

## IDENTIFICAÇÃO DE CENA ACÚSTICA UTILIZANDO CLASSIFICAÇÃO HIERÁRQUICA

Lucas Wolschick (PIC/UEM), Yandre Maldonado e Gomes da Costa (Orientador).  
E-mail: ymgcosta@uem.br.

Universidade Estadual de Maringá, Centro de Tecnologia, Maringá, PR.

**Área e subárea do conhecimento: Ciência da Computação / Metodologia e Técnicas da Computação.**

**Palavras-chave:** reconhecimento de padrões; algoritmos evolutivos; classificação hierárquica.

### RESUMO

Identificação de Cena Acústica (ICA) é um problema de reconhecimento de padrões em que o objetivo é determinar o tipo de ambiente acústico em que uma gravação sonora foi realizada. A ICA é usualmente abordada como um problema multiclasse, onde um classificador é treinado para distinguir todas as classes do problema ao mesmo tempo. A Classificação Hierárquica (CH) é um modo de realizar classificação em bases com classes organizadas hierarquicamente, e ela já foi utilizada em outros trabalhos de reconhecimento de padrões com bons resultados. Nela, classificadores são treinados em porções específicas da hierarquia e depois combinados. Neste trabalho, aplicamos a CH ao problema de ICA. Usamos uma base de dados com cenas acústicas urbanas e descritores de características acústicas e texturais. Exploramos diferentes modos de fusão de características em CH e propomos um método de geração de hierarquias usando algoritmos genéticos. Ao fim, alcançamos a melhor acurácia, de 66,6% usando um classificador hierárquico. Concluímos, assim, que o uso de CH em ICA é viável e capaz de produzir resultados positivos.

### INTRODUÇÃO

É fácil para o ser humano, dotado da audição, identificar o ambiente em que se encontra a partir dos sons e ruídos que escuta. Esse problema, quando resolvido por um computador, é conhecido como Identificação de Cena Acústica (ICA). A ICA é um problema de reconhecimento de padrões, e compreende as etapas de obtenção de dados a partir de uma base, extração de características, treinamento e classificação de amostras (DING et al., 2020). A ICA pode ser abordada como um

problema de classificação multiclasse, mas existem outras abordagens que podem ser utilizadas, como a classificação hierárquica (CH). Na CH, as classes do problema são organizadas em uma hierarquia na forma de uma árvore, cuja estrutura é considerada durante o treinamento e a classificação (SILLA; FREITAS, 2011). Na CH local, classificadores distintos são treinados em porções da hierarquia e os seus resultados são então combinados em um. Algumas formas comuns de dividir a hierarquia são por nível da árvore (LCL), um por nó (LCN) e um por conjunto de nós filhos (LCPN). A hierarquia de um problema pode ser montada de diversas formas, tanto manualmente como de maneira automatizada. Estudos anteriores alcançaram bons resultados com a aplicação de CH em diversos contextos (PEREIRA et al., 2020). Neste trabalho, aplicamos a CH ao problema de ICA, e exploramos diferentes características e técnicas de combinação delas em vários cenários, buscando determinar se a CH gera bons resultados na ICA.

## MATERIAIS E MÉTODOS

A base de dados utilizada neste trabalho é a *TAU Urban Acoustic Scenes, development dataset* (2019) (HEITTOLA; MESAROS; VIRTANEN, 2019), do desafio DCASE (*Detection and Classification of Acoustic Scenes and Events*). A base contém dez classes de cenas acústicas urbanas com 1 440 amostras de dez segundos cada, totalizando 14 400 amostras, e é previamente balanceada com uma divisão entre treino e teste de proporção 70:30. No desafio, a melhor acurácia alcançada para essa base foi de 88,4%. As características extraídas das amostras da base de dados incluem texturais e acústicas: coeficientes mel-cepstrais (MFCC), padrões rítmicos (RP), matrizes de co-ocorrência (GLCM), padrões binários locais (LBP) e quantização de fase local (LPQ), sendo estes três extraídos de espectrogramas gerados das amostras. O classificador utilizado foi o *Support Vector Machines* (SVM). Testamos duas hierarquias: uma que agrupa as classes em ambientes externos, internos e transporte, sugerida pelo DCASE de 2020; e outra, encontrada com um algoritmo genético para geração de hierarquias automatizadas de até três níveis, induzidas a partir de partições aleatórias do conjunto das classes. Como função de *fitness* desse algoritmo, foi adotada a acurácia obtida a partir da classificação de 15% da base utilizando a hierarquia avaliada com o LBP. Este trabalho foi desenvolvido da seguinte forma: primeiro, um classificador plano multiclasse foi desenvolvido e treinado sobre as dez classes do problema, para ser utilizado como *baseline* de comparação. Foram testadas 22 combinações de características em diferentes esquemas de combinação, e o esquema com melhor desempenho foi selecionado para ser estudado com CH. Foi realizada uma análise comparativa entre os três tipos de divisão da hierarquia, LCL, LCN e LCPN, as duas

abordagens de combinação de características, *internal* e *external late fusion*, e entre as duas hierarquias acima, totalizando oito combinações nesta etapa. O desempenho dos classificadores foi medido com as medidas estatísticas de acurácia, precisão, *recall* e F1-score.

