

IDENTIFICAÇÃO DE PADRÕES SONOROS DE SUÍNOS VIA TÉCNICAS DE APRENDIZADOS DE MÁQUINA

TUNG CHIUN WEN¹, CAROLINE FERREIRA FREIRE², KÉSIA OLIVEIRA DA SILVA MIRANDA³, GISELLE BORGES DE MOURA⁴

1 Mestre, Prof. Titular e Doutorando, Fatec Unidade São Roque/ Depto. de Engenharia de Biossistemas, Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ/USP), (11) 98673-7494, tungwen@usp.br

2 Engenheira Agrônoma, Mestranda, Depto. de Engenharia de Biossistemas, ESALQ USP

3 Doutora, Prof. Titular, Depto. de Engenharia de Biossistemas, ESALQ/USP

4 Doutora, Prof. Titular, Depto. de Engenharia Agrícola, Universidade Federal de Lavras (UFLA)

Apresentado no
L Congresso Brasileiro de Engenharia Agrícola - CONBEA 2021
08 a 10 de novembro de 2021 - Congresso On-line

RESUMO: A carne suína é uma das mais consumidas do mundo, atender a demanda mundial com quantidade, não é mais único objetivo dos produtores, mas também garantir uma proteína de qualidade e priorizando as questões de bem-estar animal. Nesse cenário, torna-se cada vez mais imprescindível o uso de ferramentas que auxiliem no controle da produção e gerem respostas rápidas e seguras, para que tomadas de decisão no manejo sejam mais assertivas e com caráter preventivo. Este artigo realizou a identificação de vocalizações sociais e agonísticas emitidas por suínos machos e fêmeas, em fase de creche, mantidos em baias de produção. A identificação ocorreu por meio do aprendizado de máquina, fazendo uso do software Weka e tomando como base parâmetros acústicos: amplitude, *pitch*, intensidade, energia do sinal, formantes e harmônica. Foi possível realizar a definição das duas diferentes categorias vocálicas com acurácia de até 82,8%.

PALAVRAS-CHAVE: Aprendizado de máquina; Zootecnia de precisão; Vocalização.

IDENTIFICATION OF SWINE SOUND PATTERNS VIA MACHINE LEARNING TECHNIQUES

ABSTRACT: Swine meat is one of the most consumed in the world, meeting world demand with quantity is no longer the only objective of producers, but also ensuring quality protein and prioritizing animal welfare issues. In this scenario, it is increasingly essential to use tools that help control production and generate quick and safe responses, so that decision-making in management is more assertive and with a preventive character. This article defined social and agonistic vocalizations emitted by female and male nursery swines raised in closed and climatized ambient. The identification was conducted by machine learning using Weka software and the following parameters: amplitude, *pitch*, intensity, signal energy, formant and harmonic. It was possible define two different vocalic categories with 82,8% of accuracy.

KEYWORDS: Machine learning; Precision livestock; Vocalization.

INTRODUÇÃO:

A carne suína se destaca dentre as mais produzidas e consumidas. Porém, produzir o suficiente para suprir a demanda mundial não é mais o objetivo. Com a globalização, o mercado consumidor tornou-se exigente quanto a oferta de produtos de maior qualidade. Por isso, investimentos na área de ambiência e manejo vêm aumentando nos últimos anos (Galvão, et al. 2019). Além da ética de produção, submeter os suínos a condição de estresse gera impactos financeiros, são perdas anuais de até 113 milhões de dólares por ano, segundo Souza et. al (2020). Por meio de publicações de códigos, protocolos e manuais, instalam a nova cultura de produção priorizando a ética na produção, bem-estar animal e, conseqüentemente, garantia de qualidade. Um expressivo exemplo é o protocolo europeu “*Welfare Quality*” (2004). Segundo o protocolo, quatro princípios devem ser trabalhados: boa alimentação, alojamento adequado, boa saúde e comportamento apropriado. Dentro de cada princípio alguns critérios são estabelecidos para guiar os responsáveis pela produção na análise do ambiente de produção. Esta análise, inicialmente realizada de forma visual e/ou manual, tem recebido grandes avanços na tecnologia para tornar o resultado mais rápido e confiável. Dentre os diversos indicadores possíveis de se estimar o bem-estar animal, como por meio do levantamento climático e indicadores fisiológicos, encontra-se a análise de comportamento por meio de vocalização, uma ferramenta entendida como eficiente por muitos autores para identificar sinais de doença (Ferrari, et al. 2008), distinguir sexos e idade e identificar momentos de estresse (Cordeiro, et al. 2018).

