

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DO
PESO CORPORAL E DE CARCAÇA QUENTE DE
BOVINOS NELORE**

Alexandre Cominotte

Zootecnista

2018

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DO
PESO CORPORAL E DE CARCAÇA QUENTE DE
BOVINOS NELORE**

Alexandre Cominotte

Orientador: Prof. Dr. Otávio Rodrigues Machado Neto

Dissertação apresentada a Faculdade de Ciências Agrárias e Veterinárias - UNESP, Campus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Zootecnia.

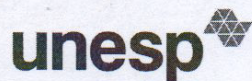
2018

C733u Cominotte, Alexandre
Uso de imagens biométricas para predição do peso corporal e de
carcaça quente de bovinos Nelore / Alexandre Cominotte. – –
Jaboticabal, 2018
vii, 73 p. : il. ; 29 cm

Dissertação (mestrado) - Universidade Estadual Paulista,
Faculdade de Ciências Agrárias e Veterinárias, 2018
Orientador: Otávio Rodrigues Machado Neto
Banca examinadora: Guilherme Luis Pereira, João Ricardo
Rebouças Dórea
Bibliografia

1. Visão computacional. 2. Bovinos de corte. 3. Predição peso
corporal. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e
Veterinárias.

CDU 636.2



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DO PESO CORPORAL E DE CARÇA QUENTE DE BOVINOS NELORE

AUTOR: ALEXANDRE COMINOTTE

ORIENTADOR: OTAVIO RODRIGUES MACHADO NETO

Aprovado como parte das exigências para obtenção do Título de Mestre em ZOOTECNIA, pela Comissão Examinadora:

Prof. Dr. OTAVIO RODRIGUES MACHADO NETO
Departamento de Produção Animal-FMVZ/UNESP / Botucatu/SP

Prof. Dr. JOÃO RICARDO REBOUÇAS DÓREA (Skype)
Department of Animal Science/University of Wisconsin / Madison/WI

Prof. Dr. GUILHERME LUIS PEREIRA
Departamento de Melhoramento e Nutrição Animal-FMVZ/UNESP / Botucatu/SP

Jaboticabal, 13 de agosto de 2018

DADOS CURRICULARES DO AUTOR

ALEXANDRE COMINOTTE – filho de Carlos Alberto Cominotte e Kelly Kassia Ferreira Cominotte, nascido em Araraquara, São Carlos, no dia 17 de outubro de 1991. Em fevereiro de 2011 ingressou no curso de Graduação em Zootecnia, na Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, *Campus* Botucatu, sendo bolsista de Treinamento Técnico (TT-1) da Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, sob orientação do Prof. Dr. André Mendes Jorge e bolsista de Iniciação Científica financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, sob orientação do Prof. Dr. Otávio Rodrigues Machado Neto, graduando em dezembro de 2015. Em março de 2006 iniciou o mestrado no programa de Pós-Graduação em Zootecnia da Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, *Campus* de Jaboticabal, onde é bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, sob orientação do Prof. Dr. Otávio Rodrigues Machado Neto. Em 2017 foi bolsista de Bolsa Estágio de Pesquisa no Exterior (BEPE) da Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, na University of Wisconsin – Madison, Estados Unidos sob orientação do Prof. Dr. Guilherme Jordão de Magalhães Rosa e supervisor Prof. Dr. Otávio Rodrigues Machado Neto.

Chalana

Lá vai uma chalana
Bem longe se vai
Navegando no remanso
Do rio Paraguai

Oh! Chalana sem querer
Tu aumentas minha dor
Nessas águas tão serenas
Vai levando meu amor

Oh! Chalana sem querer
Tu aumentas minha dor
Nessas águas tão serenas
Vai levando meu amor

E assim ela se foi
Nem de mim se despediu
A chalana vai sumindo
Na curva lá do rio

E se ela vai magoada
Eu bem sei que tem razão
Fui ingrato, eu feri
O seu pobre coração

Oh! Chalana sem querer
Tu aumentas minha dor
Nessas águas tão serenas
Vai levando meu amor

Oh! Chalana sem querer
Tu aumentas minha dor
Nessas águas tão serenas
Vai levando meu amor

DEDICO...

Ao meu pai Carlos Alberto Cominote e minha Mãe Kelly Kassia Ferreira Cominotte, pelos ensinamentos, por cuidar de mim e me instruir desde o meu nascimento. Pelos exemplos diários de caráter, respeito, honestidade, sabedoria, por acreditarem nas minhas escolhas e por fazerem dos meus sonhos os seus.

À minha irmã Mariana Aline Cominotte, por me fazer ter acreditado na carreira acadêmica, respeito, amor e principalmente companheirismo.

AGRADECIMENTOS

Agradeço em primeiro lugar a Deus pela saúde e pela família que me proporcionou.

Ao meu orientador, Prof. Dr. Otávio Rodrigues Machado Neto, pela confiança, amizade, ensinamentos e oportunidade de desenvolvimento pessoal e profissional.

Ao Prof. Dr. Guilherme Jordão de Magalhães Rosa pela co-orientação, amizade, incentivo e todo esforço para a elaboração, realização e conclusão deste projeto.

Ao Dr. João Ricardo Rebouças Dórea pela co-orientação e toda ajuda na elaboração e realização deste projeto, além da amizade, ensinamentos e disponibilidade sempre que necessário.

Aos meus pais Carlos e Kelly pelo apoio e incentivo nas mais difíceis decisões. À minha irmã Mariana pelo carinho de todos os momentos.

Aos colegas Arthur Fernandes, Fernanda, Maxwelder, Maria Jaqueline, Leonardo, Samuel, Leticia Custódio, Lucio Mota, Paloma e Wendel pela parceria durante todo o experimento e na realização das análises.

A Agência Paulista de Tecnologia dos Agronegócios (APTA) – Colina, pela ajuda e credibilidade em mim depositada para a condução deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

À Fundação de Amparo a Pesquisa do Estado de São Paulo, deixando claro meu agradecimento a todos os apoiadores e contribuintes da realização dos referentes projetos (FAPESP nº 2017/02057-0 e 2017/20812-0).

A todos, muito obrigado!

SUMÁRIO

RESUMO	iii
ABSTRACT	v
LISTA DE ABREVIACÕES	vii
CAPÍTULO 1 - CONSIDERAÇÕES GERAIS	1
1. INTRODUÇÃO.....	1
2. REVISÃO DA LITERATURA.....	2
2.1. Análise de imagens	2
2.2. Regressão Linear Múltipla	5
2.3. Regressão LASSO	6
2.4. Regressão <i>Partial Least Squares</i>	7
2.5. <i>Artificial Neural Networks</i>	8
2.6. Validação Cruzada <i>Leave-One-Out</i>	9
3. OBJETIVOS GERAIS.....	9
4. REFERÊNCIAS	10
 CAPÍTULO 2 - USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DE PESO CORPORAL E GANHO MÉDIO DIÁRIO DE BOVINOS DE CORTE NELORE	 14
RESUMO	14
1. INTRODUÇÃO.....	16
2. MATERIAL E MÉTODOS.....	17
2.1. Dados.....	18
2.2. Análise de imagens	20
2.2.1. Estimativa da distância do chão à câmera:.....	20
2.2.2. Segunda segmentação:	20
2.2.3. Extração de características:.....	21
2.3. Análise de dados	23
2.4. Predição de peso corporal e o ganho médio diário	23
2.5. Regressão linear múltipla	24
2.6. Regressão LASSO	25
2.7. Regressão <i>Partial Least Squares</i>	26
2.8. <i>Artificial Neural Network</i>	28
2.10. Avaliação do modelo.....	29
3. RESULTADOS E DISCUSSÃO.....	31
4. CONCLUSÃO.....	38
5. REFERÊNCIAS	39

CAPÍTULO 3 – USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DO PESO CORPORAL E PESO DE CARCAÇA QUENTE DE BOVINOS NELORE.....	43
RESUMO:	43
1. INTRODUÇÃO.....	45
2. MATERIAL E MÉTODOS.....	46
2.1. Conjunto de dados	46
2.2. Conjuntos Experimentais.....	48
2.3. Dispositivo Kinect®	48
2.4. Análise de imagens	49
2.4.1. Estimativa da distância do chão à câmera:.....	49
2.4.2. Segunda segmentação:	49
2.4.3. Extração de características:.....	52
2.5. Análise de dados	54
2.6. Predição de peso corporal e peso de carcaça quente.....	57
2.7. Regressão linear múltipla	57
2.8. Regressão LASSO	57
2.9. Regressão <i>Partial Least Squares</i>	58
2.10. <i>Artificial Neural Networks</i>	61
2.11. Avaliação do modelo.....	61
3. RESULTADOS E DISCUSSÃO.....	64
4. CONCLUSÃO.....	69
5. REFERÊNCIAS	70

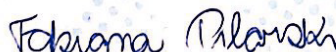
CEUA – COMISSÃO DE ÉTICA NO USO DE ANIMAIS

CERTIFICADO

Certificamos que o projeto de pesquisa intitulado “Uso de escores de avaliação visual e de imagens biométricas em bovinos Nelore”, protocolo nº 007946/18, sob a responsabilidade do Prof. Dr. Otávio Rodrigues Machado Neto, que envolve a produção, manutenção e/ou utilização de animais pertencentes ao Filo Chordata, subfilo Vertebrata (exceto o homem), para fins de pesquisa científica (ou ensino) - encontra-se de acordo com os preceitos da lei nº 11.794, de 08 de outubro de 2008, no decreto 6.899, de 15 de julho de 2009, e com as normas editadas pelo Conselho Nacional de Controle de Experimentação Animal (CONCEA), e foi aprovado pela COMISSÃO DE ÉTICA NO USO DE ANIMAIS (CEUA), da FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS, UNESP - CÂMPUS DE JABOTICABAL-SP, em reunião ordinária de 14 de junho de 2018.

Vigência do Projeto	16/06/2018 a 30/07/2018
Espécie / Linhagem	<i>Bos indicus</i> / Nelore
Nº de animais	243 animais
Peso / Idade	425 kg / 25 meses
Sexo	Macho
Origem	Fazenda do Pólo Regional de Alta Mogiana

Jaboticabal, 14 de junho de 2018.


Profª Drª Fabiana Pilarski
Coordenadora – CEUA

USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DO PESO CORPORAL E DE CARCAÇA QUENTE DE BOVINOS NELORE

RESUMO - O trabalho foi dividido em dois estudos. No estudo 1, objetivou-se prever o peso corporal (PC) e o ganho médio diário (GMD) de bovinos Nelore por meio de imagens tridimensionais e comparar quatro modelos de predição: Regressão Linear Múltipla (RLM), Regressão LASSO (LASSO), *Partial Least Squares* (PLS) e *Artificial Neural Networks* (ANN). Foram coletadas 234 imagens de bovinos Nelore. As coletas de dados foram realizadas em quatro fases da ao longo da vida do animal: Desmame aos 244 dias de idade e 202,3 kg ($\pm 27,1$), Recria aos 457 dias de idade e 213,9 kg ($\pm 25,1$), Início da Terminação aos 590 dias de idade e 334,5 kg ($\pm 29,2$) e Final da Terminação aos 763 dias de idade e 449,5 kg ($\pm 47,5$). Nas três primeiras fases foram coletadas imagens de 62 bovinos Nelore, enquanto que na última fase apenas 48 imagens foram coletadas. O GMD foi medido: 1: Desmama - Recria, 2: Desmama – Início da Terminação, 3: Desmama – Final da Terminação, 4: Recria – Início da Terminação, 5: Recria – Final da Terminação e 6: Início da Terminação – Final da Terminação. No estudo 2, quatrocentas e cinquenta imagens de bovinos Nelore foram coletadas em quatro experimentos para predição de PC e peso de carcaça quente (PCQ). Quatro conjuntos experimentais foram considerados: Set 1 inclui os experimentos 1, 2 e 3 para treinamento e experimento 4 para validação; Set 2 inclui os experimentos 1, 2 e 4 para treinamento e experimento 3 para validação; Set 3 inclui os experimentos 1, 3 e 4 para treinamento e experimento 2 para validação; Set 4 inclui os experimentos 2, 3 e 4 para treinamento e experimento 1 para validação. Os experimentos 1, 3 e 4 foram conduzidos no centro de pesquisa do Pólo Regional de Desenvolvimento Tecnológico dos Agronegócios da Alta Mogiana, em Colina - SP, Brasil. Para o experimento 1 foram utilizados 48 bovinos machos quimicamente castrados da raça Nelore, com média de 24 ± 2 meses e peso vivo médio de 449,4 ($\pm 46,9$ kg). Para o experimento 2 foram utilizados 228 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio inicial de 588,4 ($\pm 35,5$ kg). O experimento 3 foi realizado no Confinamento Experimental do Departamento de Melhoramento e Nutrição Animal da Faculdade de Medicina Veterinária e Zootecnia, UNESP – Câmpus de Botucatu em que foram utilizados 83 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio de 516,6 ($\pm 38,0$ kg). Para o experimento 4 foram utilizados 91 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio de 589,4 ($\pm 30,1$ kg). O sensor Kinect modelo 1473 da Microsoft foi utilizado para a captura das imagens de ambos estudos. Concomitantemente à aquisição de imagens, cada animal foi pesado em balança eletrônica. A análise das imagens foi dividida em três etapas: primeira, estimativa da distância do chão à câmera, segunda, uma etapa de segmentação e uma etapa final para extração das características. As características com correlação abaixo de $P < 0,05$ foram removidas do modelo. Assim para as análises estatísticas no estudo 1 foram utilizadas: Área e volume corporal, seis larguras e alturas ao longo do dorso do animal e comprimento. Para comparar os modelos avaliados, foi utilizado o método de validação cruzada Leave-One-Out. Para as análises estatísticas no estudo 2 foram utilizadas: Área e volume corporal, onze larguras e alturas ao longo do dorso do animal, comprimento, excentricidade e

duas medidas da curvatura da espinha dorsal. No estudo 1, o método de ANN aumentou a precisão e acurácia de PC comparado aos outros métodos: Desmama: (concordance correlation coefficient (CCC) = 0.94; root mean square error of prediction (RMSEP) = 8.6 kg; R^2 = 0.72). Recria: (CCC = 0.87; RMSEP = 11.4 kg; R^2 = 0.79). Início da Terminação: (CCC = 0.96; RMSEP = 7.7 kg; R^2 = 0.92). Assim, como para o PC, o uso do método ANN aumentou a precisão e acurácia das predições para GMD comparado aos outros modelos utilizados. Os resultados obtidos para foram: 1 (CCC = 0.74; RMSEP = 0.02 kg/d e R^2 = 0.67), 2 (CCC = 0.89; RMSEP = 0.02 kg/d e R^2 = 0.85), 3 (CCC = 0.84; RMSEP = 0.03 kg/d e R^2 = 0.80), 4 (CCC = 0.54; RMSEP = 0.10 kg/d e R^2 = 0.51), 5 (CCC = 0.82; RMSEP = 0.06 kg/d e R^2 = 0.76), 4 (CCC = 0.90; RMSEP = 0.09 kg/d e R^2 = 0.82). Para o estudo 2, o método preditivo de ANN para peso corporal obteve melhores resultados nos sets 1, 2, 3 e 4 comparados com PLS, LASSO e RLM (CCC = 0.73, 0.66, 0.70, 0.74; RMSEP = 19.6, 27.2, 27.2, 33.7 kg e R^2 = 0.58, 0.53, 0.53, 0.59, respectivamente). Já para peso de carcaça quente o método preditivo RLM apresentou melhores resultados para o Set 1 e 4, LASSO obteve resultados superiores para o Set 2 e para o Set 3 o método preditivo ANN apresentou melhores resultados. Set 1: (CCC = 0.14; RMSEP = 8.7 kg e R^2 = 0.05, respectivamente), Set 2 (CCC = 0.53; RMSEP = 20,1 kg, e R^2 = 0.40, respectivamente), Set 3 (CCC = 0.56; RMSEP = 20.1 kg e R^2 = 0.44, respectivamente), Set 4 (CCC = 0.69; RMSEP = 22.7 kg, e R^2 = 0.52, respectivamente). Este estudo indica que imagens tridimensionais obtidas a partir do sensor Kinect apresenta potencial para serem utilizadas para predizer peso corporal, ganho de peso e peso de carcaça quente em bovinos Nelore.

Palavras-chave: análise de imagens, bovinos de corte, visão computacional, modelos preditivos

USE OF BIOMETRIC IMAGES FOR PREDICTION OF BODY WEIGHT AND HOT CARCASS WEIGHT IN NELLORE BEEF CATTLE

ABSTRACT - The work was divided in two studies. The objective of this study was to predict the body weight (BW) and the average daily gain (ADG) of Nellore cattle by 3-D images and to compare four prediction models: Multiple Linear Regression (MLR), LASSO Regression, Partial Least Squares (PLS) and Artificial Neural Networks (ANN). A total of 234 images of bovine Nellore were collected. Data collection was performed in four stages throughout the life of the animal: Weaning at 244 days of age and 202.3 kg (± 27.1), Stocker at 457 days of age and 213.9 kg (± 25.1), Initial of Termination at 590 days of age and 334.5 kg (± 29.2) and Finish of Termination at 763 days of age and 449.5 kg (± 47.5). In the first three phases images of 62 Nellore cattle were collected, while in the last phase only 48 images were collected. The ADG was measured: 1: Weaning - Stocker, 2: Weaning - Initial of Termination, 3: Weaning - Finish of Termination, 4: Stocker - Initial of Termination, 5: Stocker - Finish of Termination and 6: Initial of Termination - Final of Termination. In study 2, four hundred and fifty images of Nellore cattle were collected in four experiments for prediction of BW and hot carcass weight (HCW). Four experimental sets were considered: Set 1 includes experiments 1, 2 and 3 for training and experiment 4 for validation; Set 2 includes experiments 1, 2 and 4 for training and experiment 3 for validation; Set 3 includes experiments 1, 3 and 4 for training and experiment 2 for validation; Set 4 includes experiments 2, 3 and 4 for training and experiment 1 for validation. Experiments 1, 3 and 4 were conducted at centro de pesquisa do Pólo Regional de Desenvolvimento Tecnológico dos Agronegócios da Alta Mogiana, em Colina - SP, Brazil. For the experiment 1, 48 male Nellore bulls, with a mean of 24 ± 2 months and a mean live weight of 449.4 (± 46.9 kg) were used. For experiment 2, 228 Nellore male bovines were used, with an average of 22 ± 2 months and initial mean live weight of 588.4 (± 35.5 kg). Experiment 3 was carried out at the Confinamento Experimental do Departamento de Melhoramento e Nutrição Animal da Faculdade de Medicina Veterinária e Zootecnia, UNESP - Câmpus de Botucatu, in which 83 male Nellore bulls were used, with an average of 22 ± 2 months and weight live weight of 516.6 (± 38.0 kg). For the experiment 4, 91 male Nellore bulls were used, with an average of 22 ± 2 months and mean live weight of 589.4 (± 30.1 kg). The Microsoft Model 1473 Kinect sensor was used to capture images from both studies. Concomitantly with the acquisition of images, each animal was weighed in an electronic scale. The analysis of the images was divided into three stages: first, the estimation of the distance from the ground to the camera, second, a segmentation stage and a final step to extract the characteristics. Characteristics with correlation below $P < 0.05$ were removed from the model. Thus, for the statistical analyzes in study 1, we used: Body area and volume, six widths and heights along the animal's back and length. To compare the evaluated models, the method of cross validation Leave-One-Out was used. For the statistical analyzes in study 2 were used: Area and body volume, eleven widths and heights along the animal's back, length, eccentricity and two measures of the curvature of the spine. In study 1, the ANN method increased PC accuracy and accuracy compared to the other methods: Weaning: (concordance correlation coefficient (CCC) = 0.94; root mean square error of prediction (RMSEP) = 8.6 kg; $R^2 = 0.72$). Stocker: (CCC = 0.87, RMSEP = 11.4 kg, $R^2 = 0.79$). Initial of Termination: (CCC = 0.96, RMSEP = 7.7 kg, $R^2 = 0.92$).

Thus, as for the BW, the use of the ANN method increased the precision and accuracy of the predictions for ADG compared to the other models used. For study 2, the predictive ANN method for body weight obtained better results in sets 1, 2, 3 and 4 compared to PLS, LASSO and RLM (CCC = 0.73, 0.66, 0.70, 0.74; RMSEP = 19.6, 27.2, 27.2, 33.7 kg and R^2 = 0.58, 0.53, 0.53, 0.59, respectively). As for warm carcass weight, the RLM predictive method presented better results for Set 1 and 4, LASSO obtained higher results for Set 2, for Set 3 the ANN predictive method presented better results. Set 1 (CCC = 0.14, RMSEP = 8.7 kg and R^2 = 0.05, respectively), Set 2 (CCC = 0.53, RMSEP = 20.1 kg, and R^2 = 0.40, respectively) = 20.1 kg and R^2 = 0.44, respectively), Set 3 (CCC = 0.56, RMSEP = 20.6 kg, and R^2 = 0.44, respectively). Set 4 (CCC = 0.69, RMSEP = 22.7 kg, and R^2 = 0.52, respectively). This study indicates that 3-D images obtained from the Kinect sensor have potential to be used to predict body weight, weight gain and hot carcass weight in Nellore cattle.

Keywords: analysis of images, beef cattle, computer vision, predictive models

LISTA DE ABREVIATÖES

ANN = Redes Neurais Artificiais (*“Artificial Neural Networks”*)

CCC = Coeficiente de Correlação de Concordância (*“Concordance Correlation Coefficient”*)

CV = Coeficiente de variação

DLNN = Redes Neurais de Aprendizado Profundo (*“Deep learning neural network”*)

DP = Desvio Padrão

GMD = Ganho Médio Diário

LASSO = Operador de Encolhimento e Seleção menos absoluto (*“Least absolute shrinkage and selection operator”*)

L1 = Regularização LASSO

L2 = Regularização RIDGE

MB = Viés Médio (*“Mean bias”*)

MSEP = Erro Quadrático Médio de Predição (*“Mean square error prediction”*)

NIPALS = Quadrados Mínimos Não-Iterativos Parciais Não-Lineares (*“Nonlinear interative partial least Squares”*)

PC = Peso Corporal

PCQ = Peso de Carcaça Quente

PLS = Mínimos Quadrados Parciais (*“Partial least Squares”*)

R^2 = Coeficiente de determinação

RMSEP = Raíz do Erro Quadrático Médio de Predição (*“Root mean square error prediction”*)

VIF = Fator de Influência de Variação (*“Variation influence fator”*)

CAPÍTULO 1 - CONSIDERAÇÕES GERAIS

1. INTRODUÇÃO

O setor agropecuário está em constante desenvolvimento, porém este setor enfrenta desafios, tais como, escassez de recursos ambientais e preocupação com danos ao meio ambiente. Assim, com a intensificação da produção, a avaliação do ganho médio diário em tempo real permite o controle do manejo nutricional e então interferência direta no rúmen. Permitindo um modelo de produção mais eficiente, economicamente viável e amplamente sustentável, como exemplo, a utilização de estratégias para diminuir a produção de metano entérico: redução na produção de H_2 e redução das populações de micro-organismos metanogênicos.

