



IDENTIFICAÇÃO AUTOMÁTICA DE PÁSSAROS COM USO DE ESPECTROGRAMAS

Gustavo Henrique Gomes Matsushita (PIBIC/CNPq/FA/UEM), Yandre Maldonado e Gomes da Costa (Orientador), e-mail: yandre@din.uem.br

Universidade Estadual de Maringá / Departamento de Informática

Ciências Exatas e da Terra / Ciência da Computação

Palavras-chave: reconhecimento de padrões, espectrograma, classificação de pássaros.

Resumo:

O presente trabalho tem por finalidade apresentar e avaliar o desempenho de um sistema de classificação automática de pássaros em três diferentes taxonomias: família, gênero e espécie. Duas bases com amostras de cantos de aves foram utilizadas. A primeira é composta por 94 espécies, 79 gêneros e 32 famílias. Já a segunda, por 547 espécies, 316 gêneros e 55 famílias. O método se baseia na extração de características presentes principalmente na textura do espectrograma, utilizando o descritor *Local Binary Pattern*. O sistema de classificação emprega *Support Vector Machine* com validação cruzada sobre a base de dados, além de técnicas para combinação de resultados. Tais combinações foram realizadas fundindo informações complementares existentes tanto em características obtidas de um espectrograma, quanto de características acústicas, resultando nas melhores taxas de acerto. Além disso, técnicas de filtragem foram aplicadas, melhorando a qualidade dos áudios e o resultado final.

Introdução

Para a conservação da biodiversidade das aves é necessário obter o conhecimento exato da identidade das mesmas, além de conhecer sua distribuição geográfica, uma vez que os pássaros possuem um papel fundamental para o ecossistema, como o controle de insetos, dispersão de sementes e polinização. Entre as técnicas utilizadas para a identificação das espécies desses animais, é muito comum a montagem de redes de neblina. Essa prática se baseia em suspender uma rede entre dois pontos, possibilitando a captura dos pássaros que se prendem a ela. Todavia,





práticas desse tipo colocam em risco a integridade das aves de uma região. Devido à preocupação com o bem-estar das aves, especialistas sugerem que técnicas não invasivas devam ser utilizadas desde a coleta de dados até o reconhecimento. Sendo assim, o objetivo desse trabalho é desenvolver um método de reconhecimento e classificação de pássaros em diferentes unidades taxonômicas, como família, gênero e espécie, com base em características presentes principalmente na textura de espectrogramas extraídos de sinais de áudio, com metodologia semelhante à utilizada em LUCIO e COSTA (2015).

Materiais e métodos

Para o desenvolvimento dos experimentos que fazem parte desta proposta, dois conjuntos de áudios com cantos de aves foram organizados, compondo as bases utilizadas no sistema de classificação. A Base I é composta por 94 espécies, 79 gêneros e 32 famílias, dividida em 10 *folds*. Já a Base II contém 547 espécies, 316 gêneros e 55 famílias distintas, separadas em 8 *folds*. Grande parte das gravações contém trechos em silêncio ou com pequenos ruídos. Portanto, foram empregados dois métodos de pré-processamento nos arquivos de áudio buscando uma melhoria na qualidade dos mesmos. A primeira técnica é conhecida como Redução de Ruído. Já a segunda técnica, trata-se de um sistema de Segmentação Automática, onde segmentos de interesse são identificados, concatenados e salvos como um único arquivo de áudio. O *software* SoX foi utilizado para a geração dos espectrogramas. Foram usados arquivos com concatenação, permitindo explorar diferentes durações de tempo. Além disso, foi aplicado um zoneamento baseado na escala Mel, onde há várias divisões não lineares na imagem. Assim, o espectrograma é dividido em várias regiões, podendo gerar um novo conjunto de características extraídos de cada zona, e permitindo a construção de um classificador específico. O descritor *Local Binary Pattern* (LBP) foi utilizado para a extração de características de espectrogramas. Além do LBP, foram aplicados descritores de características acústicas, como o RP (*Rhythm Pattern*). O algoritmo de classificação adotado foi o *Support Vector Machine* (SVM) (VAPNIK, 1982), que gera como resultado uma estimativa de probabilidade para cada classe existente no sistema de classificação. Ao extrair características de um espectrograma com o descritor LBP (OJALA et. al, 1996) e feito a classificação, essa predição resultante pode ser combinada. O mesmo acontece com as características acústicas, obtidas diretamente do arquivo de áudio. Assim, para alguns casos foram aplicadas combinações de





classificadores, fazendo com que a fusão de informações complementares existentes possibilite a obtenção de melhores resultados.