Dentre os motivos de grande aceitabilidade da análise via vocalização estão: a velocidade, acurácia, baixo investimento, além de ser uma metodologia não invasiva (Silva, et al. 2019). Acoplado à vocalização, sistemas inteligentes vêm sendo desenvolvidos para interpretar este tipo de banco de dados coletado (Cordeiro, et al. 2018), dentre eles as redes neurais. As redes neurais são ligações de neurônios artificiais, simulando o funcionamento do cérebro humano (Muller, 2019). Descobre-se a rede mais apropriada por meio de testes com submissão parcial do banco de dados e análise de resultados (acurácia alcançada).

Diante desse cenário, entende-se que a vocalização já é um método consolidado para a análise do bem-estar animal. A vocalização é o ato de produzir som com a voz ou, ainda, o ato de expressar os sentimentos, segundo o dicionário de *Cambridge*. Para a produção de suínos, trabalhos apresentam esta ferramenta integrada com análises via árvore de decisão (Cordeiro, 2018), análise de espectrogramas (Moi, M., et al. 2015), conceitos matemáticos (Silva, et al. 2019), aprendizagem de máquina (Zulato, 2010), além do uso de sistemas de monitoramento de chamadas com estresse (Puppe, et al. 2005). As redes neurais têm sido bastante utilizadas para aplicação em análise de imagens (Lee, J. et al. 2016; Chen, C. et al. 2017; Brown e Riede, 2017), porém ainda não foi amplamente explorada a aprendizagem para a análise de vocalização, a qual pode proporcionar um treinamento constante com a inserção de novos dados, permitindo trabalhar com um banco denso de dados e com resultados de alta precisão. Com isso, o presente trabalho apresenta a análise de vocalização suína em grupo, por meio de redes neurais supervisionadas e técnicas de *Machine Learning*, avaliando e diferenciando vocalizações agonísticas de vocalizações sociais.

MATERIAL E MÉTODOS:

Coleta de dados

O experimento foi realizado na Universidade de Illinois - Campus Urbana Champaign (Estados Unidos), em uma baia de produção. Foram coletadas frequências sonoras, emitidas pelos animais em quatro baias de produção de suínos.

Animais

Vinte e quatro suínos na fase de creche foram separados em quatro baias de produção sendo doze animais fêmeas e doze machos. As salas se encontram nas extremidades da instalação (conforme Figura 1), sendo a sala 1 alocada à região noroeste (NO), sala 2 à nordeste (NE), sala 3 sudeste (SE) e sala 4 à sudoeste (SO).



Figura 1 - Planta baixa da instalação experimental climatizada.

Em cada sala os animais foram alocados em uma baia contendo um bebedouro e três comedouros, o microfone para coleta dos dados sonoros foi instalado ao centro da baia a 0,8 m de distância do piso (Figura 2).



Figura 2 - Estrutura da baia: três comedouros à esquerda, um bebedouro à direita. Microfone e controladores de temperatura ao centro.

Gravação dos sinais sonoros

As frequências sonoras geradas nas quatro baias foram gravadas por microfones unidirecionais (XM8500, Behringer Inc. Bothell, USA), conectados a um amplificador de sinal e, este, a uma placa de áudio e vídeo conectada a um computador. Durante todo o período de experimento, áudios foram gravados e todos que apresentaram até dez segundos de duração foram avaliados utilizando a classificação supervisionada.

Vocalizações e ruídos

O sinal sonoro é um dado, e para que seja interpretado como informação, precisa ser transformado, antes de ser processado. O som na natureza é analógico e precisa ser digitalizado durante a coleta e armazenamento do som. Isso é realizado coletando amostras digitalizadas do som original em uma frequência duas vezes maior que a frequência natural desse som, chamado teorema de Nyquist. Um processo para extrair mais informações de um sinal sonoro, é utilizar a Transformada Rápida de Fourier, para que o som originalmente coletada no domínio do tempo, passe a ser caracterizado no domínio da frequência, de onde é possível obter informações do som, como frequência fundamental (*pitch*), harmônicas, amplitudes e energia (Zulato, 2010; Borges, 2013)

Outro tratamento importante em experimentos que envolvem coleta de sinais sonoros, é a eliminação de ruídos, em geral providos do ambiente, como: veículos, trânsito de pessoas, batidas de portas e climatizadores. E esses podem ser filtrados (passa-faixa) através de softwares ou dispositivos. A partir dos áudios processados e parâmetros acústicos serem extraídos pelos filtros, os padrões sonoros já podem ser estabelecidos. A padronização pode ocorrer analisando o espectrograma, como imagem, ou analisando os parâmetros numéricos, também extraídos do espectrograma (Borges, 2013).