A mensuração do peso corporal é um manejo produtivo importante para a gestão de sistemas de produção de bovinos de corte, seu conhecimento permite avaliar o crescimento, estado nutricional, ajustar o fornecimento de alimento e avaliar a condição de saúde animal, afim de administrar o manejo de produtos veterinários adequadamente e estabelecer o valor de mercado (Polgar et al., 1997; Ulutas et al., 2001). O método convencional para avaliação de peso corporal de bovinos através do uso balança é demorado, perigoso e estressante para os operadores e animais (Tasdemir et al., 2011). Segundo Minervino, Cardoso e Ortolani (2008), em Santarém, no Pará, PA cerca de 61,9% das fazendas não possuem balança sendo que o estado possui o quarto maior rebanho bovino brasileiro (ANUALPEC, 2016). Assim, uma maneira de superar essa limitação é através da estimativa do peso do animal usando medidas corporais, chamadas de biometria (Souza et al., 2009).

Pesquisadores tem mostrado que a análise digital de imagens pode ser utilizada para determinar medidas e peso corporal de bovinos (Ozkaya e Bozkurt 2008; Tasdemir et al., 2011; Ozkaya 2012). Portanto, 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 tridimensionais (Li et al., 2014; Ozkaya et al., 2015; Gomes et al., 2016).

O aumento da massa corporal de um animal ocorre em uma seqüência temporal: pré-natal, pós-natal até a desmama, desmama até a puberdade ou sobreano e, para animal de reprodução, da puberdade até a maturidade, proporcionando fases com velocidades de crescimento diferentes, ou seja, em cada fase, o animal apresenta taxa de crescimento característica (Gottschall, 1999). Assim nota-se a importância da avaliação de modelos preditivos para peso corporal e ganho médio diário em diferentes fases do crescimento animal. Geralmente os animais iniciam a fase de terminação com baixa condição corporal e o preço elevado de insumos tendem a diminuir esta fase. Assim análise do ganho de peso em bovinos de corte permite determinar o ponto ótimo de abate, otimizando o número de dias em confinamento, o qual acarretará em economia na utilização de insumos.

Dessa forma este trabalho teve como objetivos: 1) estimar o peso corporal, ganho médio diário e peso de carcaça quente de bovinos nelore utilizando sensor de profundidade e processamento de imagens e 2) avaliar a capacidade preditiva de diferentes modelos de predição: RLM, LASSO, PLS e ANN.

2. REVISÃO DA LITERATURA

2.1. Análise de imagens

Devido a evolução da tecnologia em captura de imagens de profundidade, técnicas de modelagem e monitoramento de imagens e sensores, recentemente diversas técnicas de visão computacional estão em desenvolvimento. A visão computacional simula a visão humana, onde há uma entrada de imagem, porém, a saída é uma interpretação como um todo, ou parcialmente.

Geralmente a visão computacional inicia-se com o processamento de imagens. Conforme Gonzalez et al. (2006) o processamento de imagens é uma técnica mais primitiva, tais como a redução de ruído ou melhoria no contraste de uma imagem, considerando este como um processo de baixo nível. Já a visão computacional apresenta processos de nível-médio, os quais são operações de segmentação ou classificação e processos de nível-alto os quais estão relacionados com as tarefas de cognição associadas com a visão humana.



Figura 1. A figura a esquerda representa uma imagem com presença de ruídos (escura), já a imagem a direita, a qual foi submetida a um processamento de imagens, podemos visualizar a identificação do animal.

A diferença entre o processamento de imagens e visão computacional (Figura 1) mostra que a imagem a esquerda (escura), não é possível visualizar a identificação do animal. Porém a imagem à direita, é uma imagem mais clara a qual foi realizado um processamento na imagem para que assim fosse permitido a identificação do animal. A visão computacional extrai as informações contidas no brinco do animal e identifica os números, possibilitando que os dados do animal possam ser encontrados em um banco de dados.

Assim como o desenvolvimento dos aparelhos eletrônicos, a visão computacional e o processamento de imagem também foram modernizados. O processamento digital de imagens foi aprimorado e suas aplicações permeiam quase todos os ramos da atividade humana, tais como: biologia, cartografia, meteorologia, geologia, entre outros, e também tem sido utilizado na medicina e na agricultura (Alves, 2013).

Segundo Gonzalez et al. (2006), uma imagem pode ser definida como uma função bidimensional, $f(x, y)$, a qual x e y são coordenadas espaciais e a amplitude de f representa a intensidade ou nível de cinza da imagem em um determinado ponto. Recentemente foram divulgadas câmeras tridimensionais de alto desempenho e baixo custo, as quais fornecem em tempo real a profundidade

e a cor da informação originalmente desenvolvidos para serem aplicados em jogos ou reconhecimento facial. Atualmente estão disponíveis os modelos Kinect® da Microsoft, Xtion Pro da ASUS e RealSense da Intel®.

O Kinect® (Microsoft Corporation, Redmond, WA) é um dispositivo de sensor de movimento, com aproximadamente 23 cm de comprimento, composto por uma câmera RGB e um sensor de profundidade (projektor laser infravermelho e uma câmera infravermelha) que permite a varredura do ambiente em três dimensões. A câmera RGB e a câmera infravermelha têm uma resolução de 640 x 480 pixels e podem gerar um fluxo de imagem de 30 *frames* por segundo. O ângulo do campo de visão do sensor de profundidade é de 57° na horizontal e 43° na vertical e pode ser operado entre 0,8 m e 3,5 m (Corrêa, 2015). Através do sensor de profundidade, é obtida uma imagem de profundidade, esta imagem representa a distância da câmera a cada valor de *pixel* na imagem (Figura 2).

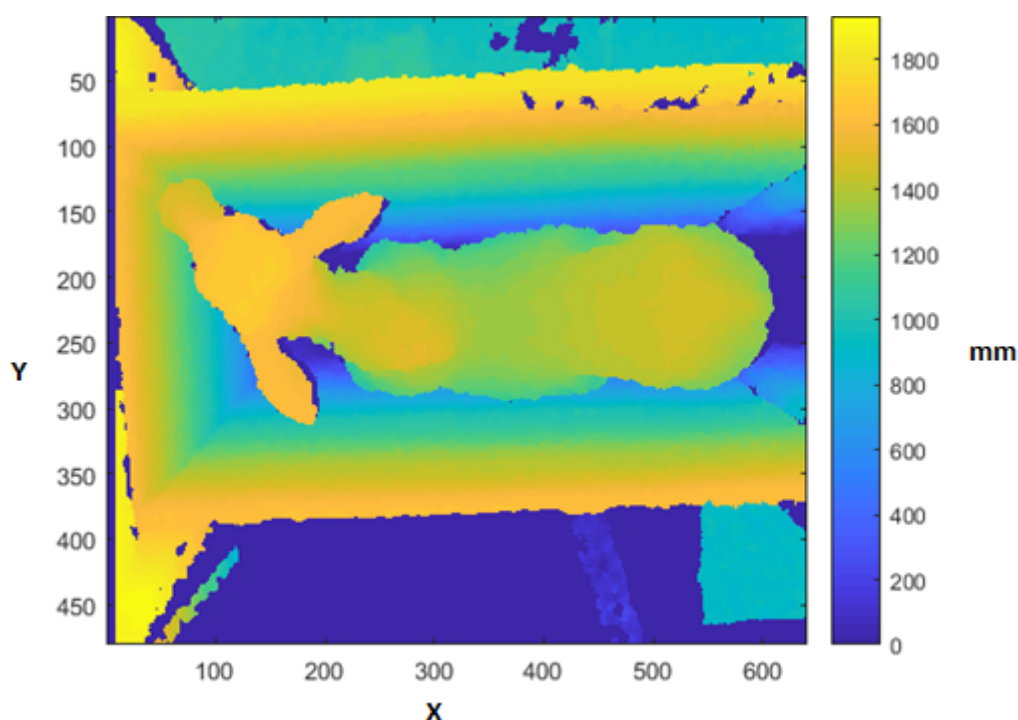


Figura 2. Imagem de profundidade de bovino Nelore obtida através do sensor Kinect®, onde X e Y corresponde as coordenadas de cada *pixel* no plano cartesiano e mm representa a distância em milímetros da câmera a cada ponto de *pixel* da imagem.

Segundo Kuzuhara et al. (2015) em um estudo sobre as propriedades leiteiras e peso corporal de vacas da raça holandesa comprovaram o potencial do uso de sistemas de câmera tridimensionais. Gomes et al. (2016) desenvolveram técnicas para obter biometria dos animais através de imagens

digitais que provaram ser um método eficiente. Portanto, há uma correlação importante entre as mensurações biométricas e o peso corporal, de modo que a massa corporal pode ser estimada por meio de índices biométricos obtidos com o uso de imagens tridimensionais (Li et al., 2014; Ozkaya et al., 2015; Gomes et al., 2016).

2.2. Regressão Linear Múltipla

Através de um estudo sobre altura de crianças, Francis Galton introduziu o termo regressão. Nesse estudo, foi constatado que a altura média dos filhos de pais de uma determinada altura tendia a “regredir” até a altura média da população. Esse fenômeno foi chamado de lei de regressão universal de Galton e foi confirmado posteriormente por Pearson (Gujarati, 2000; Neter et al., 2005). A abordagem inicial proposta por Galton é diferente do significado atual da análise de regressão, sendo definida por Montgomery et al. (2001), como “uma técnica estatística para modelar e investigar a relação entre as variáveis”. A simplicidade e utilidade dessa técnica são resultantes do conceito lógico de utilizar uma equação matemática para expressar a relação entre variáveis, sendo interessante devido à base matemática e bem desenvolvida teoria estatística (Montgomery et al., 2001). Segundo Weisberg (2005) a regressão, assim como a maioria das técnicas estatísticas, tem como meta sumarizar os dados observados da forma mais simples e útil possível. Segundo Gujarati (2000) e Neter et al. (2005), a análise de regressão estuda a dependência das características de interesse, em relação as variáveis, com o objetivo de estimar o valor médio da característica de interesse em função dos valores conhecidos das variáveis dependentes.

A análise de regressão múltipla é uma técnica estatística para investigar e modelar a relação entre variável resposta (Y) e as variáveis explanatórias ($X_1, X_2, \dots, X_k$), sendo uma das técnicas mais utilizadas na análise de dados. Pode-se citar inúmeras aplicações de análise de regressão em diversas áreas (Freedman et al., 2004, Lyles & Kupper 1997, Chen & Wang 2004). De acordo com Montgomery et al. (2001), o modelo de regressão linear múltipla com k variáveis de controle é representado pela equação [1].

$$Y = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad [1]$$

De acordo com Gujarati (2000) e Montgomery et al. (2001) os coeficientes de regressão $\beta_0, \beta_1, \dots, \beta_k$ são expressos como:

- a) O coeficiente β_0 , chamado de coeficiente de intercepto, o qual corresponde à média de Y_i quando todas as variáveis de controle são iguais a zero;
- b) Os coeficientes $\beta_1, \dots, \beta_k$ são chamados de coeficientes de regressão parciais. O coeficiente β_k pode ser interpretado como a derivada parcial de Y_i em relação a X_{ki} , ou seja, é variação de Y causada por uma variação unitária em x_k dado que as outras variáveis de controle sejam mantidas constantes.

2.3. Regressão LASSO

Diversas pesquisas têm sido desenvolvidas com o intuito de comparar diferentes métodos de regressão adequadas para situações em que há sérios problemas de multicolinearidade e/ou alta dimensionalidade. A multicolinearidade é uma condição que ocorre quando algumas variáveis preditoras estão correlacionadas entre si no modelo. Uma alta multicolinearidade é problemática pois pode aumentar a variância dos coeficientes de regressão, tornando-os instáveis. Alta dimensionalidade trata-se de quando o conjunto de dados apresenta um alto número de variáveis preditoras de forma a prejudicar o modelo. Dormann et al. (2013) estudaram casos com problemas de colinearidade, principalmente estudos ecológicos. Os resultados indicaram uma preferência para a utilização de métodos com penalização, tais como a regressão RIDGE e LASSO, com relação a outros métodos como: Regressão por Componentes Principais (RCP) e Partial Least Squares (PLS). Já para Archarjee et al. (2013) o foco do estudo é sobre casos com alta dimensionalidade, em que o número de covariáveis é maior que o número de observações, neste estudo foram selecionadas sete covariáveis simultaneamente, baseados no erro quadrático médio de predição, os quatro métodos estudados obtiveram resultados parecidos. Entretanto, a regressão

LASSO foi a que obteve maior destaque e com menor erro quadrático de predição. A regressão LASSO é um método comumente utilizado por reduzir os efeitos dos atributos que não contribuem para a identificação da variável resposta, reduzindo seus coeficientes para zero e excluindo-os do modelo (Tibshirani, 1997). Com isso a regressão LASSO realiza um tipo de processo de seleção de variáveis (Zou & Hastie, 2005). Isso torna-se necessário pois uma regressão com um grande número de variáveis independentes torna o modelo como um todo muito mais complexo e pode tirar características de interpretabilidade. Assim, a regressão LASSO pode ser uma forma de eliminar esse problema, que pode absorver o ruído dos dados e causar o *overfitting*.

2.4. Regressão *Partial Least Squares*

Assim como em regressão linear múltipla, a regressão por mínimos quadrados parciais (“PLS” – *Partial Least Squares* – sigla em inglês) tem por finalidade construir um modelo linear $Y = X\beta + E$, em que Y é uma matriz $n \times m$ de variáveis resposta, X é uma matriz $n \times p$ de variáveis preditoras, β é uma matriz $p \times m$ dos coeficientes de regressão, e E é a matriz de ruídos para o modelo, que tem a mesma dimensão de Y . Modelos de regressão por mínimos quadrados parciais (Wold et al., 2001) é uma alternativa para modelar as relações existentes entre conjuntos de variáveis observadas por meio de variáveis latentes. Quando o número de observações são inferiores em relação ao número de variáveis independentes ou ainda em situações em que ocorre multicolinearidade. Tal regressão gera um conjunto de parâmetros que fornecem informações sobre a estrutura e comportamento de X e Y , o que corrobora para sua ampla aplicação em procedimentos de seleção de variáveis (Wold et al., 2001).

Para a escolha do número de componentes a serem mantidos no modelo, avalia-se a significância em termos de predição de cada componente e a inclusão de componentes no modelo é interrompida quando eles deixam de ser significativos (Wold et al., 2001). Ferrer (2008) ressaltam que poucas ferramentas de análise estatística possuem a versatilidade da regressão PLS, a qual tem oferecido suporte em aplicações de diferentes naturezas, como discriminação e classificação de observações, modelagem e análise de processo

e identificação de desvios. As suas funções não são restritas em áreas industriais, mas também são verificadas em vários setores de negócios, finanças e marketing e áreas de saúde (Ferrer, 2008; Esposito-Vinzi, 2007).

2.5. Artificial Neural Networks

Redes neurais artificiais ou “*Artificial Neural Networks*” – ANN em inglês é uma técnica a qual tem por característica desenvolver modelos computacionais inspirados na capacidade natural do cérebro humano, e que apresenta ótimos resultados na solução de problemas complexos envolvendo alta dimensionalidade de variáveis. Assim as redes neurais artificiais são aplicadas a diversas áreas de estudo como, Ciências Agrárias, Humanas e Exatas. Esta técnica iniciou-se na década de 40 através dos estudos de McCulloch; Pitts (1943) e Hebb (1949) que estudaram a capacidade de informações das redes neurais artificiais. O modelo matemático desenvolvido por McCulloch; Pitts em 1943 tratava-se de um simples neurônio que trabalhava em um processo de decisão lógica (verdadeiro ou falso).

A rede neural pode ser utilizada para modelar ou prever a resposta correspondente a um novo padrão. Esse procedimento é semelhante a um problema de regressão, na estatística convencional, em que se deseja encontrar uma relação entre variáveis explanatórias (entradas) e resposta (saídas) (Haykin, 2001). O conhecimento é adquirido através de um processo de aprendizado ou treinamento com as respostas já conhecidas é apresentado à rede. Os dados de entrada são fornecidos ao sistema e é obtido o conjunto de respostas (saídas). As saídas da rede são comparadas com as respostas previamente conhecidas e, da diferença entre elas, é obtido o erro. Com o valor do erro, os pesos sinápticos são ajustados e o processo é repetido iterativamente, parando quando o erro atingir patamares satisfatórios (Braga, 2007). Assim, a modelagem através de redes neurais apresenta algumas vantagens, como: flexibilidade da rede (a rede pode ser adaptada para um novo ambiente através da troca de padrões de entrada → saída), facilidade de capturar novas informações do meio (dados de novos ensaios, por exemplo) e treinamento possível mesmo na presença de ruídos, o que é comum em dados experimentais (Braga, 2007).

2.6. Validação Cruzada *Leave-One-Out*

Validação cruzada é uma técnica de validação de um modelo para avaliar a robustez dos resultados de uma determinada análise. É utilizado principalmente para estimar a precisão do modelo preditivo na prática. Assim, uma das aplicações da validação cruzada é avaliar o *overfitting* e fornecer a ideia de como o modelo funcionará em um dado independente (isto é, um conjunto desconhecido de dados). Segundo Kiralj R. & Ferreira M. M. C. (2009) a validação cruzada *Leave-One-Out*, consiste em excluir cada amostra de uma vez do banco de dados que servirá para testar o modelo (construção de um modelo sem esta amostra) e a amostra removida será utilizada para o treinamento do modelo. A fase de treinamento e teste é repetida para cada número de observações em todo o conjunto de bancos de dados.

3. OBJETIVOS GERAIS

Os objetivos deste estudo foram 1) prever o peso corporal, ganho médio diário e peso de carcaça quente em bovinos Nelore, por meio de imagens tridimensionais e 2) avaliar a capacidade preditiva dos modelos de predição: RLM, LASSO, PLS e ANN.

4. REFERÊNCIAS

ALVES, N. M. E. (2013). A equação do calor aplicada ao processamento de imagens. Universidade Federal de Minas Gerais. 50p. Dissertação (Especialista em Matemática). Disponível em: http://www.mat.ufmg.br/~pgmat/monografias/Monografia_Natalia_Moreira.pdf

ANUALPEC. Anuário estatístico da pecuária de corte. São Paulo: FNP Consultoria e Comércio Ltda., 2016.

ACHARJEE, A.; FINKERS, R.; VISSER, R. GF.; MALIEPAARD, C. Comparison of Regularized Regression Methods for Omics Data, **Metabolomics**, 2013.

BRAGA, A. P; CARVALHO, A. C. P. L. F; LUDEMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2 ed2 Rio de Janeiro: LTC, 2007.

CHEN, Z. & WANG, Y. G. Efficient Regression Analysis with Ranked-Set Sampling, **Biometrics** v.60, 997–1004, 2004.

CORREA, P. B. M. Estimativa da massa corporal de bovinos por meio de sensor de profundidade Kinect. 2015. 50 f. **Dissertação (Mestrado)** - Curso de Engenharia Agrícola, Universidade Federal de Viçosa, Viçosa, 2015.

DORMANN, C. F; ELITH, J.; BACHER, S.; BUCHMANN, C.; CARL, G.; CARRÉ, G.; MARQUÉZ, J. R. G.; GRUBER, B.; LAFOURCADE, B.; LEITÃO, P. J.; MUNKMULLER, T.; McCLEAN, C.; OSBORNE, P. E.; REINEKING, B.; SCHRODER, B.; SKIDMORE, A. K.; ZURELL, D.; LAUTENBACH, S. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance, **Ecography** v.36, 027-046, 2013.

ESPOSITO-VINZI, V. Handbook of Partial Least Squares: Concepts, Methods and Applications in Marketing and Related Fields. Berlin: Springer, 2007. 850 p.

FERRER, A. PLS: A versatile tool for industrial process improvement and optimization. **Applied Stochastic Models in Business and Industry**, v.24, 551-567, 2008.

FREEDMAN, L. S.; FAINBERG, V.; KIPNIS, V.; MIDTHUNE, D.; CARROL, R. J. A new Method for Dealing with Measurement Error in Explanatory Variables of Regression Models, **Biometrics** v.60, 172–181, 2004.

GOMES, R. A.; MONTEIRO, G. R., ASSIS G. J. F., BUSATO, K. C., LADEIRA, M. M., and CHIZZOTTI, M. L. Technical note: Estimating body weight and body composition of beef cattle through digital image analysis. **Journal of Animal Science**. 2016. 94:5414-5422. doi:10.2527/jas.2016-0797

GONZALEZ, R. C. e WOODS, R. E. e EDDINS, S. L., Digital Image Processing using MATLAB, Pearson, 2006.

GOTTSCHALL, C. S. **Impacto nutricional na produção de carne e curva de crescimento**. In: LOBATO, J. F. P.; BARCELLOS, J. O. J.; KESSLER, A. M. Produção de bovinos de corte. Porto Alegre: EDIPUCRS, p. 169-192, 1999.

GUJARATI, D. N. **Econometria Básica**. 3. Ed. São Paulo: Makron Books, 2000. 846 p.

HAYKIN, S. **Redes neurais: princípios e prática**. 2 ed. Porto Alegre: Bookman, 2001.

HEBB, D. **The organization of behavior**. New York: J. Wiley, 1949.

HUI ZOU & TREVOR HASTIE. Regularization and variable selection via the elastic net. **Journal of the Royal Statistical Society**, v,67, 301–320, 2005.

KIRALJ, R. & FERREIRA M. M. C. Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. **Journal Brazilian Chemical Society**, 770-787, 2009.

KUZUHARA Y.; KAWAMURA K.; YOSHITOSHI R.; TAMAKI T.; SUGAI S.; IKEGAMI M.; KUROKAWA Y.; OBITSU T.; OKITA M.; SUGINO T.; YASUDA T. A preliminary study for predicting body weight and milk properties in lactating Holstein cows using a three-dimensional camera system. **Computers and Electronics in Agriculture**, 186-193, 2015.

LI, Z.; LUO C.; TENG G.; & LIU T. Estimation of pig weight by machine vision: A review. In: D. LI and Y. CHEN, editor, **Computer and computing technologies in agriculture VII: 7th IFIP WG 5.14 International Conference, CCTA 2013, Beijing, China, September 18–20, 2014, Revised Selected Papers, Part II**. Springer Berlin Heidelberg, Berlin, Germany. p. 42–49. doi:10.1007/978-3-642-54341-8_5

LYLES, R. H. & KUPPER, L. L. A Detailed Evaluation of Adjustment Methods for Multiplicative Measurement Error in Linear Regression with Applications in Occupational Epidemiology, **Biometrics** v.53, 1008–1025, 1997.

McCULLOCH, W; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, Chicago, 1994, 520p.

MINERVINO, A. H. H.; CARDOSO, E. da C.; & ORTOLANI, E. L. Características do sistema produtivo da pecuária no município de Santarém, Pará. **Acta Amazonica**, v. 38, n. 1, p.11-16, 2008.

MONTGOMERY, D. C; VINNING, G. C.; PECK, E. A. **Introduction to linear regression analysis**. 3rd ed. New York: J Wiley, 2001. 641 p.

NETER, J. et al. **Applied linear statistical models**. 5th ed. New York: Mc Graw-Hill/Irwin, 2005. 1396 p.

OZKAYA, S.; NEJA W.; KREZEL-CZOPECK, S.; OLER A. Estimation of bodyweight from body measurements and determination of body measurements on Limousin cattle using digital image analysis. **Animal Production. Science**. 2015, doi:10.1071/AN14943

OZKAYA, S. Accuracy of body measurements using digital image analysis in female Holstein calves. **Animal Production Science** 50, 917–920, 2012.

