jejum prévio, imediatamente antes das gravações. A análise de correlação mostrou que as características que mais se correlacionaram com o peso foram a profundidade (r = 0,94) e o comprimento corporal (r = 0,94), seguidas pela largura do peito (r = 0,92), altura de cernelha e altura de garupa (r = 0,90). Entre os algoritmos estudados, o Random Forest mostrou-se o mais preciso na predição do peso corporal, com maior coeficiente de determinação ( $R^2 = 0,94$ ) e as menores estimativas dos erros (RMSE = 5,72 e MAE = 4,14). Conclui-se que é possível prever o peso de caprinos leiteiros por meio atributos extraídos de imagens digitais e que o algoritmo Random Forest foi o algoritmo que apresentou o melhor desempenho quando comparado com as regressões Lasso, Ridge e Elastic Net.

Palavras-chave: Machine learning. Zootecnia de precisão. Morfometria por imagem.

## ABSTRACT

RICATO, Micael Barreiros, M.Sc., Universidade Federal de Viçosa, February, 2024. **Comparison of machine learning algorithms for predicting body weight of dairy goats using image processing data.** Adviser: Érica Beatriz Shultz. Co-advisers: Karina Costa Busato.

The objective was to verify the best algorithm associated with image morphometry attributes to predict the live weight of dairy goats. 186 females of the Saanen (n = 98) and Alpina (n = 88) breeds from the breeding, rearing and lactation categories were used. The 2D images were recordings and images processed to obtain morphological measurements: withers height, rump height, rump length, rump width, body length, chest width, depth and height of the leg. The animals were weighed, without prior fasting, immediately before recordings. Correlation analysis showed that the characteristics that most correlated with live weight were depth (r = 0.94) and body length (r = 0.94), followed by chest width (r = 0.92), withers height and croup height (r = 0.90). Among the algorithms studied, Random Forest proved to be the most accurate in predicting body weight, with the highest coefficient of determination ( $R^2 = 0.94$ ) and the lowest error estimates (RMSE = 5.72 and MAE = 4.14) than Lasso ( $R^2 =$  ; RMSE = and MAE= ), Ridge ( $R^2 =$  ; RMSE = and MAE= ) and Elastic Net ( $R^2 =$  ; RMSE = and MAE= ). It is concluded the Random Forest algorithm showed better performance than Lasso, Ridge, and Elastic Net regressions to predict the weight using image morphometry attributes of dairy goats.

Keywords: Machine learning, Precision livestock farming. Image morphometry.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1:</b> Passos do processamento de imagens digitais. ....	14
<b>Figura 2:</b> Representação dos locais de medição.....	22
<b>Figura 3:</b> Exemplo do processamento para extração das medidas.....	23
<b>Figura 4:</b> Correlação entre as medidas morfométricas e o peso. ....	26
<b>Figura 5:</b> Relação entre os pesos preditos e observados pelos algoritmos.....	27
<b>Figura 6:</b> Importância das características para o Random Forest. ....	29

## LISTA DE TABELAS

<b>Tabela 1:</b> Resumo do banco de dados da fase de cria.....	20
<b>Tabela 2:</b> Resumo do banco de dados da fase de recria. ....	20
<b>Tabela 3:</b> Resumo do banco de dados da fase de lactação. ....	21

## SUMÁRIO

1. INTRODUÇÃO .....	11
2. REVISÃO DE LITERATURA .....	13
3. MATERIAL E MÉTODOS .....	20
4. RESULTADOS E DISCUSSÃO.....	26
5. CONCLUSÃO.....	30
7. REFERÊNCIAS.....	31

## 1. INTRODUÇÃO

A escrituração zootécnica consiste na coleta e registro de informações advindas do sistema produtivo, algo essencial para uma boa gestão. Quando se pensa em caprinocultura de leite e quais informações coletar, é comum que sejam apontadas informações como a produção de leite por lactação, o intervalo entre partos, a persistência de lactação, a prolificidade (crias/parto) e outros índices (RIBEIRO, 1997). Entretanto, não menos importante, o peso corporal é um dos dados que deve ser registrado e analisado para o alcance de bons resultados.

Os registros dos pesos de forma regular e periódica permitem que seja realizado um acompanhamento do desenvolvimento ponderal dos animais, medida que possibilita monitorar o ganho dos animais de acordo com as fases de criação (cria, recria e lactação/reprodução). Essas informações são úteis para determinar, por exemplo, se uma cabrita poderá ser coberta, se já tiver atingido a puberdade zootécnica (60 a 70% do peso adulto) ou não (RIBEIRO, 1997). Além disso, com base nos dados de pesos corporais são calculadas dosagens de medicamentos, o ganho de peso, o ajuste da dieta e o ponto de abate.

Para aferir o peso corporal, pode-se lançar mão do uso de balanças convencionais, como é feito em muitas propriedades e cujo processo é laborioso. No entanto, com o desenvolvimento tecnológico de sensores e ferramentas de inteligência artificial (IA) a produção animal tem sido beneficiada com diversas aplicações para automatizar e otimizar os processos e tornar os sistemas mais eficientes e rentáveis (PANDORFI; ALMEIDA; GUISELINI, 2012), dentre essas tecnologias está a possibilidade de prever o peso dos animais por meio de imagens digitais.

Para que isso seja possível, é preciso extrair medidas corporais, a partir destas imagens, que serão posteriormente utilizadas para a predição. Nesse sentido, um estudo utilizando visão computacional, com caprinos e bovinos, mostrou que por meio das imagens pode-se extrair morfometrias precisas (LI; TENG, 2022). Uma conclusão semelhante foi apontada em um estudo com bovinos, que mostrou que o sistema de medição por imagens pode ser tão preciso ou até mais quando comparado aos métodos convencionais (GAUDIOSO et al., 2014).

A partir da análise de atributos (características) extraídos de imagens, algoritmos de aprendizado de máquina, associados à inteligência artificial, podem ser utilizados para prever o peso corporal de forma precisa, como ficou demonstrado em trabalhos realizados com ovinos (HAMADANI; GANAI, 2023; SANT'ANA et al., 2021). Essa forma de previsão do peso corporal, é uma estratégia que pode ser usada para reduzir a mão de obra, otimizar o tempo e tornar o processo menos estressante aos animais (ZHANG et al., 2018).

Algoritmos de aprendizado supervisionado, como o Random forest, Lasso, Ridge e Elastic Net, podem ser usados nos casos de previsão (LEITE; MORAES; LOPES, 2020; ZOU; HASTIE, 2005). Para tal, é necessário que estes algoritmos aprendam com um conjunto de dados rotulados (treinamento), por exemplo medidas do animal e peso do animal, para que quando forem apresentados a um conjunto de dados de teste, apenas as medidas do animal, este possam realizar as previsões do peso.