Tratamento e aplicação dos dados sonoros

Os dados sonoros foram gravados em formato *wav* (*wave form audio file format*). Arquivos de sons muito extensos (longa duração) trazem diferentes e somadas informações simultâneas, o que dificulta a categorização e sua acurácia, por isso apenas áudios de dez segundos foram considerados. Arquivos de curta duração apresentam captação de menor diversidade de sons gerados e este fator aumenta a acurácia. Todos os áudios foram filtrados, por meio de filtros passa-baixa (8000 Hz) e passa-alta (650 Hz), e tiveram os seguintes parâmetros acústicos extraídos: energia do sinal, *pitch*; intensidade mínima, média e máxima; amplitude mínima, média e máxima; formantes e harmônicas. Por fim, os arquivos foram separados em categorias: vocalizações sociais (*social*) e vocalizações agonísticas (*agonístico*).

Com base nos dados extraídos pelo MatLAB, foram selecionados 32 amostras da classe “social” e 32 amostras de “agonístico”, submetidos a quatro algoritmos de *machine learning* do pacote WEKA: *K-Nearest Neighbour* (*K*-KStar*); Redes Neurais Artificiais (*Multi Layer Perceptron*); *Support Vector Machine*; *Random Tree*.

Para avaliação dos algoritmos de reconhecimento vocálico, os resultados processados pelo Matlab foram planilhados no formato CSV. E para análise do desempenho dos algoritmos, foi utilizado o pacote de ferramentas do Weka (*Waikato Environment for Knowledge Analysis*), através da conversão das planilhas CSV em arquivos no formato ARFF (nativo do pacote Weka). Para a análise comparativa entre os algoritmos de aprendizado de máquina, o Weka é muito prático, por utilizar um formato padrão de dados para todos os algoritmos do pacote.

O arquivo de formato ARFF foi carregado no pacote Weka, e alguns parâmetros foram retirados (pré-processamento), como, número da sala, horário, nome do arquivo da gravação e local onde foi armazenado o arquivo. Em seguida, o conjunto de registros (amostras de classes) foi submetido aos algoritmos do pacote de mineração de dados Weka, para obter os modelos de classificação (Figura3).

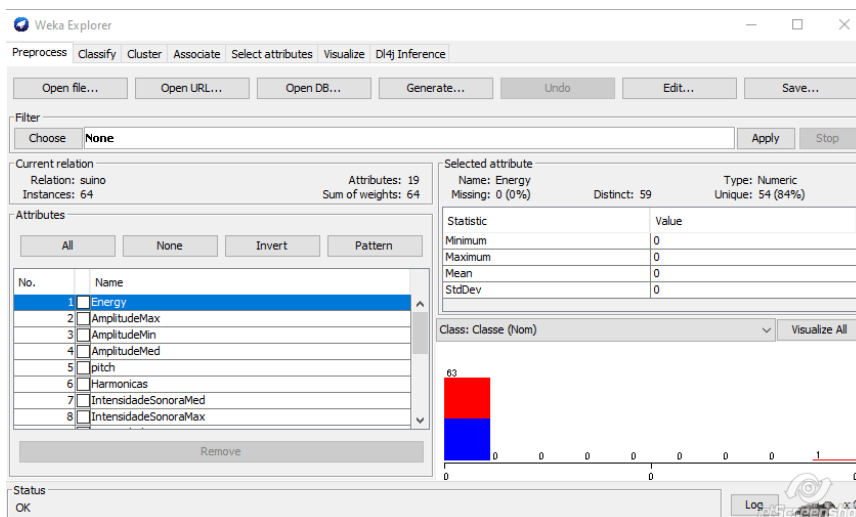


Figura 3 - Base de Dados para treinamento.

Classificação e desempenho

Para avaliar a acurácia do modelo obtido pelo treinamento, o Weka separa algumas amostras (registros) para serem testadas após o treinamento. O projeto utilizou alguns dos algoritmos de aprendizado de máquina supervisionadas (KNN(*K-Nearest Neighbor*), K^* , Redes Neurais (MLP-*Multi Layer Perceptron*), SVM (*Support Vector Machine*) e Árvores Randômicas (*Random Trees*), para treinar e avaliar as suas respectivas acurácias.

