

**UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS**

**João Paulo Lemos Escola**

**Seleção Paraconsistente de Características Acústicas  
Aplicada ao Processo Entomológico Agrícola**

**São Carlos**

**2021**



**João Paulo Lemos Escola**

**Seleção Paraconsistente de Características Acústicas  
Aplicada ao Processo Entomológico Agrícola**

Tese apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências - Programa de Pós-Graduação em Engenharia Elétrica.

Área de concentração: Sistemas Dinâmicos

Orientador: Prof. Dr. Ivan Nunes da Silva

Coorientador: Prof. Dr. Rodrigo Capobianco Guido

**São Carlos**

**2021**

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,  
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS  
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da  
EESC/USP com os dados inseridos pelo(a) autor(a).

E74s	Escola, João Paulo Lemos Seleção Paraconsistente de Características Acústicas Aplicada ao Processo Entomológico Agrícola / João Paulo Lemos Escola; orientador Ivan Nunes da Silva; coorientador Rodrigo Capobianco Guido. São Carlos, 2021.  Tese (Doutorado) - Programa de Pós-Graduação em Engenharia Elétrica e Área de Concentração em Sistemas Dinâmicos -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2021.  1. Processamento Digital de Áudio. 2. Wavelets. 3. Engenharia Paraconsistente de Características. 4. Inteligência Artificial. 5. Support Vector Machines. 6. Manejo de cigarras. I. Título.
------	--

Folha de aprovação em conformidade  
com o padrão definido  
pela Unidade.

No presente modelo consta como  
folhadeaprovacao.pdf



*Dedico este trabalho a Deus, por me permitir realizá-lo.*

*Ao meu orientador, meu co-orientador, aos meus familiares, ex-professores, alunos e ex-alunos, colegas de profissão e amigos. Sem vocês nada disso seria possível.*

*À Universidade de São Paulo e ao Instituto Federal de Educação,  
Ciência e Tecnologia de São Paulo.*

*A todos que, assim como eu, estudaram em escola pública e tiveram todas as dificuldades possíveis e imagináveis por conta da falta de estrutura, professores desmotivados e mesmo assim apertaram o cinto e foram em frente. Lhes digo que, depois do ensino médio a coisa melhora bastante, então vale a pena perseverar.*

*E a todos que acreditam no poder da educação.*



## AGRADECIMENTOS

Agradeço primeiramente a Deus, pela oportunidade de realizar este trabalho, pela saúde, inspiração e paz. Agradeço ao meu pai, José Lemos Escola (*in memoriam*), por não me deixar sair da escola na segunda série, quando disse que já sabia ler e escrever. À minha mãe, Maria da Conceição Ferreira Escola, que sempre lutou por mim e pelos meus irmãos, passando muitas dificuldades. À minha esposa, Vanessa Cristina Cassolatto, pelo apoio diário em todos os momentos da minha vida. Aos meus irmãos, Emerson Fernando Lemos Escola e Emilene Laura Lemos Escola, pelo apoio sempre que eu precisei e aos meus cunhados Alex e Bruna, sobrinhos Allan, Braylan e Júlia. Ao meu sogro, Valter Cassolatto e minha sogra Leonor Sabino Cassolatto, por terem me ensinado muito sobre a vida e sobre os relacionamentos. Aos meus professores, desde o primário até a Universidade, Deus sabe o que fizeram por mim, direta ou indiretamente. À dona Yolanda (*in memoriam*), nossa vizinha, que levou minha camisetinha para benzer quando eu estava na segunda série e estava quase reprovando. Aos meus professores do Ensino Fundamental: Ivani, Marcos, Soraya, aos meus professores do Ensino Médio e a todos os professores de escolas públicas do Brasil, por persistirem mesmo sem a condição salarial que merecem, dedicando-se por amor. Aos meus ex-colegas de trabalho na Prefeitura Municipal de São José do Rio Preto, e amigos do qual ainda fazemos animados churrascos: Erlete, Fabiane, Janete, Maria do Carmo, Rogério, que me ajudaram a crescer como profissional quando ainda não sabia que iria seguir a carreira acadêmica, especialmente a Fabiane por ter me incentivado a continuar os estudos, fazendo faculdade de computação ao invés de esperar e me aposentar na prefeitura. Aos meus professores e ex-colegas da Faculdade de Tecnologia de Rio Preto: Carlos, Djalma, Henrique, Leandro, Luciene, Marcelo, Sérgio, Valdir, Xyko e tantos outros; instituição da qual tenho muito orgulho de ter feito a graduação no curso de Informática com Ênfase em Gestão de Negócios (hoje o nome mudou um pouco), especialmente aos professores Marcelo e Leandro por me incentivarem na continuidade dos estudos em pós-graduação. Ao Prof. Dr. Lucimar Sasso que me apresentou o Prof. Rodrigo Capobianco Guido, meu querido orientador do mestrado, parceiro e amigo, que sempre esteve ao meu lado, e agora co-orientador do doutorado. Aos colegas do Speech Lab, do Instituto de Física de São Carlos/USP, onde comecei a carreira de pós-graduação: Enio, Fabrício, Fantinato, Leonardo, Luciene, Lucimar, Regiane, Sylvio e da Escola de Engenharia de São Carlos/USP, principalmente meu querido orientador Prof. Dr. Ivan Nunes da Silva, aos docentes, servidores e tantos outros. Aos meus colegas professores, servidores, alunos e ex-alunos dos cursos técnicos e cursos livres de Linux, Java, PHP, informática na Etec Philadelpho Gouvêa Netto. À minha ex-coordenadora, Valéria Donatoni Anguera, por me fazer sair da prefeitura para assumir 28 aulas, o que me fez decidir finalmente mudar de profissão; professores da área de informática: Alexei,

Camila, Fabiana, Gustavo, Humberto, Thaís, Maria Teresa, Michel, Rosana, Vitão e tantos outros, que me ajudaram muito a evoluir como profissional. Aos meus ex-alunos, professores e colegas da Faculdade de Tecnologia de Rio Preto, Sistemas de Informação na Universidade do Estado de Minas Gerais em Frutal e na Faculdade Barretos. Aos colegas docentes e ex-alunos das Engenharias, Análise e Desenvolvimento de Sistemas, Sistemas de Informação no Centro Universitário do Norte Paulista por me ajudarem na minha evolução profissional a cada aula, pergunta, avaliação, atribuição de aula e por confiarem em mim. Ao amigo Prof. Me. Uender Barbosa de Souza, do Instituto Federal de Goiás, doutorando na Universidade Federal de Goiás, matemático, cientista e amigo pesquisador de processamento digital de sinais, que me ajudou imensamente nesse caminho, parceiro de muitos artigos agora e no futuro. À reitoria do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, instituição da qual tenho muito orgulho de fazer parte. Aos meus colegas do campus Barretos, servidores, professores, alunos e ex-alunos, por apoiarem de forma grandiosa este projeto e a minha evolução profissional, em especial aos membros do grupo de pesquisa voltada ao processamento de sinais em entomologia: Alexandre de Moraes Cardoso, pai da ideia de processar os sinais de áudio das cigarras, Douglas Henrique Bottura Macagnan, Lucas Exposto Soares, Artur Kenzo Dezotti, Michel Gaglioni Rocha e aos professores da área de Informação e Comunicação, dos cursos Técnico em Informática e Análise e Desenvolvimento de Sistemas: Alvaro, Antonio, Eduardo, Mauro, Jones, Jovander, José Ricardo, Lucas, Paulo, Xyko. A todos os que passaram pela minha vida ajudando, apoiando, deixando ideias, ensinamentos e inspirações.

*“Mas, buscai primeiro o reino de Deus, e a sua justiça,  
e todas estas coisas vos serão acrescentadas.”*

*Mateus 6:33*



## RESUMO

ESCOLA, J. P. L. **Seleção Paraconsistente de Características Acústicas Aplicada ao Processo Entomológico Agrícola**. 2021. 134p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2021.

As diferentes possibilidades que têm sido utilizadas para a extração de características de sinais acústicos, visando subsequente classificação ou identificação de padrões, estão contempladas por duas modalidades distintas: *handcrafted feature extraction* e *feature learning*. Embora esta tenha ganhado força recentemente, em função do seu automatismo completo e dos seus resultados relevantes, aquela possui um diferencial de interesse: a oportunidade de compreender, do ponto de vista físico, o significado dos descritores utilizados para representar, em dimensões reduzidas, os sinais que lhes deram origem. Entretanto, para examinar tais características, objetivando selecionar as mais adequadas ao problema de interesse, critérios variados têm sido utilizados, sendo que cada um deles é sub-ótimo em determinado aspecto. Assim, visando obter critérios mais generalistas para o procedimento de *feature selection*, independentemente dos classificadores que receberão as características como entrada, a Engenharia Paraconsistente de Características (EPC), baseada na Lógica Paraconsistente, foi descrita na literatura em 2019. Diante das possibilidades que se apresentaram, este trabalho concentra-se, de modo inovador, na aplicação da EPC para a seleção de características provenientes de sinais acústicos codificados no domínio conjunto *wavelet*-Bark e com o objetivo específico de auxiliar no manejo de insetos-praga em lavouras de café, particularmente detectando a presença de cigarras em meio às plantações. Portanto, estão registrados neste documento, além dos estudos que embasam-no teoricamente, o projeto e o desenvolvimento de um sistema computacional para tal fim. De modo geral, os resultados possibilitaram que um problema que atualmente seria solucionado com estratégias do tipo *deep learning* pudesse ser resolvido com base em um classificador modesto do tipo Máquina de Vetor de Suporte (SVM), com acurácias acima de 90%, comprovando a viabilidade da estratégia proposta. Finalmente, um protótipo no âmbito *Internet of Things* está descrito para servir como possibilidade de implantação em campo.

**Palavras-chave:** Processamento Digital de Áudio. *Wavelets*. Engenharia Paraconsistente de Características. Inteligência Artificial. *Support Vector Machines*. Manejo de cigarras.



## ABSTRACT

ESCOLA, J. P. L. **Paraconsistent Selection of Acoustic Features Applied to the Agricultural Entomological Process**. 2021. 134p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2021.

Different possibilities have been used for feature extraction from acoustic signals, aiming at subsequent classification or pattern identification, encompassing two distinct modalities: handcrafted feature extraction and feature learning. Although the latter has received strong attention recently, due to its complete automatism and its relevant results, the former has still attracts attention for one reason: the opportunity to understand, from a physical point of view, the meaning of the descriptors used to represent, in reduced dimensions, the signals which originate them. Nevertheless, to analyse such features, aiming to select the most adequate ones to the problem of interest, a variety of criteria have been used, each one being sub-optimal in a certain aspect. Thus, aiming to get a more generalist criterion for the feature selection procedure, regardless of the coupled classifier, Paraconsistent Characteristics Engineering (EPC), which is based on Paraconsistent Logic, was introduced in 2019. Consequently, this work focuses, in an innovative way, on the application of EPC for feature selection from acoustic signals, encoded in the wavelet-Bark domain, and with the specific objective of helping in the management of pest insects in coffee plantations, particularly detecting the presence of cicadas. Therefore, in addition to the studies that theoretically support it, the project and the development of a computational system for that purpose are registered in this document. Generally, the results made possible a problem that would currently be solved with deep learning-related strategies to be solved based on a modest Support Vector Machine (SVM) classifier, with accuracy above 90%, proving the effectiveness of the proposed approach. Lastly, an Internet of Things prototype is described to serve as a possibility for usage in the field.

**Keywords:** Digital Audio Processing. Wavelets. Paraconsistent Feature Engineering. Artificial Intelligence. Support Vector Machines. Cicada Management.