OZKAYA, S. & BOZKURT, Y. The relationship of parameters of body measures and body weight by using digital image analysis in pre-slaughter cattle. **Archives of Animal Breeding**, Dummerstorf 51, 120–128, 2008.

POLGAR, P.; SZÜCS, E.; SZABO, F. Effect of selection for improvement of dairy performance on growth and body structure in Holstein Friesian and Hungarian Fleckvieh young bulls. (German language) **Arch. Tierz.** 40 (1997) 505-510

SOUZA, S.; LEAL, A.; BARIONI, C.; MATOS, A. Utilização de medidas biométricas para estimar peso vivo em ovinos. **Archivos Latinoamericanos Producción Animal**, v.17, p.61-66, 2009.

TASDEMIR, S. URKMEZ, A. & INAL, S. Determination of body measurements on the Holstein cows using digital image analysis and estimation of live weight with regression analysis. **Computers and Electronics in Agriculture** 76, 189–197, 2011.

TIBSHIRANI, R. Regression shrinkage and selection via the Lasso. **Journal of the Royal Statistical Society**. v.16, 385-395, 1997.

ULUTAS, Z., SAATCI, M., OZLUTURK, A. (2001). Prediction of body weights from body measurements in East Anatolian Red calves. **Journal of Atatrak University Agriculture Faculty**. 32, 61-65

WEISBERG, S. **Applied linear regression**. 3rd ed. New York: J. Wiley, 2005.

WOLD, S.; SJOSTROM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems** v.58, 109-130, 2001.

CAPÍTULO 2 - USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DE PESO CORPORAL E GANHO MÉDIO DIÁRIO DE BOVINOS DE CORTE NELORE

RESUMO: Medidas frequentes do peso corporal (PC) podem permitir a avaliação da curva de crescimento, resultando em melhores estratégias no manejo nutricional e reprodutivo do animal. Entretanto, a obtenção de PC em tempo real através de balanças de pesagens é cara e trabalhosa. Uma estratégia interessante seria a coleta de imagens tridimensionais para predição de PC. Os objetivos deste estudo foram 1) prever o PC e o ganho médio diário (GMD) de bovinos Nelore por meio de imagens tridimensionais e 2) comparar diferentes modelos de predição. Foram coletadas 234 imagens de bovinos da raça Nelore. A coleta de dados foi realizada em quatro fases da vida do animal: Desmama aos 244 dias de idade e 202,3 kg ($\pm 27,1$, $n = 62$), Recria aos 457 dias de idade e 213,9 kg ($\pm 25,1$, $n = 62$), Início da Terminação aos 590 dias de idade e 334,5 kg ($\pm 29,2$, $n = 62$) e Final da Terminação aos 763 dias de idade e 449,5 kg ($\pm 47,5$, $n = 48$). Para capturar as imagens foi utilizado o sensor Kinect® modelo 1473 da Microsoft. Concomitantemente à aquisição de imagens, cada animal foi pesado em balança eletrônica. Área e volume corporal, seis larguras e alturas ao longo do dorso do animal e comprimento foram utilizadas para predição do PC. Foram desenvolvidos quatro métodos de predição: Regressão Linear Múltipla (RLM), Least Absolute Shrinkage and Selection Operator (LASSO), *Partial Least Squares* (PLS) e *Artificial Neural Networks* (ANN). Como forma de avaliar a qualidade dos modelos preditivos uma cross-validação *leave-one-out* foi adotada. O método de ANN aumentou a precisão e acurácia para PC comparado aos outros métodos: Desmama: (concordance correlation coefficient (CCC) = 0.94; root mean square error of prediction (RMSEP) = 8.6 kg; $R^2 = 0.72$). Recria: (CCC = 0.87; RMSEP = 11.4 kg; $R^2 = 0.79$). Início da Terminação: (CCC = 0.96; RMSEP = 7.7 kg; $R^2 = 0.92$). Assim, como para o PC, o uso do método ANN aumentou a precisão e acurácia das predições para GMD comparado aos outros modelos utilizados. Os resultados obtidos para foram: 1 (CCC = 0.74; RMSEP = 0.02 kg/d e $R^2 = 0.67$), 2 (CCC = 0.89; RMSEP = 0.02 kg/d e $R^2 = 0.85$), 3 (CCC = 0.84; RMSEP = 0.03 kg/d e $R^2 = 0.80$), 4 (CCC = 0.54; RMSEP = 0.10 kg/d e $R^2 = 0.51$) , 5 (CCC = 0.82; RMSEP = 0.06 kg/d e $R^2 = 0.76$), 4

(CCC = 0.90; RMSEP = 0.09 kg/d e $R^2 = 0.82$). O uso de câmeras tridimensionais prediz com acurácia o peso corporal e ganho de peso em bovinos nelore.

Palavras-chave: análise de imagens, bovinocultura de corte, kinect®, visão computacional

1. INTRODUÇÃO

A variável mais importante mensurada na produção de bovinos de corte é o peso corporal. Através desta informação é possível a avaliação da curva de crescimento, resultando em melhores estratégias no manejo nutricional e reprodutivo do animal. Para determinar o peso corporal e desempenho dos animais, o método comumente utilizado consiste em pesagens frequentes com a utilização de balanças. No entanto, esse manejo pode causar estresse em animais (Schofield et al., 1999). Schofield et al. (1999) relataram que a pesagem convencional realizada por meio da utilização de balança tem um alto custo e requer manutenção.

Alternativamente, vários autores (Wang et al., 2008; Kashiha et al., 2014; Gomes et al., 2016) desenvolveram técnicas para obter biometria dos animais através de imagens digitais que provaram ser um método eficiente. Portanto, há uma correlação importante entre as mensurações biométricas e o peso corporal, de modo que a massa corporal pode ser estimada por meio de índices biométricos obtidos com o uso de imagens tridimensionais (Li et al., 2014; Ozkaya et al., 2015; Gomes et al., 2016).

Uma estratégia interessante para obter as imagens seria através de sensores de profundidade baseados em luz infravermelha, como o dispositivo Microsoft Kinect® (Konsgro, 2014; Microsoft Corporation, Redmond, WA). Consequentemente, a coleta de imagens é um método que evita o contato com animais e pode ser uma alternativa rápida para estimar o peso corporal (Gomes et al., 2016). As balanças, entretanto, são relativamente caras e seus componentes eletrônicos são propensos a danos em ambiente hostil coberto com esterco e urina e em contato direto com vacas (Dickinson et al., 2013). Além disso, a utilização de imagens tridimensionais permite amplo monitoramento do rebanho, devido ao maior número de informações geradas.

Para estimativa de peso corporal é importante utilizar diferentes modelos de predição para avaliar a qualidade dos modelos preditivos. O conjunto de dados pode apresentar uma colinearidade, relações não-lineares entre as variáveis. Uma alta colinearidade é problemática pois pode aumentar a variância dos coeficientes de regressão, tornando-os instáveis. A regressão linear múltipla é fundamental para fazer comparações entre a combinação de vários fatores e seus resultados. A regressão LASSO é um método comumente utilizado por

reduzir os efeitos dos atributos que não contribuem para a identificação da variável resposta, reduzindo seus coeficientes para zero e excluindo-os do modelo (Tibshirani, 1997). Assim produz melhoria do erro de predição diminuindo os grandes coeficientes de regressão para reduzir o *overfitting* através da contração das variáveis independentes. Ao contrário de RLM e LASSO, o modelo de mensuração de PLS estima as variáveis latentes como uma soma ponderada de suas variáveis manifestas. O modelo estrutural estima as variáveis latentes por meio de regressão linear simples ou múltipla entre as variáveis latentes estimadas pelo modelo de mensuração. Este algoritmo é repetido até que a convergência seja alcançada. O uso de técnicas de machine learning poderia melhorar a qualidade das previsões devido à sua capacidade de modelar relações complexas entre variáveis, como não-linearidades e interações (Gianola et al., 2011). Assim, o uso de ANN, poderia ser uma abordagem alternativa para o desenvolvimento de modelos preditivos. ANN foram desenvolvidas com base no sistema nervoso humano e consistem em neurônios interconectados (lineares ou não-lineares) organizados em camadas (Bishop, 2006).

A hipótese deste trabalho é que o peso corporal e o ganho de peso médio diário, de bovinos Nelore, podem ser preditos, por meio de imagens tridimensionais. Predições acuradas e precisas do peso corporal e do desempenho podem permitir o acompanhamento da curva de crescimento dos animais além da detecção de problemas sanitários no rebanho (desempenho sub-ótimo), em tempo real. O uso dessa ferramenta pode colaborar para tomadas de decisão que evitem a redução da rentabilidade ou prejuízos aos sistemas de produção animal. O objetivo deste estudo foi predizer peso corporal e o ganho médio diário de bovinos Nelore por meio de imagens tridimensionais e determinar o melhor modelo de predição.

2. MATERIAL E MÉTODOS

Todos os procedimentos animais foram aprovados pelo Comitê de Ética no uso de animais da Faculdade de Ciências Agrárias e Veterinárias da Universidade do Estado de São Paulo, número de protocolo 007946/18.

2.1. Dados

O experimento foi conduzido no centro de pesquisa do Pólo Regional de Desenvolvimento Tecnológico dos Agronegócios da Alta Mogiana (PRDTA - Alta Mogiana), em Colina - SP, Brasil. Foram coletadas 234 imagens de bovinos da raça Nelore, meio irmãos, os quais com aproximadamente 580 dias de idade foram quimicamente castrados. A coleta de dados foi realizada em quatro momentos ao longo do crescimento dos animais: Desmame aos 244 dias de idade e 202,3 kg ($\pm 27,1$), Recria aos 457 dias de idade e 213,9 kg, Início da Terminação aos 590 dias de idade e 334,5 kg ($\pm 29,2$) e Final da Terminação aos 763 dias de idade e 449,5 kg ($\pm 47,5$), como apresentado na Tabela 1. O desvio padrão de idade foi de $\pm 21,3$ dias.

Tabela 1. Estatística descritiva referente as quatro fases da vida dos bovinos Nelore

Fase	No. Bovinos	Média	DP ¹	CV ²	Mínimo	Máximo
Desmama (244d), kg	62	202,3	27,1	13,7	114,0	266,0
Recria (457d), kg	62	213,9	25,1	11,7	138,0	277,0
Início Terminação (590d), kg	62	334,5	29,2	8,7	274,0	391,0
Final Terminação (763d), kg	48	449,5	47,5	10,6	359,0	554,0

¹Desvio padrão

²Coefficiente de variação

Os animais foram recriados em uma área de 24 ha, constituída de capim tânzania (*Panicum maximum* cv. Tânzania) em um período de 346 dias. Durante esse período, na estação da seca (compreendida entre abril – setembro) os animais receberam suplemento mineral proteico energético, no nível 3g/kg de PC e durante a estação das águas foi ofertado suplemento mineral proteico energético, no nível de 3 g/kg de Peso Corporal, conforme Tabela 2.

A fase de terminação teve duração de 173 dias, onde os animais permaneceram em confinamento com baias individuais. Durante esse período os animais receberam uma dieta com 70% de concentrado, 14% de PB e 83% de NDT.

Nas três primeiras fases foram coletadas imagens tridimensionais de 62 bovinos de corte Nelore, enquanto que na última fase apenas 48 bovinos da raça Nelore obtiveram imagens tridimensionais coletadas. Para capturar as imagens, utilizamos o sensor Kinect[®] modelo 1473 (Microsoft Corporation, Redmond,

WA). O sensor foi posicionado no curral de manobra paralelo ao solo, como pode ser evidenciado na (Figura 1). O tempo de coleta das imagens RGB e infravermelho (640 x 480 pixels) foi de 15 segundos e 10 *frames* foram obtidos para cada animal. Concomitantemente à aquisição da imagem, cada animal foi pesado em uma balança eletrônica (Coimma LTDA, modelo RUDD 300, Dracena, São Paulo, Brasil), com capacidade mínima de 10 kg, máximo de 1.500 kg e precisão de 0,5 kg.

Tabela 2. Composição dos suplementos (%) na matéria seca durante a recria

Ingredientes	Seca ¹	Águas ²
	3g/kgPC	3g/kgPC
Milho	47,3	60,3
Farelo de Soja	37,9	24,8
Uréia	3,1	3,1
Núcleo	11,6	11,7

Composição Química Estimada

PB ³	30	20
NDT ⁴	65	67

¹Estação seca (compreendida entre abril – setembro)

²Estação das águas (compreendida entre outubro – março)

³Proteína Bruta

⁴Níveis digestíveis totais

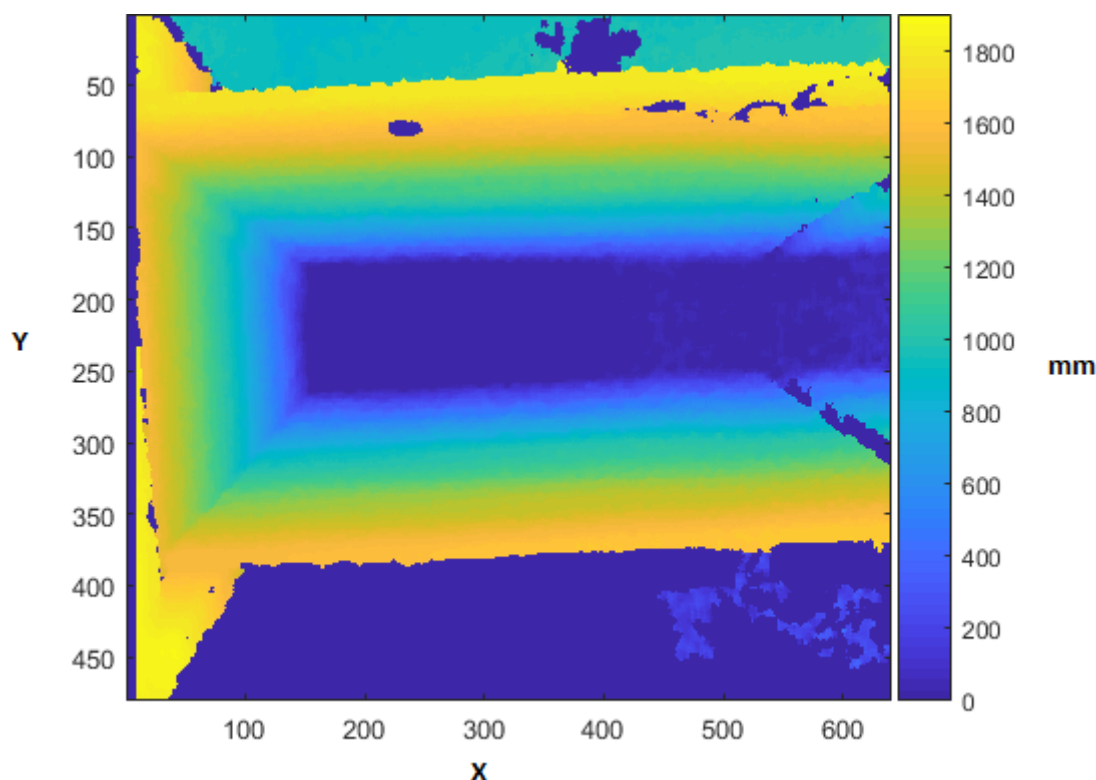


Figura 1. Imagem sem a presença do animal, utilizada para estimativa da distância da câmera ao chão sem a presença do animal, onde X e Y corresponde as coordenadas de cada *pixel* no plano cartesiano e mm representa a distância em milímetros da câmera a cada ponto de *pixel* da imagem.

2.2. Análise de imagens

A análise de imagem realizada foi uma adaptação do método automatizado desenvolvido por Fernandes et al. (comunicação pessoal). O método original foi criado para análise de vídeo, mas pode ser simplificado para imagens. Os 10 *frames* obtidos para cada animal foram divididos em três etapas: primeira, estimativa da distância do chão à câmera, segunda, uma etapa de segmentação e uma etapa final para extração das características. Para análise estatística foi utilizado o melhor *frame* segmentado adequado.

2.2.1. Estimativa da distância do chão à câmera:

Nesta etapa, o algoritmo realizou uma estimativa da distância entre a câmera e o solo com base no modo do grupo de pixels mais distante da câmera, conforme descrito na Equação [1].

$$f_{dist} = mode(maxPeak) \quad [1]$$

Onde f_{dist} é a distância do chão à câmera e $maxPeak$ é o cluster de pixels que tinha uma distância da câmera de 90% ou mais (Figura 1).

2.2.2. Segunda segmentação:

A abertura biométrica da imagem foi realizada para avaliar a inclusão do animal na cena e sua conectividade com os objetos. Então, nesta etapa, todo objeto conectado ao animal foi removido. Em seguida, a posição do corpo do animal foi avaliada por um contorno da área dorsal do animal e estimativa da posição da coluna vertebral. Para a estimativa da posição da coluna vertebral, apenas a área dorsal que estava entre o ombro e a garupa do animal foi utilizada. Ombro e garupa foram identificados através de uma adaptação da transformação de Hough para identificação de objetos redondos (Hough, 1962; Atherton e Kerbyson, 1999), (Figura 2).

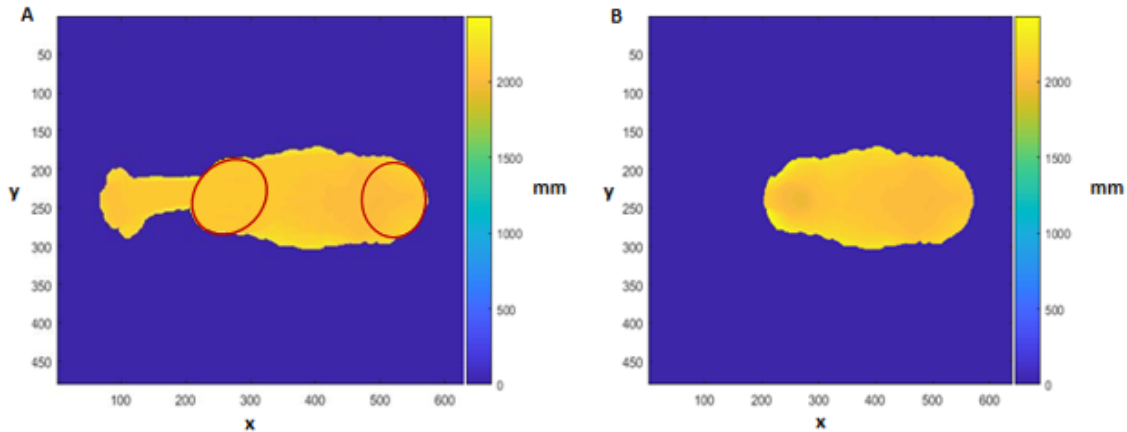


Figura 2. A: Ombro e garupa identificados por meio de uma adaptação de transformação de Hough para a identificação de objetos redondos, B: Imagem do dorso do animal segmentado após a identificação de objetos redondos. Onde X e Y corresponde as coordenadas de cada *pixel* no plano cartesiano e mm representa a distância em milímetros da câmera a cada ponto de *pixel* da imagem.

2.2.3. Extração de características:

Após a extração das características realizadas pelo algoritmo, é salvo um *frame* preto e branco (Figura 3). As características extraídas foram exportadas e salvas em um documento separado por vírgula (.csv). O conjunto de recursos extraídos pode ser dividido em duas classes: medidas corporais e descritores de forma.

As medidas corporais foram: área, volume, comprimento máximo, larguras e alturas em locais equidistantes na parte dorsal do animal dos ombros até a garupa (Figura 4). O comprimento, a largura e a área de um objeto em uma imagem podem ser estimados pela soma de pixels. Para estimar a área real de qualquer pixel, o princípio trigonométrico de ampliação de imagem pode ser usado:

$$m = f/d_o \quad [2]$$

$$a_i = 1/m_p \quad [3]$$

$$a_i = 1/m_p \quad [4]$$

$$v_o = \sum_{i=1}^k a_p * h_p \quad [5]$$

Onde m é o fator de ampliação, f a distância focal efetiva que é intrínseca à câmera e d_o é a distância que o objeto está da câmera. Assim, a área de qualquer objeto a_o pode ser estimada como a soma da área dos pixels a_p que

formam sua imagem. Para p indo de 1 a n , o índice dos n pixels que formam a representação da imagem do objeto. As alturas de pixel h_p , podem ser estimadas pela diferença entre a distância até a câmera para o pixel p na área dorsal do animal e a distância da câmera até o chão. Portanto, o volume é a soma dos volumes de cada pixel que forma a representação da imagem do objeto. Note-se que este volume não é o volume real do corpo do animal, mas um volume aparente, uma vez que apenas a visão superior do animal foi obtida.

Os descritores de formas utilizados no presente estudo foram: **[1]** Excentricidade, é uma medida de arredondamento e é estimada como a razão entre os focos e o eixo principal das elipses que possui os mesmos segundos momentos que o animal. **[2]** Coeficiente de determinação do ajuste linear para a curvatura posterior do animal. **[3]** Primeiros 20 descritores de Fourier. Descritor de Fourier é uma classe de descritores de imagem global normalmente utilizada para análise de formas e correspondência de imagens (Bowman et al., 2000; Zhang e Lu, 2002; Burger e Burge, 2016). Para robustez, cada imagem segmentada foi girada, centralizada e transformada para a representação do plano polar anterior à estimativa da transformada discreta de Fourier em uma dimensão. Assim, os descritores de Fourier para a imagem polar podem ser estimados através de uma transformação discreta de Fourier, conforme descrito na equação [6].

$$FD_k = \sum_{j=0}^{n-1} X_j w_n^{jk} \quad [6]$$

Onde FD_k é o k Descritor de Fourier para o sinal X de comprimento n (o número de colunas da imagem polar) e $w_n = e^{-(2\pi i)/n}$ com $i = \sqrt{-1}$ o imaginário constante. Como FD_0 corresponde à soma sobre o sinal, que é a área do objeto na imagem, os descritores de Fourier também foram divididos por FD_0 para invariância no tamanho do animal.

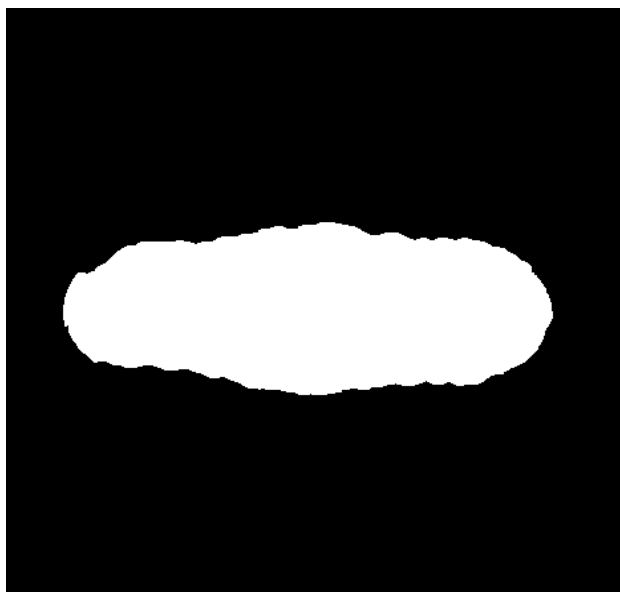


Figure 3. Frame preto e branco salvo após a extração das características da imagem.

