



ISSN: 2230-9926

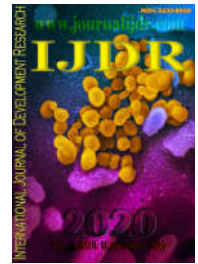
Available online at <http://www.journalijdr.com>

IJDR

International Journal of Development Research

Vol. 10, Issue, 11, pp. 42001-42007, November, 2020

<https://doi.org/10.37118/ijdr.20396.11.2020>



RESEARCH ARTICLE

OPEN ACCESS

ALGORITMOS DE MACHINE LEARNING PARA CLASSIFICAÇÃO DA CONFORMAÇÃO DE CARÇAÇAS DE CORDEIROS DESLANADOS

Manuella Paula de Mesquita Nunes¹, Gabriel Xavier Serrão², Gabriel Duarte de Carvalho¹, Alyne Cristina Sodr  Lima³, Marco Antonio Paula de Sousa², Jonas Carneiro Ara jo¹, Otavio Andre Chase^{*3}, Emerson Cordeiro Moraes³, Luciana Celi Chaves Daher¹ and Andr  Guimar es Maciel e Silva²

¹Programa de P s-gradua o em Sa de e Produ o Animal na Amaz nia (PPGSPAA), Instituto de Sa de e Produ o Animal (ISPA), Universidade Federal Rural da Amaz nia (UFRA), Campus Bel m, PA, Brasil

²Programa de P s-Gradua o em Ci ncia Animal (PPGCAN), Universidade Federal do Par  (UFPA), Campus Castanhal, PA, Brasil

³Instituto Federal de Educa o, Ci ncia e Tecnologia do Amap  (IFAP), Campus Agr cola Porto Grande, AP, Brasil

³Instituto Ciberespacial (ICIBE), Universidade Federal Rural da Amaz nia (UFRA), Campus Bel m, PA, Brasil.

ARTICLE INFO

Article History:

Received 29th August, 2020

Received in revised form

10th September, 2020

Accepted 17th October, 2020

Published online 24th November, 2020

Key Words:

Classifica o, *k*-Nearest Neighbor, Random Forest, Forma da carca a.

*Corresponding author:

Otavio Andre Chase,

ABSTRACT

Este estudo apresenta a avalia o da precis o de algoritmos de *Machine Learning* (aprendizado de m quina) na classifica o do escore de conforma o, por meio de informa es de *shape* (forma) provenientes de *Video Image Analysis* – VIA (an lise por v deo imagem) da carca a de cordeiros deslanados. Os resultados mostram que a metodologia adotada   confi vel para avaliar a conforma o das carca as.

Copyright   2020, Manuella Paula de Mesquita Nunes et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Manuella Paula de Mesquita Nunes, Gabriel Xavier Serr o, Gabriel Duarte de Carvalho et al. 2020. "Algoritmos de Machine Learning para classifica o da conforma o de carca as de cordeiros deslanados", *International Journal of Development Research*, 10, (11), 42001-42007.

INTRODUCTION

O emprego de ferramentas computacionais de controle e monitoramento de problem ticas zoot cnicas se torna uma realidade no setor produtivo, como em estudos de comportamento e bem-estar (BORCHERS et al., 2017), rastreabilidade de produtos de origem animal e avalia o de qualidade da carne e de marmoreio (BARBON et al., 2017). Estudos que envolvem an lise por v deo imagem – *Video Image Analysis* (VIA) vem se consolidando para avaliar v rias caracter sticas tanto qualitativas quanto quantitativas de carca a, pois re ne tamb m predicados como objetividade, n o ser invasivo e nem destrutivo, o que permite uma avalia o de

carca a justa e precisa de forma r pida sem interferir no fluxo da produ o (CRAIGIE et al., 2013). Sendo assim, com o grande volume de informa es, os algoritmos de *machine learning* identificam padr es que sejam uteis na solu o de problemas, gerando modelos que resumem o conhecimento humano. A aplica o dessa ferramenta na classifica o autom tica de carca as ovinas aperfei oa a avalia o de animais com aptid o para a produ o de carne e engratecem estudos sobre a rela o entre um modelo subjetivo, realizado por um avaliador como o sistema SEUROP, e um modelo objetivo, baseado na combina o multivariada de medidas para a classifica o por algoritmos. De maneira geral, para uma unidade de produ o e industrial, gerenciar esse tipo de

conhecimento é de primordial importância para o crescimento no cenário competitivo atual. Objetivou-se, com este estudo, avaliar a precisão de algoritmos de Machine Learning na classificação do score de conformação, por meio de informações de shape provenientes de *Video Image Analysis* da carcaça de cordeiros.

MATERIAL E MÉTODOS

Os procedimentos experimentais foram aprovados pelo Comitê de Ética em Pesquisa e Experimentação Animal (CEPAE protocolo nº 97-2015) da Universidade Federal do Pará e, o cuidado dos animais seguiu as diretrizes do Conselho Nacional de Controle de Experimentação Animal (CONCEA, 2015).

Obtenção das carcaças: Foram utilizadas informações de 67 carcaças de cordeiros deslanados, machos castrados, entre 8 e 11 meses de idade, provenientes de rebanhos comerciais do estado do Pará – Brasil, terminados em confinamento com o peso corporal ao abate entre 21,40 e 48,65 kg, primando pela obtenção de carcaças de conformações variadas. As informações empregadas foram provenientes das carcaças frias, obtidas após 24 horas de refrigeração a temperatura aproximada de 4°C, de onde aferiu-se primeiramente o peso de carcaça fria (PCF, em kg).