Muito embora existam diversos algoritmos que possam ser usados, o teste e validação destes devem ser realizados para verificação daquele que apresenta melhor capacidade preditiva entre seus concorrentes. Essa avaliação é feita por meio de critérios que medem a qualidade do ajuste (relação entre predito e observado), e as estimativas dos erros (IQBAL et al., 2022). Nesse sentido, trabalhos testando diferentes algoritmos têm sido realizados com ovinos (ZHANG et al., 2018) e caprinos (IQBAL et al., 2022), apontando aquele que melhor se adequa para a previsão.

Diante disso o objetivo foi verificar qual o melhor algoritmo de *machine learning* associado a atributos de imagens digitais para previsão de peso corporal de caprinos leiteiros.

## 2. REVISÃO DE LITERATURA

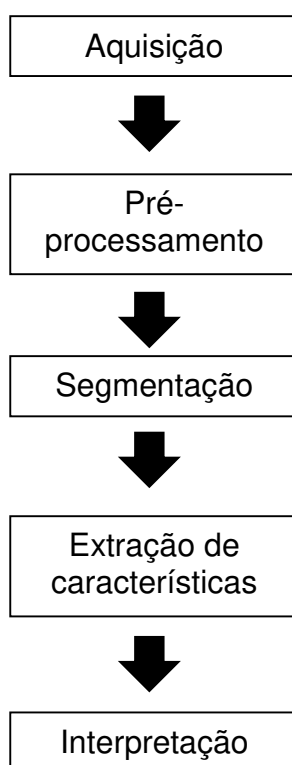
### Uso e processamento de imagens para predição do peso corporal

O uso das imagens digitais tornou-se fundamental em diversas áreas do conhecimento. O processamento e a análise de imagens digitais, por vezes associado às ferramentas da inteligência artificial (IA), estão se tornando necessários para pesquisadores e profissionais em suas atividades. As imagens digitais, e suas aplicações, vêm sendo usadas na ciência animal (FERNANDES; DÓREA; ROSA, 2020), na geografia (LI; HSU, 2022), na agricultura (SAXENA; ARMSTRONG, 2014), na medicina (KER et al., 2018; MÜLLER et al., 2004) e em muitas outras áreas, como ferramentas para interpretação de situações e tomada de decisão.

Mas, o que é uma imagem digital? Pode-se entendê-la como uma matriz de pixels. Os *pixels*, ou elementos de imagem, são as unidades básicas que compõe uma imagem digital (MACHADO; SOUKI, 2004), possuindo localização, determinada por coordenadas espaciais (x, y), e valores de intensidade específicos (GONZALEZ; WOODS, 2009). Nas imagens em preto e branco, esse valor atribuído ao *pixel* se refere a uma intensidade (nível de cinza), já nas imagens coloridas, a composição desse valor é um pouco mais complexa, pois leva em consideração uma combinação de três valores das cores vermelho (red), verde (green) e azul (blue) para formar as demais cores da imagem (QUEIROZ; GOMES, 2006).

O processamento digital de imagens possui duas vertentes de interesses quanto à aplicação. De acordo com Gonzalez e Woods (2009), em primeiro lugar está a busca por melhoria da imagem (das informações visuais) para que, por meio da interpretação humana, seja possível extrair informações úteis. Já o segundo ramo, baseia-se no processamento de dados de imagens, mas com o foco na percepção automática por meio dos computadores (máquinas), isso com o auxílio da IA.

Na Figura 1, estão representados os passos do processamento, seja para interpretação humana ou por computadores. É válido observar que esta figura traz os passos de forma resumida, pois estas etapas podem ser desdobradas e o nível de complexidade pode ser aumentado dependendo do problema a ser resolvido.



**Figura 1:** Passos do processamento de imagens digitais. Adaptado de Caliberto et al., (2008)

Em resumo, pode-se dizer que a aquisição das imagens é o passo no qual essas imagens são adquiridas, seja por um simples recebimento de uma imagem já digitalizada ou então por meio de câmeras e aparelhos especializados. O pré-processamento é a etapa na qual essas imagens são melhoradas, por exemplo para redução dos ruídos ou normalização das dimensões. A Segmentação é o momento em que são feitos recortes dos objetos de interesse nas imagens e/ou retirada do plano de fundo, uma das etapas fundamentais para a extração das características que se deseja avaliar nas imagens. Por fim, tem-se a interpretação dos resultados (CELIBERTO et al., 2008; GONZALEZ; WOODS, 2009).

A escolha das informações a serem extraídas das imagens, depende dos objetivos do estudo em questão e dos tipos de imagens. Na zootecnia, os trabalhos variam desde o uso de imagens 2D simples para extração precisa da área de objetos (características), como por exemplo a área de olho de lombo (FERREIRA et al., 2012) e área foliar de forrageiras (FERREIRA et al., 2017), até trabalhos com imagens hiperespectrais, avaliando qualidade da carne, nas quais podem ser extraídas informações de

composição química, contaminação microbiológica, características sensoriais como maciez ou coloração (JO et al., 2024).

A predição do peso corporal de animais, por meio de morfometrias extraídas de imagens, vem despertando o interesse dos pesquisadores, embora os estudos nesse sentido ainda sejam escassos. O fundamento dessas pesquisas está em trabalhos que já demonstraram a possibilidade de predizer o peso vivo, de forma precisa, por meio de morfometrias, avaliadas manualmente, e nas correlações que determinadas medidas possuem com o peso do animal (CONRADO; ARANDAS; RIBEIRO, 2015; SOUZA JUNIO, 2023).

Nos trabalhos de predição do peso a partir de imagens, o foco se dá em dois pontos. O primeiro está na extração das características (medidas corporais), seja por meio de *softwares* específicos ou então através da visão computacional (LI; TENG, 2022). Um exemplo deste tipo de estudo é o de Gaudioso et al. (2014), que mostraram o potencial de uso das imagens para se fazer medições em animais, sem que haja necessidade de contato. Ainda de acordo com estes autores, os resultados podem ser tão bons ou até melhores quando comparados com instrumentos convencionais de medição.

O segundo ponto está em usar essas medidas extraídas para predizer o peso por meio de algoritmos de aprendizado de máquina (LI; TENG, 2022; ZHANG et al., 2018). E nesse sentido os resultados têm mostrado que estes algoritmos possuem o potencial para tal finalidade, apresentando elevada capacidade preditiva e com pequena margem de erro (IQBAL et al., 2022).

## **Algoritmos de machine learning: conceito e aplicação**

O aprendizado de máquina (*machine learning*) é uma área da IA, que vem se tornando cada vez mais comum nos dias atuais (LUDERMIR, 2021). A capacidade de um computador (máquina) em realizar tarefas, sem que seja previamente programado, ou seja, a partir de informações que ele aprendeu com os dados, é um dos conceitos por trás do *machine learning* (GÉRON, 2021).