2.3. Análise de dados

Para a análise estatística dos dados foi realizada Correlação de Pearson entre o peso corporal do animal e as características extraídas das imagens tridimensionais utilizando o PROC CORR do software SAS 9.3, as características com correlação abaixo de $P < 0,05$ foram retiradas do modelo. Assim, para as análises estatísticas foram utilizadas 15 variáveis explanatórias: Área e volume corporal, seis larguras e alturas ao longo do dorso do animal e comprimento, o valor das correlações entre a variável resposta e as 15 variáveis explanatórias estão presentes na Figura 5.

2.4. Predição de peso corporal e o ganho médio diário

A predição de peso corporal e ganho médio diário foi realizada no Departamento de Ciência Animal da Universidade de Wisconsin-Madison, Estados Unidos. A predição do peso corporal foi desenvolvida utilizando quatro tipos diferentes métodos de predição. 1: *Artificial Neural Networks*, 2: *Partial Least Squares*, 3: Regressão LASSO e 4: Regressão Linear Múltipla. Foi realizado neste estudo uma validação cruzada de *Leave-One-Out* para comparar e avaliar quão robusto foram os modelos. O ganho médio diário foi calculado a partir do peso corporal predito nos quatro diferentes modelos desenvolvidos.

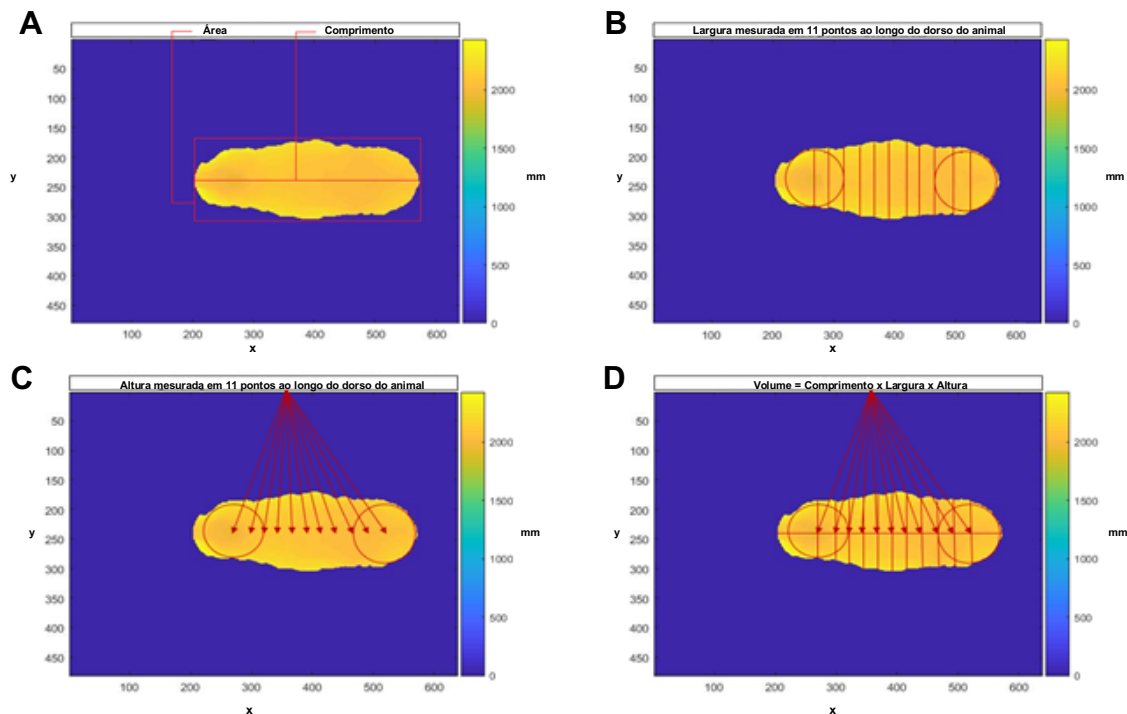


Figura 4. Extração das medidas corporais em locais equidistantes na parte dorsal do animal, dos ombros até a garupa. A: Mensuração de área e comprimento, B: 11 mensurações de largura, C: Altura medida em 11 pontos ao longo do dorso e D: volume corporal. Onde X e Y corresponde as coordenadas de cada *pixel* no plano cartesiano e mm representa a distância em milímetros da câmera a cada ponto de *pixel* da imagem.

2.5. Regressão linear múltipla

Quando temos uma variável resposta (Y) e variáveis explanatórias ($X_1, X_2, \dots, X_k$), podemos utilizar a regressão linear múltipla, onde o modelo é descrito na Equação [7].

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_kX_k + \varepsilon \quad [7]$$

Onde Y é a variável resposta e $X_1, X_2, \dots, X_k$ são as variáveis exploratórias; ε é o vetor de erros (resíduos); $B_0, B_1, \dots, B_k$ são os coeficiente do modelo.

A RLM foi realizada juntamente com a validação cruzada Leave-One-Out no software R utilizando a função TrainControl, método "LOOCV" do pacote Caret (<https://cran.r-project.org/web/packages/caret/caret.pdf>), onde Y consistia no peso corporal do animal e as variáveis independentes X foram as características obtidas através do Kinect®.

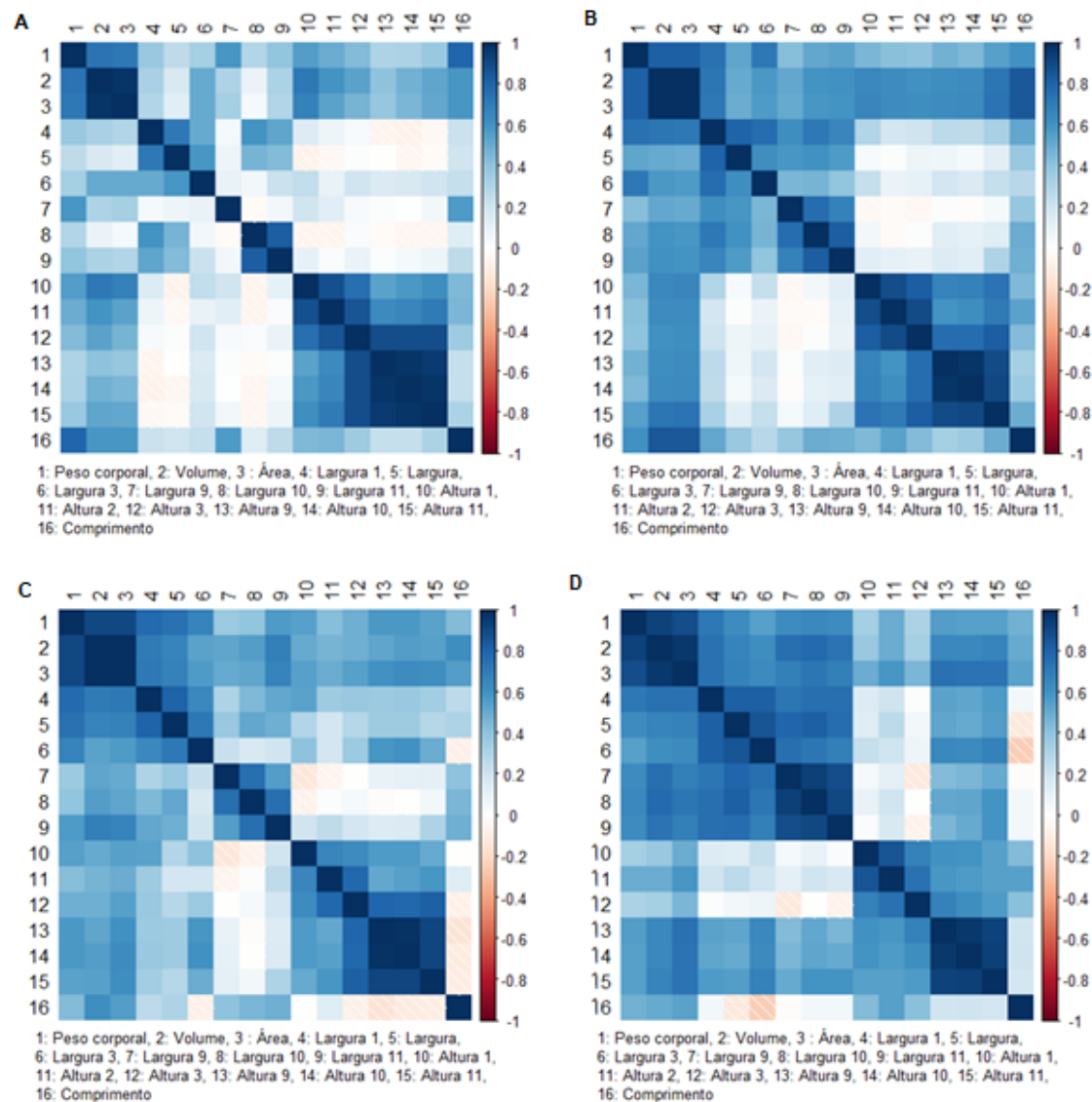


Figura 5. Correlação entre a variável resposta e as variáveis explanatórias para peso corporal obtidas através do sensor Kinect®. A: Desmama, B: Recria, C: Início da Terminação e D: Final da Terminação.

2.6. Regressão LASSO

O modelo LASSO ajusta os parâmetros por encolhimento, o que faz uma penalização nas variáveis menos significativas do modelo. Esta é uma forma de regressão regularizada ou “penalizada”, na qual a regularização L1 (lasso) é introduzida no procedimento de regressão linear múltipla padrão, usando uma função de custo composto para otimizar os coeficientes de regressão. Este modelo de regressão é indicado para situações que envolvam maior número de variáveis independentes correlacionadas (Hans, 2009), como observado no presente estudo. A regressão LASSO foi performada juntamente com a validação cruzada Leave-One-Out utilizando a função (cv.glmnet) do pacote

“glmnet” (Friedman et al., 2010) implementado no software R. Neste pacote é implementado um algoritmo que otimiza parâmetros de ajuste de modelo usando um algoritmo de descida de coordenadas. A descrição dos modelos LASSO desenvolvidos pode ser encontrada na Tabela 3.

2.7. Regressão *Partial Least Squares*

Uma regressão por mínimos quadrados parciais (PLS – “*Partial Least Squares*”) foi realizada com o JMP Pro 13 (SAS Institute Inc., Cary, NC) para prever o peso corporal, usando o algoritmo NIPALS da variável preditiva. O PLS extrai sucessivas combinações lineares dos preditores, denominados fatores, que abordam tanto a explicação da variação da resposta quanto a explicação da variação do preditor. O número de fatores retidos no modelo foi determinado usando uma validação cruzada *Leave-One-Out*, com a média prevista da raiz da PRESS calculada. O modelo que atingiu o menor valor de PRESS na validação cruzada foi escolhido. A descrição dos modelos PLS desenvolvidos podem ser encontradas na Tabela 3.

Tabela 3. Modelos desenvolvidos para ANN, PLS e LASSO

Item	Fases				
	Desmama	Recria	Início Terminação	Final Terminação	Completa
ANN					
1ª Unidade de Camadas Neurais	15	15	15	15	15
2º Tipo de Camadas de Neurais	Maxout 150	Maxout 150	Maxout 150	Maxout 150	Maxout 150
Epochs	20	20	20	20	20
L1 ¹	0,000069	0,000069	0,000069	0,000069	0,000069
L2 ²	0,000028	0,000028	0,000028	0,000028	0,000028
Proporção de dropout de entrada	0	0	0	0	0
Razão de dropout oculta	0	0	0	0	0
PLS³					
Numero de Fatores	2	3	3	14	6
LASSO					
Lambda	1,35	1,50	0,62	0,41	0,35

¹ L1: Regularização Lasso.² L2: Regularização Ridge.³ NIPALS método: Algoritmo de intervalo Parcial Interativo Não Linear utilizado para calcular os componentes da regressão PLS.

2.8. Artificial Neural Network

Redes neurais artificiais (ANN – “*Artificial Neural Network*”) foram adaptadas usando o software de código aberto para análise de *big data* chamado H2O (<https://cran.r-project.org/web/packages/h2o/h2o.pdf>), que pode ser utilizado através do software estatístico R (R Core Team, 2016). Os hiperparâmetros utilizados em *machine learning* podem afetar como o algoritmo se ajusta aos dados. Por exemplo, em *deep learning neural networks* (DLNN), o número de camadas, o número de neurônios em cada camada, a taxa de aprendizado e outros recursos são parâmetros de ajuste chamados hiperparâmetros. Como existem muitos parâmetros que podem afetar a precisão do modelo, o ajuste dos hiperparâmetros é essencial para o DLNN (Dórea et al., 2018).

Assim, neste trabalho foi utilizado uma combinação de hiperparâmetros conforme descrito por Dórea et al., 2018. Informações mais detalhadas sobre os hiperparâmetros definidos e critérios de busca podem ser encontrados na Tabela 4. Após a busca da rede neural, os hiperparâmetros para o melhor modelo foram escolhidos para cada modelo apresentado na Tabela 3. Cada modelo com suas respectivas estruturas de ANN foi treinado através da validação cruzada de *Leave-One-Out*. A função `h2o.deeplearning` foi usada para treinar os modelos. O ADADELTA foi o algoritmo de taxa de aprendizado adaptativo utilizado em todas as ANN (Zeiler, 2012).

2.9. Fator de inflação da variância

A colinearidade foi avaliada através do teste de fator de inflação da variância (VIF – “*Variation Inflation Factor*”) para cada variável explanatória utilizando a função (`vif`) do pacote “car” (<https://cran.r-project.org/web/packages/car/car.pdf>) implementado no software R, onde o modelo estatístico é descrito na fórmula 8.

$$VIF = \left[\frac{1}{1 - R_i^2} \right] \quad [8]$$

Onde R_i^2 é o coeficiente de determinação da regressão de cada variável explanatória.

Tabela 4. Hiperparâmetros e critérios de pesquisa utilizados na pesquisa de ANN. Os hiperparâmetros foram combinados para encontrar a melhor estrutura de ANN.

Pesquisa de grade: Hiperparâmetros e critérios de pesquisa				
Hiperparâmetros				
Função de Ativação	Camadas Ocultas	Prop. Drop. Ent ¹	L1 ²	L2 ³
Rectifier	2	0	0	0
Tanh	8	0,25	0,0001	0,0001
Maxout	20		0,0000	0,000001
RectifierWithDropout	40			
TanhWithDropout	50			
MaxoutWithDropout	60			
	80			
	100			
	110			
	120			
	130			
	150			
	20,20			
	30,30,30			
	25,25,25,25			
Pesquisa de rede				
estratégia	max_runtime_secs ⁵	max_models ⁶	stopping_metric ⁷	
RandomDiscrete ⁴	360	100	RMSE	

¹Proporção de dropout de entrada.

²L1 = Regularização Lasso.

³L2 = Regularização Ridge.

⁴RandomDiscrete = Obtenha uma pesquisa aleatória de todas as combinações dos seus hiperparâmetros.

⁵max_runtime_secs = Tempo de execução máximo em segundos para toda a pesquisa de grade.

⁶max_models = Número máximo de modelos pesquisados na grade.

⁷stopping_metric = Função utilizada para parada antecipada com base em nenhuma melhoria na métrica do modelo (neste caso, métrica definida como erro quadrático médio).

2.10. Avaliação do modelo

Para a avaliação, o sistema de avaliação do modelo foi conduzido conforme descrito por Tedeschi (2006). O coeficiente de determinação (R^2) foi

utilizado entre os valores preditos e observados para definir a precisão das equações desenvolvidas. O coeficiente de correlação de concordância (CCC, Lin 1989), viés médio e raiz do erro quadrático médio de predição (RMSEP – “*Root Mean Squares Error Prediction*”) foram calculados como uma medida de precisão. O coeficiente de correlação de concordância (CCC – “*Concordance Coefficient Correlation*”) foi estimado pela Equação [9] (Lin, 1989).

$$\hat{\rho}_c = \hat{\rho} x C_b \quad [9]$$

Onde componente (ρ) é a estimativa de correlação (r) que mede a precisão. O segundo componente (C_b) é o fator de correção de polarização que representa o quanto a linha de regressão se desvia da inclinação da unidade (45°).

O viés médio (MB – “*Mean Bias*”) é provavelmente a estatística mais utilizada e antiga para avaliar a precisão do modelo. É calculado pelo desvio médio entre os valores preditos e observado (Cochran e Cox, 1957). O viés médio representa o erro na tendência central (para avaliar se os valores preditos estão acima ou abaixo do predito), o viés da inclinação representa erros devidos à regressão, e o componente aleatório é a variância inexplicada que não pode ser explicada pela linear regressão (Tedeschi, 2006).

A raiz do erro quadrático médio de predição ou equação de predição de erro quadrático médio consiste na raiz quadrada da predição de erro quadrático médio (MSEP – “*Mean Square Error Prediction*”). Onde MSEP é a soma da diferença quadrática entre os valores observados e os valores preditos do modelo divididos pelo número de pontos de dados n (Bibby e Toutenburg, 1977). O MSEP foi decomposto em viés de média, viés de inclinação e erros aleatórios (Theil, 1961) para fornecer uma indicação da adequação do modelo para predição. Os modelos foram avaliados a partir do programa Model Evaluation System (MES, Colégio Station, TX; <http://nutritionmodels.tamu.edu/mes.html>) conforme descrito por Tedeschi (2006).

3. RESULTADOS E DISCUSSÃO

A média, intervalo (mínimo e máximo), desvio padrão e coeficiente de variação para o PC estão presentes na Tabela 1. O conjunto de dados variou consideravelmente, com PC ao desmame variando entre 114 a 266 kg, Recria de 138 a 277 kg, Início da Terminação de 274 a 391 kg e Final da Terminação de 359 a 554 kg. O método de ANN aumentou a precisão e acurácia para a predição de PC comparado aos outros métodos: (ANN vs. PLS vs. LASSO vs. RLM). Desmama: (CCC = 0.94 vs. 0.88 vs. 0.84 vs. 0.80; RMSEP = 8.6 vs. 12.6 vs. 14.1 vs. 16.4 kg; R^2 = 0.72 vs. 0.49 vs. 0.59 vs. 0.63). Recria: (CCC = 0.87 vs. 0.86 vs. 0.81 vs. 0.76; RMSEP = 11.4 vs. 12.2 vs. 13.6 vs. 16.3 kg; R^2 = 0.79 vs. 0.76 vs. 0.70 vs. 0.59). Início da Terminação: (CCC = 0.96 vs. 0.93 vs. 0.91 vs. 0.90; RMSEP = 7.7 vs. 10.2 vs. 11.2 vs. 12.4 kg; R^2 = 0.92 vs. 0.87 vs. 0.84 vs. 0.86). Os resultados apresentados estão presentes na Tabela 5.

O GMD foi calculado a partir da predição do PC. O GMD é uma importante ferramenta para a pecuária, pois com esta informação é possível acompanhar o crescimento animal, resultando em melhores estratégias no manejo nutricional e reprodutivo do animal. Assim, como para PC, o uso do método preditivo de ANN aumentou a precisão e acurácia das predições de GMD como evidenciado na tabela 6. O GMD foi mensurado em fases: 1: Desmame - Recria, 2: Desmame – Início da Terminação, 3: Desmame – Final da Terminação, 4: Recria – Início da Terminação, 5: Recria – Final da Terminação e 6: Início da Terminação – Final da Terminação. Os resultados obtidos para 1, 2, 3, 4, 5 e 6 foram: ANN (CCC = 0.74, 0.89, 0.84, 0.54, 0.82, 0.90; RMSEP = 0.02, 0.02, 0.03, 0.10, 0.06, 0.09 kg/d e R^2 = 0.67, 0.85, 0.80, 0.51, 0.76 e 0.82, respectivamente), PLS (CCC = 0.71, 0.86, 0.90, 0.67, 0.91, 0.90; RMSEP = 0.03, 0.03, 0.03, 0.10, 0.05, 0.10 kg/d, e R^2 = 0.52, 0.74, 0.82, 0.49, 0.85, 0.81, respectivamente), LASSO (CCC = 0.62, 0.80, 0.85, 0.60, 0.87, 0.85; RMSEP = 0.03, 0.04, 0.04, 0.11, 0.06, 0.12 kg/d e R^2 = 0.36, 0.67, 0.72, 0.40, 0.77, 0.73, respectivamente), RLM (CCC = 0.48, 0.74, 0.82, 0.46, 0.84, 0.78; RMSEP = 0.04, 0.05, 0.05, 0.14, 0.07, 0.16 kg/d, e R^2 = 0.23, 0.56, 0.68, 0.22, 0.71, 0.62, respectivamente).

A análise de regressão múltipla é uma das técnicas estatística mais utilizadas na análise de dados para investigar e modelar a relação entre variáveis resposta (Y) e as variáveis explanatórias (X_1 , X_2 , ..., X_k). Pode-se citar inúmeras aplicações de análise de regressão em diversas áreas (Lyles & Kupper 1997,

Friedman et al., 2010, Chen & Wang 2004). A regressão LASSO é um método comumente utilizado por reduzir os efeitos dos atributos que não contribuem para a identificação da variável resposta, reduzindo seus coeficientes para zero e excluindo-os do modelo (Tibshirani, 1997). Assim este tipo de modelo de regressão busca reduzir o *overfitting* através do encolhimento das variáveis explanatórias (X) pois utiliza valores absolutos na função de penalidade. Já o método preditivo PLS foi projetado como uma interseção entre a modelagem linear e a análise de componentes principais: um modelo linear capaz de determinar quais variáveis preditoras estavam relacionadas à variável resposta e lidar adequadamente com a multicolinearidade (Wold, 2004). No entanto, presume-se que o método de predição PLS apresente limitações devido à linearidade, bem como a regressão linear múltipla e a regressão LASSO. ANN pode apresentar vantagens por modelar de forma mais eficaz as relações potencialmente não-lineares entre variáveis de entrada e saída, assim as abordagens de machine learning são uma escolha poderosa e flexível (Peraí et al., 2010; Lin et al., 2012). Assim como Dórea et al. (2018), foi adotado uma estratégia para reduzir o overfitting em ANN com o uso de regularização Lasso (L1) e Ridge (L2). No ANN, a regularização de L1, também conhecida como rede flexível, restringe o valor absoluto dos pesos, levando a um vetor escasso durante a otimização (ou seja, muito próximo de zero) que tem efeito sobre a desistência (Hastie et al., 2001; e Zou e Hastie, 2005). A regularização de L2 restringe a soma dos pesos quadrados, também chamados de decaimento de peso, o que é análogo à regressão ridge usada para modelos lineares (Hastie et al., 2001; Zou e Hastie, 2005). No presente estudo, combinamos os dois métodos de penalização (L1 e L2) como forma de controlar o overfitting (Tabela 3). Para todas as fases de crescimento animal o ANN apresentou “Maxout” como função de ativação com 150 número de unidades (neurônios) nas camadas ocultas (Tabela 3). O método de predição ANN apresentou capacidade preditiva superior ao PLS, indicando que a relação entre PC, GMD e algumas variáveis obtidas através do Kinect podem ser não-lineares.