–side view (S), individualmente. Na estrutura de suporte/estúdio, buscou-se controlar a incidência de luminosidade por meio de uma fonte única de luz; emitida por uma fita dupla de LED de 24 V_{CC} (Volts em corrente contínua), disposta nas bordas internas (laterais e superior) da estrutura recoberta por um anteparo negro.

Realizou-se o processamento das imagens via *software* ImageJ 1.05i (ABRÀMOFF, MAGALHÃES, & RAM, 2004), determinando-se 15 projeções na carcaça (seleções da imagem (2D) efetuadas no objeto de estudo, originariamente 3D no cenário realista) da qual extraem-se informações de interesse. As projeções delimitaram a carcaça inteira e regiões, que foram baseadas tanto em cortes comerciais (LAGE et al., 2014) quanto em regiões-alvo utilizadas para reconhecimento, a exemplo da análise pelos avaliadores de tipificação, chamados de terços dorsais da carcaça (COLOMER-ROCHER et al., 1988), tendo como base pontos anatômicos previamente demarcados e pontos arbitrários (PA). As projeções foram demarcadas manualmente tendo como referência uma escala em centímetros na imagem (30 cm) gerando a seleção do contorno da carcaça inteira na (Figura 2 -A e B) e das regiões (Figura 2 -C e D). A leitura das projeções de visualização, conforme mostra a Figura 2 foram, Dorsal View (A): D1= projection of the entire carcass on the dorsal view.

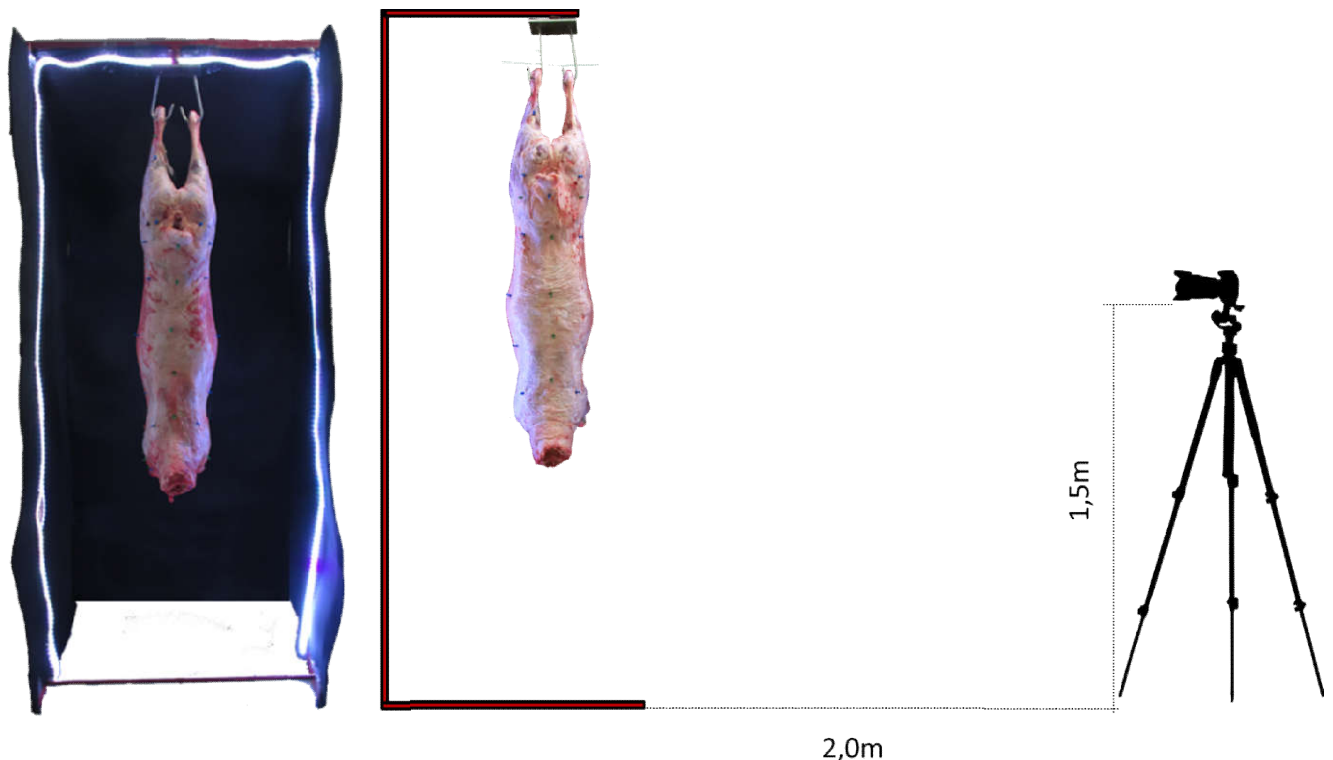


Figura 1. Estrutura de suporte com o controle a incidência de luminosidade e posicionamento da carcaça para captura de imagens digitais

Obtenção das imagens da carcaça e Video Image Analysis (VIA): A Figura 1 ilustra a configuração do experimento. Para obtenção das imagens, foi utilizada uma câmera digital com resolução de 16 MP (*Mega Pixels*), fabricada pela empresa Nikon, modelo *PowerShot SX160 IS*[®], sem utilização de flash posicionada a uma distância de 2 metros de distância da estrutura de suporte de carcaça/estúdio e 1,5 metro de altura, onde as carcaças frias foram posicionadas longitudinalmente por meio das articulações tarso-metatarsianas para serem fotografadas: na vista dorsal –dorsal view (D) e na vista lateral