O aprendizado de máquina pode ser dividido em duas vertentes: aprendizado não supervisionado e aprendizado supervisionado. No aprendizado não supervisionado os dados não são rotulados, ou seja, não há uma informação prévia associada

ao dado. A ideia nesse tipo de aprendizado é descobrir categorias e identificar padrões semelhantes nos dados (GÉRON, 2021). Como exemplo deste tipo de problema pode-se citar a identificação de padrões semelhantes de clientes de uma determinada loja virtual por meio da análise das últimas compras. Essa informação pode ser útil para que campanhas de marketing sejam mais direcionadas de acordo com as preferências daquelas pessoas, por exemplo.

O aprendizado supervisionado consiste em fornecer ao algoritmo um conjunto de dados de treinamento que possuam rótulos associados a esses dados. Após a fase de treinamento, na qual a máquina já aprendeu, são fornecidos dados não rotulados para testar a capacidade de fazer novas previsões com base no conhecimento adquirido (GÉRON, 2021). Pode-se citar como exemplo do aprendizado supervisionado, a previsão do peso dos animais com base em morfometrias.

No exemplo a cima, são fornecidas para o algoritmo as informações de medidas corporais de vários animais e os pesos (rótulos) correspondentes dos mesmos. Feito isso o algoritmo será treinado para prever o peso com base nos dados que lhe foram entregues. Após essa fase, novas informações de medidas corporais serão inseridas, mas sem que haja rótulos (pesos) associados, nessa etapa o objetivo é avaliar como foi o aprendizado e a capacidade de fazer previsões do peso a partir de informações novas (medidas).

Dentro do aprendizado supervisionado, podem ser utilizados diversos algoritmos para solucionar os problemas. O Random Forest é um destes algoritmos de *machine learning*, que pode ser usado tanto em casos de classificação quanto de regressão (LEITE; MORAES; LOPES, 2020). O mecanismo deste modelo consiste na construção de uma floresta aleatória com diversas árvores de decisão, cujos resultados de cada árvore colaboram para a geração de um resultado final preciso (LORENZETT; TELÖCKEN, 2016). Uma vantagem do Random Forest é a sua flexibilidade e robustez quanto ao ruído e a ocorrência de overfitting (sobreajuste), ou seja, um ajuste excessivo aos dados o que leva a uma dificuldade em fazer generalizações (SARKAR et al., 2015).

Algoritmos de regressão penalizada como o Ridge, Lasso, Elastic Net, podem ser usados nos estudos de previsão. Diferentemente do mecanismo do Random Forest, estes trabalham adicionando uma penalidade, que vai variar de acordo com o

algoritmo escolhido, à função de regressão (GÉRON, 2021). Esse processo de penalização dos parâmetros tem como vantagens a redução da ocorrência de um sobreajuste aos dados (overfitting), a seleção das variáveis mais importantes para o modelo, descartando as de menor relevância e a redução da complexidade do modelo, algo que tem despertado a atenção por esclarecer melhor as relações entre a resposta e as variáveis estudadas (ZOU; HASTIE, 2005).

### **Avaliação da precisão e acurácia dos modelos de *machine learning***

Assim como a escolha dos modelos para o problema que se procura solucionar, a avaliação destes modelos é uma das etapas fundamentais. Muito embora existam diversos algoritmos que possam ser usados, o teste e validação destes devem ser realizados para verificação daquele que apresenta melhor capacidade preditiva entre seus concorrentes. Essa avaliação é feita por meio de critérios que medem a qualidade do ajuste (relação entre predito e observado), e as estimativas dos erros (IQBAL et al., 2022).

As métricas utilizadas para avaliar os modelos variam de acordo com a natureza do problema, regressão ou classificação, assim sendo, esta revisão se concentrou nas principais métricas utilizadas para avaliar os algoritmos de regressão.

#### **Mean absolute error (MAE)**

O erro médio absoluto mede a diferença absoluta entre os valores preditos pelo modelo e os valores observados no banco de dados. Quanto menor for a diferença entre o predito e o observado, melhor será o modelo, logo, quanto menor o valor de MAE, melhor. O MAE é considerado mais robusto, comparado ao RMSE (próxima métrica), quando se tem *outliers* no banco de dados. Essa diferença se dá, pois, o MAE trabalha com a diferença absoluta dos valores, já no RMSE esses valores são elevados ao quadrado, o que acaba por elevar o resultado final.

O MAE pode ser calculado pela seguinte fórmula:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - p_i|$$

Onde:

- $n$  é o número de amostras;
- $y_i$  é o valor observado;
- $p_i$  é o valor predito.

### Root Mean Squared Error (RMSE)

A raiz do erro quadrático médio, assim como o MAE, é uma forma de medir o desvio dos valores observados e preditos. Entretanto, por sua forma de cálculo, essa métrica é mais sensível a valores mais extremos, isso porque considera o quadrado dos erros. Se os erros, ponto a ponto, forem pequenos, o quadrado destes não será tão significativo, mas se esses forem muito grandes, ao serem elevados ao quadrado a diferença será maior no resultado final da métrica.

O RMSE pode ser calculado pela seguinte fórmula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2}$$

Onde:

- $n$  é o número de amostras;
- $y_i$  é o valor observado;
- $p_i$  é o valor predito.

### Coeficiente de determinação ( $R^2$ )

O coeficiente de determinação é usado para explicar o quanto da variável dependente é explicada pelo modelo de regressão. O valor do  $R^2$  varia de 0 a 1, sendo que valores mais próximos de 1 indicam um melhor ajuste do modelo, enquanto que os valores próximos a 0 mostram que o modelo não é tão bom para explicar a variabilidade da variável resposta.

O  $R^2$  pode ser calculado pela seguinte fórmula:

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

Onde:

- $n$  é o número de amostras;
- $y_i$  é o valor observado;
- $\hat{y}_i$  é o valor predito;
- $\bar{y}$  representa o valor médio das amostras.

### 3. MATERIAL E MÉTODOS

O estudo foi conduzido no setor de caprinocultura da Universidade Federal de Viçosa - UFV, localizada no município de Viçosa, Minas Gerais, Brasil. O projeto foi aprovado pela Comissão de Ética no Uso de Animais de Produção (CEUAP), da UFV, sob o protocolo 60/23. Foram utilizadas 186 fêmeas caprinas, sendo 98 da raça Saanen, e 88 da raça Alpina. Os animais estavam divididos nas seguintes categorias: cria (n= 62, idade 1-3 meses) (Tabela 1), recria (n= 24, idade 4-14 meses) (Tabela 2) e lactante/não gestante (primíparas e múltíparas) (n= 100, idade >15 meses) (Tabela 3).