Nosso estudo mostrou que para estimar o PC de bovinos de corte é importante utilizar diferentes modelos de predição, pois pode apresentar colinearidade entre as variáveis. Assim, a fim de investigar a colinearidade entre as variáveis explanatórias foi determinado que o valor de VIF ≥ 10 fossem

excluídas do modelo e então a capacidade preditiva do modelo de RLM avaliado novamente, conforme tabela 7. Além do teste de VIF, ao analisar a matriz de correlação entre as variáveis explanatórias (Figura 5), fica evidente a alta correlação entre as alturas e/ou larguras, mostrando que este teste é de extrema importância para diminuir o erro de predição. Este estudo mostrou que ao utilizar a câmera de infravermelho pode-se obter mensurações biométricas em tempo real de um determinado animal ou categoria para fins de predição de peso corporal, pois foi encontrado uma boa correlação entre a variável resposta e as variáveis explanatórias. Para realizar a predição de PC e avaliar a capacidade preditiva do modelo, é necessário segmentar o banco de dados de acordo com a fase de crescimento animal, sexo, idade. Quando não há uma segmentação no banco de dados os valores preditos tendem a ficar mais próximos aos valores observados (figura 6), aumentando assim o valor do R^2 . Gomes et al. (2016) relataram que quando as mensurações biométricas de bovinos Angus e Nelore foram analisados separadamente, a largura do peito foi correlacionada com PC para ambas as raças, entretanto, outras medidas corporais como: largura do tórax e área dorsal, não apresentaram o mesmo comportamento apresentando maiores correlações para uma raça do que para a outra.

Tabela 5. Parâmetros de avaliação de predições de peso corporal em diferentes fases e modelos preditivos

Item	Inclinação (±SE)	H0: b = 1	R²	Viés	CCC	RMSEP (Kg)	RMSEP / Média (%)	Decomposição MSEP (%)		
				médio (Kg)				Viés médio	Viés inclinação	Erro aleatório
Desmama										
ANN	0,82 (±0,03) ¹	0,01	0,91	0,09	0,94	8,63	4,26	0,01	31,49	68,50
PLS	0,79 (±0,05) ¹	0,01	0,79	0,00	0,88	12,61	6,23	0,00	20,91	79,09
LASSO	0,72 (±0,05) ¹	0,02	0,73	-0,41	0,84	14,18	7,00	0,69	24,48	71,43
RLM	0,74 (±0,06) ¹	0,04	0,65	-0,23	0,80	16,49	8,15	0,02	18,61	81,37
Recria										
ANN	0,75 (±0,04) ¹	0,01	0,79	-0,37	0,87	11,40	5,32	0,11	29,74	70,15
PLS	0,76 (±0,05) ¹	0,06	0,76	0,00	0,86	12,20	5,70	0,00	23,80	76,20
LASSO	0,66 (±0,05) ¹	0,01	0,70	-0,12	0,81	13,63	6,37	0,01	38,94	61,05
RLM	0,69 (±0,07) ¹	0,01	0,59	-0,09	0,76	16,39	7,66	0,00	21,20	78,80
Início										
Terminação										
ANN	0,91 (±0,03)	0,07	0,92	-0,97	0,96	7,78	2,32	1,57	11,06	87,37
PLS	0,87 (±0,04) ¹	0,04	0,87	0,00	0,93	10,26	3,06	0,00	12,57	87,43
LASSO	0,83 (±0,04) ¹	0,07	0,84	0,04	0,91	11,26	3,36	0,00	17,31	82,69
RLM	0,86 (±0,05)	0,01	0,86	-0,01	0,90	12,40	3,70	0,00	9,92	90,08
Final										
Terminação										
ANN	0,85 (±0,05) ¹	0,01	0,85	-1,26	0,91	18,14	4,48	0,48	13,46	86,06
PLS	0,90 (±0,04)	0,03	0,90	0,00	0,95	14,18	3,15	0,00	9,12	90,88
LASSO	0,85 (±0,05) ¹	0,01	0,79	-0,63	0,90	19,83	4,41	0,10	12,28	87,62
RLM	0,94 (±0,08)	0,49	0,78	1,73	0,88	23,44	5,21	0,55	1,00	98,45

¹Teste de inclinação (b = 1, P > 0,05)

Tabela 6. Parâmetros de avaliação de predições de ganho médio diário em diferentes fases e modelos preditivos

Item	Inclinação (±SE)	H0: b = 1	R ²	Viés médio (Kg/d)	CCC	RMSEP (Kg/d)	RMSEP / Média (%)	Decomposição MSEP (%)		
								Viés médio	Viés inclinação	Erro aleatório
D – R¹										
ANN	0,74 (±0,06) ⁷	0,02	0,67	0,00	0,81	0,02	37,03	0,12	20,15	79,73
PLS	0,62 (±0,07) ⁷	0,01	0,52	0,00	0,71	0,03	55,55	0,00	29,59	70,41
LASSO	0,49 (±0,07) ⁷	0,01	0,36	0,00	0,62	0,03	55,55	0,03	41,63	58,34
RLM	0,44 (±0,10) ⁷	0,01	0,23	0,00	0,48	0,04	74,07	0,05	52,13	47,82
D – IT²										
ANN	0,89 (±0,04) ⁷	0,03	0,85	0,00	0,92	0,02	5,68	1,11	7,19	91,70
PLS	0,83 (±0,06) ⁷	0,09	0,74	0,00	0,86	0,03	7,85	0,00	10,75	89,25
LASSO	0,74 (±0,06) ⁷	0,03	0,65	0,00	0,80	0,04	10,47	0,08	19,03	80,89
RLM	0,73 (±0,08) ⁷	0,03	0,56	0,00	0,74	0,05	13,08	0,01	13,63	86,36
D – FT³										
ANN	0,84 (±0,06) ⁷	0,01	0,80	0,01	0,89	0,03	6,39	0,11	12,09	87,80
PLS	0,87 (±0,05) ⁷	0,03	0,82	0,00	0,90	0,03	6,39	0,16	8,89	90,95
LASSO	0,80 (±0,07) ⁷	0,00	0,72	0,00	0,85	0,04	8,52	0,08	13,62	86,30
RLM	0,91 (±0,09)	0,33	0,68	0,00	0,82	0,05	10,66	0,47	2,02	97,51
R – IT⁴										
ANN	0,54 (±0,06) ⁷	0,01	0,51	0,00	0,69	0,10	11,01	0,20	42,47	57,33
PLS	0,52 (±0,06) ⁷	0,01	0,49	0,02	0,67	0,10	11,01	0,00	44,04	55,96
LASSO	0,46 (±0,07) ⁷	0,01	0,40	0,01	0,60	0,11	12,11	0,01	48,30	51,69
RLM	0,40 (±0,09) ⁷	0,01	0,22	0,00	0,46	0,14	15,41	0,00	37,57	62,43
R – FT⁵										
ANN	0,82 (±0,06) ⁷	0,01	0,76	0,00	0,87	0,06	9,28	0,29	13,22	86,49
PLS	0,83 (±0,05) ⁷	0,02	0,85	0,00	0,91	0,05	6,63	0,19	18,07	81,74
LASSO	0,79 (±0,06) ⁷	0,02	0,77	0,00	0,87	0,06	7,96	0,58	18,65	80,77
RLM	0,90 (±0,08) ⁷	0,28	0,71	0,01	0,84	0,07	7,71	2,08	2,45	95,47
IT – FT⁶										
ANN	0,90 (±0,06)	0,13	0,82	0,00	0,90	0,09	14,04	0,00	4,80	95,20
PLS	0,92 (±0,06)	0,24	0,81	0,05	0,90	0,10	15,60	0,08	3,97	95,95
LASSO	0,87 (±0,07) ⁷	0,12	0,73	0,00	0,85	0,12	18,72	0,01	5,01	94,98
RLM	0,93 (±0,10) ⁷	0,55	0,62	0,01	0,78	0,16	24,96	1,06	0,75	98,19

¹Desmama – Recria.²Desmama – Início da Terminação.³Desmama – Final da Terminação.⁴Recria – Início da Terminação.⁵Recria – Final da Terminação.⁶Início da Terminação – Final de Terminação.⁷Teste de inclinação (b = 1, P > 0,05).

A validação cruzada pode ser utilizada para avaliar a capacidade preditiva de diferentes modelos de predição. No estudo de Gomes et al. (2016) não foi realizada validação dos dados e então o alto valor de R^2 é inerente ao modelo e não a validação como apresentado em nosso estudo. Assim não se sabe se os modelos possuem boa capacidade preditiva. Com isso é importante a utilização da validação para avaliar os métodos de predição. É importante notar que a validação cruzada realizada neste trabalho foi com animais do mesmo rebanho e obtiveram o mesmo regime ambiental e alimentar, isto faz com que a validação seja mais crítica e ajuda a reduzir ou evitar *overfitting* no modelo. Portanto, é necessário ter cuidado como a aplicação desses modelos será em outro conjunto de dados. Os resultados obtidos mostram que é de suma importância utilizar diferentes métodos de predição conjuntamente com métodos de validação para poder avaliar a capacidade preditiva dos modelos em prever o PC a partir de imagens tridimensionais.

Tabela 7. Parâmetros de avaliação de predições de peso corporal em diferentes fases e modelos, utilizando a avaliação de VIF para o modelo preditivo RLM

Avaliação de VIF para o modelo preditivo RLM										
Item	Inclinação (±SE)	H0: b = 1	R²	Viés médio (Kg)	CCC	RMSEP (Kg)	RMSEP / Média (%)	Decomposição MSEP (%)		
								Viés médio	Viés inclinação	Erro aleatório
Desmama										
RLM	0,74 (±0,06) ¹	0,04	0,65	-0,23	0,80	16,49	8,15	0,02	18,61	81,37
Desmama (VIF)										
RLM	0,74 (±0,08) ¹	0,03	0,67	-0,46	0,81	16,00	7,90	0,09	19,50	80,41
Recria										
RLM	0,69 (±0,07) ¹	0,01	0,59	-0,09	0,76	16,39	7,66	0,00	21,20	78,80
Recria (VIF)										
RLM	0,70 (±0,06) ¹	0,08	0,63	-0,15	0,79	15,19	7,10	0,01	22,91	77,08
Início Terminação										
RLM	0,86 (±0,05)	0,01	0,86	-0,01	0,90	12,40	3,70	0,00	9,92	90,08
Início Terminação (VIF)										
RLM	0,82 (±0,05) ¹	0,03	0,76	0,10	0,87	14,23	4,25	0,01	13,14	86,85
Final Terminação										
RLM	0,94 (±0,08)	0,49	0,78	1,73	0,88	23,44	5,21	0,55	1,00	98,45
Final Terminação (VIF)										
RLM	0,75 (±0,07) ¹	0,03	0,66	0,03	0,81	27,68	6,15	0,00	17,32	82,68

¹Teste de inclinação (b = 1, P > 0,05)

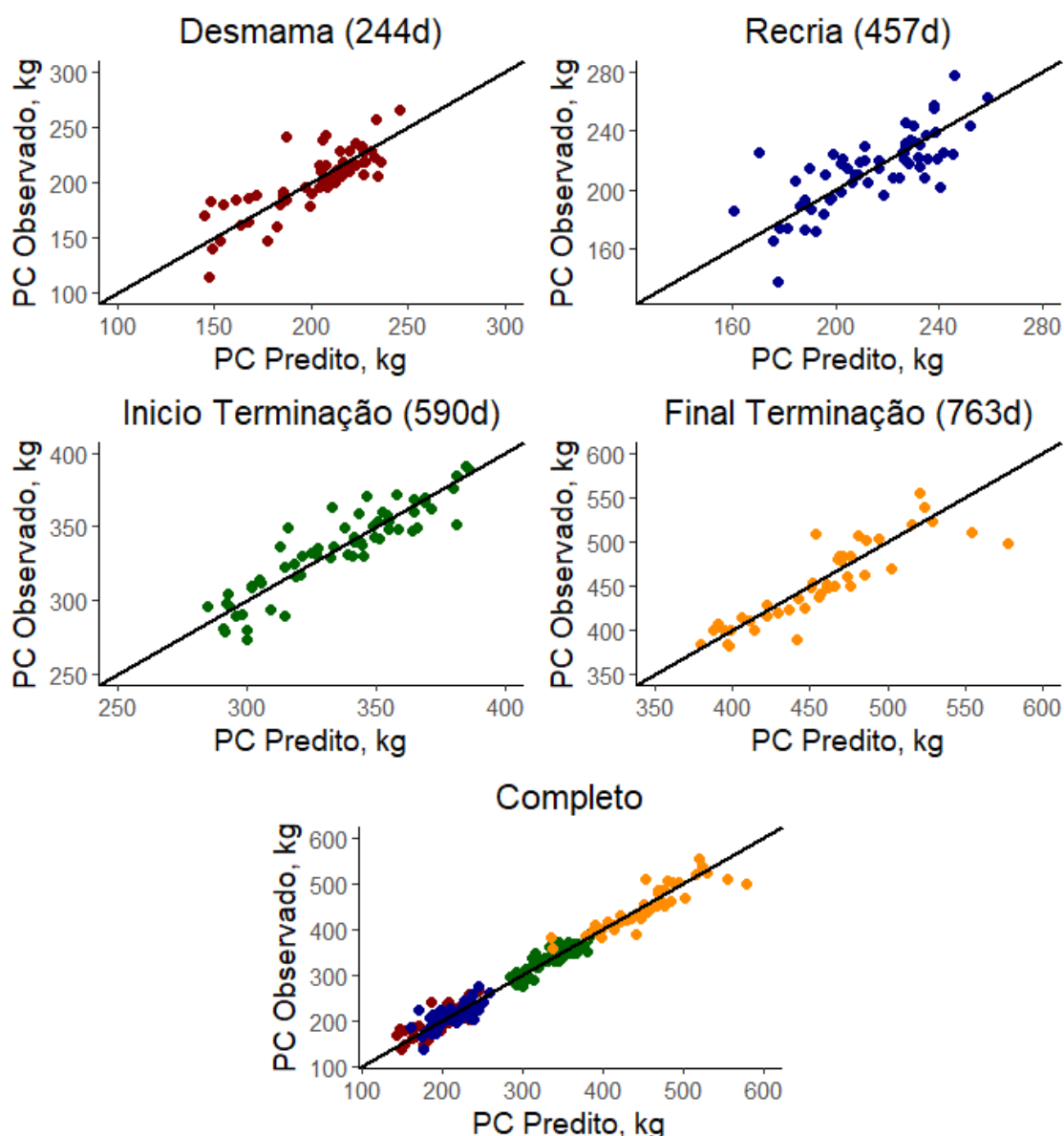


Figura 6 Gráfico gerado a partir dos valores de PC Observado vs. Predito com ou sem segmentação do banco de dados por fase do crescimento animal.

4. CONCLUSÃO

O uso de imagens tridimensionais foi capaz de prever o peso corporal com boa acurácia e precisão e consequentemente o ganho médio diário em bovinos Nelore e, assim, monitorar a curva de crescimento dos animais em tempo real auxilia no monitoramento da saúde do rebanho, além de reduzir a tensão gerada pela pesagem por métodos convencionais, como o método convencional de pesagem em balanças e, assim, proporcionar maior lucratividade à produção.

5. REFERÊNCIAS

ATHERTON, T. J. & KERBYSON, D. J. Size in variant circle detection. **Image and Vision Computing**, v.17, 795–803, 1999.

BIBBY, J., & TOUTENBURG H. Prediction and Improved Estimation in Linear Models. **John Wiley and Sons**, Berlin, Germany. 1977.

BISHOP, C. M. 2006. Pattern Recognition and Machine Learning. S 432 pringer, New York, New York, USA.

BOWMAN, E.T.; SOGA, K.; & DRUMMOND, W. 2016. Fourier Shape Descriptor, pp. 665 - 711.

BURGER E.T. & BURGE, M. J. Particle shape characterization using Fourier descriptor analysis, **Géotechnique**, v.51, 545-554, 2016.

Chen, Z. & Wang, Y. G. (2004), 'Efficient Regression Analysis with Ranked-Set Sampling', **Biometrics** 60(4), 997–1004.

COCHRAN, W. G. & COX, G. M. 1957. **Experimental Design**, 2nd edition. John Wiley and Sons, New York, NY, USA.

DICKSON, R. A.; MORTON J. M.; BEGGS D. S.; ANDERSON G. A.; PYMAN M. F.; MANSELL P. D; BLACKWOOD C. B. An automated walk-over weighing system as a tool for measuring live weight change in lactating dairy cows. **Journal of Dairy Sci.** v.96, 4477–4486, 2013.

DÓREA, J. R. R.; WELD, K.; ARMENTANO, L. E. Mining data from milk infrared spectroscopy to improve feed intake predictions in lactating dairy cows. **Journal of Dairy Sci.** 2018.

FRIEDMAN, J.; HASTIE, T.; & TIBSHIRANI R. Regularization paths for generalized linear models via coordinate descent. **Journal Statistics Software** v.33, 1–22, 2010.

GIANOLA, D. H.; OKUT, K. A.; WEIGEL, ROSA, G. J. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. **BMC Genetics** v.12, 87-99, 2011.

GOMES, R. A.; MONTEIRO, G. R., ASSIS G. J. F., BUSATO, K. C., LADEIRA, M. M., and CHIZZOTTI, M. L. Technical note: Estimating body weight and body composition of beef cattle through digital image analysis. **Journal of Animal Science**. 2016. 94:5414-5422. doi:10.2527/jas.2016-0797

HANS, C. Bayesian lasso regression. *Biometrika*, 96(4):835–845. URL <http://biomet.oxfordjournals.org/content/96/4/835.abstract>.doi:10.1093/biomet/as p047. 2009.

Hough.P. V. C, Method and means for recognizing complex patterns, U. S. Patent, 4746348, 1962.

KASHIHA, M.; BAHR, C; OTT, S.; MOONS, C. P.; NIEWOLD, T. A.; ÖDBERG, F. O.; BERCK-MANS, D. Automatic weight estimation of individual pigs using image analysis. **Computers and Electronics in Agriculture**. v.107, 38-44. 2014.

HASTIE, T.; R. TIBSHIRANI, and J. FRIEDMAN. 2001. The Elements of Statistical Learning: Data Mining, Inference and Prediction. **Springer**. New York, New York, USA.

KONSGRO, J. Estimation of pig weight using a Microsoft Kinect prototype imaging system. **Computers and Electronics in Agriculture**. 109:32–35. doi:10.1016/j.compag.2014.08.008. 2014.

LI, Z.; LUO C.; TENG G.; & LIU T. 2014. Estimation of pig weight by machine vision: A review. In: D. LI and Y. CHEN, editor, **Computer and computing technologies in agriculture VII: 7th IFIP WG 5.14 International Conference, CCTA 2013, Beijing, China, September 18–20, 2013, Revised Selected Papers, Part II**. Springer Berlin Heidelberg, Berlin, Germany. p. 42–49. doi:10.1007/978-3-642-54341-8_5

LIN, L. I. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* v.45, 255–268, 1989.

LIN, M. I. B.; W. A. GROVES, A.; FREIVALDS, E. G.; LEE, and M. HARPER. 2012. Comparison of artificial neural network (ANN) and partial least squares (PLS) regression models for predicting respiratory ventilation: an exploratory story. **European Journal of Applied Physiology**, 112, 1603-1611. <https://dx.doi.org/10.1007/s00421-011-2118-6>.

LYLES, R. H. & KUPPER, L. L. A Detailed Evaluation of Adjustment Methods for Multiplicative Measurement Error in Linear Regression with Applications in Occupational Epidemiology, **Biometrics** v.53, 1008–1025, 1997.

OZKAYA, S.; NEJA W.; KREZEL-CZOPECK, S.; OLER A. Estimation of bodyweight from body measurements and determination of body measurements on Limousin cattle using digital image analysis. **Animal Production Science**. 2015, doi:10.1071/AN14943

PERAI, A. H.; H. NASSIRI MOGHADDAM, S.; ASADPOUR, J.; BAHRAMPOUR, & MANSOORI, G. A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. *Poult. Sci.* v.89, 1562–1568, 2010. <https://dx.doi.org/10.3382/ps.2010-01005>.

R Core Team. 2016. R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

SCHOFIELD, C. P.; MARCHANT J. A.; WHITE R. P.; BRANDL N.; WILSON, M. Monitoring pig growth using a prototype imaging system. **Journal of Agricultural Engineering Research** 72:205–210. doi:10.1006/jaer.1998.0365. 1999.
TEDESCHI, L. O. 2006. Assessment of the adequacy of mathematical models. *Agric. Syst.* 89:225–247. <https://dx.doi.org/10.1016/j.agsy.2005.11.004>.

TEDESCHI, L. O. 2006. Assessment of the adequacy of mathematical models. *Agric. Syst.* 89:225–247. <https://dx.doi.org/10.1016/j.agsy.2005.11.004>.

THEIL, H. 1961. Economic forecasts and policy. Pages 6–48 in *Contributions to Economic Analysis*. R. Strotz, J. Tinbergen, P. Verdoorn, and H. J. Witteveen, ed. North-Holland Publishing Company, Amsterdam, the Netherlands.

TIBSHIRANI, R. Regression shrinkage and selection via the Lasso. **Journal of the Royal Statistical Society**. v.16, 385-395, 1997.

WANG, Y.; YANG W.; WINTER P.; and WALKER, L. Walk-through weighing of pigs using machine vision and an artificial neural network. **Biosystems Engineering** v.100, 117–125, 2008. doi:10.1016/j.biosystemseng.2007.08.008.

WOLD, H. 2004. Partial Least Squares. in Encyclopedia of Statistical Sciences. **John Wiley & Sons, Inc.**

ZHANG D. S. & LU. G. J. A Shape-based image retrieval using generic Fourier descriptor. **Signal Processing: Image Communication**. Doi: 10.1016/S0923-5965(02)00084-X

ZEILER, M. D. 2012. ADADELTA: an adaptive learning rate method. arXiv. 1212.5701:1:6. <http://arxiv.org/abs/1212.5701>.

ZOU, H.; and T. HASTIE. 2005. Regularization and variable selection via the elastic net. **Journal of the Royal Statistical Society**. 67:301–320.