Side View (B): S1=projection of the entire carcass on the side view; Dorsal View (C): D2=dorsal shoulder; D3=dorsal rib I; D4=dorsal rib II; D5=dorsal loin; D6=dorsal croup - leg; D7=dorsal right leg; D8=dorsal left leg. Side View (D): S2= side shoulder; S3= side rib I; S4= side rib II; L5=side loin; S6= side croup - leg; S7= side leg. De cada projeção, obtiveram-se os seguintes descritores de forma (Figura 3) (Leibrandt e Le Pennec, 2015): area (A, cm²), perimeter (P, cm), length (L, cm), width (W, cm); circularity (CIR- parâmetro adimensional obtido pela equação $4\pi \times [\text{área}/(\text{perímetro})^2]$, cujos valores vão

de 0, representando uma forma extremamente alongada, a 1, uma forma perfeitamente circular); solidity (SOL- razão entre a área da seleção e área do fecho convexo da seleção); aspect ratio (AR- parâmetro adimensional obtido pela razão entre a largura e o comprimento da seleção) e; convexity (CON- parâmetro adimensional obtido pela razão entre o perímetro e o perímetro do fecho convexo da seleção).

O conjunto das informações dos descritores de forma (atributos) das 15 projeções da carcaça (contorno da carcaça inteira e regiões) compôs o banco de dados denominado de SHAPE_VIA1, contabilizando 108 atributos. Ao banco de dados composto apenas pelos descritores das projeções do contorno da carcaça inteira, denominou-se SHAPE_VIA2, com 16 atributos. Ainda, a partir das imagens digitais das carcaças frias, obteve-se a pontuação para conformação segundo o Sistema SEUROP (European Union, 1994) de classificação de carcaças ovinas por um avaliador treinado, seguindo os escores de Superior (S), Excelente (E); Muito boa (U); Regular/Boa (R), Ordinária/Subretilínea (O); Mediocre/Pobre (P). O conjunto das informações dos descritores de forma (atributos) das 15 projeções da carcaça (contorno da carcaça inteira e regiões) compôs o banco de dados denominado de SHAPE_VIA1, contabilizando 108 atributos.

Ao banco de dados composto apenas pelos descritores das projeções do contorno da carcaça inteira, denominou-se SHAPE_VIA2, com 16 atributos. Ainda, a partir das imagens digitais das carcaças frias, obteve-se a pontuação para conformação segundo o Sistema SEUROP (European Union, 1994) de classificação de carcaças ovinas por um avaliador treinado, seguindo os escores de Superior (S), Excelente (E); Muito boa (U); Regular/Boa (R), Ordinária/Subretilínea (O); Mediocre/Pobre (P).

Abordagem estatística: Todos os dados foram padronizados pelo Teste Z. Para redução da dimensionalidade dos bancos de dados primários, utilizou-se a técnica de RGCCA escolhendo o modelo reflexivo (*‘mode a’*), de acordo com a metodologia de (TENENHAUS e TENENHAUS, 2011), onde avaliadas as relações sistêmicas entre as informações de forma da carcaça obtida por VIA (descritores de forma) e dados referentes à composição tecidual da carcaça, dos cortes comerciais e de qualidade da carne (Tabela 1) Serviu-se dos pacotes RGCCA do software R *version 3.2.4* (R CORE TEAM, 2016) para processamento dessas análises. Utilizou-se o *software* WEKA versão 3.8.1 (1999-2016) que possui o módulo *iBk* referente ao algoritmo KNN cuja métrica é a distância euclidiana.

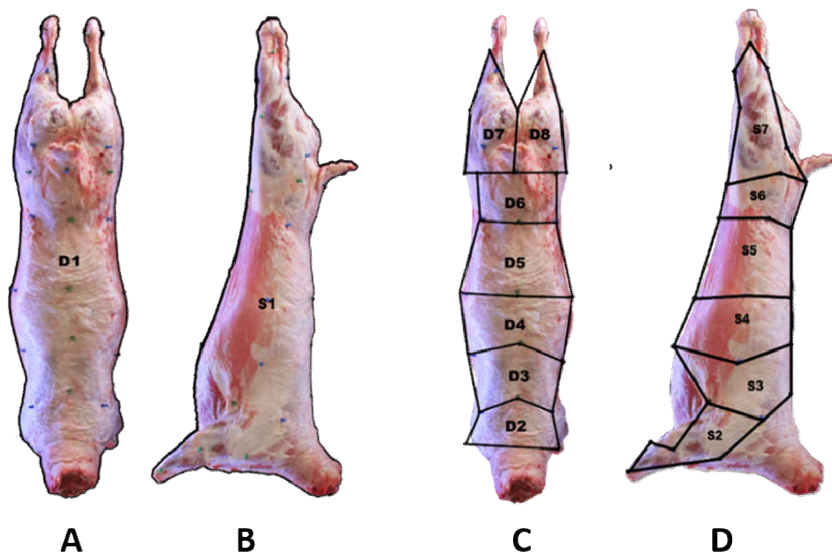


Figura 2. Projeções do contorno da carcaça inteira e das regiões obtidas por meio do software ImageJ 1.05i, nas vistas dorsal e lateral