## LISTA DE FIGURAS

Figura 1 – Plano Paraconsistente (Adaptado de (GUIDO, 2018a)) . . . . .	44
Figura 2 – Atributos de entrada das classes $C_1$ , $C_2$ , $C_3$ e $C_4$ em vermelho, azul, laranja e verde, respectivamente. Claramente são linearmente separáveis.	46
Figura 3 – $P$ , representado como um ponto vermelho no plano paraconsistente. . .	48
Figura 4 – DWT: sinal $s[\cdot]$ de $n$ amostras discretas e máxima frequência $\pi$ , decomposto até o terceiro nível. . . . .	49
Figura 5 – Relação entre os filtros de análise e síntese (adaptado de (JÚNIOR, 2007))	52
Figura 6 – Formato das respostas ao impulso dos filtros <i>wavelet</i> de <i>Haar</i> , <i>Daubechies</i> , <i>Vaidyanathan</i> , <i>Beylkin</i> , <i>Coiflet</i> e <i>Symmlet</i> , respectivamente (extraído de (JÚNIOR, 2007)) . . . . .	53
Figura 7 – Formatos das funções <i>scaling</i> dos filtros <i>wavelet</i> de <i>Haar</i> , <i>Daubechies</i> , <i>Vaidyanathan</i> , <i>Beylkin</i> , <i>Coiflet</i> e <i>Symmlet</i> , respectivamente (extraído de (JÚNIOR, 2007)) . . . . .	54
Figura 8 – Formatos das funções <i>wavelet</i> dos filtros <i>wavelet</i> de <i>Haar</i> , <i>Daubechies</i> , <i>Vaidyanathan</i> , <i>Beylkin</i> , <i>Coiflet</i> e <i>Symmlet</i> , respectivamente (extraído de (JÚNIOR, 2007)) . . . . .	54
Figura 9 – Espectro de frequências em escala logarítmica de um sinal de <i>Quesada gigas</i> . . . . .	58
Figura 10 – Estrutura da SVM. Os coeficientes do hiperplano, determinados durante a fase de treinamento supervisionado são $\{w_0, w_1, \dots, w_{X-1}\}$ . O elemento de saída é resultado da combinação linear das saídas da camada oculta com os referidos coeficientes. . . . .	60
Figura 11 – Estrutura do sistema proposto, no qual cada sinal acústico $s_i[\cdot]$ é convertido para o seu vetor de características $f_i[\cdot]$ . . . . .	62
Figura 12 – Estrutura básica proposta para o módulo MPP. As siglas estão explicadas ao longo do texto. . . . .	63
Figura 13 – Estrutura básica proposta para o módulo MEP. As siglas estão explicadas ao longo do texto. . . . .	65
Figura 14 – Estrutura básica proposta para o módulo MTT. As siglas estão explicadas ao longo do texto. . . . .	66
Figura 15 – Estrutura básica proposta para o módulo MAP. As siglas estão explicadas ao longo do texto. . . . .	67
Figura 16 – Exemplos de sinais capturados na lavoura de café: ruído, cigarra, DWTP ruído, DWTP cigarra, respectivamente. . . . .	73
Figura 17 – Exemplos de energias dos sinais de cigarra (esquerda) e ruído (direita).	74

Figura 18 – EPC em GR1 com 25 <i>features</i> , com sinais de 10 segundos (441.000 amostras). Acima os planos paraconsistentes, com os melhores resultados (esquerda) e piores resultados (direita), em ordem crescente de distância em relação a VO. Abaixo, as distâncias em relação a VO estão plotadas em diferentes cores de acordo com as famílias <i>wavelet</i> . . . . .	75
Figura 19 – EPC em GR1 com 25 <i>features</i> em sinais de 5 segundos. . . . .	76
Figura 20 – EPC em GR1 com 25 <i>features</i> , em sinais de 2 segundos. . . . .	76
Figura 21 – EPC em GR1 com 25 <i>features</i> , em sinais de 1 segundo. . . . .	77
Figura 22 – PPs em GR1 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) <i>features</i> . . . . .	77
Figura 23 – EPC em GR1: distâncias de VO. . . . .	78
Figura 24 – Melhores <i>features</i> eleitas por EPC em GR1, com segmentos de 1s (esfera), 2s (quadrado), 5s (triângulo) e 10s (círculo). As cores representam o número de <i>features</i> utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). . . . .	80
Figura 25 – PPs em GR1 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores <i>features</i> . . . . .	81
Figura 26 – EPC em GR1: distâncias de VO com as melhores <i>features</i> . . . . .	81
Figura 27 – EPC em GR2 com 25 <i>features</i> , em sinais de 10 segundos. . . . .	82
Figura 28 – EPC em GR2 com 25 <i>features</i> , em sinais de 5 segundos. . . . .	82
Figura 29 – EPC em GR2 com 25 <i>features</i> , em sinais de 2 segundos. . . . .	83
Figura 30 – EPC em GR2 com 25 <i>features</i> , em sinais de 1 segundo. . . . .	83
Figura 31 – PPs em GR2 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) <i>features</i> . . . . .	84
Figura 32 – EPC em GR2: distâncias de VO. . . . .	84
Figura 33 – Melhores <i>features</i> eleitas por EPC em GR2, com segmentos de 1s (esfera), 2s (quadrado), 5s (triângulo) e 10s (círculo). As cores representam o número de <i>features</i> utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). . . . .	86
Figura 34 – PPs em GR2 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores <i>features</i> . . . . .	87
Figura 35 – EPC em GR2: distâncias de VO com as melhores <i>features</i> . . . . .	87
Figura 36 – EPC em GR3 com 25 <i>features</i> , com sinais de 1 segundo. . . . .	88
Figura 37 – PP em GR3 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) <i>features</i> . . . . .	89

Figura 38 – EPC em GR3: distâncias de VO. . . . .	89
Figura 39 – Melhores <i>features</i> eleitas por EPC em GR3. As cores representam o número de <i>features</i> utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). . . . .	90
Figura 40 – PPs em GR3 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores <i>features</i> . . . . .	91
Figura 41 – EPC em GR3: distâncias de VO com as melhores <i>features</i> . . . . .	91
Figura 42 – EPC em GR4 com 25 <i>features</i> , em sinais de 1 segundo. . . . .	92
Figura 43 – PP em GR4 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) <i>features</i> . . . . .	92
Figura 44 – EPC em GR4: distâncias de VO. . . . .	93
Figura 45 – Melhores <i>features</i> eleitas por EPC em GR4. As cores representam o número de <i>features</i> utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). . . . .	94
Figura 46 – PPs em GR4 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores <i>features</i> . . . . .	95
Figura 47 – EPC em GR4: distâncias de VO com as melhores <i>features</i> . . . . .	95
Figura 48 – Acurácia resultante dos testes utilizando 25 <i>features</i> . . . . .	97
Figura 49 – Matrizes de confusão resultantes dos testes utilizando 25 <i>features</i> . . . . .	97
Figura 50 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 25 <i>features</i> . . . . .	98
Figura 51 – Acurácia resultante dos testes utilizando 20 <i>features</i> . . . . .	98
Figura 52 – Matriz de confusão resultante dos testes utilizando 20 <i>features</i> . . . . .	99
Figura 53 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 20 <i>features</i> . . . . .	99
Figura 54 – Acurácia resultante dos testes utilizando 15 <i>features</i> . . . . .	100
Figura 55 – Matriz confusão resultante dos testes utilizando 15 <i>features</i> . . . . .	100
Figura 56 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 15 <i>features</i> . . . . .	100
Figura 57 – Acurácia resultante dos testes utilizando 10 <i>features</i> . . . . .	101
Figura 58 – Matriz confusão resultante dos testes utilizando 10 <i>features</i> . . . . .	101
Figura 59 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 10 <i>features</i> . . . . .	102
Figura 60 – Acurácia resultante dos testes utilizando 5 <i>features</i> . . . . .	102
Figura 61 – Matriz confusão resultante dos testes utilizando 5 <i>features</i> . . . . .	102
Figura 62 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 5 <i>features</i> . . . . .	103
Figura 63 – Acurácia resultante dos testes utilizando 2 <i>features</i> . . . . .	103
Figura 64 – MC resultante dos testes utilizando 2 <i>features</i> . . . . .	104

Figura 65 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 2 <i>features</i> . . . . .	104
Figura 66 – Acurácia resultante dos testes utilizando 25 <i>features</i> . . . . .	105
Figura 67 – Matrizes confusão resultantes dos testes utilizando 25 <i>features</i> . . . . .	105
Figura 68 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 25 <i>features</i> . . . . .	106
Figura 69 – Acurácia resultante dos testes utilizando 20 <i>features</i> . . . . .	106
Figura 70 – MC resultante dos testes utilizando 20 <i>features</i> . . . . .	107
Figura 71 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 20 <i>features</i> . . . . .	107
Figura 72 – Acurácia resultante dos testes utilizando 15 <i>features</i> . . . . .	107
Figura 73 – Matriz confusão resultante dos testes utilizando 15 <i>features</i> . . . . .	108
Figura 74 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 15 <i>features</i> . . . . .	108
Figura 75 – Acurácia resultante dos testes utilizando 10 <i>features</i> . . . . .	109
Figura 76 – Matriz confusão resultante dos testes utilizando 10 <i>features</i> . . . . .	109
Figura 77 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 10 <i>features</i> . . . . .	109
Figura 78 – Acurácia resultante dos testes utilizando 5 <i>features</i> . . . . .	110
Figura 79 – Matriz confusão resultante dos testes utilizando 5 <i>features</i> . . . . .	110
Figura 80 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 5 <i>features</i> . . . . .	111
Figura 81 – Acurácia resultante dos testes utilizando 2 <i>features</i> . . . . .	111
Figura 82 – Matriz de confusão resultante dos testes utilizando 2 <i>features</i> . . . . .	112
Figura 83 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 2 <i>features</i> . . . . .	112
Figura 84 – Acurácia resultante em GR3 . . . . .	113
Figura 85 – Matrizes confusão resultantes em GR3. . . . .	114
Figura 86 – Curvas ROC resultantes em GR3: melhor resultado (azul) e pior resultado (vermelho). . . . .	115
Figura 87 – Acurácia resultante em GR4. . . . .	116
Figura 88 – Matrizes confusão resultantes em GR4. . . . .	116
Figura 89 – Curvas ROC resultantes em GR4: melhor resultado (azul) e pior resultado (vermelho). . . . .	117
Figura 90 – Dispositivo INS acondicionado em uma caixa de plástico ABS. . . . .	118
Figura 91 – Dispositivo “Guido”: disposição dos componentes internos (esquerda) e implantação na lavoura (direita). . . . .	119
Figura 92 – Esquema de integração dos módulos físicos do INS. . . . .	119

Figura 93 – Área de plantio de café no estado de São Paulo em 2021 (Fonte: (IEA..., 2021)). . . . .	120
Figura 94 – Protótipos de suportes para posicionamento experimental de dispositivos INS. . . . .	121



## LISTA DE TABELAS

Tabela 1 – Trabalhos Correlatos: técnicas manuais (TM), uso de <i>softwares</i> no campo (SC) e uso de <i>softwares</i> em laboratório (SL). . . . .	41
Tabela 2 – Características das famílias de <i>wavelets</i> utilizadas no presente trabalho, incluindo a quantidade de momentos da função <i>wavelet</i> . . . . .	54
Tabela 3 – As 25 bandas críticas do sistema auditivo humano que são aproveitadas no presente trabalho para obtenção das energias predominantes do sinal.	55
Tabela 4 – Parâmetros utilizados na bateria de testes em MTT . . . . .	67
Tabela 5 – Grupos de testes. . . . .	72
Tabela 6 – Exemplos de coeficientes de energia resultantes do módulo MPP. . . . .	73
Tabela 7 – Filtros <i>wavelet</i> empregados nos testes do módulo MEP. . . . .	74
Tabela 8 – Possibilidades na busca das N melhores <i>features</i> . . . . .	79
Tabela 9 – Melhores resultados de testes em GR1 com 25 <i>features</i> . . . . .	98
Tabela 10 – Melhores resultados de testes em GR1 com 20 <i>features</i> . . . . .	99
Tabela 11 – Melhores resultados de testes em GR1 com 15 <i>features</i> . . . . .	100
Tabela 12 – Melhores resultados de testes em GR1 com 10 <i>features</i> . . . . .	101
Tabela 13 – Melhores resultados de testes em GR1 com 5 <i>features</i> . . . . .	102
Tabela 14 – Melhores resultados de testes em GR1 com 2 <i>features</i> . . . . .	103
Tabela 15 – Melhores resultados de testes em GR2 com 25 <i>features</i> . . . . .	105
Tabela 16 – Melhores resultados de testes em GR2 com 20 <i>features</i> . . . . .	106
Tabela 17 – Melhores resultados de testes em GR2 com 15 <i>features</i> . . . . .	108
Tabela 18 – Melhores resultados de testes em GR2 com 10 <i>features</i> . . . . .	109
Tabela 19 – Melhores resultados de testes em GR2 com 5 <i>features</i> . . . . .	110
Tabela 20 – Melhores resultados de testes em GR2 com 2 <i>features</i> . . . . .	111
Tabela 21 – Melhores resultados de testes em GR3. . . . .	114
Tabela 22 – Melhores resultados de testes em GR4 . . . . .	115