Os animais foram criados em regime intensivo, em baias coletivas, com acesso ao solário (exceto as crias), além de serem equipadas com comedouro, bebedouro (água *ad libitum*), saleiro e piso com cobertura de cepilho de madeira. A alimentação consistiu de concentrado e volumoso (silagem de milho), balanceada para atender às exigências de cada categoria.

**Tabela 1:** Resumo do banco de dados da fase de cria.

	<b>MÉDIA (cm)</b>	<b>MÍN. (cm)</b>	<b>MÁX. (cm)</b>	<b>SD (cm)</b>
<b>ACIMG</b>	48,58	37,23	59,30	4,19
<b>AGIMG</b>	49,74	38,20	65,02	4,63
<b>LGIMG</b>	10,05	7,35	13,06	1,36
<b>CGIMG</b>	12,38	7,84	17,42	2,57
<b>CCIMG</b>	45,76	35,30	56,07	3,50
<b>PIMG</b>	22,40	17,24	26,67	1,71
<b>LPIMG</b>	13,00	8,94	18,77	2,26
<b>APIMG</b>	27,28	18,46	35,71	3,68

SD= desvio padrão. AC= altura da cernelha; AG= altura da garupa; LG= largura da garupa; CG= comprimento da garupa; CC= comprimento corporal; P= profundidade; LP= largura do peito; e AP= altura da pata. Variáveis seguidas por IMG representam as medidas por imagens

**Tabela 2:** Resumo do banco de dados da fase de recria.

	<b>MÉDIA (cm)</b>	<b>MÍN. (cm)</b>	<b>MÁX. (cm)</b>	<b>SD (cm)</b>
<b>ACIMG</b>	76,71	67,08	85,88	5,94

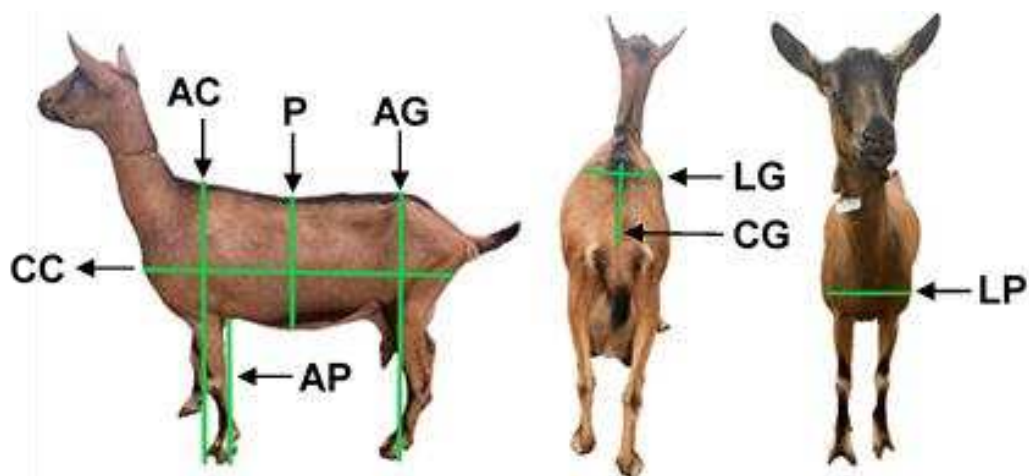
<b>AGIMG</b>	78,93	69,90	88,70	5,55
<b>LGIMG</b>	21,21	16,83	26,41	2,61
<b>CGIMG</b>	26,79	20,06	32,30	3,50
<b>CCIMG</b>	71,45	55,74	83,74	7,53
<b>PIMG</b>	34,76	27,68	39,04	2,78
<b>LPIMG</b>	21,78	17,44	25,55	2,16
<b>APIMG</b>	43,81	38,67	49,72	3,54

SD= desvio padrão. AC= altura da cernelha; AG= altura da garupa; LG= largura da garupa; CG= comprimento da garupa; CC= comprimento corporal; P= profundidade; LP= largura do peito; e AP= altura da pata. Variáveis seguidas por IMG representam as medidas por imagens.

**Tabela 3:** Resumo do banco de dados da fase de lactação.

	<b>MÉDIA (cm)</b>	<b>MÍN. (cm)</b>	<b>MÁX. (cm)</b>	<b>SD (cm)</b>
<b>ACIMG</b>	88,8	78,8	101,5	4,6
<b>AGIMG</b>	90,9	79,5	102,2	4,9
<b>LGIMG</b>	24,6	17,7	30,9	2,6
<b>CGIMG</b>	29,1	18,7	40,0	4,1
<b>CCIMG</b>	90,0	73,1	109,4	6,2
<b>PIMG</b>	43,0	35,1	55,0	3,6
<b>LPIMG</b>	26,5	21,7	34,8	2,6
<b>APIMG</b>	47,6	40,1	55,0	3,1

SD= desvio padrão. AC= altura da cernelha; AG= altura da garupa; LG= largura da garupa; CG= comprimento da garupa; CC= comprimento corporal; P= profundidade; LP= largura do peito; e AP= altura da pata. Variáveis seguidas por IMG representam as medidas por imagens.

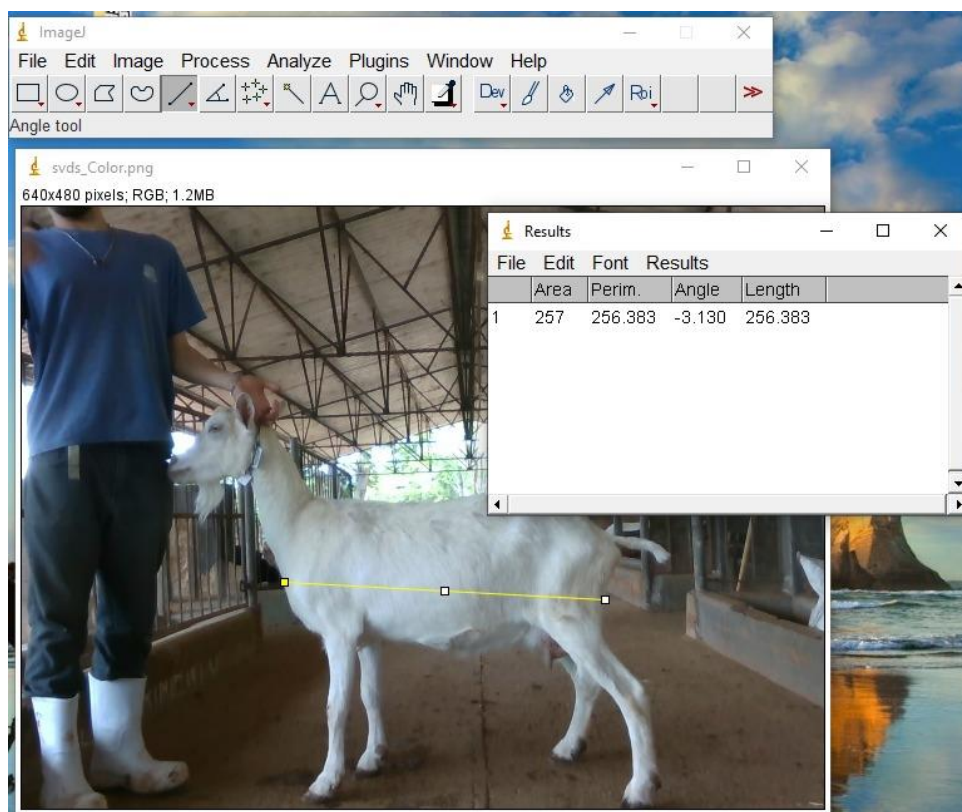


**Figura 2:** Representação dos locais de medição.

AC= altura da cernelha; AG= altura da garupa; LG= largura da garupa; CG= comprimento da garupa; CC= comprimento corporal; P= profundidade; LP= largura do peito; e AP= altura da pata.