## RESULTADOS E DISCUSSÃO

Após uma otimização prévia dos hiper-parâmetros do classificador SVM, testamos várias características e métodos de combinação destas. Testamos individualmente o LPQ, LBP, RP, GLCM e MFCC, e depois testamos *early fusion* e *late fusion* – métodos de combinação de características que ocorrem antes e depois da classificação, respectivamente – dessas características com regras de fusão da soma e do produto. A Tabela 1 destaca os melhores resultados encontrados para cada tipo de fusão. Individualmente, o MFCC obteve a melhor acurácia com 58,6%, e o melhor tipo de fusão foi a *late fusion*, usando o produto, com 66,3% de acurácia. Ao todo, testamos 22 configurações. A melhor delas, LPQ+MFCC+RP com *late fusion* e regra do produto, foi usada como base para a classificação hierárquica.

**Tabela 1 – Melhores resultados encontrados para o classificador plano**

Tipo de fusão	Características	Acurácia	Precisão	Recall	F1-Score
--	MFCC	0.586	0.596	0.586	0.589
<i>Early</i>	LPQ+MFCC	0.646	0.652	0.656	0.647
<b><i>Late (regra do produto)</i></b>	<b>LPQ+MFCC+RP</b>	<b>0.663</b>	<b>0.666</b>	<b>0.663</b>	<b>0.663</b>
<i>Late (regra da soma)</i>	LPQ+MFCC+RP	0.648	0.650	0.648	0.647

Os melhores resultados obtidos para a classificação hierárquica estão dispostos na Tabela 2. Foram testados os esquemas LCN, LCL e LCPN com *internal* e *external late fusion* para duas hierarquias, totalizando 12 experimentos. A Hierarquia 1 é a hierarquia sugerida na descrição da base de dados, contendo três grandes classes englobando ambientes externos, ambientes internos e no interior de meios de transporte. A Hierarquia 2 foi a encontrada durante a otimização e possui quatro grandes divisões, duas delas com duas classes e as outras duas com três. Nestes experimentos, obtivemos resultados ligeiramente melhores que o da classificação plana, sendo o melhor resultado o do *external late fusion* LCN na hierarquia 1, com uma acurácia de 66,6% (75% da melhor acurácia conhecida para essa base). Os resultados sugerem que há espaço para exploração de hierarquias alternativas que podem obter resultados ainda melhores.

**Tabela 2 – Melhores resultados encontrados para o classificador hierárquico**

Hierarquia	Tipo de <i>late fusion</i>	Melhor divisão	Acurácia	Precisão	Recall	F1-Score
1	<i>external</i>	LCN	0.666	0.667	0.666	0.665
	<i>internal</i>	LCN	0.660	0.662	0.660	0.660
2	<i>external</i>	LCL	0.656	0.660	0.656	0.656
	<i>internal</i>	LCN	0.644	0.650	0.644	0.644

## CONCLUSÕES

Este trabalho investigou o uso da classificação hierárquica no problema de identificação de cena acústica. Em nossos experimentos, alcançamos a maior acurácia das configurações testadas, 66,6%, com um classificador hierárquico. As diferenças de desempenho observadas entre as hierarquias sugerem que podem existir outras com desempenhos ainda maiores. Concluímos que a CH consegue obter melhores resultados que a classificação plana na ICA. No futuro, exploraremos o uso de mais descritores de características, como dissimilaridade, aprendizagem contrastiva, aprendizagem profunda, e de outras técnicas de geração de hierarquias em busca de melhorias no desempenho na tarefa de ICA.

## AGRADECIMENTOS

Agradecemos ao PIC/UEM pela oportunidade desta pesquisa e ao Ministério da Educação pelo Programa de Educação Tutorial, do qual o autor é bolsista.

## REFERÊNCIAS

DING, Biyun et al. Acoustic scene classification: a comprehensive survey. **Expert Systems with Applications**, p. 121902, 2023.

HEITTOLA, Toni; MESAROS, Annamari; VIRTANEN, Tuomas. **TAU urban acoustic scenes 2019, development dataset**. Zenodo, 11 mar. 2019.

PEREIRA, Rodolfo M. et al. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. **Computer methods and programs in biomedicine**, v. 194, p. 105532, 2020.

SILLA, Carlos N.; FREITAS, Alex A. A survey of hierarchical classification across different application domains. **Data mining and knowledge discovery**, v. 22, p. 31-72, 2011.