## LISTA DE ABREVIATURAS E SIGLAS

AUC	<i>Area Under the Curve</i>
APE	Armazena Parâmetros de Energia
ASCII	<i>American Standard Code for Information Interchange</i>
CCT	Consulta Classes de Treinamento
CDI	Cálculo de Dissimilaridade Interclasses
CNC	Consulta Nome da Classe
CLT	Consulta Limiar de Treinamento
CSA	Cálculo de Similaridade Intraclases
CSI	Cálculo de Sobreposições Interclasses
DFT	<i>Discrete Fourier Transform</i>
DNN	<i>Deep Neural Network</i>
DTW	<i>Dynamic Time Warping</i>
DWT	<i>Discrete Wavelet Transform</i>
DWTP	<i>Discrete Wavelet Transform Packet</i>
EB	Escala Bark
EMR	Escolha da Melhor Resposta
EPC	Engenharia Paraconsistente de Características
EPE	Extração dos Parâmetros de Energia
FFT	<i>Fast Fourier Transform</i>
FIR	<i>Finite Impulse Response</i>
GCC	Graus de Certeza e Contradição
GMM	<i>Gaussian Mixture Model</i>
GR1	Grupo de Testes 1
GR2	Grupo de Testes 2

GR3	Grupo de Testes 3
GR4	Grupo de Testes 4
IoT	<i>Internet of Things</i>
KBC	<i>Knowledge-Based Classifier</i>
LDB	Leitura dos Dados Brutos
LPE	Leitura dos Parâmetros de Entrada
LP	Lógica Paraconsistente
MC	Matriz de Confusão
MEP	Módulo de Engenharia Paraconsistente de Características
MLP	<i>Multi-Label Perceptron</i>
MPP	Módulo de Pré-Processamento
MRA	<i>Multi-Resolution Analysis</i>
PP	Plano Paraconsistente
PMC	<i>Pattern-Matching Classifier</i>
PNN	<i>Probabilistic Neural Network</i>
PRFB	<i>Perfect Reconstruction Filter Bank</i>
ROC	<i>Receiver Operating Characteristic</i>
SVM	<i>Support Vector Machine</i>
STFT	<i>Short Time Fourier Transform</i>
UAD	Último Arquivo de Dados
VO	Vértice Ótimo

## LISTA DE SÍMBOLOS

$\alpha$	Letra grega alfa
$\beta$	Letra grega beta
$\gamma$	Letra grega gama
$\pi$	Letra grega pi
$\phi$	Letra grega fi
$\psi$	Letra grega psi
$\omega$	Letra grega omega



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>29</b>
1.0.1	Seleção de Características em Sistemas de Reconhecimento de Padrões	29
1.0.2	Cigarras: pragas prejudiciais aos cafezais	29
1.0.3	Objetivos	31
1.0.4	Lastro da Pesquisa	32
1.0.4.1	Artigos publicados e periódicos	32
1.0.4.2	Capítulos de livros	33
1.0.4.3	Trabalhos publicados em anais de congressos	33
1.0.5	Organização do Trabalho	33
<b>2</b>	<b>REVISÃO DOS CONCEITOS</b>	<b>35</b>
<b>2.1</b>	<b>Trabalhos Correlatos no Âmbito da Detecção de Pragas em Lavouras</b>	<b>35</b>
2.1.1	Aspectos gerais da bioecologia das cigarras	35
2.1.2	Monitoramento	36
2.1.2.1	Técnicas manuais	36
2.1.2.2	Técnicas com software	37
2.1.3	Reflexões sobre o estado da arte	40
<b>2.2</b>	<b>Engenharia Paraconsistente de Características (EPC): selecionando</b>	
	<b><i>features</i></b>	<b>43</b>
2.2.1	Etapas para o uso da EPC	44
2.2.2	Exemplo numérico do uso da EPC	45
<b>2.3</b>	<b>Wavelets e Escala Bark</b>	<b>47</b>
2.3.1	Cálculo da DWT	52
2.3.2	Famílias de Transformadas <i>Wavelet</i>	53
2.3.3	Escala Bark (EB)	55
<b>3</b>	<b>A ABORDAGEM PROPOSTA</b>	<b>57</b>
<b>3.1</b>	<b>A Estrutura do Sistema</b>	<b>57</b>
<b>3.2</b>	<b>Implementação do Sistema</b>	<b>62</b>
3.2.1	Módulo de Pré-Processamento (MPP)	62
3.2.2	Módulo de Engenharia Paraconsistente de Características (MEP)	64
3.2.3	Módulo de Treinamento e Teste (MTT)	65
3.2.4	Módulo de Aplicação (MAP)	67
<b>3.3</b>	<b>Detalhes Adicionais</b>	<b>68</b>
<b>4</b>	<b>TESTES E RESULTADOS</b>	<b>69</b>

<b>4.1</b>	<b>Procedimentos Gerais</b>	<b>69</b>
<b>4.2</b>	<b>Materiais e Métodos</b>	<b>69</b>
4.2.1	Módulo de Pré-Processamento - MPP	70
4.2.2	Módulo de Engenharia Paraconsistente de Características - MEP	71
4.2.3	Módulo de Treinamento e Teste - MTT	71
<b>4.3</b>	<b>Bateria de Testes</b>	<b>72</b>
4.3.1	Testes do módulo MPP	72
4.3.2	Testes do módulo MEP	74
4.3.2.1	Grupo GR1	75
4.3.2.2	Grupo GR2	81
4.3.2.3	Grupo GR3	87
4.3.2.4	Grupo GR4	91
4.3.3	Testes do módulo MTT	96
4.3.3.1	Grupo GR1	96
4.3.3.2	Grupo GR2	104
4.3.3.3	Grupo GR3	112
4.3.3.4	Grupo GR4	115
<b>4.4</b>	<b>Protótipo em IoT</b>	<b>117</b>
<b>4.5</b>	<b>Comentários Gerais Sobre os Resultados</b>	<b>121</b>
<b>5</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS</b>	<b>123</b>
	<b>REFERÊNCIAS</b>	<b>125</b>

## 1 INTRODUÇÃO

*A introdução deste trabalho concentra-se na observação de dois problemas distintos, isto é, seleção generalizada de características em problemas de reconhecimento de padrões e combate às cigarras que atuam como pragas daninhas aos cafezais, os quais fundem-se para caracterizar o objetivo da pesquisa.*

### 1.0.1 Seleção de Características em Sistemas de Reconhecimento de Padrões

As variadas técnicas que têm sido utilizadas para a extração de características de sinais acústicos, objetivando subsequente classificação ou identificação de padrões, estão contempladas por duas modalidades distintas: *handcrafted feature extraction* e *feature learning* (GUIDO, 2018a). Embora esta tenha ganhado força recentemente, em função do seu automatismo completo e dos seus resultados relevantes, aquela possui um diferencial de interesse: a oportunidade de compreender, do ponto de vista físico, o significado dos descritores utilizados para representar, em dimensões reduzidas, os sinais que lhes deram origem. Entretanto, para examinar tais características, objetivando selecionar as mais adequadas ao problema de interesse, critérios variados têm sido utilizados, sendo que cada um deles é sub-ótimo em determinado aspecto. Assim, visando obter critérios mais generalistas para o procedimento de *feature selection*, independentemente dos classificadores que receberão as características como entrada, a Engenharia Paraconsistente de Características (EPC) (GUIDO, 2018a), baseada na Lógica Paraconsistente (CARNIELLI; CONIGLIO, 2016), foi descrita na literatura em 2019. Desde então, inúmeras possibilidades tem surgido para aplicar tal técnica e, assim sendo, explora-se, nesta tese, o *gap* que consiste em utilizá-la no âmbito da seleção de características acústicas extraídas por processos *handcrafted-based*.

### 1.0.2 Cigarras: pragas prejudiciais aos cafezais

Com o aumento da população humana, também é crescente a necessidade de produção de alimentos e de outros produtos advindos de campos agrícolas. Visando suprir tal necessidade, é comum encontrar monocultivos em áreas extensas, fato que pode contribuir com a diminuição da biodiversidade e favorecer a ocorrência de espécies de insetos-praga (ZHAO *et al.*, 2018; SUNDERLAND *et al.*, 2019).

Geralmente o manejo é feito com o uso de defensivos químicos, objeto de estudo da Entomologia, aplicados muitas vezes até em áreas onde a ocorrência das espécies de insetos-praga ainda sequer atingiu a densidade populacional capaz de promover algum prejuízo, o que onera o custo de produção e pode causar impactos ao ambiente e à saúde humana. O desenvolvimento de *hardwares* e *softwares* específicos para a detecção da praga no campo pode fornecer subsídios para o desenvolvimento de uma forma de produção de

menor impacto. Nesse sentido, a captação, registro e análise de sinais acústicos emitidos por insetos pode ser uma alternativa para otimizar a produção de certos cultivos (MANKIN *et al.*, 2011; MONTGOMERY *et al.*, 2021).

Um bom exemplo de inseto capaz de emitir sinais acústicos é a cigarra (Hemiptera: Cicadidae). Em condições brasileiras, a cultura do cafeeiro pode ser atacada por diversos artrópodes e, entre eles, *Quesada gigas*, a de maior tamanho no país, podendo ultrapassar os 70 mm de comprimento (incluindo as asas) e 20 mm de largura no caso dos machos, é considerada praga-chave em todo o estado de Minas Gerais e na região nordeste do estado de São Paulo (MARTINELLI; ZUCCHI, 1997). O fato de que há relatos de ocorrência de cigarras em cafezais paulistas desde o período entre 1900 e 1904 (D'UTRA, 1908) interfere na forma de condução da lavoura, o que praticamente obriga os cafeicultores a adotarem práticas de aprimoramento do sistema produtivo, como o aumento do espaçamento entre plantas para permitir a mecanização da lavoura, visando o uso de defensivos para manejo das cigarras e de outras pragas e doenças que afetam a produtividade da cultura (ANDRADE, 2018).

A emissão de sons da cigarra ocorre pelo órgão cimbálico dos machos. Esses sons são produzidos pela contração de uma membrana chamada tímalo, que por sua sucessão constitui o canto da cigarra e que pode ser empregado para diferenciação das espécies, já que cada espécie produz um som diferente (MACCAGNAN, 2008). O acasalamento é o principal fator de emissão de sons pelos machos nas diversas espécies de cigarras, dessa forma também há diferenciação do sinal acústico emitido de acordo com o tamanho do inseto e o contexto no qual ele se encontra, como, por exemplo, disputas de território ou captura por predador. Tais sinais sonoros também são frequentemente empregados em pesquisas para análise taxonômica (MACCAGNAN, 2008; ANDRADE, 2018). Nas regiões onde atacam a cultura de café, os machos costumam emitir som de outubro a dezembro. Na década de 1970, como não havia método eficiente de controle das cigarras, muitos cafeicultores não tinham outra opção senão erradicar as lavouras infestadas, o que culminava no abandono da cultura por diversos produtores (SOUZA; REIS; SILVA, 2007).

As cigarras sugam a seiva no xilema da raiz das plantas em estágio de ninfa (MACCAGNAN; MATUO; MARTINELLI, 2017) e, em estágio adulto, sugam a parte aérea da planta. São consideradas, no Brasil, uma praga-chave do café. Nas últimas décadas, o controle recomendado é por meio de defensivos químicos sistêmicos (MARTINELLI *et al.*, 1998). Mais recentemente, mostrando-se importante alternativa tecnológica, o uso de uma armadilha sonora que atrai a *Quesada gigas* para um sistema fechado de pulverização tem sido uma opção (MACCAGNAN, 2008), entretanto, sua distribuição pode ser realizada de forma mais inteligente e econômica, se for conhecida a incidência das pragas em cada região do terreno. Esse monitoramento, quando é realizado, normalmente é feito manualmente por trabalhadores da lavoura (PEREIRA, 2017).