Simultaneamente a pesagem foram realizadas as coletas de imagens bidimensionais (2D), sob iluminação natural. Para isso utilizou-se um tripé profissional como auxílio e uma câmera Intel® RealSense™ Depth Câmera (D435) para condução das gravações individuais das seguintes regiões: lateral, frente (esterno) e superior (dorso e garupa). A câmera estava disposta em duas alturas em relação ao solo, a altura foi de 67 cm, para coletar as imagens da lateral e de frente, enquanto para a região superior estava posicionada a 1,35 m. Além disso, a distância da câmera até o animal era de 1,80 m, e para demarcar essa distância foi utilizado um giz como auxílio.

Assim, foram obtidos vídeos de 1 min/animal e em seguida no *software* da câmera profissional, foram realizados *frames* das gravações para obtenção das melhores imagens. Posteriormente, estas foram processadas no *software* ImageJ para obtenção dos valores em pixels de comprimento (ABRÀMOFF; MAGALHÃES; RAM, 2004) dos parâmetros morfométricos (Figura 3). As imagens apresentaram 640 x 480 *pixels* e foram salvas em formato PNG. No imageJ, a ferramenta *Straight* foi utilizada para mensurar medidas morfológicas. Um objeto de medida conhecida foi posto nas imagens para que posteriormente fossem feitos os cálculos de transformação de *pixels* para centímetros. Posteriormente a coleta dos vídeos, foi aferido o peso vivo dos animais, sem jejum prévio.



**Figura 3:** Exemplo do processamento para extração das medidas.

As medidas morfológicas avaliadas na imagem foram: altura da cernelha (AC): medida do ponto mais elevado da escápula até o solo; altura da garupa (AG): medida do ponto mais alto da garupa até o solo; comprimento corporal (CC): medido entre o ponto do ombro anterior até o ponto de inserção da cauda; profundidade (P): medida da linha dorsal até o ponto mais profundo do tórax; altura da pata (AP): medida do ponto mais alto da pata dianteira até o solo; largura do peito (LP): medida do ponto mais largo do peito e seu correspondente no lado oposto; largura da garupa (LG): medida entre os pontos onde a garupa é mais larga; e comprimento da garupa (CG): medido do ponto frontal até o ponto mais posterior da garupa (adaptado de American Dairy Goat Association, 2022).

Para as análises estatísticas, foi realizada uma análise de correlação de Pearson, a 5% de significância, entre as variáveis e o peso vivo. Para as predições do peso corporal o banco de dados foi dividido, de forma ponderada por categoria, em 70% de treinamento e 30% de teste. Foram utilizados os algoritmos Ridge, Elastic Net, Lasso e Random Forest (RF), com as *features* de morfologia por imagem. O modelo foi otimizado selecionando variáveis com base na redução do erro quadrático médio (RMSE) pelo método de validação cruzada (5 k-folds).

Nesse método, os dados foram divididos em cinco subconjuntos distintos e o modelo então foi treinado e testado repetidamente, alternando quais subconjuntos foram utilizados para treinamento e teste. Isso permitiu uma avaliação ampla do desempenho do modelo em diferentes conjuntos de dados, proporcionando aumentar a confiabilidade dos resultados e redução do overfitting.

Após os ajustes a precisão e acurácia do modelo foram obtidas por interpretação do coeficiente de determinação ( $R^2$ ), RMSE e erro médio absoluto (MAE).

### **Mean absolute error (MAE)**

O erro médio absoluto mede a diferença absoluta entre os valores preditos pelo modelo e os valores observados no banco de dados. Quanto menor for a diferença entre o predito e o observado, melhor será o modelo, logo, quanto menor o valor de MAE, melhor. O MAE pode ser calculado pela seguinte fórmula:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - p_i|$$

Onde,  $n$  é o número de amostras,  $y_i$  é o valor observado e  $p_i$  é o valor predito.

### **Root Mean Squared Error (RMSE)**

A raiz do erro quadrático médio, assim como o MAE, é uma forma de medir o desvio dos valores observados e preditos. Entretanto, por sua forma de cálculo, essa métrica é mais sensível a valores mais extremos, isso porque considera o quadrado dos erros. O RMSE pode ser calculado pela seguinte fórmula:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2}$$

Onde  $n$  é o número de amostras,  $y_i$  é o valor observado e  $p_i$  é o valor predito.

### **Coeficiente de determinação ( $R^2$ )**

O coeficiente de determinação é usado para explicar o quanto da variável dependente é explicada pelo modelo de regressão. O valor do  $R^2$  varia de 0 a 1, sendo

que valores mais próximos de 1 indicam um melhor ajuste do modelo, enquanto que os valores próximos a 0 mostram que o modelo não é tão bom para explicar a variabilidade da variável resposta. O  $R^2$  pode ser calculado pela seguinte fórmula:

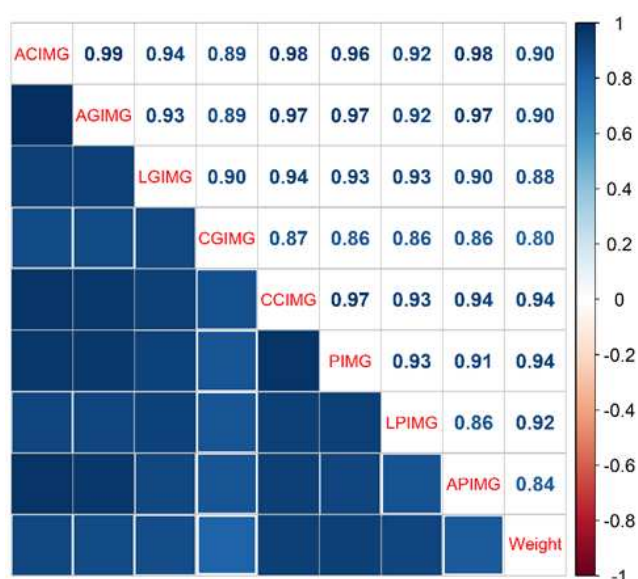
$$R^2 = 1 - \frac{\text{Variância Residual}}{\text{Variância Total}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Onde  $n$  é o número de amostras,  $y_i$  é o valor observado,  $\hat{y}_i$  é o valor predito e  $\bar{y}$  representa o valor médio das amostras.