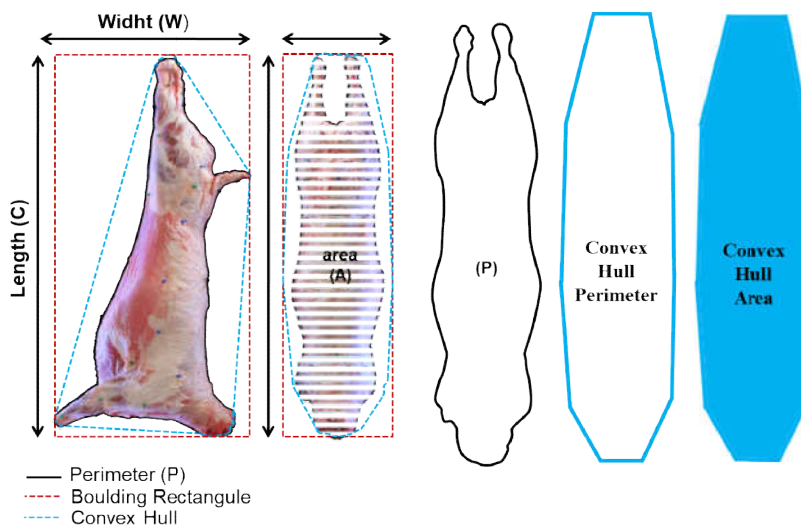


Figura 3. Contornos das projeções para a obtenção dos descritores de forma

Para a mensuração desse algoritmo utiliza-se o número de vizinhos mais próximos, identificado pelo parâmetro k , que possui influências no resultado da classificação. Optou-se pela escolha de valores de k ímpares como critério de desempate, seguiram a metodologia de Witten (2011) e Anchalía e Roy (2014).

Utilizou-se a técnica de amostragem dos exemplos foi utilizada a validação cruzada *leave-one-out* (Devijver & Kittler, 1982), dado o número limitado de exemplos disponíveis, treinando-se o classificador com $N-1$ exemplos e testando com 1 exemplo, processo que se repete até que todos os exemplos tenham sido testados. Para o a utilização do classificador Random Forest utilizou-se o pacote "RandomForest" no modo "Classificador" no software R version 3.2.4 (R CORE TEAM, 2016). Baseado na predição de um conjunto de árvores de decisão, para obter uma única resposta como saída (Breiman, 2001). Os hiperparâmetros utilizados foram baseados no número de árvores que diferiram de (5,10, 25,100, 300, 500 e 1000), e critérios padronizados seguindo a metodologia de Breiman (2001) e Girolamo Neto (2014). Para a técnica de amostragem dos exemplos utilizou-se o método *Bagging*, ou *Bootstrap Aggregating*, onde dois terços da amostras do conjunto de dados são utilizados no treinamento e o restante das amostras para teste (Breiman, 1996a ; Breiman, 1996b).

Abordagens da utilização do Machine Learning e banco de dados utilizados: Os bancos de dados inicial (SHAPE_VIA1 e SHAPE_VIA2) foram estabelecidos a partir de 67 carcaças, denominadas pela terminologia como exemplos, contendo áreas e descritores da carcaça e referente aos cortes comerciais, que compõem os atributos, e o escore de conformação, tomado como classe ou rótulo associado a cada exemplo. Os bancos de dados primários originaram outros dois, a partir da redução de seus atributos pela técnica RGCCA, com a finalidade de avaliar a acurácia e o comportamento de classificação dos algoritmos de *learning machine* (*k-Nearest Neighbor* and *Random Forest*) em diferentes situações (Tabela 1). Para tanto, foi avaliada adicionalmente a inclusão do atributo peso de carcaça fria (PCF) nos bancos de dados.

Métricas de avaliação da predição pelos algoritmos de Machine Learning: Para verificar o relacionamento entre as classificações realizadas pelos algoritmos, ou seja, a obtenção do número de classificações corretas em relação ao número de classificações feitas pelo algoritmo, utilizando-se a métrica de avaliação dita matriz de confusão segundo a metodologia de Syed, (2014). Assim, a partir dessas métricas se faz possível para mensurar a acurácia dos algoritmos, ou seja, a medida de qualidade de um classificador (ROCHA, 2008; FERREIRA, 2017) utilizando-se da métrica porcentagem de exemplos corretamente classificados (PECC= "exemplos corretamente classificados" / todos os exemplos", em %).

RESULTADOS E DISCUSSÃO

O banco de dados inicial foi reduzido seguindo a técnica de redução de dimensionalidade RGCCA, utilizou-se o critério de remoção de variáveis com menores valores de *outerloadings* com AVE <0,60 (CHIN, 1998; HAIR et al., 2005), no conjunto SHAPE_VIA1 (0,54) e SHAPE_VIA2 (0,54). Sendo assim, o BD1_(I) que possuía 108 atributos com a redução ficou com 46

atributos (BD1_(R)); e o BD2_(I) que possuía 16 atributos ficou com 9 atributos (BD2_(R)). Para o aprimoramento de estudos relacionados a conformação, as projeções foram pré-determinadas para ressaltar a atuação do o avaliador devidamente treinado, pois o mesmo divide mentalmente a carcaça em três planos musculares de vista frontal: o terço anterior (cranial), o terço médio (torax) e o terço posterior (caudal). or valorização comercial dos cortes (Cezar e Souza, 2007). Ou seja, as variáveis/atributos estão intimamente ligadas ao shape da carcaça (forma da carcaça). Neste estudo, não foram classificadas carcaças com escore 'Superior' (S) e 'Excelente' (E), conforme descreve o sistema de tipificação que foi adotado. Quanto ao número e à proporção de carcaças por classe, para análises realizadas com 67 exemplos, obteve-se a distribuição e porcentagem descritas na Figura 3. Houve uma prevalência da classe Subretlína de 40,29% conjuntos. Em contrapartida, a classe Muito Boa (U) tem o menor quantitativo de carcaças (4,4%), o que dificultaria a classificação da mesma pelos algoritmos. Um fato ocorrido no presente estudo, foi a presença de classes desbalanceadas, conforme apresenta a Figura 4, onde essa prevalência de determinadas classes, podem interferir na classificação já que o algoritmo KNN se baseia no número de vizinhos mais próximos ao classificar um novo exemplo, principalmente quando se utiliza a distância euclidiana. Segundo Batista (2003), banco de dados que possuem essa característica, falham na predição de classes menos frequentes em detrimento das mais frequentes, que apresentam maior precisão.