O processamento digital de sinais (PDS) permite vasta gama de aplicações digitais, incluindo o processamento de áudio (HAYKIN; VEEN, 2001; GUPTA *et al.*, 2021; SOUZA; ESCOLA; BRITO, 2022). Dentro dessa área, a Transformada Discreta *Wavelet* (*Discrete Wavelet Transform* - DWT) se mostra um dos meios mais eficientes de análise tempo-frequência em um sinal de áudio (GUIDO, 2017; GUIDO, 2021). Técnicas como a bioacústica (DEPRAETERE *et al.*, 2012; ULLOA *et al.*, 2018; VALLEE, 2018) e a ecoacústica (SUEUR; FARINA, 2015; FARINA; GAGE, 2017) integram ciências como a biologia, ecologia e a acústica para estudar a relação entre os seres vivos na produção e recepção de sinais de áudio. A implementação de soluções tecnológicas para aprimorar, otimizar e solucionar problemas da agricultura está em constante ascensão, o conceito de *Smart Farm* (KODALI; YERROJU; SAHU, 2018; ALVES *et al.*, 2019; ALONSO *et al.*, 2020; GUPTA *et al.*, 2020) remete ao uso de tecnologia na lavoura, assim como a Agricultura de precisão (ARTUZO; FOGUESATTO; SILVA, 2017; MOLIN, 2017), Agricultura 4.0 (RIBEIRO; MARINHO; ESPINOSA, 2018; SANTOS; ESPERIDIÃO; AMARANTE, 2019) ou Agricultura Digital (BASSOI *et al.*, 2019; CARDOSO, 2020), termos mais adotados no Brasil.

Assim sendo, acredita-se que as soluções tecnológicas como sistemas inteligentes, Internet das Coisas (Internet of Things - IoT) (OLIVEIRA, 2017; JIANG *et al.*, 2020), redes neurais (SILVA; SPATTI; FLAUZINO, 2010; ANITescu *et al.*, 2019), inteligência artificial (NILSSON, 2014; ACEMOGLU; RESTREPO, 2019; HAENLEIN; KAPLAN, 2019), deep learning (GOODFELLOW; BENGIO; COURVILLE, 2016; PRINCE *et al.*, 2019; MASTELINI *et al.*, 2019) e aprendizagem de máquinas (LIAKOS *et al.*, 2018; MOLNAR, 2020) podem auxiliar no monitoramento da lavoura, detectando a presença de pragas em cada região, bem como fazendo a classificação da espécie e catalogando sua incidência por meio da análise da densidade de indivíduos presentes em cada região. Esse processo pode ser realizado pela equipe de trabalhadores da lavoura (SOUZA; REIS; SILVA, 2007), mostrando-se, entretanto, penoso.

### 1.0.3 Objetivos

Diante do exposto em ambas as sub-seções anteriores, o objetivo fundamental deste trabalho é o de, de modo inovador, aplicar a EPC para selecionar características provenientes de sinais acústicos codificados no domínio conjunto *wavelet*-Bark, visando auxiliar no manejo de insetos-praga em lavouras de café, particularmente detectando a presença de cigarras em meio às plantações. De modo mais específico, os objetivos implicam:

- a captura de sinais acústicos das cigarras em uma lavoura de café, utilizando um gravador digital, criando uma base de dados original e que pode ser reaproveitada para futuras pesquisas;

- o desenvolvimento de um algoritmo para codificação dos sinais acústicos, em formato WAV, com base em *wavelets* e na escala acústica Bark (EB);
- o desenvolvimento de uma estratégia original para selecionar as melhores *features* com base na EPC, a qual constitui uma técnica recentemente apresentada na literatura científica e que não havia ainda sido utilizada com o mesmo propósito;
- a viabilização do sistema de detecção dos padrões proposto, utilizando um classificador modesto, utilizando uma Máquina de Vetor de Suporte (SVM) (CORTES; VAPNIK, 1995), ao invés de um complexo tal como os baseados em *deep learning*, em função dos promissores resultados proporcionados pela EPC;
- o detalhamento dos testes realizados para a constatação da viabilidade da estratégia proposta, assim como a disseminação dos mesmos na literatura;
- o desenvolvimento de um protótipo em *Internet of Things* (IoT), utilizando um micro-controlador e módulos adicionais de captura de áudio para implantação em lavoura de café.

Destaca-se, desse modo, que a contribuição deste trabalho é **original**, pois não há registro na literatura científica do uso da EPC para a análise de sinais acústicos representados no domínio conjunto *wavelet*-Bark para a finalidade em questão. Assim, acredita-se que este trabalho, o qual encontra-se publicado em periódico especializado e com alto fator de impacto, conforme indicado adiante, tenha efetivamente contribuído com as pesquisas da área.

#### 1.0.4 Lastro da Pesquisa

Durante este trabalho de doutorado, foram publicados artigos científicos diversos, possibilitando avaliações por outros fóruns:

##### 1.0.4.1 Artigos publicados e periódicos

- ESCOLA, J. P. L.; GUIDO, R.C.; SILVA, I.N.; SOARES, L.E. Análise de Área de Cobertura de Dispositivo IoT para Monitoramento em Smart Farm. *Revista Ibérica de Sistemas e Tecnologias de Informação* (RISTI), n. 42, p. 1-12, 2021.
- ESCOLA, J.P.L.; GUIDO, R.C.; SILVA, I.N. ; MACCAGNAN, D.H.B.; CARDOSO, A.M. ; SOUZA, U.B. Estado da Arte no Monitoramento Acústico de Cicadidae em Lavouras de Café. *Revista Macambira*, v. 5, n. 1, 2021.
- ESCOLA, J.P.L.; GUIDO, R.C.; SILVA, I.N. ; CARDOSO, A.M. ; MACCAGNAN, D.H.B.; DEZOTTI, A.K. Automated Acoustic Detection of a Cicadid Pest in Coffee Plantations. *Computers and electronics in Agriculture*, v. 169, p. 105215, 2020.

DEZOTTI, A. K. ; ESCOLA, J. P. L. ; CARDOSO, A. M. ; SOARES, L. E. ; MACCAGNAN, D. H. B. Dispositivo de Monitoramento de Densidade Populacional de Insetos a partir de Sinais Acústicos Emitidos. *Brazilian Journal of Animal and Environmental Research*, v. 2, p. 1781-1785, 2019.

#### 1.0.4.2 Capítulos de livros

- ESCOLA, J.P.L.; GUIDO, R. C.; CARDOSO, A. M.; MACCAGNAN, D. H. B.; RIBEIRO, J. M.; CARDOSO, J. R. F.; A Case Study of Wavelets and SVM Application in Coffee Agriculture: Detecting Cicadas Based on their Acoustic and Image Patterns. In: SILVA, I.N.; (Org.). *Application of Expert Systems - Theoretical and Practical Aspects*. 1ed. London: IntechOpen, 2019.
- DEZOTTI, A. K.; ESCOLA, J. P. L.; SOARES, L. E.; MACCAGNAN, D. H. B.; CARDOSO, A. M. Software capaz de reconhecer cigarras no cafeeiro. In: Eduardo Soares Calixto e Helena Maura Torezan Silingardi. (Org.). *Temas Atuais em Ecologia Comportamental e Interações*. 1ed. Uberlândia, MG: Editora Composer, 2017, v. 1, p. 657-674.

#### 1.0.4.3 Trabalhos publicados em anais de congressos

- ESCOLA, J.P.L.; GUIDO, R.C.; SILVA, I.N.; FONSECA, E.S. Wavelet Transform Applied to Coffee Entomology. In: *SPSympo 2021 - Signal Processing Symposium*, 2021, Lodz, Poland.
- ESCOLA, J.P.L.; GUIDO, R.C.; SILVA, I.N.; CARDOSO, A.M.; MACCAGNAN, D.H.B. ; DEZOTTI, A.K. ; SAMPAIO, I. M.; SOARES, L. E. Classificação Digital de Cicadidae com Wavelets e Support Vector Machines. In: *XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2017, São Pedro-SP.
- DEZOTTI, A. K. ; ESCOLA, J. P. L. ; SOARES, L. E. ; MACCAGNAN, D. H. B. ; CARDOSO, A. M. Software capaz de reconhecer cigarras no cafeeiro. In: *II BecInt - Behavioral Ecology and Interactions Symposium*, 2017.

Além das publicações listadas, outros artigos científicos foram submetidos e encontram-se, no momento da confecção deste documento, em processo de revisão por pares.

#### 1.0.5 Organização do Trabalho

Visando prover explicações detalhadas, o texto a seguir está organizado como segue. O Capítulo 2 contém a descrição de alguns trabalhos que envolvem manejo de cigarras e tecnologias empregadas no processo, assim como uma breve revisão da literatura contemplando a EPC e os detalhes usados para extração de características no domínio

conjunto *wavelet*-Bark. O Capítulo 3 foi dedicado à técnica proposta e aos algoritmos desenvolvidos para a sua aplicação. No Capítulo 4, encontram-se os testes e resultados, além de materiais, métodos e discussões. No último Capítulo, estão as conclusões e menções aos possíveis trabalhos futuros.

## 2 REVISÃO DOS CONCEITOS

*Em função da importância de uma revisão da literatura sobre as metodologias empregadas no manejo de insetos-praga na lavoura cafeeira e sobre as tecnologias que utilizam sinais acústicos para lidar com os diversos processos envolvendo insetos em lavouras, este Capítulo contém a descrição de alguns trabalhos correlatos e dos conceitos utilizados nesta tese.*

### 2.1 Trabalhos Correlatos no Âmbito da Detecção de Pragas em Lavouras

#### 2.1.1 Aspectos gerais da bioecologia das cigarras

A característica mais marcante das cigarras é a emissão de sinais acústicos. Apenas o macho tem órgão específico para a geração desses sinais. Pringle (1954) é o primeiro a fornecer detalhadamente uma análise anatômica e fisiológica do processo de geração de sinais acústicos por cigarras. A onda sonora é gerada pela deformação do tímalo, uma placa com propriedade elástica localizada dorso-lateralmente em cada uma das laterais do abdome. No interior do abdome, há espaços preenchidos por ar que atuam sobre a onda sonora como uma caixa de ressonância. A morfologia do tímalo, a frequência que este é deformado, aspectos morfológicos e movimentos do abdome no momento da emissão do sinal propiciarão as características do sinal emitido, que é único para cada espécie.

Segundo Sueur (2001), há exemplos de espécies de cigarras que emitem sinais com frequência pouco abaixo de 1kHz até espécies com frequências ultrassônicas próximas à 25kHz. A alta intensidade em que os sinais são emitidos também é uma característica marcante das cigarras, sendo que em algumas espécies ela pode ser superior à 100 dB (VILLET, 1987). O fato de o sinal emitido pela cigarra ser espécie-específico faz dele um importante carácter taxonômico para fins de identificação, levantamento de riqueza de espécie e estimativa de tamanho populacional (SUEUR, 2001).

Embora o sinal acústico possa ter outras funções, como de espantar predadores, sua principal função é de atrair a fêmea para o acasalamento (CLARIDGE, 1985). Após a cópula, a fêmea deposita os seus ovos no interior de ramos da parte aérea da planta (JUNIOR *et al.*, 2012). A ninfa recém eclodida vai ao solo, onde se enterra em busca da raiz da planta a qual sugará seiva. É no subsolo, agregada à raiz da planta, que a ninfa permanecerá por vários anos até pouco tempo antes de completar seu desenvolvimento, momento em que ela volta à superfície e fixa-se a um substrato vertical, onde passará pelo processo de metamorfose e se tornará adulta, deixando fixada ao substrato sua última exúvia (TRIPLEHORN; JOHNSON, 2011).