A eficiência dos modelos classificadores (qualitativos), são extraídos da matriz confusão (Figura 4), que expressa a performance de um modelo obtido pelo algoritmo de aprendizado supervisionado. Onde: TP-*TruePositive* (Positivo Verdadeiro); TN-*TrueNegative* (Negativo Verdadeiro); FP-*FalsePositive* (Falso Positivo); FN-*FalseNegative* (Falso Negativo) e N-número de amostras testadas

| | | |
|-------------|----------------|----|
| | Classe Predita | |
| Classe Real | TP | FN |
| | FP | TN |

Figura 4 - Matriz Confusão.

As métricas utilizadas para avaliar o desempenho dos métodos, serão: acurácia, precisão e sensibilidade (*recall*). Onde:

$$\text{ACURACIA} = (TP+TN)/N; \quad (1)$$

$$\text{PRECISÃO} = TP/(TP+FP); \quad (2)$$

$$\text{SENSIBILIDADE} = TP/(TP+FN) \quad (3)$$

No projeto, os classificadores interpretam duas classes comportamentais de suínos: Social e Agonístico, e a adaptação para a matriz confusão foi convencionado como Positivo=Social e Negativo=Agonístico.

RESULTADOS E DISCUSSÃO

Tabela1: Resultados obtidos nos algoritmos de *Machine Learning*

| Indicadores | KStar | RNA(MLP) | SVM | RT |
|-----------------------------|--------------|-----------------|------------|-----------|
| TP (social) | 25 | 18 | 24 | 24 |
| TN (agonistico) | 28 | 22 | 23 | 23 |
| FP (falso social) | 4 | 10 | 9 | 9 |
| FN (falso agonistico) | 7 | 14 | 8 | 8 |
| Acurácia | 82,81% | 62,5% | 73,44% | 73,43% |
| PrecisãoS (social) | 0,862 | 0,643 | 0,727 | 0,727 |
| PrecisãoA (agonistico) | 0,800 | 0,611 | 0,742 | 0,742 |
| SensibilidadeS (social) | 0,781 | 0,563 | 0,750 | 0,750 |
| SensibilidadeA (agonistico) | 0,875 | 0,688 | 0,730 | 0,719 |

Comparando os algoritmos nos quais os dados foram submetidos, o algoritmo “Vizinho Mais Próximo” (KStar) obteve melhor resultado de classificação (82,81%), seguido de SVM e “Arvore Randômica” (RT), que obtiveram resultados próximos (73,44% e 73,43%, respectivamente), finalizando com o RNA (62,5%). Comparados aos outros algoritmos, o KStar apresentou uma acurácia, precisão e sensibilidade superior.

A acurácia $[(TP + TN) / (TP + TN + FP + FN)]$ consiste no desempenho, ou precisão geral, da categorização das classes previamente definidas, no caso: vocalização social e vocalização agonística. Enquanto que a precisão $[TP / (TP + FP)]$ consiste no desempenho de cada classe de forma individual. Com isso, podemos identificar qual classe foi melhor definida em cada técnica utilizada. Este parâmetro permite identificar qual das classes é definida de forma mais eficaz em cada rede supervisionada utilizada. No presente estudo, a precisão se apresentou de forma semelhante para ambas as categorias, com isso, a rede Kstar permanece adequada para fazer uso da classificação de vocalizações sociais e agonísticas. Por fim, a sensibilidade $[TP / (TP + FN)]$, também conhecida como *recall*, traduz o acerto da categorização de cada classe levando em consideração o número de arquivos atribuídos previamente a cada uma. O que pode ser muito útil quando atribui-se números diferentes de arquivos para cada categoria, nestes casos, o *recall* indica se há um desbalanceamento devido ao número de registros atribuídos a cada classe. Para a presente análise, atribuiu-se o mesmo número de vocalizações para ambas as categorias, na intenção de não provocar o desbalanceamento ou o vício da rede em categorizar apenas uma classe. Dessa forma, a sensibilidade permite avaliar os acertos em cada classe de uma forma normalizada. Ainda há outros parâmetros (F1 e *Score*) utilizados para casos em que há desbalanceamento entre classes, o que não está aplicado a este estudo.

