

Detecção de Eventos Sonoros para Sistemas de Segurança

Tito Caco Curimbaba Spadini e Ricardo Suyama

Resumo—O presente trabalho visa investigar técnicas para classificação de eventos sonoros atípicos para sistemas de segurança baseados no monitoramento por meio do som ambiente. Para isso, são comparadas diferentes técnicas de classificação, considerando a acurácia e o tempo de classificação, com o intuito de encontrar as mais vantajosas para cenários de vigilância, minimizando a possibilidade de eventos relevantes passarem despercebidos. Foram confrontadas técnicas de LDA, QDA, *Decision Tree* e KNN. Os resultados obtidos indicam que o KNN é a técnica mais indicada para as finalidades descritas.

Palavras-Chave—Processamento Digital de Sinais, Processamento Digital de Áudio, Aprendizagem de Máquinas, Reconhecimento de Padrões, Sistema de Vigilância.

Abstract—This work aims to investigate different techniques for the classification of atypical sound events, considering the accuracy and classification time, in order to find which of these techniques are most advantageous for surveillance scenarios, minimizing the possibility of relevant events to go unnoticed. LDA, QDA, *Decision Tree* and KNN techniques were compared. The results obtained indicate that the KNN is the most indicated classification method for the purposes described.

Keywords—Digital Signal Processing, Digital Audio Processing, Machine Learning, Pattern Recognition, Surveillance System.

I. INTRODUÇÃO

A área de segurança demanda aprimoramentos que culminem em maiores precisão e velocidade na identificação de eventos de risco. A partir de técnicas de reconhecimento de padrões é possível identificar tais situações através do áudio do ambiente vigiado. Nesse sentido, o objetivo do presente trabalho é investigar técnicas de classificação de padrões que permitam a classificação correta de eventos sonoros - como *Linear Discriminant Analysis* (LDA) [1], *Quadratic Discriminant Analysis* (QDA), *Decision Tree* [2] e *k-Nearest Neighbors* (KNN) [3] - a fim de encontrar a soluções que proporcionem maiores precisões e velocidades.

II. DESENVOLVIMENTO

Os testes foram realizados com uma base de milhares de amostras de áudios, obtidas de diferentes bases gratuitas, como *freesound* [4], *soundbible* [5] e *freesfx* [6], com 4 categorias: “casual” (sem qualquer evento de risco), “explosão”, “tiro” e “sirene”. Os áudios foram gravados em padrão Mono, 16 kHz, 8-bit *Signed* em WAV e separados em grupos de treinamento, validação e teste. Com o auxílio da *MIRtoolbox* [7], foram

Tito Caco Curimbaba Spadini e Ricardo Suyama. Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas, Universidade Federal do ABC, Santo André-SP, Brasil, E-mails: tito.caco@aluno.ufabc.edu.br, ricardo.suyama@ufabc.edu.br. Trabalho parcialmente financiado pelo CNPq.

extraídas as *Features Zero-Crossing Rate* (ZCR), *Root Mean Square* (RMS), *Brightness*, *Spread*, *Skewness*, *Kurtosis*, entropia, *Spectral Flatness Measure* (SFM) e *Mel Frequency Cepstral Coefficient* (MFCC) de cada áudio, totalizando 21 *Features*, visto que o MFCC envolve 13 coeficientes.

Para a classificação dos trechos de áudio foram utilizadas 4 abordagens: LDA, QDA, *Decision Tree* e KNN em conjunto com um *Wrapper* [8] baseado em *Greedy Forward Search* para selecionar a melhor composição de *Features* para o *Input*. Após o *Wrapper* encontrar a melhor composição de *Features*, foram feitos testes com diferentes composições de áudios de entrada que não compunham as bases utilizadas nas etapas de treinamento e validação.

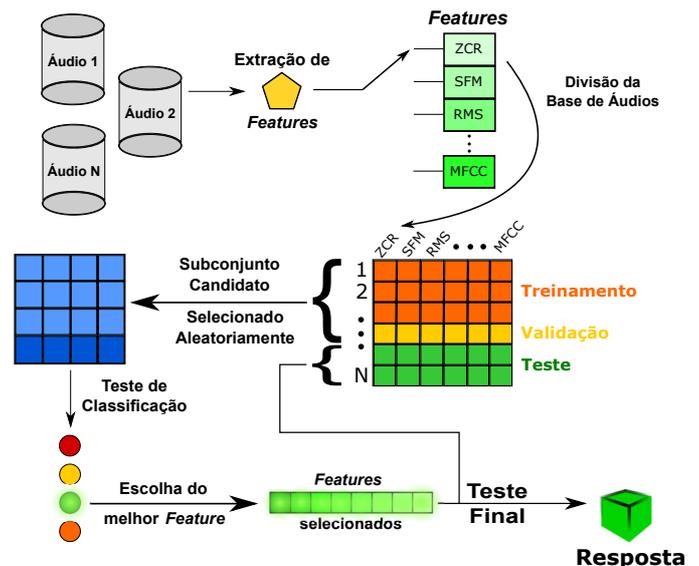


Fig. 1. Passos do trabalho, desde a base de áudios até a resposta (classificador treinado com os *Features* selecionados).

No diagrama da Fig. 1 é possível ver os passos ao longo do desenvolvimento do projeto.

III. RESULTADOS E DISCUSSÃO

A Tabela I oferece uma visão global dos classificadores abordados neste trabalho nos quesitos acurácia e tempo de classificação. Por ela, podemos constatar a discrepante diferença de acurácia (de mais de 20%) entre os classificadores LDA e KNN, bem como o *Tree* ter se apresentado como aproximadamente 82% mais rápido que o KNN.