### 2.1.2 Monitoramento

A ação de monitorar, vigiar ou verificar algo, visando determinado fim, permite acompanhar um processo com o intuito de exercer controle, a partir de um dado momento até a sua finalização ou até um ponto de tempo possível de ser atingido (MICHAELIS, 2015).

Foram analisados diversos artigos que abordam assuntos relacionados ao monitoramento e detecção das cigarras em seus mais diversos ambientes. A partir destes, buscou-se, principalmente, relacionar os processos com o uso de tecnologias, como *softwares* e dispositivos tecnológicos, a fim de descobrir se os agricultores empregam tais artifícios em suas rotinas de cultivo de espécies vegetais que têm incidência de cigarras como praga e também de fazer um levantamento do estado da arte em relação a busca de soluções para o problema.

#### 2.1.2.1 Técnicas manuais

No trabalho de Gonçalves e Faria (1989), que busca analisar a eficiência de inseticidas granulados no controle de ninfas móveis em cafeeiros, o levantamento populacional, a exemplo de outros aqui citados, é feito por meio de observação direta e contagem não automatizada de insetos. Particularmente nesse caso, de forma aleatória nos talhões, foram abertas trincheiras (1.5m X 1.5m X 0.5m) no centro da cova e o solo retirado foi peneirado, sendo os insetos coletados e contados pelos pesquisadores.

De acordo com Reis e Souza (1991) e Martinelli *et al.* (1998) é possível detectar a presença de cigarras na lavoura analisando o estado das plantas, cujos sintomas são: “clorose das folhas das extremidades dos ramos, queda precoce de folhas, com permanência de folhas no ápice” (MARTINELLI *et al.*, 1998), confirmando que o modo mais eficiente de manejo é realizado por meio de controle químico com inseticidas.

Em Martinelli e Zucchi (1997), o monitoramento é feito por armadilha constituída de uma armação de ferro coberta com tela de nylon, sob a copa do café e, complementarmente, também foram realizadas coletas manuais.

No artigo de Fornazier (2000), é analisada a eficiência de diversas formas de aplicação de inseticidas, para tanto são feitas covas sob as plantas a fim de realizar a contagem manual das ninfas para estimativa de sua população inicial. A avaliação dos resultados também foi realizada de forma manual, analisando não só o número de ninfas ainda vivas após o procedimento, mas também a medida da produtividade de 50 das plantas submetidas ao processo. Neste artigo, fica registrado que o incremento da produtividade ocorreu somente após o terceiro ano de aplicação de defensivos.

Em Almeida (2004), também é citado o processo manual de coleta e contagem (por observação direta) de ninfas das cigarras. Nesse caso, o autor se refere a várias

visitas realizadas em diversas regiões produtoras do estado de São Paulo, atividades que indubitavelmente consumiram grande dispêndio de tempo.

Em Santos (2007), é feito o registro de uma espécie de cigarras no município de Tapiratiba, estado de São Paulo, onde os exemplares foram coletados manualmente com rede entomológica, que consiste em um aro com rede, preso a um cabo de madeira ou de aço.

Segundo Souza, Reis e Silva (2007) o monitoramento da infestação de cigarras na cultura do café é feito manualmente e consiste em entrincheirar covas nas plantações à procura de ninfas nas raízes ou em verificar a presença de exúvias ao longo da base do caule. Esse processo é realizado escolhendo-se 10 plantas aleatórias em um talhão, entrincheirando-as de um único lado, de fora para dentro. Terminado o processo, as ninfas capturadas devem ser contadas e o resultado multiplicado por 2, em razão de o processo ser realizado somente de um dos lados da árvore. O autor recomenda que, caso o resultado seja de 35 ninfas ou mais, deve-se realizar o controle químico.

O método do entrincheiramento também é empregado em Lunz *et al.* (2010), utilizando implemento tratorizado para raspagem de solo. Com auxílio de hastes metálicas são feitas verificações à procura de ninfas vivas. Posteriormente, são aplicados inseticidas e é feita a comparação de populações entre amostragens antes e depois de sua aplicação.

No artigo de Maccagnan *et al.* (2014), são apresentadas evidências da detecção de cigarras, por meio da constatação a olho nu da presença de exúvias presas ao tronco das árvores bem como galerias de ninfas junto ao sistema radicular. A catação manual foi realizada para os indivíduos adultos presentes no tronco das árvores.

Em Hou *et al.* (2015), o monitoramento é realizado pelo processo de escavação em três tipos de plantas: *Pinus tabuliformis* Carr. (Pinaceae), *Populus tomentosa* Carr. (Salicaceae) e *Pyrus xerophila* Yu (Rosaceae). A distribuição das ninfas em solo é feita pela contagem manual a cada 10cm de profundidade do solo. A coleta de exúvias foi realizada diariamente às 9h da manhã durante os anos 2013 e 2014.

No trabalho de Sato e Sato (2015), foi realizado o monitoramento duas espécies de cigarras no Japão a fim de verificar o impacto da mudança de ambiente. A coleta das exúvias foi realizada manualmente e por uma rede com cabo de 4 metros para coleta em lugares mais elevados. A contagem foi realizada em laboratório.

#### 2.1.2.2 Técnicas com software

Em Soares *et al.* (2008), foi analisada a distribuição de ninfas de cigarras em plantações de paricá (*Schizolobium amazonicum* Huber ex Ducke) através de interpolação de unidades amostrais. A partir do mapa da localidade, foi utilizado um software para dividir a área a fim de encontrar pontos para coletas de amostragens. A contagem das

cigarras foi feita de forma não automatizada, ou seja, pela observação direta dos insetos e os dados geraram mapas de dispersão dos indivíduos no solo a fim de conhecer sua incidência dentro da região de estudo.

Em Chen e Shiao (2008), são apresentadas duas novas espécies de cigarras. A metodologia consistiu na captura, em campo, que foi realizada por um gravador digital Foster FR-2 e um microfone Sennheiser MKH 70 P48. O microfone foi colocado a cerca de 5m dos insetos em geral. O espectro de frequências é analisado a fim de diferenciar as espécies.

No trabalho publicado por Cole (2008), relata-se a captura de sons de cigarras com um microfone modelo C451E e um gravador DAT. A maioria das gravações aconteceu a 1 metro ou mais de distância. As frequências foram analisadas por espectrogramas gerados a partir da Transformada Rápida de Fourier (*Fast Fourier Transform* - FFT). Nesse caso, uma comparação é feita entre duas espécies, analisando os gráficos dos sons dos machos.

O autor do artigo Lee (2008) apresenta uma pesquisa na qual é feito o levantamento das espécies de cigarras na península coreana. Os sons das cigarras são capturados e analisados por oscilogramas e sonogramas. Os dados são comparados, não para detectar, mas para diferenciar as espécies.

Outro trabalho que apresenta conceitos importantes para esse foco de estudo é Sueur *et al.* (2008), cuja proposta é uma solução econômica e não invasiva para o desenvolvimento de inventários RBA (*Rapid Biodiversity Assessments*) utilizando o som da comunicação das espécies para estimar a população em determinada localidade. Foram aplicados recursos como o índice de Shannon e Transformada de Fourier de Tempo Reduzido (*Short Time Fourier Transform* - STFT). As gravações foram realizadas utilizando um microfone omni-direcional Sennheiser Me62/K6 conectado a um gravador digital Edirol R-09, com sinais de 30 segundos cada. O estudo foi aplicado em 45 espécies, sendo 15 de aves, 15 de anfíbios e 15 de insetos. Nos experimentos, foram criados sons juntando coros de diversas espécies com o objetivo de verificar a eficácia do sistema em ambientes onde elas vivem em conjunto.

Em Seabra, Quartau e Bruford (2009), descreve-se uma pesquisa sobre a variação de espécies de cigarras na região do Mar Mediterrâneo. São citadas questões importantes, como a similaridade morfológica e acústica entre as espécies *Cicada barbara* e *Cicada orni*. Entretanto, a pesquisa ficou restrita à variação genética e geográfica das espécies, não sendo aplicados estudos relacionados ao som. Os espécimes foram coletados manualmente ou encaminhados por instituições parceiras e estudados em laboratório. Os estudos estatísticos foram realizados por meio de *softwares* específicos para esse fim.

Em Sanborn, Heath e Heath (2009), relata-se que os sons emitidos pela cigarra *Beameria venosa* são analisados cuidadosamente. O inseto emite uma alta frequência, o que

---

se torna uma vantagem, pois seus predadores (pássaros e lagartos) não ouvem, resultando em baixa pressão predatória. Os sons foram capturados em fitas, em um dispositivo Uher 4000. O microfone foi colocado o mais próximo possível, para de evitar interferências externas. As gravações foram realizadas em diferentes cidades dos EUA.

De acordo com o artigo Dardar, Belal e Basheer (2013), as cigarras adultas foram coletadas manualmente das macieiras (*Malus domestica*) ao por do sol e ao amanhecer, por se mostrarem mais fáceis de coletar nesses períodos. Nesse estudo, foi utilizada uma armadilha com uma estrutura em formato de “tenda” com tela de pano. As ninfas foram coletadas por escavação de 10 a 15cm. Também nesse trabalho foram coletados os sons, posteriormente analisados por meio do *software Speech Analyzer*. Os ciclos de frequência também foram analisados a fim de registrar as características dos indivíduos frente ao ambiente de estudo.

No caso de Hill e Marshall (2013), registra-se a pesquisa sobre um gênero de cigarras comum nos EUA. Foi utilizado um gravador digital e um microfone para capturar o som dos insetos e um *Global Positioning System* (GPS) manual foi utilizado para precisar a posição geográfica. Os sons foram analisados em laboratório utilizando *softwares*, sendo que a temperatura também foi medida manualmente.

A identificação de insetos pelo uso do sons emitidos por eles é efetivamente aplicada por Lampson *et al.* (2013), com o objetivo de diminuir tempo e custos com defensivos agrícolas, entretanto o processo é realizado por meio de um acelerômetro, pois o sinal emitido pelo inseto estudado é o sinal vibratório propagado através do substrato. Aqui são aplicados os métodos de identificação GMM (*Gaussian Mixture Model*) e PNN (*Probabilistic Neural Network*). Os resultados obtidos foram de 92% de detecção com testes envolvendo pares de macho-fêmea e 86% com conjuntos de 30 indivíduos. Os testes foram realizados em laboratório com cerca de 100 indivíduos, alguns criados em laboratório e outros previamente capturados. As gravações foram realizadas por meio de um acelerômetro (Modelo 4370, Bruel & Kjaer).

Importante trabalho apresentado em relação ao objeto de estudo aqui desenvolvido está em documentado em Aide *et al.* (2013). Os autores desenvolvem um sistema de coleta de dados, monitoramento e identificação baseado em gravações de áudio. O sistema é composto por *softwares*, além de uma estação de monitoramento remota alimentada por energia solar. Essa estação envia 1 minuto de gravação a cada 10 minutos de intervalo. As gravações de áudio são coletadas por um microfone com resposta em frequência de 20 Hz a 20 kHz conectado a um dispositivo iPod. O processamento dos arquivos é feito por meio de uma STFT. A detecção é realizada por meio de um algoritmo que utiliza *Hidden Markov Models* (HMMs).

No caso de Chiavacci, Bednarz e Mckay (2015), a detecção das cigarras foi realizada por meio de armadilhas com câmeras, com o objetivo de analisar a densidade populacional

de cigarras na região, além de estudar a relação entre a incidência da população de cigarras e escolha da localização dos ninhos pelos papagaios, que são seus predadores naturais. A contagem foi realizada manualmente a cada 3 dias, sendo tanto as cigarras adultas quanto as exúvias armazenadas separadamente. A posição de cada armadilha foi registrada manualmente pelos pesquisadores utilizando um dispositivo GPS.

O trabalho de Prince *et al.* (2019) apresenta um conjunto de algoritmos de detecção acústica com intuito de reduzir consumo de energia e armazenamento em dispositivos de baixo custo. O foco principal dos autores está no monitoramento da ecolocalização de morcegos, de espécies de cigarras ameaçadas de extinção, além de buscar evidências de caça predatória em reservas naturais protegidas.