Tabela 1. Descrição dos bancos de dados/estratégias avaliados pelos algoritmos Learning Machine

Banco de dados	Descrição	Número de Atributos
BD1 _(I)	Formado pelo banco de dados primário SHAPE_VIA1	108 atributos
BD1 _(R)	Formado pelo banco de dados SHAPE_VIA1 reduzido por RGCCA	46 atributos
BD2 _(I)	Formado pelo banco de dados primário SHAPE_VIA2	16 atributos
BD2 _(R)	Formado pelo banco de dados SHAPE_VIA2 reduzido por RGCCA	9 atributos

No apêndice (Tabela A) estão descritas as projeções obtidas e a estatística descritiva do banco de dados inicial obtidas in situ em carcaças resfriadas de cordeiros que compuseram o SHAPE_VIA1 e SHAPE_VIA2 estão dispostas na Tabela 2.

Para as classificações com o algoritmo KNN, optou-se desconsiderar o valor de $k=1$, haja vista que, quando se considera apenas 1 vizinho mais próximo, a flexibilidade e a variância da classificação aumentam ao passo que se utiliza apenas uma classe para a predição, o que não é interessante por desconsiderar a maioria dos outros exemplos contidos no treinamento (Gimenez, 2003).

Na Tabela 3 têm-se os resultados de exemplos corretamente classificados (PECC%) nos diferentes banco de dados/estratégias e pela inclusão ou não do peso da carcaça fria com seus respectivos valores de k ímpares. Para os bancos de dados com áreas e descritores que envolve a não inclusão de peso de carcaça fria, foi observado maiores valores de PECC% no SHAPE_VIA2 para melhores valores de $k=17$ (55,22% de PECC) para BD2_(I) e $k=21, 23$ e 25 (55,22% de PECC) para BD2_(R).

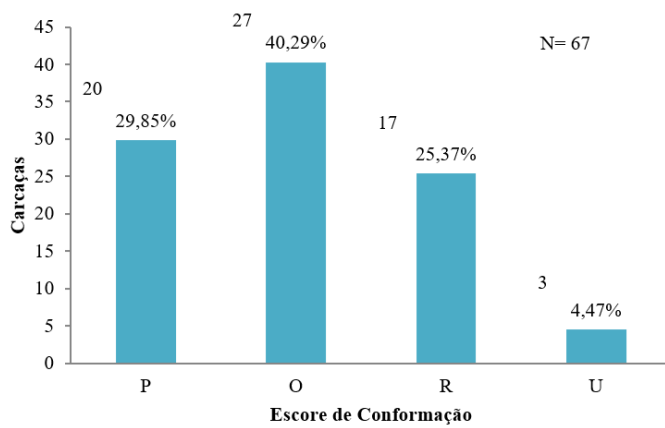


Figura 4. Distribuição e porcentagem relativa ao total por classe de carcaças ovinas do banco de dados de carcaças de cordeiros (n=67). Escore de Tipificação de carcaças: P – Médio/Pobre (Poor); O – Subretilínea (Ordinary); R – Boa/Regular (Regular); U – Muito Boa (Ultra Good).

Tabela 2. Média, desvio padrão (DP), coeficiente de variação (CV) e amplitude do banco de dados primários obtidas in situ em carcaças resfriadas de cordeiros que SHAPE_VIA1 e SHAPE_VIA2

Vistas	Variável	Média	DP	CV(%)	Amplitude
D1	A1	3179,55	369,7692	11,6296	(2314,93-4032,78)
	P1	430,55	22,97671	5,336628	(381,75- 499,61)
	L1	29,77	2,632843	8,842482	(24,57- 39,81)
	C1	158,72	7,504049	4,727861	(143,49- 178,35)
	CIR1	0,22	0,023581	10,57637	(0,17-0,28)
	SOL1	0,82	0,024064	2,947294	(0,73- 0,86)
	AR1	0,19	0,014412	7,68099	(0,16- 0,24)
CON1	0,8	0,019192	2,386149	(0,77-0,85)	
D2	A2	290,94	49,70157	17,08292	(171,73- 405,26)
	P2	69,61	5,870231	8,432598	(52,08- 81,84)
	L2	22,65	2,558029	11,29534	(17,14- 27,94)
	C2	18,09	2,285292	12,63057	(12,25- 24,7)
	AR2	1,27	0,209919	16,51578	(0,73- 1,92)
D3	A3	244,5	70,61365	28,88065	(109,24- 412,2)
	P3	65,98	8,580122	13,00413	(47,54- 85,31)
	L3	22,99	2,053668	8,931475	(17,97- 27,56)
	C3	16,43	3,156805	19,21672	(9,78- 23,62)
	AR3	1,44	0,257115	17,7972	(1- 2,46)
D4	A4	598,33	87,09885	14,55701	(384,96- 835,71)
	P4	107,14	6,60507	6,164858	(89,07- 123,85)
	L4	28,44	3,014468	10,59845	(22,41- 36,7)
	C4	27,76	2,432689	8,764108	(22,16- 32,38)
	CIR4	0,67	0,048877	7,258359	(0,54- 0,79)
	SOL4	0,9	0,040566	4,506221	(0,8- 1)
	AR4	1,03	0,136047	13,26749	(0,751,38)
CON4	0,98	0,01323	1,352027	(0,941)	
D5	A5	453,89	64,88228	14,29478	(331,7- 644,07)
	P5	88,08	6,976466	7,920457	(75,14- 106,34)
	L5	28,32	2,928708	10,34137	(22,29- 36,7)
	C5	21,52	2,077793	9,656863	(17,59- 26,48)
	AR5	1,32	0,160276	12,12827	(0,86- 1,7)
D6	A6	331,74	49,65095	14,96674	(227,52- 439,53)
	P6	75,46	4,866778	6,449128	(61,99- 86,03)
	L6	23,3	3,100698	13,30789	(12,7- 28,06)
	C6	18,24	2,781893	15,24983	(14,48- 32,32)
	CIR6	0,75	0,053394	7,101346	(0,54- 0,83)
	AR6	1,3	0,22605	17,32381	(0,41- 1,79)
D7	A7	334,98	53,91203	16,09395	(192,93- 468,02)
	P7	85,71	9,692952	11,3086	(70,73- 109,38)
	L7	14,41	2,548066	17,6828	(11,11- 32,25)
	C7	35,69	5,534859	15,50769	(12,57- 47,05)