CAPÍTULO 3 – USO DE IMAGENS BIOMÉTRICAS PARA PREDIÇÃO DO PESO CORPORAL E PESO DE CARÇA QUENTE DE BOVINOS NELORE

RESUMO: A variável mais importante mensurada na produção animal é o peso corporal (PC). Comumente a avaliação do ganho de peso corporal é realizada por meio de balança. No entanto, esse manejo pode causar estresse nos animais. Uma estratégia interessante é a coleta de imagens, um método que evita a contenção dos animais e pode ser uma alternativa rápida para estimar o PC. Os objetivos deste estudo foram 1) prever o PC e o peso de carcaça quente (PCQ) de bovinos Nelore por meio de imagens tridimensionais e 2) comparar diferentes modelos de predição. Quatrocentas e cinquenta imagens foram coletadas para predição de peso corporal e peso de carcaça quente divididas em quatro conjuntos experimentais. Os experimentos 1, 3 e 4 foram conduzidos no centro de pesquisa do Pólo Regional de Desenvolvimento Tecnológico dos Agronegócios da Alta Mogiana (PRDTA - Alta Mogiana), em Colina - SP, Brasil. Para o experimento 1 foi utilizado 48 bovinos machos quimicamente castrados da raça Nelore, com média de 24 ± 2 meses e peso vivo médio de $449,47 \pm 46,97$ kg. Para o experimento 2 foi utilizado 228 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio inicial de $588,49 \pm 35,53$ kg. O experimento 3 foi realizado no Confinamento Experimental do Departamento de Melhoramento e Nutrição Animal da Faculdade de Medicina Veterinária e Zootecnia, UNESP – Câmpus de Botucatu em que foram utilizados 83 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio de $516,69 \pm 38,00$ kg. Para o experimento 4 foi utilizado 91 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio de $589,49 \pm 30,10$ kg. Para capturar as imagens foi utilizado o sensor Kinect® modelo 1473 da Microsoft. Concomitantemente à aquisição de imagens, cada animal foi pesado em balança eletrônica. A análise de imagem foi dividida em três etapas: primeira, estimativa da distância do chão à câmera, segunda, uma etapa de segmentação e uma etapa final para extração das características. As características com correlação abaixo de $P < 0,05$ foram removidas do modelo. Foram desenvolvidos quatro métodos de predição: Regressão Linear Múltipla (RLM), *Least Absolute Shrinkage and Selection Operator* (LASSO), *Partial Least Squares* (PLS) e *Artificial Neural Networks*

(ANN). Os resultados das predições para peso corporal nos Set 1, 2, 3 e 4 foram: ANN (CCC = 0.73, 0.66, 0.70, 0.74; RMSEP = 19.6, 27.2, 27.2, 33.7 kg e R^2 = 0.58, 0.53, 0.53 e 0.59, respectivamente), PLS (CCC = 0.72, 0.65, 0.54, 0.69; RMSEP = 21.6, 28.0, 36.8, 39.3 kg, e R^2 = 0.55, 0.53, 0.30, 0.51, respectivamente), LASSO (CCC = 0.70, 0.60, 0.58, 0.76; RMSEP = 20.9, 28.3, 32.3, 40.6 kg e R^2 = 0.58, 0.53, 0.39, 0.50, respectivamente), RLM (CCC = 0.74, 0.65, 0.61, 0.67; RMSEP = 20.9, 27.9, 31.3, 42.6 kg, e R^2 = 0.58, 0.53, 0.39, 0.51, respectivamente). O uso do método de ANN melhorou a acurácia e precisão das predições de peso corporal para os conjuntos experimentais. Diferentemente da predição de peso corporal, o uso do método de regressão LASSO e RLM melhorou a precisão e acurácia das predições de peso de carcaça quente. Os resultados das predições para peso de carcaça quente nos Set 1, 2, 3 e 4 foram: ANN (CCC = 0.65, 0.19, 0.56, 0.64; RMSEP = 13.0, 30.0, 20.0, 23.2 kg e R^2 = 0.45, 0.17, 0.44 e 0.52, respectivamente), PLS (CCC = 0.60, 0.54, 0.51, 0.65; RMSEP = 14.6, 20.3, 23.0, 25.3 kg, e R^2 = 0.44, 0.42, 0.32, 0.45, respectivamente), LASSO (CCC = 0.67, 0.53, 0.55, 0.66; RMSEP = 13.1, 20.1, 20.8, 24.1 kg e R^2 = 0.49, 0.40, 0.46, 0.48, respectivamente), RLM (CCC = 0.14, 0.53, 0.40, 0.69; RMSEP = 8.75, 20.2, 24.6, 22.7 kg, e R^2 = 0.05, 0.40, 0.20, 0.52, respectivamente). O uso de câmeras tridimensionais prediz com boa acurácia o peso corporal e peso de carcaça quente em bovinos nelore.

Palavras-chave: análise de imagens, bovinocultura de corte, kinect®, modelos, visão computacional

1. INTRODUÇÃO

O peso corporal é uma característica econômica importante para a gestão de sistemas de produção de bovinos de corte, seu conhecimento permite avaliar o crescimento, o estado nutricional, ajustar o fornecimento de alimento e avaliar a condição de saúde animal para administrar o manejo de produtos veterinários adequadamente e estabelecer o valor de mercado (Polgar et al., 1997; Ulutas et al., 2001). No entanto, esse manejo pode causar estresse em animais (Schofield et al., 1999). A mensuração do peso de carcaça quente é uma característica de grande importância econômica para a pecuária de bovinos de corte, pois esta representa o retorno econômico para os produtores, visto que o produtor é remunerado pela indústria frigorífica de acordo com o peso de carcaça quente dos animais, assim um sistema capaz de prever este peso permite a tomada de decisões para aumentar a lucratividade do sistema produtivo.

Alternativamente, alguns autores (Wang et al., 2008; Kashiha et al., 2014) desenvolveram técnicas para obter mensurações biométrica de animais através de imagens digitais que provaram ser um método eficiente, Segundo Gomes et al. (2016), a coleta de imagens é um método que evita o contato com animais e pode ser uma alternativa rápida para estimar o peso corporal, devido à facilidade de instalação e movimentação de equipamentos. As balanças, entretanto, são relativamente caras e seus componentes eletrônicos são propensos a danos em ambiente hostil coberto com esterco e urina e em contato direto com vacas (Dickinson et al., 2013). Por isso, é necessário um sistema de pesagem automatizado de baixo custo e robusto.

Para estimativa de peso corporal e peso de carcaça quente é importante utilizar diferentes modelos de predição, pois o conjunto de dados pode apresentar uma multicolinearidade entre as variáveis. Uma alta multicolinearidade é problemática pois pode aumentar a variância dos coeficientes de regressão, tornando-os instáveis. A regressão linear múltipla é fundamental para fazer comparações entre a combinação de vários fatores e seus resultados. A regressão LASSO é um método de análise de regressão que executa tanto a seleção quanto a regularização de variáveis para aumentar a precisão e a capacidade de interpretação do modelo estatístico, o qual produz a melhoria do erro de predição diminuindo os grandes coeficientes de regressão para reduzir o *overfitting* através da contração das variáveis independentes. Ao

contrário de RLM e LASSO, o modelo de mensuração de PLS estima as variáveis latentes como uma soma ponderada de suas variáveis manifestas. O modelo estrutural estima as variáveis latentes por meio de regressão linear simples ou múltipla entre as variáveis latentes estimadas pelo modelo de mensuração. Este algoritmo é repetido até que a convergência seja alcançada. O uso de técnicas de aprendizado de máquina poderia melhorar a qualidade das previsões devido à sua capacidade de modelar relações complexas entre variáveis, como não-linearidades e interações (Gianola et al., 2011). Assim, o uso de ANN, por exemplo, poderia ser uma abordagem alternativa para o desenvolvimento de modelos preditivos. O modelo de redes neurais artificiais foram desenvolvidos com base no sistema nervoso humano e consistem em neurônios interconectados (lineares ou não-lineares) organizados em camadas (Bishop, 2006).

A hipótese deste trabalho é que o peso corporal e o peso de carcaça quente, de bovinos Nelore, podem ser preditos, por meio de imagens tridimensionais. Predições acuradas e precisas do peso corporal podem permitir o acompanhamento da curva de crescimento dos animais além da detecção de problemas sanitários no rebanho (desempenho sub-ótimo), em tempo real. O uso dessa ferramenta pode colaborar para tomadas de decisão que evitem a redução da rentabilidade ou prejuízos aos sistemas de produção animal. O objetivo deste estudo foi prever peso corporal e o peso de carcaça de bovinos Nelore por meio de imagens tridimensionais e determinar o melhor modelo de predição através de validação.

2. MATERIAL E MÉTODOS

2.1. Conjunto de dados

Quatrocentas e cinquenta imagens foram coletadas para predição de peso corporal e peso de carcaça quente divididas em quatro experimentos. Os experimentos 1, 3 e 4 foram conduzido no centro de pesquisa do Pólo Regional de Desenvolvimento Tecnológico dos Agronegócios da Alta Mogiana (PRDTA - Alta Mogiana), em Colina - SP, Brasil. Para o experimento 1 foi utilizado 48 bovinos machos quimicamente castrados da raça Nelore, com média de 24 ± 2

meses e peso vivo médio de $449,47 \pm 46,97$ kg. Para o experimento 2 foi utilizado 228 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio inicial de $588,49 \pm 35,53$ kg. O experimento 3 foi realizado no Confinamento Experimental do Departamento de Melhoramento e Nutrição Animal da Faculdade de Medicina Veterinária e Zootecnia, UNESP – Câmpus de Botucatu em que foram utilizados 83 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio de $516,69 \pm 38,00$ kg. Para o experimento 4 foi utilizado 91 bovinos machos não castrados da raça Nelore, com média de 22 ± 2 meses e peso vivo médio de $589,49 \pm 30,10$ kg.

Para capturar as imagens, utilizamos o sensor Kinect® modelo 1473 da Microsoft. O sensor foi posicionado no curral de manobra paralelo ao solo, como pode ser evidenciado na (Figura 1). O tempo de coleta das imagens RGB e infravermelho (640x480 pixels) foi de 15 segundos e 10 *frames* foram obtidos para cada animal. Concomitantemente à aquisição da imagem, cada animal foi pesado em uma balança eletrônica (Coimma LTDA, modelo RUDD 300, Dracena, São Paulo, Brasil), com capacidade mínima de 10 kg, máximo de 1.500 kg e precisão de 0,5 kg.

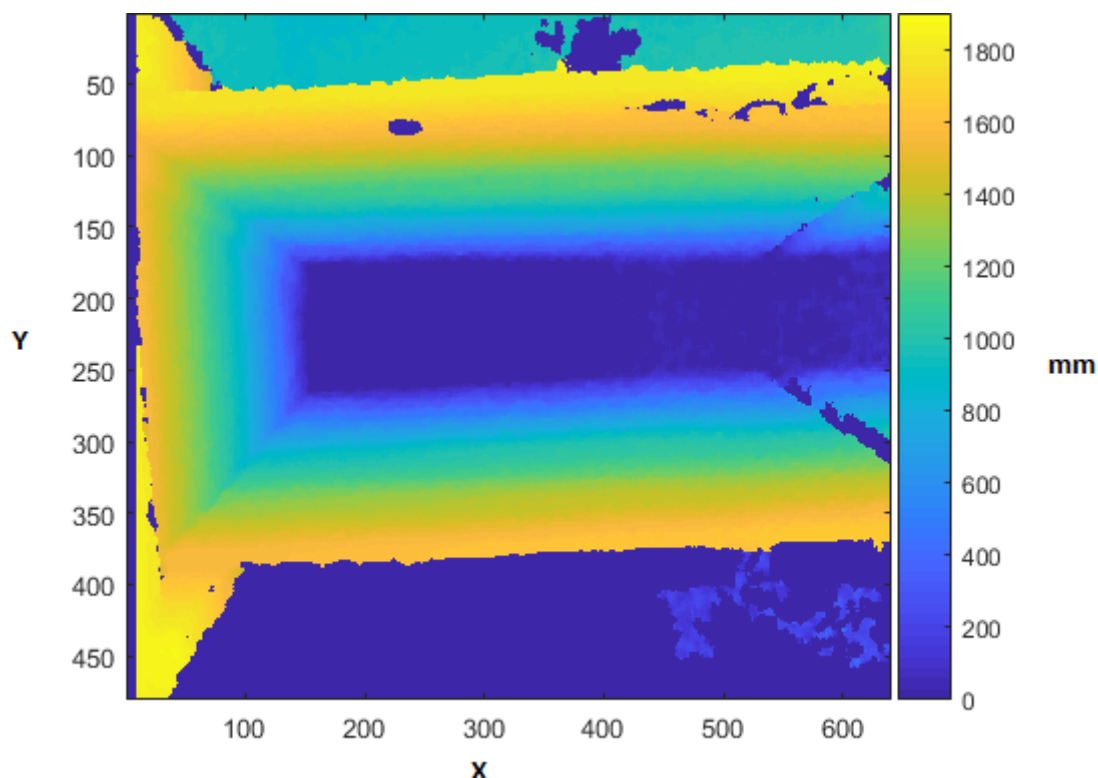


Figura 1. Imagem sem a presença do animal, utilizada para estimativa da distância da câmera ao chão sem a presença do animal, onde X e Y corresponde as coordenadas de cada *pixel* no plano cartesiano e mm representa a distância em milímetros da câmera a cada ponto de *pixel* da imagem.

2.2. Conjuntos Experimentais

Quatro alternativos conjuntos experimentais foram considerados para predição de peso corporal e peso de carcaça quente, RLM, LASSO, PLS e ANN, como apresentado na tabela 1. Os conjuntos experimentais considerados foram: Set 1 inclui o experimento 1, 2 e 3 com média de peso corporal de $553,30 \pm 63,08$ kg e média de $315,34 \pm 37,32$ kg de peso de carcaça quente para treino do modelo e o experimento 4 com média de peso corporal de $589,94 \pm 30,26$ kg e média de peso de carcaça quente de $334,22 \pm 16,79$ kg para validação do modelo de predição. Set 2 inclui o experimento 1, 2 e 4 com média de peso corporal de $532,18 \pm 65,90$ kg e média de $301,70 \pm 36,43$ kg de peso de carcaça quente para treino do modelo e o experimento 3 com média de peso corporal de $558,49 \pm 35,61$ kg e média de peso de carcaça quente de $336,17 \pm 23,17$ kg para validação do modelo de predição. Set 3 inclui o experimento 1, 3 e 4 com média de peso corporal de $570,67 \pm 59,30$ kg e média de $325,74 \pm 34,02$ kg de peso de carcaça quente para treino do modelo e o experimento 2 com média de peso corporal de $516,69 \pm 38,22$ kg e média de peso de carcaça quente de $290,08 \pm 22,19$ kg para validação do modelo de predição. Set 4 inclui o experimento 2, 3 e 4 com média de peso corporal de $573,99 \pm 45,59$ kg e média de $449,48 \pm 47,46$ kg de peso de carcaça quente para treino do modelo e o experimento 1 com média de peso corporal de $326,21 \pm 28,49$ kg e média de peso de carcaça quente de $260,12 \pm 28,10$ kg para validação do modelo de predição, como apresentado na tabela 2 para peso corporal e tabela 3 para peso de carcaça quente.

Tabela 1. Descrição dos diferentes conjuntos experimentais utilizados para predição de peso corporal e peso de carcaça quente

Conjuntos	Experimentos	
	Treino	Validação
Set 1	1, 2 e 3	4
Set 2	1, 2 e 4	3
Set 3	1, 3 e 4	2
Set 4	2, 3 e 4	1

2.3. Dispositivo Kinect®

O Kinect® (Microsoft Corporation, modelo 1473, Redmond, WA, USA) é um dispositivo de sensor de movimento, com aproximadamente 23 cm de comprimento, composto por uma câmera RGB e um sensor de profundidade

(projektor laser infravermelho e uma câmera infravermelha) que permite a varredura do ambiente em três dimensões. A câmera RGB e a câmera infravermelha têm uma resolução de 640 x 480 pixels e podem gerar um fluxo de imagem de 30 *frames* por segundo. O ângulo do campo de visão do sensor de profundidade é de 57° na horizontal e 43° na vertical e pode ser operado entre 0,8 m e 3,5 m (Corrêa, 2015).

2.4. Análise de imagens

A análise de imagem realizada foi uma adaptação do método automatizado desenvolvido por Fernandes et al., (comunicação pessoal). O método original foi criado para análise de vídeo, mas pode ser simplificado para imagens. Os 10 *frames* obtidos para cada animal foram divididos em três etapas: primeira, estimativa da distância do chão à câmera, segunda, uma etapa de segmentação e uma etapa final para extração das características. Para a análise estatística foi utilizado o melhor *frame* segmentado adequado.

2.4.1. Estimativa da distância do chão à câmera:

Nesta etapa, o algoritmo realizou uma estimativa da distância entre a câmera e o solo com base no modo do grupo de pixels mais distante da câmera, conforme descrito na Equação [1].

$$f_{dist} = mode(maxPeak) \quad [1]$$

Onde f_{dist} é a distância do chão à câmera e $maxPeak$ é o cluster de pixels que tinha uma distância da câmera de 90% ou mais (Figura 1).

2.4.2. Segunda segmentação:

A abertura biométrica da imagem foi realizada para avaliar a inclusão do animal na cena e sua conectividade com os objetos. Então, nesta etapa, todo objeto conectado ao animal foi removido. Em seguida, a posição do corpo do animal foi avaliada por um contorno da área dorsal do animal e estimativa da posição da coluna vertebral. Para a estimativa da posição da coluna vertebral,

apenas a área dorsal que estava entre o ombro e a garupa do animal foi utilizada. Ombro e garupa foram identificados através de uma adaptação da transformação de Hough para identificação de objetos redondos (Hough, 1962; Atherton e Kerbyson, 1999), (Figura 2),

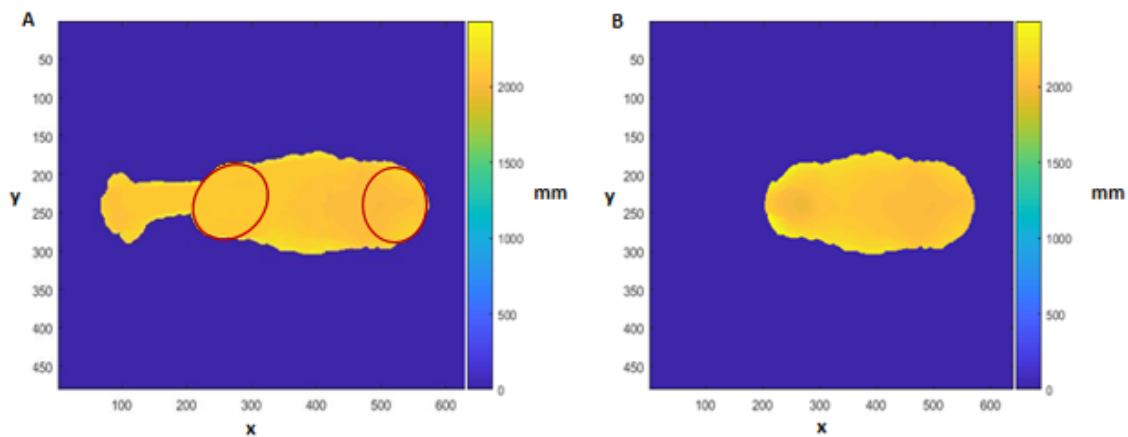


Figura 2. A: Ombro e garupa identificados por meio de uma adaptação de transformação de Hough para a identificação de objetos redondos, B: Imagem do dorso do animal segmentado após a identificação de objetos redondos. Onde X e Y corresponde as coordenadas de cada *pixel* no plano cartesiano e mm representa a distância em milímetros da câmera a cada ponto de *pixel* da imagem.

Tabela 2. Estatística descritiva de peso corporal de bovinos de corte Nelore para os conjuntos experimentais

Dados	Treino						Validação					
	No. Animais	Média	DP	CV	Mínimo	Máximo	No. Animais	Média	DP	CV	Mínimo	Máximo
Set 1 ¹	359	553,30	63,08	11,40	359,00	665,00	91	589,94	30,26	5,12	505,00	653,00
Set 2 ²	222	532,18	65,90	12,38	359,00	653,00	228	558,49	35,61	6,05	485,00	665,00
Set 3 ³	367	570,67	59,30	10,39	359,00	665,00	83	516,69	38,22	7,39	450,00	612,00
Set 4 ⁴	402	573,99	45,59	7,94	450,00	665,00	48	449,48	47,46	10,55	359,00	554,00

¹Set 1: Conjunto experimental composto por experimento 1, 2 e 3 para treino e experimento 4 para validação do modelo preditivo.

²Set 2: Conjunto experimental composto por experimento 1, 2 e 4 para treino e experimento 3 para validação do modelo preditivo.

³Set 3: Conjunto experimental composto por experimento 1, 3 e 4 para treino e experimento 2 para validação do modelo preditivo.

⁴Set 4: Conjunto experimental composto por experimento 2, 3 e 4 para treino e experimento 1 para validação do modelo preditivo.

Tabela 3. Estatística descritiva de peso de carcaça quente de bovinos de corte Nelore para os conjuntos experimentais

Dados	Treino						Validação					
	No. Animais	Média	DP	CV	Mínimo	Máximo	No. Animais	Média	DP	CV	Mínimo	Máximo
Set 1 ¹	359	315,34	37,32	11,83	205,00	394,00	91	334,22	16,79	5,02	277,60	376,80
Set 2 ²	222	301,70	36,42	12,07	205,00	376,80	228	336,17	23,17	6,92	259,20	394,40
Set 3 ³	367	325,74	34,02	10,44	204,80	394,40	83	290,08	22,19	7,65	248,80	343,60
Set 4 ⁴	402	326,21	28,49	8,73	248,80	394,40	48	260,12	28,10	10,80	205,00	317,00

¹Set 1: Conjunto experimental composto por experimento 1, 2 e 3 para treino e experimento 4 para validação do modelo preditivo.

²Set 2: Conjunto experimental composto por experimento 1, 2 e 4 para treino e experimento 3 para validação do modelo preditivo.

³Set 3: Conjunto experimental composto por experimento 1, 3 e 4 para treino e experimento 2 para validação do modelo preditivo.

⁴Set 4: Conjunto experimental composto por experimento 2, 3 e 4 para treino e experimento 1 para validação do modelo preditivo.

2.4.3. Extração de características:

Após a extração das características realizadas pelo algoritmo, um *frame* preto e branco (Figura 3) é armazenado. As características extraídas foram exportadas e salvas em um documento separado por vírgula (.csv). O conjunto de recursos extraídos pode ser dividido em duas classes: medidas corporais e descritores de forma.

As medidas corporais foram: área, volume, comprimento máximo, larguras e alturas em locais equidistantes na parte dorsal do animal dos ombros até a garupa (Figura 4). O comprimento, a largura e a área de um objeto em uma imagem podem ser estimados pela soma de pixels. Para estimar a área real de qualquer pixel, o princípio trigonométrico de ampliação de imagem pode ser usado:

$$m = f/d_o \quad [2]$$

$$a_i = 1/m_p \quad [3]$$

$$a_i = 1/m_p \quad [4]$$

$$v_o = \sum_{i=1}^k a_p * h_p \quad [5]$$

Onde m é o fator de ampliação, f a distância focal efetiva que é intrínseca à câmera e d_o é a distância que o objeto está da câmera. Assim, a área de qualquer objeto a_o pode ser estimada como a soma da área dos pixels a_p que formam sua imagem. Para p indo de 1 a n , o índice dos n pixels que formam a representação da imagem do objeto. As alturas de pixel h_p , podem ser estimadas pela diferença entre a distância até a câmera para o pixel p na área dorsal do animal e a distância da câmera até o chão. Portanto, o volume é a soma dos volumes de cada pixel que forma a representação da imagem do objeto. Note-se que este volume não é o volume real do corpo do animal, mas um volume aparente, uma vez que apenas a visão superior do animal foi obtida.

Os descritores de formas utilizados no presente estudo foram: **[1]** Excentricidade, é uma medida de arredondamento e é estimada como a razão entre os focos e o eixo principal das elipses que possui os mesmos segundos momentos que o animal, **[2]** Coeficiente de determinação do ajuste linear para a curvatura posterior do animal, **[3]** Primeiros 20 descritores de Fourier. Descritor de Fourier é uma classe de descritores de imagem global normalmente utilizada para análise de formas e correspondência de imagens (Bowman et al., 2000; Zhang e Lu, 2002; Burger e Burge, 2016). Para robustez, cada imagem segmentada foi girada, centralizada e transformada para a representação do plano polar anterior à estimativa da transformada discreta de Fourier em uma dimensão. Assim, os descritores de Fourier para a imagem polar podem ser estimados através de uma transformação discreta de Fourier, conforme descrito na equação [6].

$$FD_k = \sum_{j=0}^{n-1} X_j w_n^{jk} \quad [6]$$

Onde FD_k é o k Descritor de Fourier para o sinal X de comprimento n (o número de colunas da imagem polar) e $w_n = e^{-(2\pi i)/n}$ com $i = \sqrt{-1}$ o imaginário constante. Como FD_0 corresponde à soma sobre o sinal, que é a área do objeto na imagem, os descritores de Fourier também foram divididos por FD_0 para invariância no tamanho do animal.