### 2.1.3 Reflexões sobre o estado da arte

Conforme pode ser observado na Tabela 1, na maioria dos trabalhos não há utilização de *softwares* no processo de monitoramento. Além disso, somente dois trabalhos utilizam software em campo e, desses, apenas um se foca no desenvolvimento de um dispositivo automatizado direcionado para o problema proposto. Todos os trabalhos relacionados utilizam algum tipo de técnica manual na lavoura.

Nos trabalhos de Chen e Shiao (2008), Cole (2008), Lee (2008) e Hill e Marshall (2013), registra-se a aplicação de processamento de áudio para fins de identificação de espécies, entretanto esse procedimento é realizado em laboratório e não há evidência do uso de redes neurais, nem outro procedimento para o desenvolvimento de um possível processo automatizado de identificação, nem da aplicação dos recursos diretamente em campo.

Uma das possibilidades para um manejo natural dos insetos-praga pode estar no uso de predadores naturais. O artigo de Almeida (2004) contém um destaque para um dos inimigos naturais mais comuns das cigarras adultas, o gavião *Miluago chimchima*. Em Sazima (2009), também há relatos de predadores de cigarras adultas, como o gavião-sovi (*Ictinia plumbea*) e Cuculidae (*Crotophaga ani*, *Guira guira* e *Piaya cayana*). No caso das ninfas, os predadores são o frango-d'água (*Gallinula chloropus*) e o socozinho (*Butorides striata*). Outros pássaros são considerados potenciais predadores das cigarras, como comentado sobre o Gavião do Mississipi por Chiavacci, Bednarz e McKay (2015).

Na maioria dos artigos referentes ao controle de cigarras, a questão ambiental, como a contaminação do solo por defensivos, é enfatizada. Em Fornazier (2000), quando a aplicação dos defensivos para controle de cigarras no café é analisada em solo em declive, como montanhas, os autores recomendam a abolição dos procedimentos de aplicação de granulado sistêmico devido ao perigo de contaminação do ambiente e do trabalhador.

Muitos dos artigos aqui apresentados utilizam o trabalho manual no processo de

Tabela 1 – Trabalhos Correlatos: técnicas manuais (TM), uso de *softwares* no campo (SC) e uso de *softwares* em laboratório (SL).

Artigo	TM	SC	SL
(GONÇALVES; FARIA, 1989)	Sim	Não	Não
(REIS; SOUZA, 1991)	Sim	Não	Não
(MARTINELLI; ZUCCHI, 1997)	Sim	Não	Não
(FORNAZIER, 2000)	Sim	Não	Não
(SOUZA; REIS; SILVA, 2007)	Sim	Não	Não
(SANTOS, 2007)	Sim	Não	Não
(LEE, 2008)	Sim	Não	Sim
(COLE, 2008)	Sim	Não	Sim
(CHEN; SHIAO, 2008)	Sim	Não	Sim
(SUEUR <i>et al.</i> , 2008)	Sim	Não	Sim
(SEABRA; QUARTAU; BRUFORD, 2009)	Sim	Não	Sim
(SANBORN; HEATH; HEATH, 2009)	Sim	Não	Sim
(LUNZ <i>et al.</i> , 2010)	Sim	Não	Não
(DARDAR; BELAL; BASHEER, 2013)	Sim	Não	Sim
(HILL; MARSHALL, 2013)	Sim	Não	Sim
(LAMPSON <i>et al.</i> , 2013)	Sim	Não	Sim
(AIDE <i>et al.</i> , 2013)	Sim	Sim	Sim
(MACCAGNAN <i>et al.</i> , 2014)	Sim	Não	Não
(HOU <i>et al.</i> , 2015)	Sim	Não	Não
(SATO; SATO, 2015)	Sim	Não	Não
(CHIAVACCI; BEDNARZ; MCKAY, 2015)	Sim	Não	Sim
(PRINCE <i>et al.</i> , 2019)	Sim	Sim	Sim

Fonte: Elaborada pelo autor.

detecção e monitoramento das espécies de cigarras, pela catação, pela escavação do solo na região da projeção da copa das plantas para coleta de ninfas, por observação direta a olho nu ou, simplesmente, por ouvir seu som. A partir da utilização desses métodos, considera-se, que existe incidência naquela região e que a aplicação dos defensivos químicos deve ser realizada, mesmo sem conhecer sua densidade populacional. Em alguns trabalhos como o de Lampson *et al.* (2013), o sinal acústico propagado através do substrato é empregado com intuito de detectar insetos, entretanto, não foram realizados testes com cigarras. A infraestrutura de coleta e processamento de dados apresentada em Aide *et al.*

(2013) se mostra importante exemplo a ser seguido para os objetivos do presente estudo. Em Sueur *et al.* (2008) e Prince *et al.* (2019), também há uso de sons para detecção de insetos, mostrando-se trabalhos promissores na utilização de procedimentos não invasivos e automatizados para detecção e identificação que podem ser empregados no presente problema.

Uma possível solução para o processo de monitoramento de cigarras na lavoura de café pode estar na elaboração de um dispositivo que seja capaz de detectar a presença de cigarras na lavoura. Esse dispositivo pode ser desenvolvido utilizando técnicas de processamento digital de áudio utilizando DWT (FONSECA *et al.*, 2007; GUIDO *et al.*, 2007; JUNIOR *et al.*, 2007; CHEN; CHEN; GUIDO, 2010; ESCOLA, 2014) e reconhecimento de padrões com redes neurais artificiais (HAYKIN *et al.*, 2009; SILVA; SPATTI; FLAUZINO, 2010).

Em Escola (2014), o conjunto de tecnologias aplicadas ao reconhecimento de assinaturas por meio do ruído emitido pela caneta no processo de assinatura de documentos manuscritos pode ser empregado para detectar e identificar as diversas espécies de cigarras. Nesta dissertação foi realizado o processamento dos sinais de áudio pela DWT e o classificador SVM, mostrando-se importante ferramenta para o desenvolvimento de uma solução que possibilite a detecção e identificação de cigarras no ambiente cafeeiro sem a necessidade de intervenção manual do ser humano. Futuramente, essa solução pode ser utilizada para auxílio no manejo de insetos, mais especificamente, para auxiliar o homem na tomada de decisão em relação ao melhor momento de implementar a medida de controle.

Dispositivos eletrônicos como em Aide *et al.* (2013) ou ainda como *Raspberry Pi* (RICHARDSON; WALLACE, 2012) podem ser empregados com a finalidade de acondicionar o *software* de rede neural na lavoura, que, por se tratar de dispositivo de razoável capacidade de processamento e prover baixo custo de energia, pode se mostrar solução compatível com a aplicação aqui pretendida. Como citado em (RICHARDSON; WALLACE, 2012), o consumo de energia e sua manutenção deve ser um problema, podendo ser aplicados recursos como substituição periódica de baterias ou ainda a coleta de energia solar por meio de placas fotovoltaicas, caso o fornecimento de energia elétrica seja um problema por conta da localização do dispositivo na lavoura.

O uso do processamento digital de imagens também pode ser empregado para detecção não só de exúvias ou de adultos no tronco das plantas, mas também das ninfas no solo. Neste caso, seria necessário incluir o processo de captura das imagens, que pode ser feito manualmente, utilizando uma câmera fotográfica, empregando um drone (JESUS; SILVA; ROCHA, 2015) ou robô (CARDOSO *et al.*, 2020a; CARDOSO *et al.*, 2020b).

## 2.2 Engenharia Paraconsistente de Características (EPC): selecionando *features*

A qualidade dos sistemas de classificação constitui busca constante por parte de engenheiros e pesquisadores da área de PDS. Garantir que o sistema de classificação desenvolvido obterá o melhor resultado possível com o mínimo custo computacional é desejo permanente. Muitas vezes, um algoritmo de inteligência artificial é desenvolvido, de forma a resolver determinado problema, mas não é realizado nenhum tipo de análise para verificar se tal esforço é compatível ou exagerado em relação ao problema a que se propõe a resolver (BLUMER *et al.*, 1987).

A EPC (GUIDO, 2018a; FONSECA *et al.*, 2020), baseada na Lógica Paraconsistente (CARNIELLI; CONIGLIO, 2016), permite realizar uma análise da qualidade dos elementos entregues ao classificador, de forma que se possa entender o quão importantes são as características utilizadas. Assim, pode-se otimizar o sistema de classificação proposto e, possivelmente, reduzir a dimensão dos vetores de características (comumente chamadas de *features*), diminuindo portanto o custo computacional e o dispêndio energético do sistema.

A escolha da EPC na análise de características se justifica por se tratar de uma técnica recentemente apresentada na literatura científica, que não havia ainda sido utilizada para a finalidade aqui proposta. Além disso, se mostrando recurso de comprovada eficiência, sendo implementada em estudos recentes de PDS (GHOSH; PAL; DAS, 2020; CAOBIANCO; GUIDO; SILVA, 2021; GUIDO *et al.*, 2021).

De acordo com Guido (2018a), pode-se exemplificar o uso da EPC considerando-se um problema de classificação de padrões cujas classes  $\{C_1, C_2, C_3, \dots, C_N\}$  são representadas por vetores de características com  $T$  amostras. Considerando que um determinado conjunto de características de entrada gere um resultado satisfatório em um classificador modesto, é de se esperar que gere também um resultado equivalente em um classificador mais elaborado e devidamente sintonizado. Assim, a seleção das melhores características de entrada baseada em um método modesto permite a generalização, ou seja, os vetores serão adequados tanto para uso com classificadores modestos, quanto os mais sofisticados. Esse é o princípio norteador para utilização de estratégias modestas de análise das características.

Pode-se considerar um classificador menos robusto (modesto) como sendo do tipo Pattern-Matching Classifier (PMC): limiar, distância euclidiana, *Dynamic Time Warping* (DTW), por exemplo; e não modesto o classificador do tipo Knowledge-Based Classifier (KBC): SVM, *Multi-Label Perceptron* (MLP), *Deep Neural Network* (DNN), por exemplo.

Numa situação ideal, vetores de características extraídos de determinada classe apresentam considerável similaridade entre si, denotando, por outro lado, marcante distinção quando extraídos de classes diferentes. Assim, a EPC busca um método quantitativo para investigar os conjuntos de entrada, observando possíveis conflitos intraclasses e interclasses.

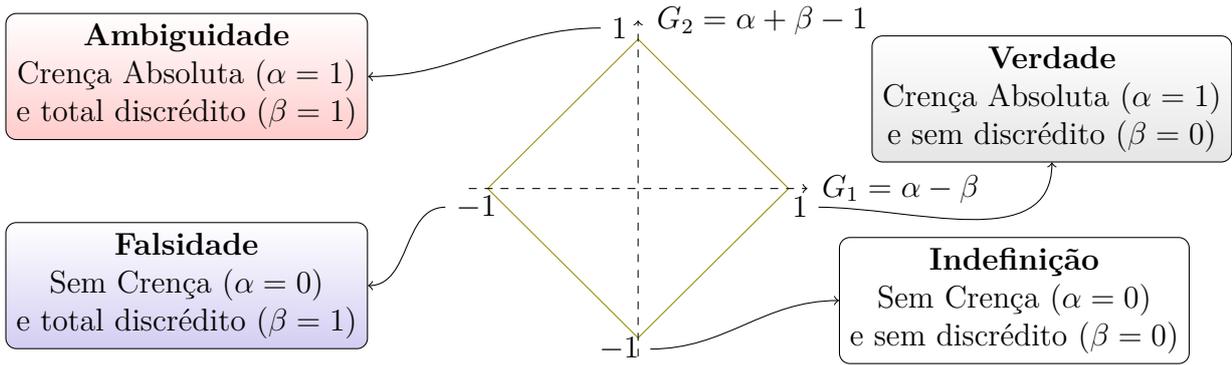


Figura 1 – Plano Paraconsistente (Adaptado de (GUIDO, 2018a))

### 2.2.1 Etapas para o uso da EPC

1. Normalizar os vetores de características de modo que todos os elementos permaneçam no intervalo de 0 a 1.
2. Calcular o número de sobreposições interclasses ( $R$ ). O ideal é que não existam sobreposições, pois, não será possível separar linearmente as classes.
3. Calcular o nível de similaridade intraclasse ( $\alpha$ );

$$\alpha = \min\{\bar{Y}(C_1), \bar{Y}(C_2), \dots, \bar{Y}(C_N)\} \quad (2.1)$$

onde:

$$\bar{Y}(C_1) = \frac{1}{T} \sum_{i=0}^{T-1} \text{sv}C_1[i], \quad (2.2)$$

$$\bar{Y}(C_2) = \frac{1}{T} \sum_{i=0}^{T-1} \text{sv}C_2[i], \quad (2.3)$$

⋮

$$\bar{Y}(C_N) = \frac{1}{T} \sum_{i=0}^{T-1} \text{sv}C_N[i], \quad (2.4)$$

4. Calcular o nível de dissimilaridade interclasse ( $\beta$ ), isto é,  $\beta = \frac{R}{F}$ , sendo  $F = N \cdot (N - 1) \cdot X \cdot T$ , onde  $N$  é o número de classes,  $X$  é número de vetores e  $T$  é tamanho de cada vetor.
5. Calcular os graus de certeza e de contradição, isto é,  $G_1 = \alpha - \beta$  e  $G_2 = \alpha + \beta - 1$ , respectivamente, e plotar o ponto  $P = (G_1, G_2)$  no plano paraconsistente, como exibido na Figura 1.