D8	CIR7	0,59	0,063112	10,61449	(0,4- 0,71)
	AR7	0,43	0,268084	62,35673	(0,26- 2,57)
	A8	328,47	59,04787	17,97684	(190,5- 510,58)
	P8	85,25	9,742383	11,42847	(69,61- 106,77)
	L8	14,11	2,77134	19,64437	(10,41- 33,02)
	C8	35,69	5,557272	15,57048	(12,57- 46,29)
	AR8	0,59	0,062758	10,6712	(0,41- 0,71)
CON8	0,42	0,276276	65,52936	(0,25- 2,63)	
Vista Lateral					
L1	A9	3737,95	388,9993	10,40675	(2861,35- 4603,09)
	P9	420,33	20,68962	4,922291	(373,15-477,89)
	L9	61,7	3,957925	6,415274	(50,73- 69,71)
	C9	164,96	8,400761	5,092606	(143,62- 186,29)
	CIR9	0,27	0,023315	8,494345	(0,22- 0,32)
	SOL9	0,63	0,035874	5,658319	(0,52- 0,69)
	AR9	0,37	0,025329	6,762279	(0,31- 0,44)
CON9	0,92	0,023794	2,580498	(0,85- 0,95)	
L2	A10	373,39	75,94579	20,3396	(240,81- 541,44)
	P10	98,62	22,32464	22,63644	(0,09- 117,39)
	L10	38,22	3,58835	9,389056	(30,86- 46,35)
	C10	28,93	5,264914	18,20142	(16- 41,33)
	CIR10	0,45	0,070164	15,44188	(0,31- 0,61)
	SOL10	0,96	0,050985	5,337805	(0,79- 1)
	AR10	1,38	0,33192	24,08649	(0,78- 2,2)
CON10	1	0,029458	2,950522	(0,97- 1,23)	
L3	A11	508,92	196,6784	38,64616	(209,4- 906,5)
	P11	107,01	13,52578	12,6392	(81,61- 131,69)
	L11	36,92	2,293161	6,210816	(32,51- 42,29)
	C11	26,46	7,84427	29,65099	(13,08- 39,68)
	CIR11	0,55	0,097713	17,79118	(0,37- 0,75)
	SOL11	0,84	0,099852	11,95257	(0,62- 1)
	AR11	1,52	0,449823	29,53256	(0,85- 2,57)
CON11	0,98	0,01261	1,281951	(0,95- 1,02)	
L4	A12	812,32	95,81974	11,79576	(615,24- 1120,82)
	P12	115,26	7,108403	6,167313	(99,76- 140,16)
	L12	37,85	1,921278	5,076314	(32,13- 42,16)
	C12	29,16	3,137186	10,75716	(23,11- 37,9)
	AR12	1,31	0,152297	11,6056	(0,94- 1,68)
L5	A13	463,04	70,83575	15,29809	(341,28- 683,45)
	P13	88,51	6,032778	6,81594	(76,67- 108,31)
	L13	27,66	2,151671	7,780123	(22,92- 34,03)
	C13	20,97	2,089093	9,963911	(16,19- 25,71)
	AR13	1,33	0,152196	11,43875	(0,96- 1,63)
L6	A14	355,23	60,75964	17,10423	(234,54- 529,85)
	P14	77,87	6,983466	8,968005	(61,18- 92,96)
	L14	26,2	3,043091	11,61298	(19,05- 32,76)
	C14	17,48	1,622002	9,276559	(14,86- 24,57)
	AR14	1,5	0,168224	11,21515	(1- 1,94)
L7	A15	552,15	90,91447	16,46545	(365,7- 795,45)
	P15	118,06	7,807717	6,613153	(97,23- 135,13)
	L15	23,44	2,886574	12,31576	(16,89- 29,71)
	C15	45,72	2,890162	6,322097	(36,19- 50,73)
	AR15	0,51	0,048074	9,4054	(0,41- 0,64)
CON15	0,99	0,005839	0,58836	(0,97- 1)	

*Com a redução pela técnica RGCCA, temos para o banco de dados de dados reduzidos para: Shape_VIA1(N=46):Dorsal= D1(A1, L1, W1,CIR1, SOL1); D2 (A2, P2, L2); D3 (A3, P3, L3); D4 (A4, L4, AR4); D5 (A5, P5, L5, AR5); D6 (A6, P6, L6); D7 (A7, P7, L7); D8 (A8, P8, W8). Lateral= L1 (A9, CIR9); L2 (A10); L3 (A11, P11, L11, CIR11); L5 (A13, P13, L13, SOL13); L6 (A14, P14, L14, AR14); L7 (A15, P15, L15, AR15). Shape_VIA2 (N=11):Dorsal: A1, L1, C1,CIR, SOL1, AR1,CON1; Lateral: A9, C9, CIR9, SOL9.