#### 4. RESULTADOS E DISCUSSÃO

Todas as medidas morfológicas por imagem estiveram altamente correlacionadas ( $r > 80$ ) com o peso corporal (Figura 3). As características de maior correlação foram a profundidade ( $r = 0,94$ ) e o comprimento corporal ( $r = 0,94$ ), seguidas pela largura do peito ( $r = 0,92$ ), altura de cernelha e altura de garupa ( $r = 0,90$ ). O que converge com os resultados de correlações encontrados por Berhe (2017), para a altura de cernelha, de garupa e comprimento corporal ( $r = 0,93$ ,  $r = 0,68$  e  $r = 0,90$ , respectivamente) de caprinos da raça Maefur. O autor ainda concluiu seu estudo chamando a atenção para a importância destas medidas dentro de um programa de seleção no rebanho.

Trabalhos na literatura têm apontado o perímetro torácico como uma das características mais correlacionadas com o peso corporal de caprinos da raça Canindé ( $r = 0,69$ ) (CONRADO; ARANDAS; RIBEIRO, 2015b), da raça Saburai ( $r = 0,95$ ) (DAKH-LAN et al., 2021) e ovinos Morada Nova ( $r = 0,92$  para fêmeas e  $r = 0,95$  para machos) (SOUZA JUNIOR, 2023). Embora neste trabalho essa característica não tenha sido avaliada, os resultados revelaram a importância de duas outras, a profundidade ( $r = 0,94$ ) e largura do peito ( $r = 0,92$ ). Essa observação desperta a atenção para uma possível equivalência ou associação entre essas duas e o perímetro torácico, o que pode ser útil em casos em que este último não possa ser mensurado.

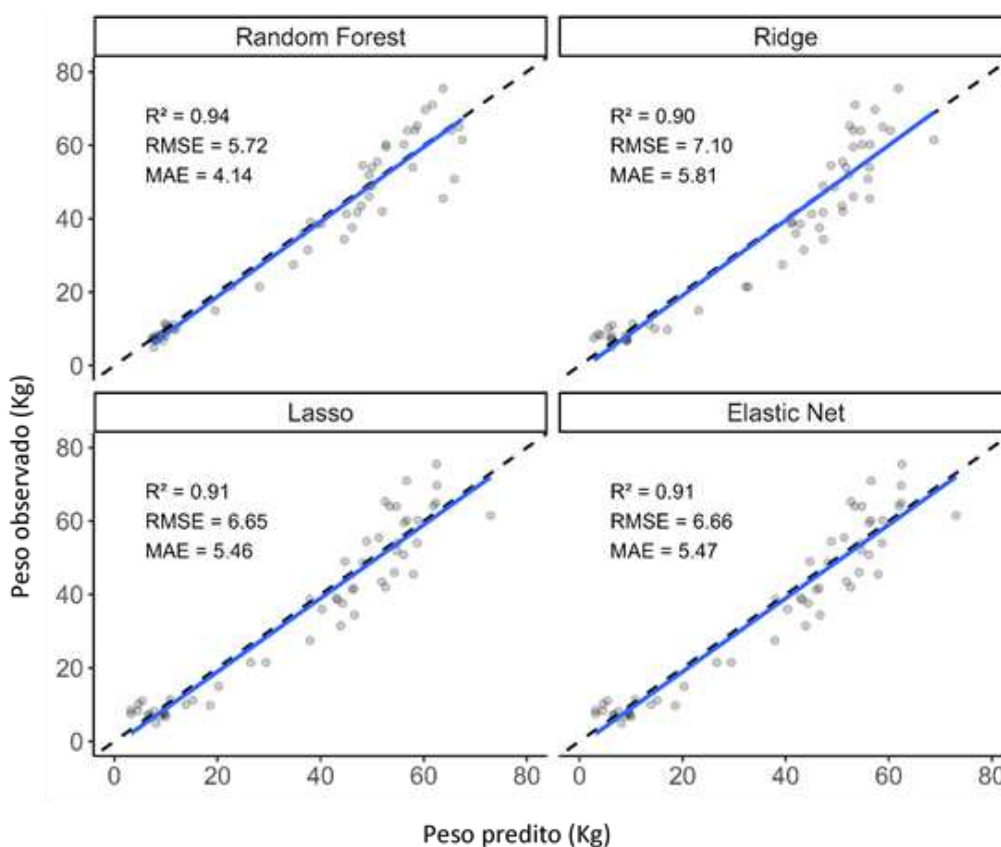


**Figura 4:** Correlação entre as morfometrias e o peso.

AC= altura da cernelha; AG= altura da garupa; LG= largura da garupa; CG= comprimento da garupa; CC= comprimento corporal; P= profundidade; LP= largura do peito; e AP= altura da pata. A sigla IMG refere-se a variável medida na imagem

O comprimento de garupa ( $r = 0,80$ ) e a largura de garupa ( $r = 0,88$ ), se mostraram bem correlacionadas com o peso corporal (Figura 3), embora tenham tido menor importância na predição (Figura 6). Costa et al., (2020), trabalhando com caprinos Saanen e Anglo Nubiana, encontraram uma elevada correlação da largura de garupa ( $r = 0,92$ ) para as fêmeas. Valores elevados de correlação entre comprimento de garupa e peso foram encontrados, também, para novilhas ( $r = 0,90$ ) e machos ( $r = 0,82$ ), mas com baixa correlação em vacas ( $r = 0,19$ ) (REIS et al., 2008). Entretanto, contrastando com os resultados, um trabalho recente com ovelhas mostrou que a largura de garupa foi a que menos se correlacionou ( $r = 0,37$ ) com o peso (SOUZA JUNIOR, 2023).

Com base nos gráficos de correlação entre predito e observado (Figura 5), é possível dizer que os modelos de *machine learning* estudados foram capazes de prever o peso das cabras com elevada precisão ( $R^2 > 0,90$ ). Entretanto, observa-se certa superioridade do Random Forest ( $R^2 = 0,94$ ; RMSE = 5,72 e MAE = 4,14) em relação aos demais, não apenas pelo maior coeficiente de determinação, mas, também, pelas métricas que estimam os erros associados (Figura 5).