Notavelmente,  $(G_1, G_2) = (-1, 0)$ ,  $(G_1, G_2) = (1, 0)$ ,  $(G_1, G_2) = (0, -1)$  e  $(G_1, G_2) = (0, 1)$  implicam em Falsidade, Verdade, Indefinição e Ambiguidade, respectivamente. Nesse contexto, conforme  $P = (G_1, G_2)$  se aproxima dos vértices, pode-se considerar que:

- $(G_1, G_2) = (-1, 0)$  (**Falsidade**): “um classificador não linear profundo é requerido por conta de que os vetores de entrada intraclasse são notavelmente espalhados e os vetores interclasse significativamente sobrepostos” (GUIDO, 2018a).
- $(G_1, G_2) = (1, 0)$  (**Verdade**): “um classificador linear é possivelmente suficiente para resolver o problema, pois, os vetores de entrada estão agrupados e os vetores interclasse não se sobrepõem” (GUIDO, 2018a).
- $(G_1, G_2) = (0, -1)$  (**Indefinição**): “os dados de entrada causam indefinição, ou seja, tanto os dados de entrada quanto os vetores interclasse são muito diferentes” (GUIDO, 2018a).
- $(G_1, G_2) = (0, 1)$  (**Ambiguidade**): “os dados de entrada causam ambiguidade, pois, tanto os dados de entrada quanto os vetores interclasse são consideravelmente similares” (GUIDO, 2018a).

Quanto mais próximo do vértice  $(G_1=1, G_2=0)$  o ponto  $P$  estiver, mais favoráveis serão considerados os vetores de características apresentados para a resolução do problema e mais simples será o classificador necessário.

Caso não seja possível encontrar um conjunto de entradas refinado o suficiente após o referido estudo dos possíveis conjuntos de entrada, ainda há a possibilidade de se utilizar métodos como *Auto Machine Learning* (FEURER *et al.*, 2019) e *Deep Learning* (LECUN; BENGIO; HINTON, 2015; GOODFELLOW; BENGIO; COURVILLE, 2016; WANG *et al.*, 2021), pois costumam convergir satisfatoriamente mesmo com um conjunto desfavorável de características de entrada, entretanto implicam em custo computacional.

### 2.2.2 Exemplo numérico do uso da EPC

1. Considera-se as  $N=4$  classes com  $X=5$  vetores de tamanho  $T=2$ :

classe  $C_1$ :  $\{0.98, 0.94\}$ ,  $\{0.96, 0.99\}$ ,  $\{0.98, 0.97\}$ ,  $\{0.96, 0.92\}$  e  $\{0.97, 0.94\}$

classe  $C_2$ :  $\{0.87, 0.83\}$ ,  $\{0.87, 0.88\}$ ,  $\{0.89, 0.87\}$ ,  $\{0.86, 0.83\}$  e  $\{0.87, 0.85\}$

classe  $C_3$ :  $\{0.76, 0.72\}$ ,  $\{0.78, 0.77\}$ ,  $\{0.77, 0.79\}$ ,  $\{0.75, 0.74\}$  e  $\{0.77, 0.72\}$

classe  $C_4$ :  $\{0.65, 0.61\}$ ,  $\{0.64, 0.66\}$ ,  $\{0.67, 0.68\}$ ,  $\{0.65, 0.65\}$  e  $\{0.67, 0.61\}$

2. Calcula-se os vetores de intervalos para cômputo do número de sobreposições interclasses:

classe  $C_1$ :  $\{0.96, 0.92\}$ , como menores componentes, e  $\{0.98, 0.99\}$ , sendo os maiores elementos;

classe  $C_2$ :  $\{0.86, 0.83\}$ , como menores componentes, e  $\{0.89, 0.88\}$ , sendo os maiores elementos;

classe  $C_3$ :  $\{0.76, 0.72\}$ , como menores componentes, e  $\{0.78, 0.79\}$ , sendo os maiores elementos;

classe  $C_4$ :  $\{0.64, 0.61\}$ , como menores componentes, e  $\{0.67, 0.68\}$ , sendo os maiores elementos;

Comparando cada elemento dos vetores iniciais  $C_1$  a  $C_4$  com os vetores de intervalos  $C_1$  a  $C_4$ , nota-se que não há sobreposição, ou seja,  $R = 0$ , conforme a Figura 2.

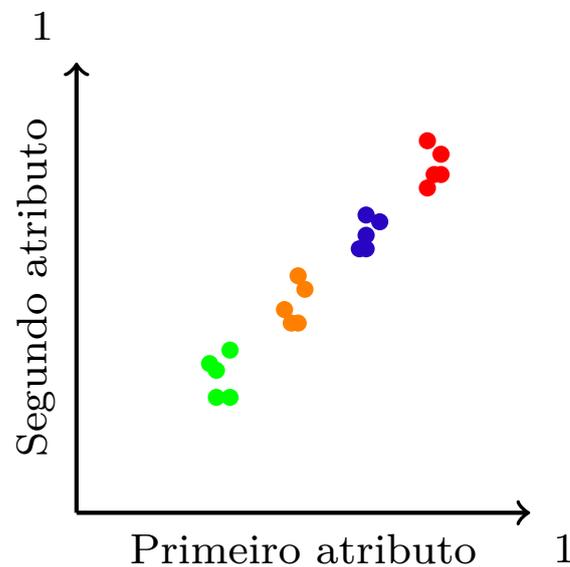


Figura 2 – Atributos de entrada das classes  $C_1$ ,  $C_2$ ,  $C_3$  e  $C_4$  em vermelho, azul, laranja e verde, respectivamente. Claramente são linearmente separáveis.

3. Computa-se o nível de similaridade intraclasse ( $\alpha$ ):

- Computa-se os vetores de amplitude:

$$\text{classe } C_1: \{0.98 - 0.96, 0.99 - 0.92\} = \{0.02, 0.07\}$$

$$\text{classe } C_2: \{0.89 - 0.86, 0.88 - 0.83\} = \{0.03, 0.05\}$$

$$\text{classe } C_3: \{0.78 - 0.76, 0.79 - 0.72\} = \{0.02, 0.07\}$$

$$\text{classe } C_4: \{0.67 - 0.64, 0.68 - 0.61\} = \{0.03, 0.07\}$$

- Computa-se os vetores de similaridade:

$$\text{classe } C_1: \{1 - 0.02, 1 - 0.07\} = \{0.98, 0.93\}$$

$$\text{classe } C_2: \{1 - 0.03, 1 - 0.05\} = \{0.97, 0.95\}$$

classe  $C_3$ :  $\{1 - 0.02, 1 - 0.07\} = \{0.98, 0.93\}$

classe  $C_4$ :  $\{1 - 0.03, 1 - 0.07\} = \{0.97, 0.93\}$

$$\bar{Y}(C_1): \frac{0.98+0.93}{2} = 0.955$$

$$\bar{Y}(C_2): \frac{0.97+0.95}{2} = 0.960$$

$$\bar{Y}(C_3): \frac{0.98+0.93}{2} = 0.955$$

$$\bar{Y}(C_4): \frac{0.97+0.93}{2} = 0.950$$

Resultando em  $\alpha = \min\{0.955, 0.960, 0.955, 0.950\} = 0.950$ .

4. Computa-se o nível de dissimilaridade interclasse:

$$F = 4 \cdot (4 - 1) \cdot 5 \cdot 2 = 120$$

$$\beta = \frac{R}{F} = \frac{0}{120} = 0,$$

5. Computa-se o grau de certeza e contradição, localizando  $P$  no plano paraconsistente:

$$G_1 = \alpha - \beta = 0.950 - 0 = 0.95$$

$$G_2 = \alpha + \beta - 1 = 0.950 + 0 - 1 = -0.05$$

- Plotando o gráfico  $P = (G_1, G_2)$  e calculando a distância do ponto aos quatro cantos do plano paraconsistente, tem-se:

$$d((G_1, G_2), (-1, 0)) = \sqrt{(0.95 + 1)^2 + (-0.05)^2} = 1.95$$

$$d((G_1, G_2), (1, 0)) = \sqrt{(0.95 - 1)^2 + (-0.05)^2} = 0.07$$

$$d((G_1, G_2), (0, -1)) = \sqrt{(0.95)^2 + (-0.05 + 1)^2} = 1.34$$

$$d((G_1, G_2), (0, 1)) = \sqrt{(0.95)^2 + (-0.05 - 1)^2} = 1.41$$

Assim, considerando que o ponto  $P$  na Figura 3 está mais próximo do vértice  $(1, 0)$ , que representa “Verdade”, sendo o Vértice Ótimo (VO), a adoção de um sistema de classificação mais elaborado não é necessária para o referido caso.

### 2.3 Wavelets e Escala Bark

O PDS de áudio consiste na realização de cálculos computacionais em dados acústicos provenientes de diversos meios, como um arquivo digital ou um microfone. Inúmeras são as ferramentas utilizadas na ciência para PDS de áudio, tal como a Transformada Discreta de Fourier (DFT) e a Transformada Rápida de Fourier (FFT). Essas técnicas, entretanto, possibilitam somente a conversão de um sinal do domínio do tempo para o domínio da frequência. Além de permitir filtragens em sinais de áudio, a DWT (GUIDO,

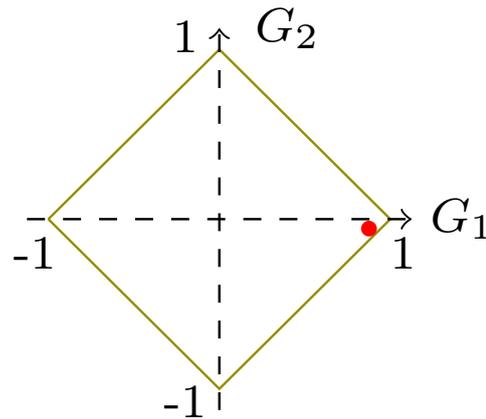


Figura 3 –  $P$ , representado como um ponto vermelho no plano paraconsistente.

2017) revela o suporte de tempo de cada frequência do sinal, de forma eficiente (GUIDO; ADDISON; WALKER, 2009).

A DWT é considerada uma ferramenta poderosa de análise tempo-frequência, ou seja, a partir de um sinal no domínio do tempo, pode-se obter um sinal no domínio tempo-frequência (GUIDO; ADDISON; WALKER, 2009), e é utilizada em diversas pesquisas (GUIDO, 2017; GUIDO, 2021; WANG *et al.*, 2020; CAOBIANCO; GUIDO; SILVA, 2021) envolvendo processamento de sinais. Com a DWT, é possível produzir filtragens e separações de um sinal em sub-bandas do seu espectro de frequências. Além disso, segundo (JENSEN; COUR-HARBO, 2001; WALKER, 2002; ADDISON, 2017), a DWT realiza essas tarefas de maneira ainda mais eficiente do que a Transformada de Fourier de Tempo Reduzido (STFT) (HAYKIN; VEEN, 2001).