Resultados e Discussão

Diferentes abordagens foram utilizadas para a construção dos classificadores. Para os melhores resultados, a frequência de 11 kHz e amplitude de -60 dB foram definidos como parâmetros para originar os espectrogramas, sendo que para a Base I, eles foram zoneados com escala Mel. Devido à baixa qualidade das amostras, filtros de Redução de Ruído seguido de Segmentação Automática foram aplicados. As predições geradas nos melhores testes com uso de LBP e RP foram combinadas, e os resultados obtidos estão representados na Tabela 1.

Tabela 1 - Melhores resultados obtidos

Classes	Descritores	Regra	F-measure	Acurácia
94 espécies	LBP (mel), RP	Soma	46,31%	45,19%
79 gêneros	LBP (mel), RP	Produto	44,71%	40,43%
32 famílias	LBP (mel), RP	Máximo	36,72%	45,59%
547 espécies	LBP, RP	Produto	17,78%	22,80%
316 gêneros	LBP, RP	Produto	20,87%	31,67%
55 famílias	LBP, RP	Produto	24,89%	40,32%

Ao observar os resultados obtidos, nota-se acurácias que poderiam ser consideradas baixas para tarefas de classificação. Entretanto, não é correto fazer uma comparação com outros sistemas para classificação de pássaros desenvolvidos na literatura, já que as bases utilizadas são diferentes. Mesmo assim, vale ressaltar que poucas pesquisas utilizam quantidades tão grande de classes. Muitos ruídos se fizeram presentes em grande parte dos arquivos, sendo visível o ganho de qualidade aplicando filtros de redução de ruídos e segmentação automática. Além disso, zonar os espectrogramas antes de extrair suas características, bem como aproveitar, via regras de fusão, a existência de informações complementares existentes em outras predições, como as geradas com o descritor acústico RP, foram necessárias para conseguir as taxas de acertos mais altas. Por fim, após essa série de experimentos, percebe-se pelos resultados, que a classificação em diferentes taxonomias não resulta necessariamente em um resultado melhor, mesmo reduzindo o número de classes. Ao se montar um sistema com menos classes, a expectativa é que a complexidade para





classificação diminua. Por outro lado, organizar animais em uma mesma categoria taxonômica, não impede que pássaros de mesma classe tenham cantos totalmente diferentes, dificultando a identificação dos padrões. Apesar de tudo, em casos com um número muito grande de espécies, a organização em gêneros ou famílias, apresenta-se como uma melhor alternativa dependendo da finalidade do estudo.

Conclusões

O método com utilização de espectrogramas é algo recente, e ainda há muito a ser estudado. Sendo assim, este trabalho explorou principalmente a utilização de características visuais para o desenvolvimento de um sistema de classificação de pássaros. Um dos atributos visuais presente nos espectrogramas é a textura, a qual é analisada com o uso do descritor LBP. Entretanto, devido a quantidade de ruídos externos e trechos ociosos existentes nas amostras, é necessário realizar determinadas etapas de pré-processamento, como filtragens, para a obtenção de melhores resultados. A aplicação da técnica de redução de ruídos, seguida de uma segmentação automática, apresentaram-se como alternativas de filtragens promissoras, obtendo melhoras significativas nas taxas de acertos. Dentre esses resultados, estão algumas das melhores acurácias obtidas sem o uso de combinação de classificadores. Além dos filtros, a aplicação de zoneamento com escala Mel nas imagens, bem como a combinação de classificadores, favoreceram a obtenção de acurácias maiores. A combinação, inclusive, evidenciou a existência de informações complementares importantes entre as características obtidas nos domínios visual e acústico.

Agradecimentos

Ao CNPq pelo apoio financeiro para a realização deste trabalho.

Referências

- LUCIO, D. R.; COSTA, Y. M. G. Bird Species Classification Using Spectrograms. **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications**. p. 543-550, 2015.
- OJALA, T.; PIETIKÄINEN, M.; HARWOCK, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, 29(1):51-59, 1996.
- VAPNIK, V. N. **Estimation of Dependences Based on Empirical Data**. Springer-Verlag, New York, USA, 1982.