**Figura 5:** Relação entre os pesos preditos e observados pelos algoritmos.  $R^2$ = coeficiente de determinação, RMSE= raiz do erro quadrático médio e MAE= erro médio absoluto

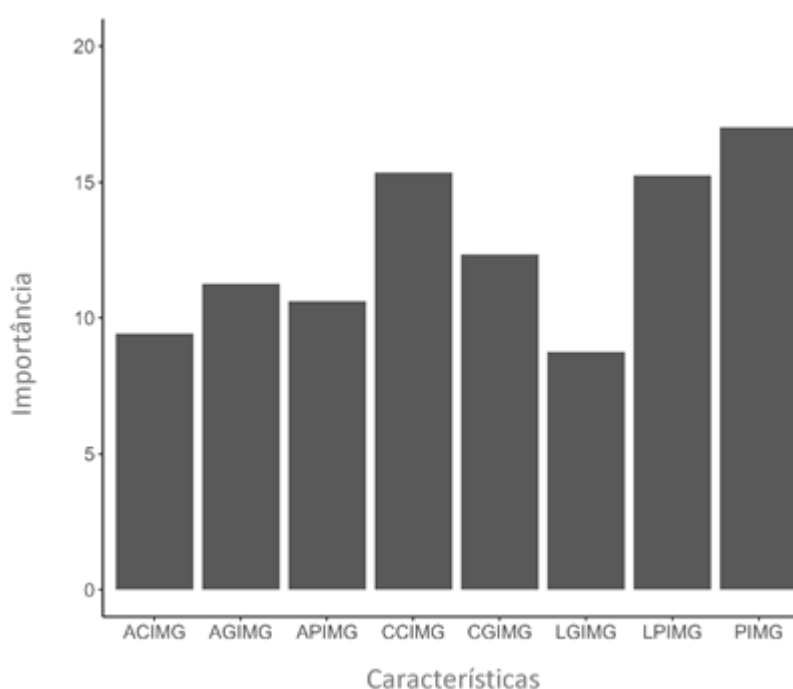
O Lasso, Elastic Net, e o Ridge são algoritmos de regressão penalizada, ou seja, são modelos que adicionam uma penalidade, que vai variar de acordo com o algoritmo escolhido, à função de regressão. Esse processo de penalização dos parâmetros tem como vantagens a redução da ocorrência de um sobreajuste aos dados (overfitting), a seleção das variáveis mais importantes para o modelo, descartando as de menor relevância e a redução da complexidade do modelo, algo que tem despertado a atenção dos pesquisadores por esclarecer melhor as relações entre a resposta e as variáveis estudadas (ZOU; HASTIE, 2005).

O Random Forest, por sua vez, é um algoritmo poderoso que trabalha com a construção de uma floresta aleatória com diversas árvores de decisão independentes, cujos resultados de cada árvore colaboram para a geração de um resultado final preciso (LORENZETT; TELÖCKEN, 2016). Este tem se mostrado com melhor desempenho quando comparado com os algoritmos de regressão penalizada (Ridge, Lasso e Elastic Net) (SARKAR et al., 2015), muito embora demande maior poder computacional. Ainda neste trabalho os autores chamaram atenção para a flexibilidade e robustez, do Random Forest, quanto ao ruído e a ocorrência de overfitting.

Resultados semelhantes aos deste estudo foram encontrados por Çakmakçı (2022), em um estudo de predição do peso de ovelhas, no qual o Random Forest obteve MAE = 3,66 e RMSE = 4,47, valores menores quando comparados com o Support Vector Machines com Radial Basis Function Kernel (MAE = 3,99 e RMSE = 5,07), Classification and Regression Trees (MAE = 4,97 e RMSE = 5,69) e o Model Average Neural Networks (MAE = 4,06 e RMSE = 4,97). Ainda em um trabalho recente com ovinos, o Random Forest ( $R^2 = 0,735$ ) foi capaz de prever o peso com maior precisão no conjunto de teste quando comparado com Support Vector Machines ( $R^2 = 0,714$ ) e o Classification and Regression Trees ( $R^2 = 0,578$ ) (TIRINK et al., 2023).

Iqbal et al., (2022), trabalhando com cabras Beetal, encontraram melhores resultados para o coeficiente de determinação e as menores estimativas dos erros para os algoritmos Random Forest ( $R^2 = 0,785$ ; MAE = 2,784 e RMSE = 3,583) e o Gradient Boosting Machine ( $R^2 = 0,794$ ; MAE = 2,866 e RMSE = 3,536), em comparação com Regression Trees ( $R^2 = 0,770$ ; MAE = 2,909 e RMSE = 3,721), Support Vector Machine ( $R^2 = 0,778$ ; MAE = 2,787 e RMSE = 3,695). Ressalta-se, porém, a escassez de trabalhos com caprinos, testando estes algoritmos a partir de dados oriundos de imagens, o que dificulta a comparação dos resultados.

De acordo o gráfico de importância das características (*features*) para o Random Forest (Figura 6), três das oito apresentaram maior peso para o algoritmo, sendo estas a profundidade (17,01%), o comprimento corporal (15,35%) e a largura do peito (15,23%). Resultados semelhantes, foram encontrados em um trabalho com ovelhas, no qual, para o Random Forest, as três características mais importantes em ordem decrescente foram a largura do peito, profundidade do peito e o comprimento corporal (ÇAKMAKÇI, 2022). A largura de garupa e a altura de cernelha foram as morfometrias que mostraram-se de menor relevância na predição, com 8,76% e 9,42%, respectivamente.



**Figura 6:** Importância das características para o Random Forest.

AC= altura da cernelha; AG= altura da garupa; LG= largura da garupa; CG= comprimento da garupa; CC= comprimento corporal; P= profundidade; LP= largura do peito; e AP= altura da pata. A sigla IMG refere-se a variável medida na imagem.

## 5. CONCLUSÃO

Todas as características avaliadas por imagens estiveram altamente correlacionadas com o peso corporal. É possível prever o peso de caprinos leiteiros por meio atributos extraídos de imagens digitais e o Random Forest foi o algoritmo que apresentou o melhor desempenho quando comparado com os com as regressões Lasso, Ridge e Elastic Net.