Para o banco de dados que incluem áreas e descritores com a inclusão do peso de carcaça fria, observou-se maiores valores de PECC, nos bancos de dados relacionados ao SHAPE_VIA2, com $k=13$ (53,73% de PECC) no BD2₍₁₎ e $k=13$ (56,72% de PECC) no BD2_(R). De forma geral, os maiores valores de PECC%, dentro de cada banco de dados, foram obtidos de bancos de dados referentes ao SHAPE_VIA2. Sugerindo que banco de dados com menos variáveis/atributos podem favorecer uma maior acurácia do que os banco de dados vinculados ao VIA1 (sem redução pelo RGCCA).

Tabela 3. Acurácia do algoritmo KNN (PECC%) utilizando a metodologia de validação cross validation leave-one-out (LOO) para a classificação do escore de conformação de carcaças de cordeiros com diferentes dimensionalidades

K	Porcentagem de exemplos corretamente classificados (PECC%)							
	Banco de dados com Áreas e descritores				Banco de dados com Áreas e descritores + PCF			
	VIA1		VIA2		VIA1		VIA2	
	BD1 _(I)	BD1 _(R)	BD2 _(I)	BD2 _(R)	BD1' _(I)	BD1' _(R)	BD2' _(I)	BD2' _(R)
3	32,84	37,31	47,76	44,78	37,31	43,28	50,75	53,73
5	40,30	37,31	47,76	52,24	40,3	41,79	46,27	53,73
7	41,79	40,30	47,76	52,24	40,3	44,78	47,76	53,73
9	41,79	41,79	53,73	50,75	35,82	41,79	52,24	43,28
11	46,27	43,28	52,24	50,75	44,78	43,28	52,24	50,75
13	44,78	43,28	53,73	52,24	46,27	40,3	53,73	56,72
15	44,78	44,78	50,75	50,75	44,78	43,28	50,75	52,24
17	46,27	41,79	55,22	53,73	47,76	43,28	49,25	52,24
19	47,76	41,79	49,25	53,73	44,78	40,3	50,75	53,73
21	46,27	43,28	52,24	55,22	47,76	43,28	47,76	55,22
23	43,28	43,28	50,75	55,22	44,78	44,78	50,75	53,73
25	43,28	46,27	46,27	55,22	44,78	44,78	47,76	53,73
27	43,28	43,28	43,28	53,73	43,28	44,78	43,28	52,24
29	46,27	43,28	44,78	53,73	46,27	44,78	47,76	53,73

* (I)= Inicial Dataset; (R): Reduced Dataset; bancos seguidos de ', são aqueles que possuem PCF.

Tabela 4. Matrizes de confusão para as estratégias de classificação do escore de conformação de carcaças ovinas processadas com conjuntos numéricos por k de maior acurácia na melhor estratégia. (K= 13; BD2_(R)- PECC(%)= 56,72)

	Pobre	Subretilínea	Boa	Muito Boa
Pobre	17	3	0	0
Subretilínea	6	19	2	0
Boa	1	14	2	0
Muito Boa	0	3	0	0

Tabela 5. Acurácia do algoritmo Random Forest (PECC%) para a classificação do escore de conformação de carcaças de cordeiros com diferentes banco de dados e diferentes número de árvores

B	Porcentagem de exemplos corretamente classificados (PECC%)							
	Banco de dados com Áreas e descritores				Banco de dados com Áreas e descritores + PCF			
	VIA1		VIA2		VIA1		VIA2	
	BD1 _(I)	BD1 _(R)	BD2 _(I)	BD2 _(R)	BD1' _(I)	BD1' _(R)	BD2' _(I)	BD2' _(R)
5	44,07	42,188	44,26	45,000	43,33	36,66	52,83	37,70
10	42,42	39,39	44,78	45,45	46,26	45,45	44,77	37,31
25	35,82	50,74	41,79	40,299	43,28	52,23	43,28	38,80
50	46,27	44,77	43,28	46,269	41,79	43,28	47,76	43,28
100	50,74	46,26	46,26	38,805	50,74	44,77	35,82	40,29
300	53,73	40,29	43,28	41,79	40,29	40,29	43,28	40,29
500	49,25	44,77	43,28	43,284	50,74	43,28	43,28	44,77
1000	47,76	41,79	43,28	38,806	50,74	44,77	40,29	38,80

Na Tabela 4, está disposta a matriz de confusão relacionado ao maior valor de acurácia obtido. Verifica-se que a classe menor quantificada, “muito boa”, não houve acertos dentro da mesma, os erros da mesma foram classificados de acordo com a classe majoritária (Subretilínea). Assim como ocorreu na classe “Boa”, que teve o mesmo comportamento. Em geral, podemos dizer que o algoritmo conseguiu aprender com mais exemplos obtidos, o que ocorreu com a classe majoritária (subretilínea), que obteve mais acertos dentro da matriz. Com propósito de aperfeiçoar a acurácia de classificação foi realizada uma comparação entre os algoritmos utilizados com o algoritmo Floresta Aleatória – Random Forest (PECC%). Os resultados adquiridos, em sua maiorias foram menores do que 50%, apenas os valores 52,83% (BD2'_(I)) com 5 árvores (B) e 53,73% (BD1_(I)) com 300 árvores foram tidos como maiores dentro da análise, conforme apresenta a Tabela 5.