Conforme mostra a Tabela I, o classificador *Tree* se destacou pela sua velocidade de classificação e o classificador KNN se destacou por sua acurácia elevada.

TABELA I

DESEMPENHO DOS CLASSIFICADORES EM TERMOS DE ACURÁCIA E INTERVALO MÉDIO DE TEMPO PARA REALIZAR UMA CLASSIFICAÇÃO (ΔT).

Classificador	ΔT [ms]	Acurácia [%]
LDA	3.5 ± 0.30	70.12 ± 0.88
QDA	3.1 ± 0.26	76.19 ± 0.93
Tree	1.8 ± 0.17	81.72 ± 0.74
KNN	9.9 ± 0.83	90.47 ± 0.64

O classificador *Decision Tree* exibiu acurácias relativamente baixas para as categorias 0 (“casual”) e 1 (“explosão”), além de uma elevada taxa de *misclassifications* para o grupo em que a categoria “casual” é a predição, sendo que este é o grupo de maior preocupação. As *misclassifications* deste grupo representam os eventos que realmente se tratam de riscos, mas que foram classificados como inócuos, o que pode ser entendido como o pior cenário possível. Porém, esse classificador se mostrou muito eficiente na classificação da categoria 2 (“tiro”).

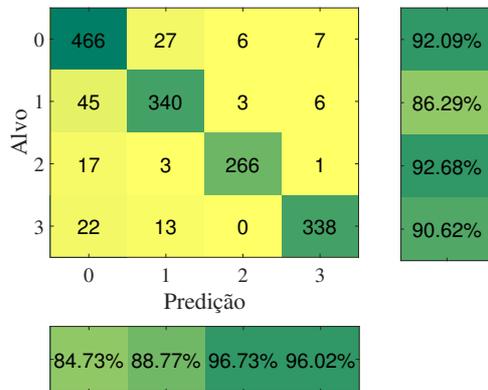


Fig. 2. Matriz de confusão do classificador KNN.

A matriz de confusão do KNN (Fig. 2) mostra um classificador bem mais confiável por conta de suas maiores acurácias, sobretudo quanto ao grupo de maior risco, onde eventos de verdadeiro risco são erroneamente classificados como “casual”.

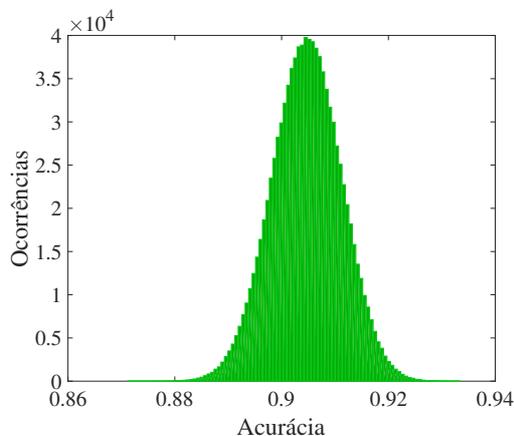


Fig. 3. Histograma de distribuição da acurácia do classificador KNN.

se conferir o histograma de distribuição da acurácia (Fig. 3), que exibe um comportamento de considerável estreitamento, indicando elevada confiabilidade de sua acurácia.

IV. CONCLUSÕES

Este trabalho indica que, para fins de classificação de situações de potencial risco à segurança com base no áudio ambiente, após terem sido confrontadas as técnicas de LDA, QDA, *Decision Tree* e KNN, o classificador KNN exibiu a maior acurácia, enquanto a *Decision Tree* exibiu o menor tempo de classificação. Em futuros trabalhos, pretende-se estudar o comportamento de misturas de classificadores e confrontar os resultados obtidos com os deste trabalho.

AGRADECIMENTOS

Agradecemos ao CNPq e à UFABC pelo financiamento e pelo apoio ao projeto.

REFERÊNCIAS

- [1] Y. Guo, T. Hastie, e R. Tibshirani, “Regularized linear discriminant analysis and its application in microarrays,” *Biostatistics*, vol. 8, n.º 1, pag. 86, 2007. [Online]. Disponível em: [+http://dx.doi.org/10.1093/biostatistics/kxj035](http://dx.doi.org/10.1093/biostatistics/kxj035)
- [2] L. Breiman, J. Friedman, C. J. Stone, e R. A. Olshen, *Classification and regression trees*. CRC press, 1984.
- [3] N. S. Altman, “An introduction to kernel and nearest-neighbor nonparametric regression,” *The American Statistician*, vol. 46, n.º 3, pags. 175–185, 1992. [Online]. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/00031305.1992.10475879>
- [4] “Freesound.org - freesound.org,” <https://www.freesound.org/>, (Acessado em 04/13/2017).
- [5] “Free sound clips — soundbible.com,” <http://soundbible.com/>, (Acessado em 04/13/2017).
- [6] “freesfx.co.uk - download free sound effects,” <http://www.freesfx.co.uk/>, (Acessado em 04/13/2017).
- [7] O. Lartillot, P. Toivainen, e T. Eerola, *MIRtoolbox 1.6.1 User's Manual*, Finnish Centre of Excellence in Interdisciplinary Music Research, dec 2014.
- [8] I. Guyon e A. Elisseeff, “An introduction to variable and feature selection,” *J. Mach. Learn. Res.*, vol. 3, pags. 1157–1182, Mar. 2003. [Online]. Disponível em: <http://dl.acm.org/citation.cfm?id=944919.944968>

Para comprovar a confiabilidade do classificador KNN, pode-