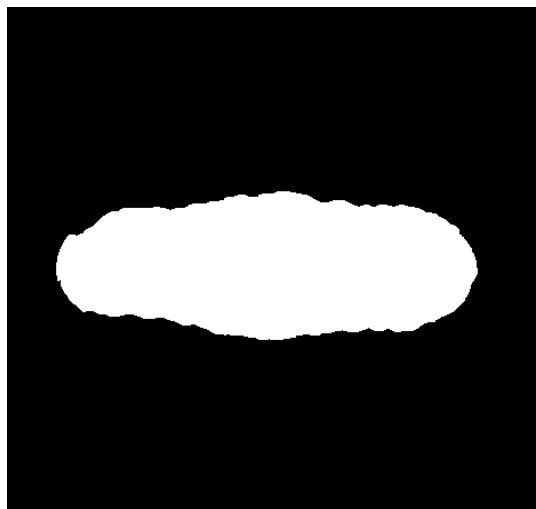


Figure 3. Frame preto e branco salvo após a extração dos recursos da imagem

2.5. Análise de dados

Para a análise estatística dos dados foi realizada Correlação de Pearson entre o peso corporal do animal, peso de carcaça quente e as características extraídas das imagens tridimensionais utilizando o PROC CORR do software SAS 9.4, as características com correlação abaixo de $P < 0,05$ foram retiradas do modelo. Assim, para as análises estatísticas de peso corporal foram utilizados: Área e volume corporal, onze larguras e alturas ao longo do dorso do animal, comprimento, excentricidade e duas medidas da curvatura da espinha dorsal. Já para as análises estatísticas de peso de carcaça quente foram utilizados: Área e volume corporal, onze larguras e alturas ao longo do dorso do animal, comprimento, excentricidade, duas medidas da curvatura da espinha dorsal e quatro medidas descritivas de fourrier. O valor das correlações entre a variável resposta e as 28 variáveis explanatórias para predição de peso corporal estão presentes na Figura 5 e o valor das correlações para predição de peso de carcaça quente estão presentes na figura 6.

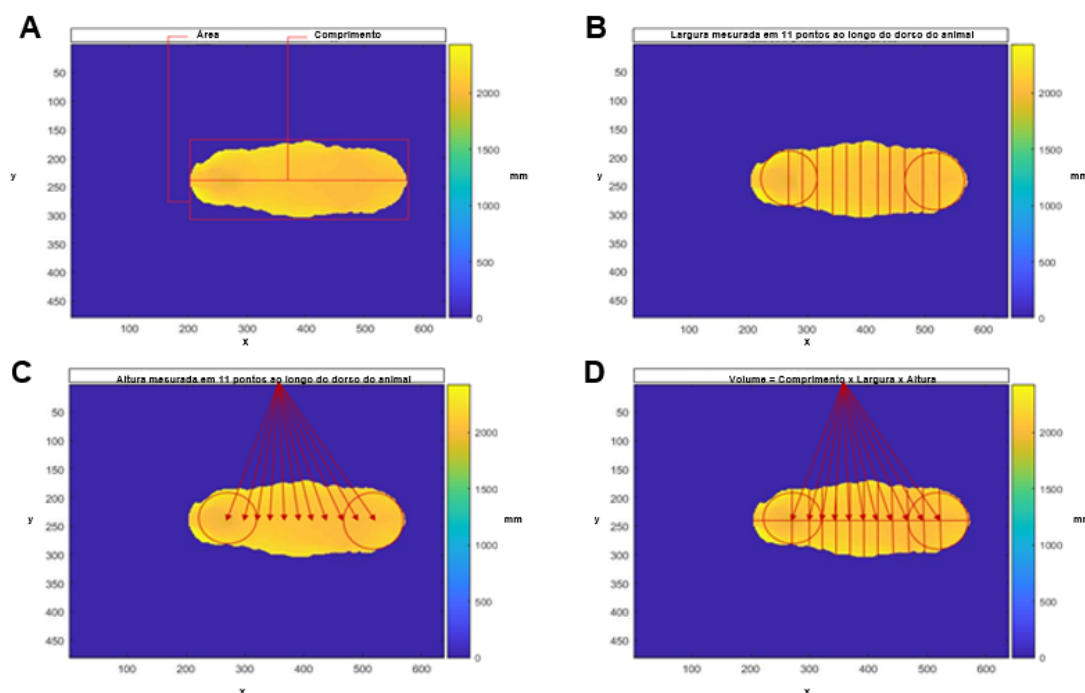


Figura 4. Extração das medidas corporais em locais equidistantes na parte dorsal do animal, dos ombros até a garupa. A: Mensuração de área e comprimento, B: 11 mensurações de largura, C: Altura medida em 11 pontos ao longo do dorso e D: volume corporal. Onde X e Y corresponde as coordenadas de cada *pixel* no plano cartesiano e mm representa a distância em milímetros da câmera a cada ponto de *pixel* da imagem.

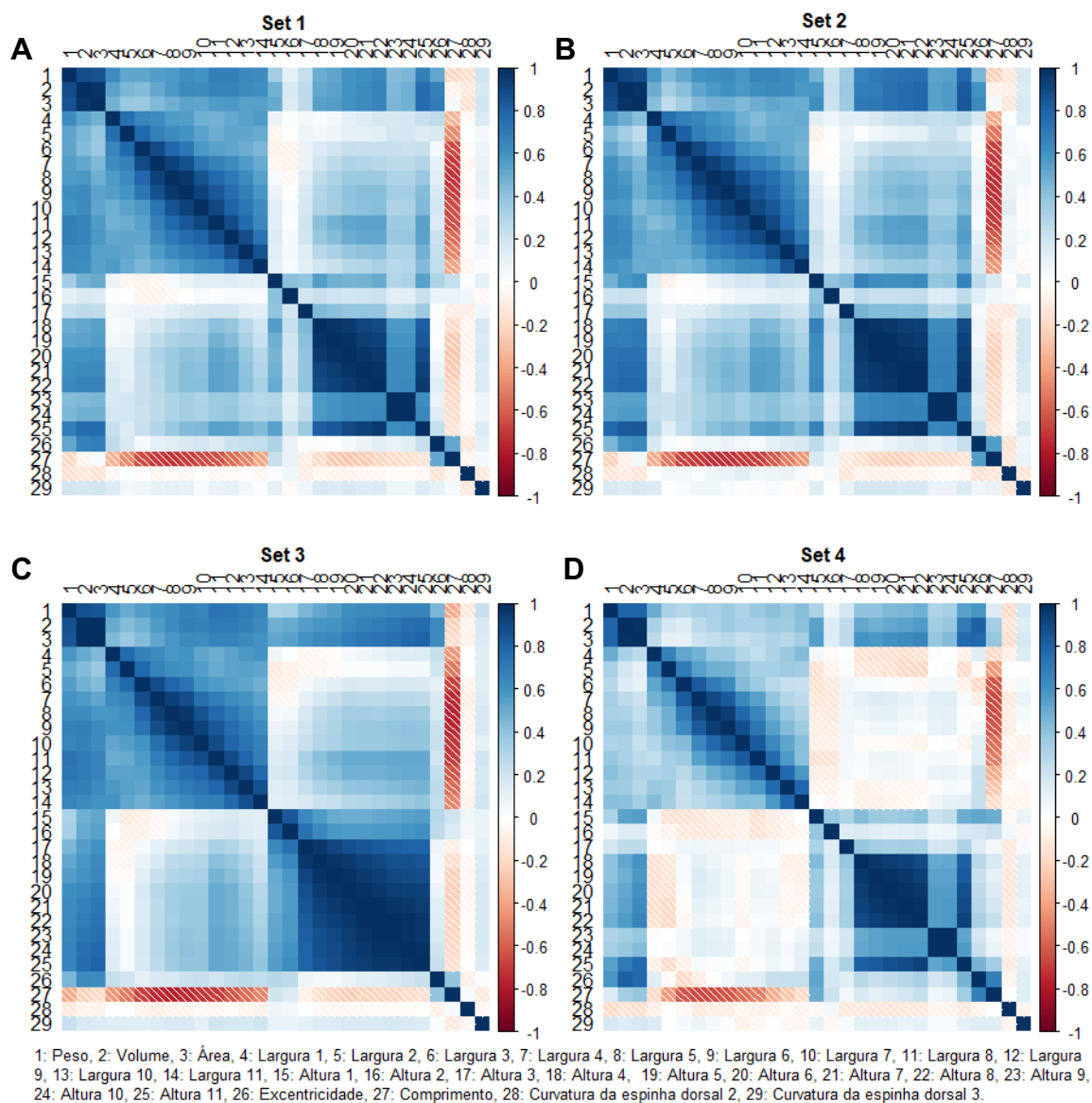


Figura 5. Correlação entre a variável resposta e as variáveis explanatórias para peso corporal obtidas através do sensor Kinect®. A: Desmama, B: Recria, C: Início da Terminação e D: Final da Terminação.

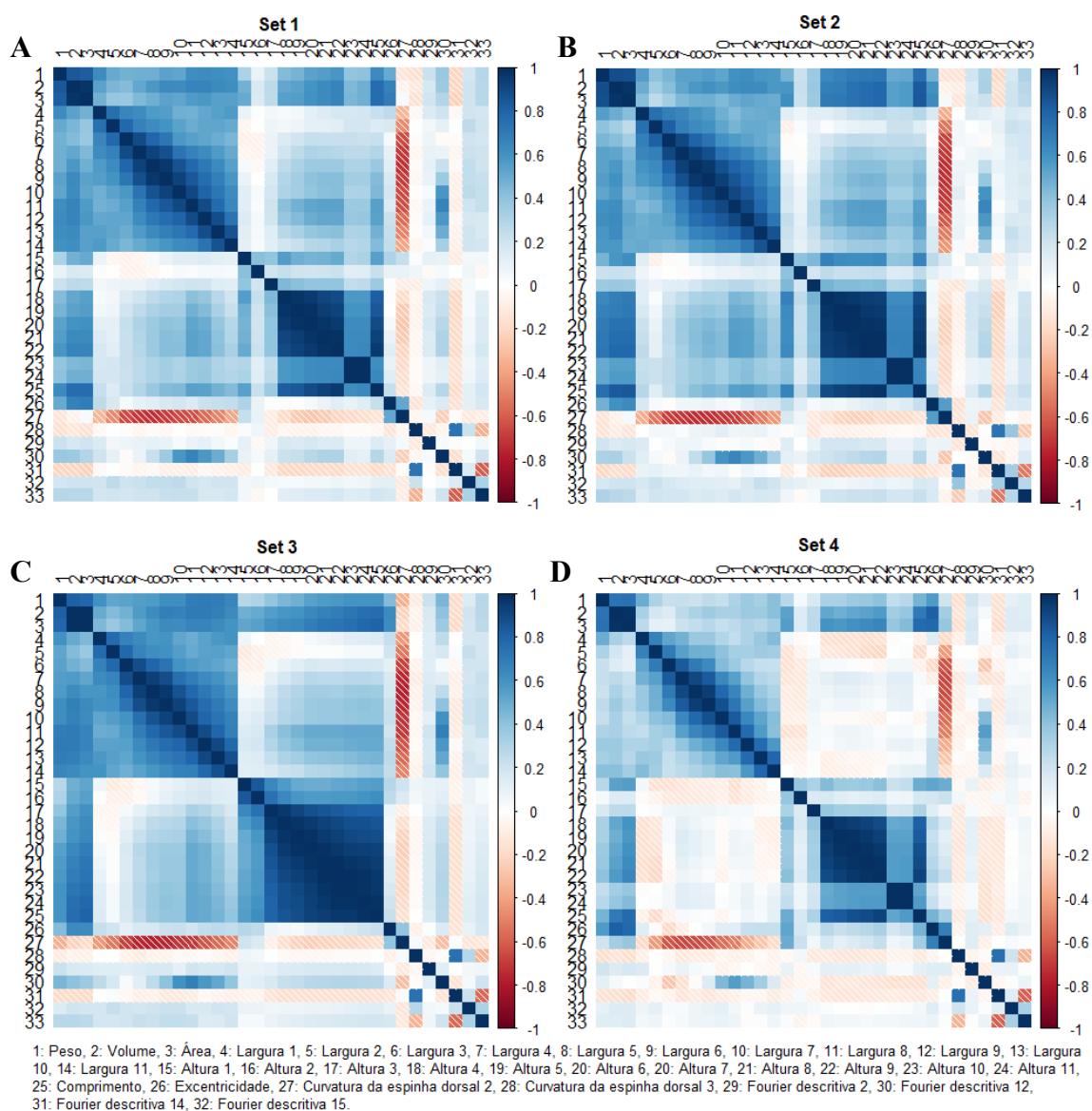


Figura 6. Correlação entre a variável resposta e as variáveis explanatórias para peso corporal obtidas através do sensor Kinect®. A: Desmama, B: Recria, C: Início da Terminação e D: Final da Terminação.

2.6. Predição de peso corporal e peso de carcaça quente

A predição de peso corporal e peso de carcaça quente foi realizada no Departamento de Ciências Animais da Universidade de Wisconsin-Madison, Estados Unidos. A predição do peso corporal e peso de carcaça quente foi desenvolvida utilizando quatro tipos diferentes métodos de predição, 1: Artificial Neural Network, 2: Partial Least Squares, 3: Regressão LASSO e 4: Regressão Linear Múltipla. Foi realizado neste estudo uma validação 10 k-fold para determinar o número de fatores retidos no modelo e para comparar e avaliar quão robusto foram os modelos foi utilizado uma validação externa.

2.7. Regressão linear múltipla

Quando temos uma variável resposta (Y) e variáveis explanatórias ($X_1, X_2, \dots, X_k$), podemos utilizar a regressão linear múltipla, onde o modelo é descrito na Equação [7].

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_kX_k + \varepsilon \quad [7]$$

Onde Y é a variável resposta e $X_1, X_2 \dots X_k$ são as variáveis explanatórias; ε é o vetor de erros (resíduos); $B_0, B_1, \dots B_k$ são os coeficiente do modelo.

A RLM foi realizada juntamente com a validação cruzada Leave-One-Out no software R utilizando a função TrainControl, método "LOOCV" do pacote Caret (<https://cran.r-project.org/web/packages/caret/caret.pdf>), onde Y consistia no peso corporal do animal e as variáveis independentes X foram as características obtidas através do Kinect®.

2.8. Regressão LASSO

O modelo LASSO ajusta os parâmetros por encolhimento, o que faz uma penalização nas variáveis menos significativas do modelo, esta é uma forma de regressão regularizada ou “penalizada”, na qual a regularização L1 (lasso) é introduzida no procedimento de regressão linear múltipla padrão, usando uma função de custo composto para otimizar os coeficientes de regressão. Este modelo de regressão é indicado para situações que envolvam maior número de

variáveis independentes correlacionadas (Hans, 2009), como observado no presente estudo. A regressão LASSO foi performada juntamente com a validação 10 k-fold utilizando a função `cv.glmnet` (`nfolds=10`) do pacote “`glmnet`” (Friedman et al., 2010) implementado no software R. Neste pacote é implementado um algoritmo que otimiza parâmetros de ajuste de modelo usando um algoritmo de descida de coordenadas. A descrição dos modelos LASSO desenvolvidos para predição de peso corporal pode ser encontrada na Tabela 4 e a descrição dos modelos para predição de peso de carcaça quente pode ser encontrada na tabela 5.

2.9. Regressão *Partial Least Squares*

Uma regressão por mínimos quadrados parciais (PLS – “*Partial Least Squares*”) foi realizada com o JMP Pro 13 (SAS Institute Inc., Cary, NC) para prever o peso corporal, usando o algoritmo NIPALS da variável preditiva. O PLS extrai sucessivas combinações lineares dos preditores, denominados fatores, que abordam tanto a explicação da variação da resposta quanto a explicação da variação do preditor. O número de fatores retidos no modelo foi determinado usando uma validação 10 k-fold, com a média prevista da raiz da PRESS calculada. O modelo que atingiu o menor valor de PRESS na validação cruzada foi escolhido. A descrição dos modelos PLS desenvolvidos para peso corporal podem ser encontradas na Tabela 4 e os modelos desenvolvidos para peso de carcaça quente está presente na tabela 5.

Tabela 4. Modelos desenvolvidos para Artificial Neural Networks, Partial Least Squares e Least absolute shrinkage and selection operator para predição de peso corporal

Item	Conjuntos experimentais ¹			
	Set 1	Set 2	Set 3	Set 4
ANN				
1ª Unidade de Camadas Neurais	28	28	28	28
2º Tipo de Camadas de Neurais	RectifierWithDropout	RectifierWithDropout	RectifierWithDropout	RectifierWithDropout
	100	100	100	100
Epochs	20	20	20	20
L1 ¹	0,000024	0,000024	0,000024	0,000024
L2 ²	0,000019	0,000019	0,000019	0,000019
Proporção de dropout de entrada	0	0	0	0
Razão de dropout oculta	0	0	0	0
PLS⁴				
Número de Fatores	8	5	8	13
LASSO				
Lambda	0,70	0,32	0,30	0,13

¹Conjuntos experimentais: Set 1 = Conjunto experimental composto por experimento 1, 2 e 3 para treino e experimento 4 para validação do modelo preditivo; Set 2 = Conjunto experimental composto por experimento 1, 2 e 4 para treino e experimento 3 para validação do modelo preditivo; Set 3 = Conjunto experimental composto por experimento 1, 3 e 4 para treino e experimento 2 para validação do modelo preditivo; Set 4 = Conjunto experimental composto por experimento 2, 3 e 4 para treino e experimento 1 para validação do modelo preditivo.

²L1 = Regularização Lasso.

³L2 = Regularização Ridge.

⁴NIPALS método: Algoritmo de intervalo parcial interativo não linear utilizado para calcular os componentes da regressão PLS.

Tabela 5. Modelos desenvolvidos para Artificial Neural Networks, Partial Least Squares e Least absolute shrinkage and selection operator para predição de peso de carcaça quente

Item	Conjuntos experimentais ¹			
	Set 1	Set 2	Set 3	Set 4
ANN				
1ª Unidade de Camadas Neurais	32	32	32	32
2º Tipo de Camadas de Neurais	RectifierWithDropout	RectifierWithDropout	RectifierWithDropout	RectifierWithDropout
	100	100	100	100
Epochs	20	20	20	20
L1 ¹	0,000024	0,000024	0,000024	0,000024
L2 ²	0,000019	0,000019	0,000019	0,000019
Proporção de dropout de entrada	0	0	0	0
Razão de dropout oculta	0	0	0	0
PLS⁴				
Número de Fatores	1	5	4	13
LASSO				
Lambda	0,57	1,80	0,16	0,18

¹Conjuntos experimentais: Set 1 = Conjunto experimental composto por experimento 1, 2 e 3 para treino e experimento 4 para validação do modelo preditivo; Set 2 = Conjunto experimental composto por experimento 1, 2 e 4 para treino e experimento 3 para validação do modelo preditivo; Set 3 = Conjunto experimental composto por experimento 1, 3 e 4 para treino e experimento 2 para validação do modelo preditivo; Set 4 = Conjunto experimental composto por experimento 2, 3 e 4 para treino e experimento 1 para validação do modelo preditivo.

²L1 = Regularização Lasso.

³L2 = Regularização Ridge.

⁴NIPALS método: Algoritmo de intervalo parcial interativo não linear utilizado para calcular os componentes da regressão PLS.

2.10. Artificial Neural Networks

Redes neurais artificiais (ANN – “*Artificial Neural Network*”) foram adaptadas usando o software de código aberto para análise de *big data* chamado H2O (<https://cran.rproject.org/web/packages/h2o/h2o.pdf>) que pode ser utilizado através do software estatístico R (R Core Team, 2016). Os hiperparâmetros utilizados em machine learning podem afetar como o algoritmo se ajusta aos dados. Por exemplo, em deep learning neural networks (DLNN), o número de camadas, o número de neurônios em cada camada, a taxa de aprendizado e outros recursos são parâmetros de ajuste chamados hiperparâmetros. Como existem muitos parâmetros que podem afetar a precisão do modelo, o ajuste dos hiperparâmetros é essencial para o DLNN (Dórea et al., 2018).

Assim, neste trabalho foi utilizada uma combinação de hiperparâmetros para predição de peso corporal e peso de carcaça quente conforme descrito por Dórea et al., 2018. Informações mais detalhadas sobre os hiperparâmetros definidos e critérios de busca podem ser encontrados na Tabela 6. Após a busca da rede neural, os hiperparâmetros para peso corporal escolhido pode ser encontrado na Tabela 4 e para peso de carcaça quente os modelos escolhidos então presente na tabela 5. Cada modelo com suas respectivas estruturas de redes neurais artificiais foi treinado através da validação 10 k-fold. A função `h2o.deeplearning` foi usada para treinar os modelos. O ADADELTA foi o algoritmo de taxa de aprendizado adaptativo utilizado em todas as ANN (Zeiler, 2012).

2.11. Avaliação do modelo

Para a avaliação, o sistema de avaliação do modelo foi conduzido conforme descrito por Tedeschi (2006). O coeficiente de determinação (R^2) foi utilizado entre os valores preditos e observados para definir a precisão das equações desenvolvidas. O coeficiente de correlação de concordância (CCC, Lin 1989), viés médio e root mean square error prediction (RMSEP) foram calculados como uma medida de precisão. O coeficiente de correlação de concordância foi estimado pela Equação [9] (Lin, 1989).

$$\hat{\rho}_c = \hat{\rho}x C_b$$

[9]

Onde componente (ρ) é a estimativa de correlação (r) que mede a precisão. O segundo componente (C_b) é o fator de correção de polarização que representa o quanto a linha de regressão se desvia da inclinação da unidade (45°).

O mean bias (MB) ou viés médio é provavelmente a estatística mais utilizada e antiga para avaliar a precisão do modelo. É calculado pelo desvio médio entre os valores preditos e observado (Cochran e Cox, 1957). O viés médio representa o erro na tendência central (para avaliar se os valores preditos estão acima ou abaixo do predito), o viés da inclinação representa erros devidos à regressão, e o componente aleatório é a variância inexplicada que não pode ser explicada pela linear regressão (Tedeschi, 2006).

O root mean square error prediction (RMSEP) ou equação de predição de erro quadrático médio consiste na mean square error prediction (MSEP) ou raiz quadrada da predição de erro quadrático médio. Onde MSEP é a soma da diferença quadrática entre os valores observados e os valores preditos do modelo divididos pelo número de pontos de dados n (Bibby e Toutenburg, 1977). O MSEP foi decomposto em viés de média, viés de inclinação e erros aleatórios (Theil, 1961) para fornecer uma indicação da adequação do modelo para predição. Os modelos foram avaliados a partir do software Model Evaluation System (MES, Colégio Station, TX; <http://nutritionmodels.tamu.edu/mes.html>) conforme descrito por Tedeschi (2006).