As vocalizações que compõe o presente banco de dados apresentam características distintas quanto à sua composição, formando um grupo de vocalizações graves e de curta duração, os grunhidos, e um segundo grupo de vocalizações agudas e de maior tempo de duração, se assemelhando à gritos. Essas distinções permitiram realizar a distinção das vocalizações em sociais (graves) e agonísticas (agudas), o que está de acordo com artigos relacionados à temática em que são definidos sons sociais como aqueles de menor intensidade, energia de sinal e duração; e sons relacionados à dor ou desconforto com maior intensidade, energia de sinal e duração (Cordeiro, et al. 2018; da Silva, et al. 2019). Como os suínos são animais sociais e apresentam hierarquia quando submetidos em grupo (Massari, et al. 2015), é comum que sob mudanças de grupos os animais se submetam a brigas até que a classe dominante seja definida, o que justifica sons agonísticos em baias de produção. Além disso, os animais apresentam repertório de vocalização bastante extenso, com objetivo de atrair parceiros até para avisar possíveis perigos ou intimidadas potenciais agressores (Cordeiro, et al. 2008). Visto tamanha riqueza, é natural que haja muitas vocalizações, sociais e agonísticas, em baias de produção, podendo revelar a situação de um ambiente e sua compatibilidade com o bem-estar dos animais, o que acarreta em índices produtivos.

CONCLUSÕES: Foi possível fazer uso de técnicas de aprendizado supervisionadas para classificar diferentes tipos de vocalização de suínos criados em grupo, destacando vocalizações sociais e vocalizações agonísticas. A técnica que apresentou melhor desempenho foi a *Kstar*, a qual, até o momento, não havia sido aplicada para este tipo de análise (vocalização para suínos).

A metodologia proposta para esse projeto se mostrou viável para aplicação em suínos, e com chances de ser viável em outras espécies produtivas. Na análise comparativa com os cinco algoritmos de aprendizado de máquina, os dois melhores resultados obtidos (KNN e K*) são da mesma categoria (algoritmos preguiçosos), que não geram modelos, os demais algoritmos permitem a implementação em dispositivos embarcados e dispositivos móveis.

REFERÊNCIAS:

Borges, G. **Vocalização de suínos em grupo sob diferentes condições térmicas**. 2013. Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, 2013.

Cordeiro, A.F. da S.; et al. Use of vocalisation to identify sex, age, and distress in pig production. **Biosystems engineering**, v. 173, p. 57-63, set., 2018.

Ferrari, S.M. et al. Analysis of Cough Sounds for Diagnosis of Respiratory Infections in Intensive Pig Farming. **American Society of Agricultural and Biological Engineers**, v. 51, n. 3, p. 1051 – 1055, 2008.

Galvão, A. T.; et al. Bem-estar animal na suinocultura. **Putbvet – Medicina Veterinária e Zootecnia**, v. 13, n. 3, a289, p. 1-6, mar., 2019.

Lee, J. et al. Automatic Recognition of Aggressive Behaviour in Pigs Using a Kinect Depth Sensor. **Sensors**, v. 16, n. 5, p. 631 –641, mai. 2016.

Massari, J. M. et al. Características comportamentais de suínos em crescimento e terminação em sistema “*wean to finish*”. **Engenharia Agrícola**, v. 35, n. 4, p. 646-656, jul./ago. 2015.

Moi, M., et al. Vocalização como indicativo do bem-estar de suínos submetidos a situações de estresse. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 67, n. 3, p. 837-845, 2015.

Muller, T.B. **Aplicação de redes neurais artificiais na avicultura comercial**. Trabalho de Conclusão de Curso em Medicina Veterinária. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

Puppe, B., et al. Castration-induced vocalization in domestic piglets, *Sus scrofa*: Complex and specific alterations of the vocal quality. **Applied Animal Behaviour Science**, v. 95, p. 67-78, nov., 2005.

Silva, J.P. da et al. Classification of piglet (*Sus Scrofa*) stress conditions using vocalization pattern and applying paraconsistent logic Π . **Computer and electronics in agriculture**. São Paulo, 166 (2019), p. 1 – 6.

Souza, A. V.; et al. Quanto custa o estresse por calor na produção de aves e suínos. **Nutritime Revista Eletrônica**, Viçosa, v. 17, n. 01, p. 8647 – 8653, jan./fev., 2020.

Welfare Quality, 2009. Disponível em: <http://www.welfarequalitynetwork.net/media/1018/pig_protocol.pdf>. Acesso em: 01, jul. 2020.

Zulato, P. R. F. **Análise de padrões de vocalização para determinar o bem-estar de suínos**. 2010. 51 f. TCC (Graduação) - Curso de Engenharia Elétrica, Eletrônica, Universidade de São Paulo, São Carlos, 2010.