CONCLUSÃO

Os algoritmos k-NN (k- vizinhos mais próximos e floresta aleatória – *random forest*), foram capazes de prever os escores de conformação a partir de medidas de shape. Comparativamente, o algoritmo (k-NN) demonstrou ser mais eficiente que o Floresta Aleatória. Entretanto, os resultados não foram satisfatórios isto se deve ao volume de dados não

ser muito grande, bem como o desbalanceamento das classes obtidas. Como a metodologia adotada se demonstrou confiável para avaliar a conformação de carcaças, há possibilidade que no futuro se desenvolva uma melhoria dos algoritmos adotados com novas entradas (carcaças). Em uma pesquisa futura, estudos serão feitos com número maior de carcaças, assim como a heterogeneidade dos escores de tipificação para que desta maneira haja um melhor aprendizado por parte dos algoritmos, o que permitirá melhor classificação das carcaças. Também, novos testes serão realizados com outros algoritmos de aprendizado de máquina para futuras comparações. A Tabela A apresenta os Acrônimos e descrições das áreas obtidas na carcaça de cordeiros através da análise de imagens pelo software ImageJ 1.05i; que pode ajudar na compreensão do processamento das imagens.

AGRADECIMENTOS

Os autores agradecem a todos que participaram da pesquisa. À CAPES Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) /PRÓ-AMAZÔNIA pelo auxílio bolsa e experimento. Universidade Federal Rural da Amazônia (UFRA), Universidade Federal do Pará (UFPA) e ao Instituto Federal de Educação, Ciência e Tecnologia do Pará (IFPA) Campus Castanhal. Aos Programas de Pós-Graduação Saúde e Produção Animal na Amazônia e Ciência Animal.

Apêndice

Tabela A. Acrônimos e descrições das áreas obtidas na carcaça de cordeiros através da análise de imagens pelo software ImageJ 1.05i

Acrônimos	Área	Descrição
Dorsal Figura 2- A e B		
D1	Carcaça	Seleção do contorno da carcaça inteira em sua vista dorsal.
D2	Paleta	Compreende a região composta pelos pontos entre as articulações escápulas umerais, as espinhas da escápula e a demarcação da cernelha.
D3	Costela I	Compreende a região composta pelos pontos entre as espinhas da escápula, extremos das 5ª costelas e as demarcações da cernelha e 5ª vértebra torácica.
D4	Costela II	Compreende a região composta pelos pontos extremos das 5ª costelas, extremos das 13ª costelas e as demarcações da 5ª e 13ª vértebras torácicas.
D5	Lombar	Compreende a região composta pelos pontos extremos das 13ª costelas, tuberosidade do coxal e as demarcações da 13ª vértebra torácica e a 1ª vértebra Sacral)
D6	Garupa-pernil	Compreende a região composta pelos pontos das tuberosidades do coxal, as pontas dos íleos e as demarcações da 1ª sacral e inserção da cauda.
D7	Pernil direito	Compreende a região composta pelos pontos PA1, PA2, extremo do bordo interno da perna, PA5 e inserção da cauda
D8	Pernil esquerdo	Compreende a região composta pelos pontos PA5, extremo do bordo interno da perna, PA3, PA4 e inserção da cauda
Lateral Figura 2- C e D		
L1	Carcaça	Seleção do contorno da carcaça inteira em seu painel lateral.

REFERÊNCIAS

- Abramoff, M. D., Magalhães, P. J., & Ram, S. J. 2004. Image processing with ImageJ. *Biophotonics International*, 11(7), 36-42.
- Anchalia, P. P., & Roy, K. 2014. The k-Nearest Neighbor Algorithm using map reduce paradigm. In *Intelligent Systems, Modelling and Simulation (ISMS)*. In: 5th International Conference on IEEE. p. 513-518
- Brett.L. 2013. *Machine learning with R*. Packt Publishing Ltd.
- CEC (Commission of the European Communities). (2002). Report from the Commission to the Council on the implementation of Council Regulation (EEC) No. 2137/92 concerning the Community scale for the classification of carcasses of ovine animals. From: (http://europa.eu.int/eur-lex/lex/LexUriServ/site/en/com/2002/com2002_0295en01.pdf)
- Devijver, P. A., & Kittler, J. 1982. *Pattern recognition: a statistical approach*. Englewood Cliff's, New Jersey: Prentice/Hall.
- Ferreira, Alessandro dos Santos. *Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja*. 2017. Master's Thesis.
- Kaiser, H. F. 1960. The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141-151.
- Moreira, A. C. 2007. Comparação da análise de componentes principais e da CATPCA na avaliação da satisfação do passageiro de uma transportadora aérea. *Investigação Operacional*, 27(2), 165-178.
- Rasband, W. S. 2016. ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA. From (<https://imagej.nih.gov/ij/>)
- Santos, C. L. D., Pérez, J. R. O., Muniz, J. A., Geraseev, L. C., & Siqueira, E. R. D. 2001. Desenvolvimento relativo dos tecidos ósseo, muscular e adiposo dos cortes da carcaça de cordeiros Santa Inês. *Revista Brasileira de Zootecnia*, 487-492.
- SYED, M. E. Attribute weighting in *K-nearest neighbor* classification. Dissertação (mestrado). Escola de Ciências da Informação – School of Information Sciences. Universidade de Tampere, 50 p. 2014
- Witten, I. H., Eibe, F., Hall, M. A. 2011. *Data mining: practical machine learning tools and techniques*. 3° ed. Elsevier: 665 p.
- Young, G. O. 1964. "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15-64.