## 7. REFERÊNCIAS

- ABRÀMOFF, M. D.; MAGALHÃES, P. J.; RAM, S. J. Image processing with ImageJ. **Biophotonics International**, v. 11, n. 7, p. 36-42, 2004.
- BERHE, W. G. Relationship and Prediction of Body Weight from Morphometric Traits in Maefur Goat Population in Tigray, Northern Ethiopia. **Journal of Biometrics & Biostatistics**, v. 8, n. 5, p. 1–6, 2017.
- ÇAKMAKÇI, C. Live Weight Prediction in Norduz Sheep Using Machine Learning Algorithms. **Turkish Journal of Agriculture - Food Science and Technology**, v. 10, p. 587–594, 5 maio 2022.
- CELIBERTO, L. et al. **Quantificação de regiões cerebrais utilizando processamento de imagens**. 1 jan. 2008.
- CONRADO, V. C.; ARANDAS, J. G.; RIBEIRO, M. Modelos de regressão para predição do peso da raça Canindé através de medidas morfométricas. **Archivos de Zootecnia**, v. 64, n. 247, p. 277–280, 16 set. 2015.
- COSTA, E. DE O. et al. Medidas morfométricas associadas à predição matemática do peso corporal como ferramenta de manejo em rebanhos de caprinos. **Archivos de Zootecnia**, v. 69, n. 268, p. 454–460, 15 out. 2020.
- DAKHLAN, A. et al. Short Communication: Prediction of body weight based on body measurements in female Saburai goat. **Biodiversitas Journal of Biological Diversity**, v. 22, n. 3, 24 fev. 2021.
- FERNANDES, A. F. A.; DÓREA, J. R. R.; ROSA, G. J. DE M. Image Analysis and Computer Vision Applications in Animal Sciences: An Overview. **Frontiers in Veterinary Science**, v. 7, 21 out. 2020.
- FERREIRA, O. G. L. et al. Measurement of rib-eye area by the method of digital images. **Revista Brasileira de Zootecnia**, v. 41, p. 811–814, mar. 2012.
- FERREIRA, O. G. L. et al. Determinação de área foliar por análise de imagens digitais. **Archivos de Zootecnia**, v. 66, n. 256, p. 593–597, 15 out. 2017.
- GAUDIOSO, V. et al. “Photozoometer”: A new photogrammetric system for obtaining morphometric measurements of elusive animals. **Livestock Science**, v. 165, p. 147–156, 1 jul. 2014.
- GERÓN, A. **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn, Keras & TensorFlow: Conceitos, Ferramentas e Técnicas para a Construção de Sistemas Inteligentes**. Alta Books, 2021.
- GONZALEZ, R.C; WOODS, R. E. **Processamento de imagens digitais**. São Paulo: Edgard Blucher – 2000.
- HAMADANI, A.; GANAI, N. A. Artificial intelligence algorithm comparison and ranking for weight prediction in sheep. **Scientific Reports**, v. 13, n. 1, p. 13242, 15 ago. 2023.
- IQBAL, F. et al. Comparing the Predictive Ability of Machine Learning Methods in Predicting the Live Body Weight of Beetal Goats of Pakistan. **Pakistan Journal of Zoology**, v. 54, p. 231–238, 1 mar. 2022.

JO, K. et al. Hyperspectral imaging–based assessment of fresh meat quality: Progress and applications. **Microchemical Journal**, v. 197, p. 109785, 1 fev. 2024.

KER, J. et al. Deep Learning Applications in Medical Image Analysis. **IEEE Access**, v. 6, p. 9375–9389, 2018.

LEITE, D. R. A.; MORAES, R. M. DE; LOPES, L. W. Método de Aprendizagem de Máquina para Classificação da intensidade do desvio vocal utilizando Random Forest. **Journal of Health Informatics**, v. 12, 2020.

LI, K.; TENG, G. Study on Body Size Measurement Method of Goat and Cattle under Different Background Based on Deep Learning. **Electronics**, v. 11, n. 7, p. 993, jan. 2022.

LI, W.; HSU, C.-Y. GeoAI for Large-Scale Image Analysis and Machine Vision: Recent Progress of Artificial Intelligence in Geography. **ISPRS International Journal of Geo-Information**, v. 11, n. 7, p. 385, jul. 2022.

ZHANG, A. L. et al. Algorithm of sheep body dimension measurement and its applications based on image analysis. **Computers and Electronics in Agriculture**, v. 153, p. 33–45, 1 out. 2018.

LORENZETT, CASSIO DAL CASTEL; TELOKEN, A. **Estudo Comparativo entre os algoritmos de Mineração de Dados Random Forest e J48 na tomada de Decisão**. 11 maio 2016. Disponível em: <<https://www.semanticscholar.org/paper/Estudo-Comparativo-entre-os-algoritmos-de-Minera%C3%A7%C3%A3o-lorenzett-Teloken/e7b51ebfffebe2ccc5e90bc663cd1ee06932959e>>. Acesso em: 6 maio. 2024

LUDERMIR, T. B. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, v. 35, p. 85–94, 19 abr. 2021.

MACHADO, A. W.; SOUKI, B. Q. Simplificando a obtenção e a utilização de imagens digitais: scanners e câmeras digitais. **Revista Dental Press de Ortodontia e Ortopedia Facial**, v. 9, p. 133–156, ago. 2004.

MÜLLER, H. et al. A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. **International Journal of Medical Informatics**, v. 73, n. 1, p. 1–23, 1 fev. 2004.

PANDORFI, H.; ALMEIDA, G. L. P.; GUISELINI, C. Zootecnia de precisão: princípios básicos e atualidades na suinocultura. **Revista Brasileira de Saúde e Produção Animal**, v. 13, p. 558–568, jun. 2012.

QUEIROZ, J.; GOMES, H. Introdução ao Processamento Digital de Imagens. **RITA**, v. 13, p. 11–42, 1 jan. 2006.

REIS, G. L. et al. Predição do peso vivo a partir de medidas corporais em animais mestiços Holandês/Gir. **Ciência Rural**, v. 38, p. 778–783, jun. 2008.

RIBEIRO, S. D. de A. **Caprinocultura: Criação Racional de Caprinos** / Silvio Doria de Almeida Ribeiro. – São Paulo: Nobel, 1997.

SANT'ANA, D. A. et al. Weighing live sheep using computer vision techniques and regression machine learning. **Machine Learning with Applications**, v. 5, p. 100076, 15 set. 2021.

SARKAR, R. K. et al. Evaluation of random forest regression for prediction of breeding value from genomewide SNPs. **Journal of Genetics**, v. 94, n. 2, p. 187–192, jun. 2015.

SAXENA, L.; ARMSTRONG, L. A Survey of Image Processing Techniques for Agriculture; **Australian Society of Information and Communication Technologies in Agriculture**: Perth, Australia, 2014.

SOUSA JUNIOR, Silvio de. **Predição do peso de ovinos da raça Morada Nova a partir de medidas corporais**. 2023. 24 f. Trabalho de Conclusão de Curso (Graduação em Zootecnia) – Centro de Ciências Agrárias, Universidade Federal do Ceará, Fortaleza, 2023.

TIRINK, C. et al. Estimation of Body Weight Based on Biometric Measurements by Using Random Forest Regression, Support Vector Regression and CART Algorithms. **Animals: an open access journal from MDPI**, v. 13, n. 5, p. 798, 22 fev. 2023.

ZOU, H.; HASTIE, T. Regularization and Variable Selection Via the Elastic Net. **Journal of the Royal Statistical Society Series B: Statistical Methodology**, v. 67, n. 2, p. 301–320, 1 abr. 2005.