Tabela 6. Hiperparâmetros e critérios de pesquisa utilizados para peso corporal e peso de carcaça quente na pesquisa de ANN. Os hiperparâmetros foram combinados para encontrar a melhor estrutura de ANN.

Pesquisa de grade: Hiperparâmetros e critérios de pesquisa				
<i>Hiperparâmetros</i>				
Função de Ativação	Camadas Ocultas	Prop. Drop.	L1 ²	L2 ³
Rectifier	2	0	0	0
Tanh	8	0,25	0,0001	0,0001
Maxout	20		0,00000	0,00000
RectifierWithDropo	40			
TanhWithDropout	50			
MaxoutWithDropout	60			
	80			
	100			
	110			
	120			
	130			
	150			
	20,20			
	30,30,30			
	25,25,25,25			
<i>Pesquisa de rede</i>				
estratégia	max_runtime_sec	max_models	stopping_metric ⁷	
RandomDiscrete ⁴	360	100	RMSE	

¹Proporção de dropout de entrada.

²L1 = Regularização Lasso.

³L2 = Regularização Ridge.

⁴RandomDiscrete = Obtenha uma pesquisa aleatória de todas as combinações dos seus hiperparâmetros.

⁵max_runtime_secs = Tempo de execução máximo em segundos para toda a pesquisa de grade.

⁶max_models = Número máximo de modelos pesquisados na grade.

⁷stopping_metric = Função utilizada para parada antecipada com base em nenhuma melhoria na métrica do modelo (neste caso, métrica definida como erro quadrático médio).

3. RESULTADOS E DISCUSSÃO

A média, intervalo (mínima e máxima), desvio padrão e coeficiente de variação do peso corporal estão presentes na Tabela 2. O conjunto de dados variou consideravelmente, com peso corporal no Set 1 variando entre de 359 e 665 kg para o treino do modelo e 505 a 653 kg para a validação do modelo preditivo. Set 2 variando entre 359 e 653 kg para o treino e 485 a 665 kg para a validação. Set 3 de 359 a 665 kg para o treinamento e 450 a 612 kg para a validação. Set 4 de 450 a 665 kg para o teste do modelo e 359 a 554 kg para a validação do modelo preditivo, com observado na tabela 2.

Assim como para peso corporal o conjunto de dados para peso de carcaça quente variou consideravelmente. Set 1 variando entre de 205 e 394 kg para o treino do modelo e 277 a 376 kg para a validação do modelo preditivo. Set 2 variando entre 205 e 376 kg para o treino e 259 a 394 kg para a validação. Set 3 de 204 a 394 kg para o treinamento e 248 a 343 kg para a validação. Set 4 de 248 a 394 kg para o teste do modelo e 205 a 317 kg para a validação do modelo preditivo, como apresentado na tabela 3.

Os resultados das predições para peso corporal nos Set 1, 2, 3 e 4 foram: ANN (CCC = 0.73, 0.66, 0.70, 0.74; RMSEP = 19.6, 27.2, 27.2, 33.7 kg e R^2 = 0.58, 0.53, 0.53 e 0.59, respectivamente), PLS (CCC = 0.72, 0.65, 0.54, 0.69; RMSEP = 21.6, 28.0, 36.8, 39.3 kg, e R^2 = 0.55, 0.53, 0.30, 0.51, respectivamente), LASSO (CCC = 0.70, 0.60, 0.58, 0.76; RMSEP = 20.9, 28.3, 32.3, 40.6 kg e R^2 = 0.58, 0.53, 0.39, 0.50, respectivamente), RLM (CCC = 0.74, 0.65, 0.61, 0.67; RMSEP = 20.9, 27.9, 31.3, 42.6 kg, e R^2 = 0.58, 0.53, 0.39, 0.51, respectivamente). O uso do método de ANN melhorou a acurácia e precisão das predições de peso corporal para os conjuntos experimentais: Set 1, 2, 3 e 4 como apresentado na Tabela 8. ANN pode apresentar vantagens por modelar de forma mais eficaz as relações potencialmente não-lineares entre variáveis de entrada e saída, assim as abordagens de aprendizado de máquina são uma escolha poderosa e flexível (Peraí et al., 2010; Lin et al., 2012). No presente estudo, foi encontrado que o método ANN obteve melhor desempenho do que os outros modelos preditivos para os quatro conjuntos experimentais.

Diferentemente da predição de peso corporal, o uso do método de regressão LASSO e RLM melhorou a precisão e acurácia das predições de peso

de carcaça quente como evidenciado na tabela 9. Os resultados das predições para peso de carcaça quente nos Set 1, 2, 3 e 4 foram: ANN (CCC = 0.65, 0.19, 0.56, 0.64; RMSEP = 13.0, 30.0, 20.0, 23.2 kg e R^2 = 0.45, 0.17, 0.44 e 0.52, respectivamente), PLS (CCC = 0.60, 0.54, 0.51, 0.65; RMSEP = 14.6, 20.3, 23.0, 25.3 kg, e R^2 = 0.44, 0.42, 0.32, 0.45, respectivamente), LASSO (CCC = 0.67, 0.53, 0.55, 0.66; RMSEP = 13.1, 20.1, 20.8, 24.1 kg e R^2 = 0.49, 0.40, 0.46, 0.48, respectivamente), RLM (CCC = 0.14, 0.53, 0.40, 0.69; RMSEP = 8.75, 20.2, 24.6, 22.7 kg, e R^2 = 0.05, 0.40, 0.20, 0.52, respectivamente). A partir dos resultados expostos, a predição do peso de carcaça quente foi mais robusta e acurada para os modelos lineares como regressão linear múltipla e LASSO indicando que diferentemente dos resultados para predição de peso corporal onde pode ser encontrada características não-lineares a predição de peso de carcaça quente pode ser possível apresentar caráter linear.

O método PLS apresentou RMSEP inferior para os Set 1, 2, 3 e 4 comparado as ANN, pois o método de PLS foi projetado como uma interseção entre a modelagem linear e a análise de componentes principais: um modelo linear capaz de determinar quais variáveis preditoras estavam relacionadas à variável resposta e lidar adequadamente com a multicolinearidade (Wold, 2004). No entanto, presume-se que o método de predição PLS apresente limitações devido à linearidade, bem como a regressão linear múltipla e a regressão LASSO. A regressão linear múltipla é fundamental para fazer comparações entre a combinação de vários fatores e seus resultados. Já a regressão LASSO é um método de análise de regressão que executa tanto a seleção quanto a regularização de variáveis, a fim de aumentar a precisão e a capacidade de interpretação do modelo estatístico que produz. A regressão LASSO melhora o erro de predição diminuindo os grandes coeficientes de regressão para reduzir o *overfitting* através do encolhimento das variáveis explanatórias (X) porque a regressão LASSO utiliza valores absolutos na função de penalidade.

A validação independente pode ser utilizada para comparar os desempenhos de diferentes procedimentos de modelagem preditiva. Ao calcular o peso corporal sem um método de validação como estudo de Gomes et al., 2016, os resultados apresentam um R^2 elevado, pois trata-se de um R^2 calculado para o modelo e assim mais propenso ao *overfitting* comparado quando utiliza-se um método de validação. Ao validar modelos com um conjunto de dados

externo, é comum encontrar estudos que excluam animais ou alguns registros de cada animal do conjunto de dados de treinamento (Klaffenböck et al., 2017; McParland et al., 2014; Shetty et al., 2017). Shetty et al. (2017) comparou a validação de um modelo usando conjuntos de dados externos determinados de 2 maneiras diferentes. Um conjunto de dados foi criado pela exclusão aleatória de vacas do conjunto de treinamento, e o segundo conjunto de dados foi criado pela exclusão aleatória de registros. Esses autores concluíram que as predições do modelo são superestimadas quando os registros da mesma vaca em ambos os conjuntos de dados (treinamento e validação) são mantidos.

Ao realizar o procedimento de validação independente conforme descrito por Dórea et al. (2018), foi constatado que o erro de predição obteve resultados superiores ao treinamento para peso corporal e peso de carcaça quente (tabela 9). De acordo com Dórea et al. (2018) a validação dos modelos excluindo todo o conjunto, está mais próximo da realidade, pois os modelos desenvolvidos devem ser utilizados para prever novos conjuntos de dados provenientes de fazendas ou rebanhos com diferentes dietas, clima e práticas de manejo. Os resultados obtidos mostram que é importante realizar validação para os modelos de predição, e assim, prever o peso corporal e peso de carcaça quente a partir de imagens tridimensionais torna-se uma importante ferramenta a ser utilizada.

Tabela 7. Parâmetros de avaliação de predições de peso corporal em diferentes conjuntos experimentais

Item	Inclinação (±SE)	H0: b = 1	R²	Viés Médio (Kg)	CCC	RMSEP (Kg)	RMSEP / Média (%)	MSEP decomposição (%)		
								Viés Médio	Viés Médio	Erro Aleatório
Set 1										
ANN	0.59 (±0.05) ¹	0.01	0.58	3.22	0.73 (±0.04)	19.68	3.34	2.68	37.90	59.42
PLS	0.69 (±0.06) ¹	0.01	0.55	5.33	0.72 (±0.04)	21.60	3.66	6.09	18.11	75.80
LASSO	0.70 (±0.06) ¹	0.02	0.58	6.52	0.74 (±0.04)	20.95	3.55	9.71	18.55	71.74
RLM	0.69 (±0.06) ¹	0.01	0.58	6.62	0.74 (±0.04)	20.95	3.55	10.00	18.95	71.05
Set 2										
ANN	0.57 (±0.03) ¹	0.01	0.53	-12.19	0.66 (±0.03)	27.22	5.11	20.05	30.33	49.62
PLS	0.59 (±0.03) ¹	0.01	0.53	-13.80	0.65 (±0.03)	28.07	5.03	24.18	26.17	49.65
LASSO	0.60 (±0.03) ¹	0.01	0.53	-14.12	0.65 (±0.03)	28.31	5.07	24.90	24.70	50.40
RLM	0.59 (±0.03) ¹	0.01	0.53	-13.55	0.66 (±0.03)	27.95	5.00	23.53	25.86	50.61
Set 3										
ANN	0.62 (±0.06) ¹	0.01	0.53	6.04	0.70 (±0.05)	27.23	5.27	4.91	27.11	67.98
PLS	0.56 (±0.09) ¹	0.02	0.30	-0.65	0.54 (±0.07)	36.87	7.14	0.03	20.22	79.75
LASSO	0.58 (±0.08) ¹	0.01	0.39	5.81	0.61 (±0.06)	32.34	6.26	3.23	24.05	72.72
RLM	0.56 (±0.07) ¹	0.01	0.39	1.96	0.62 (±0.06)	31.36	6.07	0.40	27.65	71.95
Set 4										
ANN	0.74 (±0.09) ¹	0.08	0.59	12.40	0.74 (±0.06)	33.75	5.87	12.51	12.14	74.35
PLS	0.78 (±0.11) ¹	0.06	0.51	11.88	0.69 (±0.07)	39.34	8.75	9.12	6.35	84.53
LASSO	0.76 (±0.11) ¹	0.04	0.50	16.77	0.67 (±0.07)	40.63	9.04	17.04	7.33	75.63
RLM	0.72 (±0.06) ¹	0.04	0.51	13.77	0.67 (±0.07)	42.65	9.49	7.35	17.06	75.59

¹Teste de inclinação (b = 1, P > 0,05)

Tabela 8. Parâmetros de avaliação de predições de peso de carcaça quente em diferentes conjuntos experimentais

Item	Inclinação (±SE)	H0: b = 1	R²	Viés Médio (Kg)	CCC	RMSEP (Kg)	RMSEP / Média (%)	MSEP decomposição (%)		
								Viés Médio	Viés Médio	Erro Aleatório
Set 1										
ANN	0.56 (±0.08) ¹	0.01	0.45	3.16	0.65 (±0.05)	13.06	3.90	5.87	30.78	63.35
PLS	0.59 (±0.06) ¹	0.01	0.44	6.89	0.60 (±0.06)	14.62	4.37	22.23	21.77	56.00
LASSO	0.63 (±0.06) ¹	0.01	0.49	4.43	0.67 (±0.05)	13.17	3.94	11.31	21.37	67.32
RLM	0.08 (±0.03) ¹	0.01	0.05	0.01	0.14 (±0.06)	8.75	2.62	0.00	86.95	13.05
Set 2										
ANN	0.19 (±0.03) ¹	0.01	0.17	-21.27	0.19 (±0.03)	30.00	8.92	50.25	38.55	11.20
PLS	0.44 (±0.03) ¹	0.01	0.42	-10.06	0.54 (±0.03)	20.33	6.05	24.49	39.51	36.00
LASSO	0.43 (±0.03) ¹	0.01	0.40	-9.11	0.53 (±0.03)	20.13	5.99	20.50	42.12	37.38
RLM	0.43 (±0.03) ¹	0.01	0.40	-9.51	0.53 (±0.03)	20.29	6.04	21.96	41.55	36.49
Set 3										
ANN	0.52 (±0.06) ¹	0.01	0.44	11.33	0.56 (±0.06)	20.06	6.91	31.87	26.99	41.14
PLS	0.57 (±0.09) ¹	0.02	0.32	10.49	0.51 (±0.07)	23.01	7.93	20.77	16.57	62.66
LASSO	0.55 (±0.06) ¹	0.01	0.46	12.89	0.55 (±0.06)	20.89	7.20	38.10	22.31	39.59
RLM	0.42 (±0.09) ¹	0.01	0.20	9.82	0.40 (±0.08)	24.61	8.48	16.25	26.83	56.92
Set 4										
ANN	0.61 (±0.08) ¹	0.05	0.52	12.67	0.64 (±0.07)	23.20	8.91	29.83	21.14	49.03
PLS	0.78 (±0.12) ¹	0.09	0.45	5.18	0.65 (±0.08)	25.35	9.75	4.19	5.55	90.26
LASSO	0.75 (±0.11) ¹	0.03	0.48	8.53	0.66 (±0.07)	24.16	9.29	12.48	8.13	79.39
RLM	0.75 (±0.10) ¹	0.02	0.52	8.49	0.69 (±0.07)	22.76	8.75	13.93	8.91	77.16

¹Teste de inclinação (b = 1, P > 0,05)

Tabela 9. Root mean square error prediction (RMSEP) calculado através dos modelos preditivos: ANN, PLS, LASSO e RLM para peso corporal e peso de carcaça quente.

Item	Peso Corporal				Peso Carcaça Quente			
	RMSEP (Kg)				RMSEP (kg)			
	Set 1 ¹	Set 2 ²	Set 3 ³	Set 4 ⁴	Set 1 ¹	Set 2 ²	Set 3 ³	Set 4 ⁴
Conjunto de treinamento 10-kfold em validação cruzada								
ANN	20.16	20.91	18.69	20.27	15.71	13.42	15.25	14.82
PLS	21.28	20.72	20.53	18.23	16.77	13.77	15.98	14.70
LASSO	21.54	20.33	20.43	18.06	16.67	13.94	15.67	14.82
RLM	23.62	21.83	22.21	21.00	7.33	13.95	15.04	15.08
Validação independente								
ANN	19.68	27.22	27.23	33.75	13.06	30.00	20.06	23.20
PLS	21.60	28.07	36.87	39.34	14.62	20.33	23.01	25.35
LASSO	20.95	28.31	32.34	40.63	13.17	20.13	20.89	24.16
RLM	20.95	27.95	31.36	40.63	8.75	20.29	24.61	22.76

¹Conjuntos experimentais: Set 1 = Conjunto experimental composto por experimento 1, 2 e 3 para treino e experimento 4 para validação do modelo preditivo; Set 2 = Conjunto experimental composto por experimento 1, 2 e 4 para treino e experimento 3 para validação do modelo preditivo; Set 3 = Conjunto experimental composto por experimento 1, 3 e 4 para treino e experimento 2 para validação do modelo preditivo; Set 4 = Conjunto experimental composto por experimento 2, 3 e 4 para treino e experimento 1 para validação do modelo preditivo.

4. CONCLUSÃO

O uso de imagens tridimensionais foi capaz de prever o peso corporal e o peso de carcaça quente com boa acurácia e precisão em bovinos Nelore e, assim, monitorar a curva de crescimento dos animais em tempo real auxiliando no monitoramento da saúde do rebanho e reduzindo a tensão gerada pela pesagem por métodos convencionais, pode proporcionar maior lucratividade à produção.

O uso da validação independente ajudou a reduzir o efeito *overfitting* do modelo, mas é necessário estudos futuros com um maior banco de dados e animais de diferentes fazendas, a fim de oferecer maior robustez e precisão nas predições.

5. REFERÊNCIAS

ATHERTON, T. J. & KERBYSON, D. J. Sizeinvariantcircle detection. **Image and Vision Computing**, v.17, 795–803, 1999.

BISHOP, C. M. 2006. Pattern Recognition and Machine Learning. S 432 pringer, New York, New York, USA.

BIBBY, J., & TOUTENBURG H. Prediction and Improved Estimation in Linear Models. **John Wiley and Sons**, Berlin, Germany. 1977.

BOWMAN, E.T.; SOGA, K.; & DRUMMOND, W. 2016. Fourier Shape Descriptor, pp. 665 - 711.

BURGER E.T. & BURGE, M. J. Particle shape characterization using Fourier descriptor analysis, **Géotechnique**, v.51, 545-554, 2016.

CORREA, P. B. M. Estimativa da massa corporal de bovinos por meio de sensor de profundidade Kinect. 2015. 50 f. **Dissertação (Mestrado)** - Curso de Engenharia Agrícola, Universidade Federal de Viçosa, Viçosa, 2015.

COCHRAN, W. G. & COX, G. M. 1957. **Experimental Design**, 2nd edition. John Wiley and Sons, New York, NY, USA.

DICKSON, R. A.; MORTON J. M.; BEGGS D. S.; ANDERSON G. A.; PYMAN M. F.; MANSELL P. D; BLACKWOOD C. B. An automated walk-over weighing system as a tool for measuring liveweight change in lactating dairy cows. **J. Dairy Sci.** v.96, 4477–4486, 2013.

DÓREA, J. R. R.; WELD, K.; ARMENTANO, L. E. Mining data from milk infrared spectroscopy to improve feed intake predictions in lactating dairy cows. **J. Dairy Sci.** 2018.

FRIEDMAN, J.; HASTIE, T.; & TIBSHIRANI R. Regularization paths for generalized linear models via coordinate descent. **J Statist Software** v.33, 1–22, 2010.

GIANOLA, D. H.; OKUT, K. A.; WEIGEL, ROSA, G. J. Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. **BMC Genet** v.12, 87-99, 2011.

GOMES, R. A.; MONTEIRO, G. R., ASSIS G. J. F., BUSATO, K. C., LADEIRA, M. M., and CHIZZOTTI, M. L. Technical note: Estimating body weight and body composition of beef cattle through digital image analysis. **Journal of Animal Science**. 2016. 94:5414-5422. doi:10.2527/jas.2016-0797

HANS, C. Bayesian lasso regression. *Biometrika*, 96(4):835–845. URL <http://biomet.oxfordjournals.org/content/96/4/835.abstract>. doi:10.1093/biomet/asp047. 2009.

KASHIHA, M.; BAHR, C; OTT, S.; MOONS, C. P.; NIEWOLD, T. A.; ÖDBERG, F. O.; BERCK-MANS, D. Automatic weight estimation of individual pigs using image analysis. **Computers and Electronics in Agriculture**. v.107, 38-44. 2014.

KLAFFENBOCK, M.; A. STEINWIDDER, C.; FASCHING, G.; TERLER, L.; GRUBER, G.; MÉSZÁROS, and J. SOLKNER. 2017. The use of mid-infrared spectrometry to estimate the ration composition of lactating dairy cows. **J. Dairy Sci**. 100:5411–5421. <https://dx.doi.org/10.3168/jds.2016-12189>.

Hough.P. V. C, Method and means for recognizing complex patterns, U. S. Patent, 4746348, 1962.

LIN, L. I. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* v.45, 255–268, 1989.

LIN, M. I. B.; W. A. GROVES, A.; FREIVALDS, E. G.; LEE, and M. HARPER. 2012. Comparison of artificial neural network (ANN) and partial least squares (PLS) regression models for predicting respiratory ventilation: an exploratory study. **European Journal of Applied Physiology**, 112, 1603-1611. <https://dx.doi.org/10.1007/s00421-011-2118-6>.

McPARLAND, S.; E. LEWIS, E. KENNEDY, S. G.; MOORE, B. McCARTHY, M. O'DONOVAN, S. T.; BUTLER, J. E.; PRYCE, and D. P. BERRY. 2014. Mid-infrared spectrometry of milk as a predictor of energy intake and efficiency in lactating dairy cows. **Journal of Dairy Science**. 97:5863-5871.

PERAI, A. H.; H. NASSIRI MOGHADDAM, S.; ASADPOUR, J.; BAHRAMPOUR, & MANSOORI, G. A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. *Poult. Sci.* v.89, 1562–1568, 2010. <https://dx.doi.org/10.3382/ps.2010-01005>.

POLGAR, P.; SZÜCS, E.; SZABO, F. Effect of selection for improvement of dairy performance on growth and body structure in Holstein Friesian and Hungarian Fleckvieh young bulls. (German language) **Arch. Tierz.** 40 (1997) 505-510

R Core Team. 2016. R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

SHETTY, N., P. LOVENHAHL, M. LUND, and A. BUITENHUIS. 2017. Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid infrared spectroscopy of milk. **J. Dairy Sci.** 100:253–264.

SCHOFIELD, C. P.; MARCHANT J. A.; WHITE R. P.; BRANDL N.; WILSON, M. Monitoring pig growth using a prototype imaging system. **J. Agric. Eng. Res.** 72:205–210. doi:10.1006/jaer.1998.0365. 1999.

TEDESCHI, L. O. 2006. Assessment of the adequacy of mathematical models. *Agric. Syst.* 89:225–247. <https://dx.doi.org/10.1016/j.agsy.2005.11.004>.

THEIL, H. 1961. Economic forecasts and policy. Pages 6–48 in *Contributions to Economic Analysis*. R. Strotz, J. Tinbergen, P. Verdoorn, and H. J. Witteveen, ed. North-Holland Publishing Company, Amsterdam, the Netherlands.

ULUTAS, Z., SAATCI, M., OZLUTURK, A. (2001). Prediction of body weights from body measurements in East Anatolian Red calves. **Journal of Atatrk Universty Agriculture Faculty.** 32, 61-65

WANG, Y.; YANG W.; WINTER P.; and WALKER, L. Walk-through weighing of pigs using machine vision and an artificial neural network. **Biosystems Eng.** v.100, 117–125, 2008. doi:10.1016/j.biosystemseng.2007.08.008.

WHITTEMORE, C. T. & SSCHOFIELD, C. P. A case for size and shape scaling for understanding nutrient use in breeding sows and growing pigs. **Livestock Production Science**, v.65, 203-208, 2000.

WOLD, H. 2004. Partial Least Squares. in Encyclopedia of Statistical Sciences. **John Wiley & Sons, Inc.**

ZHANG D. S. & LU. G. J. **A Shape-based image retrieval using generic Fourier descriptor.** Signal Processing: Image Communication. **Doi:** 10.1016/S0923-5965(02)00084-X

ZEILER, M. D. 2012. ADADELTA: an adaptive learning rate method. arXiv. 1212.5701:1:6. <http://arxiv.org/abs/1212.5701>.