Duas são as abordagens, equivalentes, para aplicação da DWT: a primeira é baseada em conceitos da álgebra linear e a segunda está mais intimamente relacionada aos conceitos elementares de PDS, sendo portanto adotada neste trabalho. Em tal caso, a DWT funciona como um par de filtros, sendo um deles passa-baixas ( $h[\cdot]$ ) e o outro, passa-altas ( $g[\cdot]$ ), de modo que, em geral, as frequências de corte (-3dB) concentram-se na metade da máxima frequência presente no sinal de entrada.

O processo de aplicação da DWT em um sinal discreto consiste em submetê-lo a ambos os filtros via convolução (HAYKIN; VEEN, 2001; JÚNIOR, 2007). Esse processo pode ser aplicado  $\frac{\log(n)}{\log(2)}$  vezes para um sinal de comprimento  $n$  e, para cada aplicação desse processo, tem-se um nível de decomposição e obtêm-se dois novos sinais, um com as frequências abaixo da metade da máxima frequência e o outro com as frequências acima da metade da máxima frequência do sinal original. Quando da passagem do sinal original pelo filtro passa-altas, os termos obtidos são chamados coeficientes de *detalhamento*. Contrariamente, quando da passagem do sinal original pelo filtro passa-baixas, os termos chamados coeficientes de *aproximação* designam o sinal obtido. No caso da DWT regular,

após cada nível de decomposição, somente o sinal obtido pela aplicação do filtro passa-baixas é aproveitado para continuidade do processo recursivo de decomposição. Entretanto, no caso da *DWT-packet* (DWTP) (WANG *et al.*, 2020), ambos os sinais resultantes em um nível são utilizados para produzir o nível seguinte de decomposição.

É importante frisar que cada vez que um nível da transformação é realizado, os dois novos sinais obtidos são subamostrados por 2, pois eles contêm apenas metade da faixa de frequências do sinal original, de acordo com o Teorema da Amostragem (HAYKIN; VEEN, 2001) e como ilustra a Figura 4. Um sinal de  $n$  amostras tem a sua DWT com a mesma quantidade de amostras, compostas por uma sequência de coeficientes, iniciando-se com aqueles provenientes da aplicação do filtro passa-baixas no último nível, seguidos pelos coeficientes resultantes da aplicação dos filtros passa-altas nos níveis intermediários e terminando com os coeficientes resultantes da aplicação do filtro passa-altas do primeiro nível de decomposição.

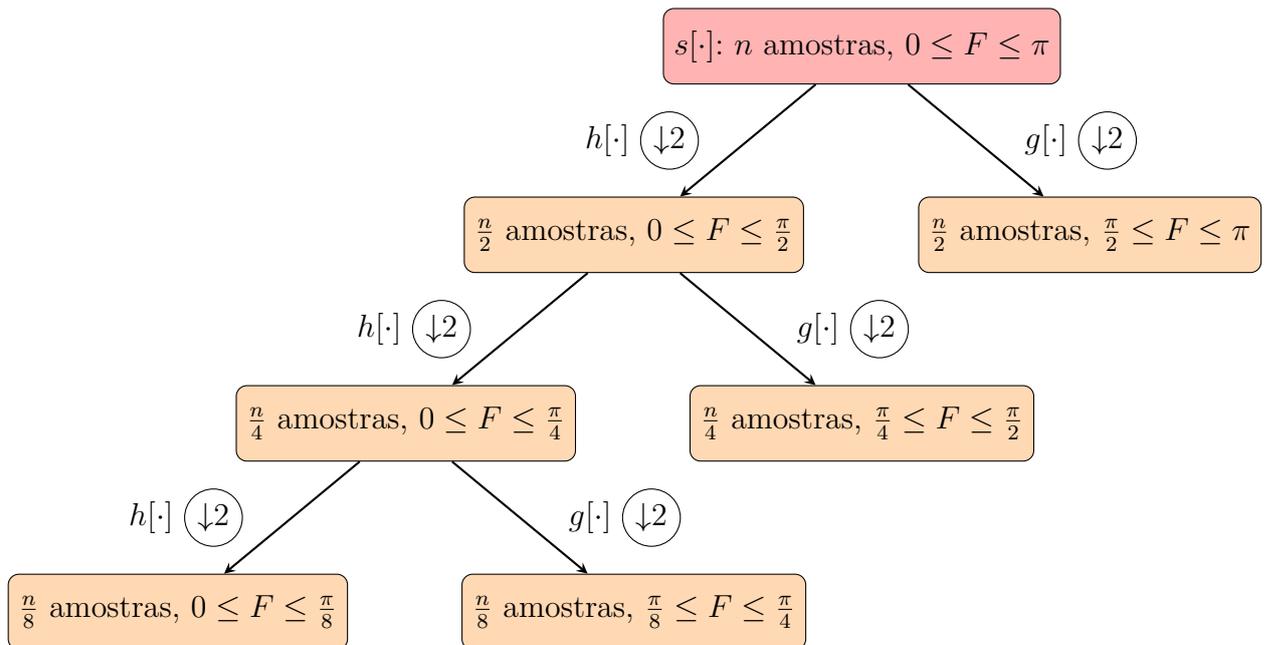


Figura 4 – DWT: sinal  $s[\cdot]$  de  $n$  amostras discretas e máxima frequência  $\pi$ , decomposto até o terceiro nível.

Para se obter a melhor resolução em frequência, é necessário realizar a decomposição até o último nível possível, por isso, é necessário que o sinal discreto tenha comprimento equivalente a uma potência de 2. Além disso, não se pode utilizar qualquer filtro para realizar a DWT, pois, para que um filtro digital seja considerado um filtro *wavelet*, é necessário que a resposta em frequência do filtro passa-baixas seja 0 em  $\omega = \pi$ .

Como observado, o cálculo da DWT é um processo de filtragem e subamostragem por 2, realizado conjuntamente nos sinais da DWT em cada nível, o que pode ser representado

por uma “convolução adaptada” da seguinte forma:

$$y[\cdot] = x[\cdot] * t[\cdot] = \sum_{k=0}^{n-1} t_k x_{2n-k} \quad , \quad (2.5)$$

ou, mais especificamente:

$$y_{passa-baixas}[\cdot] = x[\cdot] * h[\cdot] = \sum_{k=0}^{n-1} h_k x_{2n-k} \quad (2.6)$$

e

$$y_{passa-altas}[\cdot] = x[\cdot] * g[\cdot] = \sum_{k=0}^{n-1} g_k x_{2n-k} \quad . \quad (2.7)$$

Apenas para complementar essa revisão, é interessante mencionar que a possibilidade de aplicação da DWT que está mais intimamente conexa com os conceitos da álgebra linear consiste na análise de multi-resolução (*Multi-Resolution Analysis* - MRA), proposta por Mallat (MALLAT, 1989; WILLIAMS; AMARATUNGA, 1994; AKAY, 1998). Essa técnica consiste em decompor o sinal sob análise  $\vec{f}$ , que é um vetor, em uma soma de outros vetores, que são pertencentes a uma sequência de subespaços vetoriais (LIPSCHUTZ; ABELLANAS; ONTALBA, 1992). Em outras palavras, isso significa representar um sinal em vários níveis de resolução. Assim, de acordo com a MRA, no caso de um vetor  $\vec{f}$  de  $n$  pontos:

$$\vec{f} = \vec{A} + \vec{D} \quad (2.8)$$

onde

$$\vec{A} = \sum_{k=0}^{\frac{n}{2}-1} \langle \vec{f}, \vec{v}_k \rangle \vec{v}_k \quad ,$$

e

$$\vec{D} = \sum_{k=0}^{\frac{n}{2}-1} \langle \vec{f}, \vec{w}_k \rangle \vec{w}_k \quad ,$$

ou seja:

- $\vec{A}$  é a projeção de  $\vec{f}$  num sub-espaço  $V$ , com uma base de  $\frac{n}{2}$  vetores;
- $\vec{D}$  é a projeção de  $\vec{f}$  num sub-espaço  $W$ , com uma base de  $\frac{n}{2}$  vetores;
- $V \perp W \leftrightarrow \vec{A} \perp \vec{D}$  ;
- $\vec{v}_i \perp \vec{w}_i \leftrightarrow \langle \vec{v}_i, \vec{w}_i \rangle = 0$  .

Como o processo acima consiste em uma primeira decomposição do sinal, ou seja, decomposição de nível 1, o vetor  $\vec{A}$  (aproximação) é novamente decomposto na soma de dois outros vetores ortogonais, podendo, esse processo, como foi mencionado, ser repetido  $\frac{\log(n)}{\log(2)}$  vezes. Assim, generalizando, para uma decomposição de nível  $j$ , temos:

$$\vec{f} = \vec{A}_j + \sum_{i=1}^j \vec{D}_i. \quad (2.9)$$

sendo que:

- $\vec{A}_j$  é a projeção de  $\vec{f}$  num subespaço  $V_j$ , com uma base contendo  $\frac{n}{2^j}$  vetores;
- $\vec{D}_i$  é a projeção de  $\vec{f}$  num subespaço  $W_i$ , com uma base contendo  $\frac{n}{2^i}$  vetores;
- $V_j \perp W_j \leftrightarrow \vec{A}_j \perp \vec{D}_j$  ;
- $v_{i,j}^{\vec{v}} \perp w_{i,j}^{\vec{w}} \leftrightarrow \langle v_{i,j}^{\vec{v}}, w_{i,j}^{\vec{w}} \rangle = 0$  .

O processo anterior equivale a escrever (STRANG; NGUYEN, 1996):

$$f[n] = \sum_{k=0}^{\frac{n}{2^j}-1} H_{j,k}[n] \phi_{j,k}[n] + \sum_{t=1}^j \sum_{k=0}^{\frac{n}{2^j}-1} G_{t,k}[n] \psi_{t,k}[n] \quad (2.10)$$

onde

- $\phi[n]$  e  $\psi[n]$  formam uma base de Riesz (STRANG; NGUYEN, 1996) para escrever  $\vec{f}$ ;
- $\phi[n] = \sum_k h_n \phi[2n - k]$ ;
- $\psi[n] = \sum_k g_n \phi[2n - k]$ ;
- $H_{j,k}[n] = \langle f, \phi_{j,k}[n] \rangle$ ;
- $G_{t,k}[n] = \langle f, \psi_{t,k}[n] \rangle$ ;
- $\{0\} \leftarrow \dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \rightarrow L^2$ ;
- se  $f[n] \in V_j \rightarrow f[2n] \in V_{j+1}$ ;
- $V_{j+1} = V_j \oplus W_j$ ;
- os coeficientes  $h_k$  correspondem ao filtro passa-baixas (aproximação);
- os coeficientes  $g_k$  correspondem ao filtro passa-altas (detalhe);
- $h[\cdot]$  e  $g[\cdot]$ , que são chamados filtros de análise, formam um par de *Quadrature Mirror Filters* (QMF);
- um filtro com  $k$  coeficientes é dito filtro de suporte  $k$ .

A função  $\psi$  é definida recursivamente, chamada função *wavelet* e é ortogonal à função *scaling* ( $\phi$ ), que é definida recursivamente por dilatações e translações de si mesma (STRANG; NGUYEN, 1996).

Cada par de filtros de análise,  $h[\cdot]$  e  $g[\cdot]$ , possui uma única função *scaling* e uma única função *wavelet* associadas. A forma de obtenção dessas funções a partir dos filtros é detalhada por Jensen e Cour-Harbo (2001).

Existem outros filtros associados com  $h[\cdot]$  e  $g[\cdot]$  chamados *filtros de síntese*. Representados por  $\bar{h}[\cdot]$  e  $\bar{g}[\cdot]$ , são utilizados para inverter a transformada, permitindo recuperar o sinal original a partir de um sinal transformado (JÚNIOR, 2007). Tais filtros obedecem as relações das Equações 2.11, 2.12 e 2.13, para  $k = 0, \dots, n - 1$ , que ficam mais claras por meio do exemplo na Figura 5, para filtros de suporte 4.

$$g_k = (-1)^k h_{N-k-1} \quad , \quad (2.11)$$

$$\bar{h}_k = h_{N-k-1} \quad , \quad (2.12)$$

$$\bar{g}_k = (-1)^{k+1} h_k \quad . \quad (2.13)$$

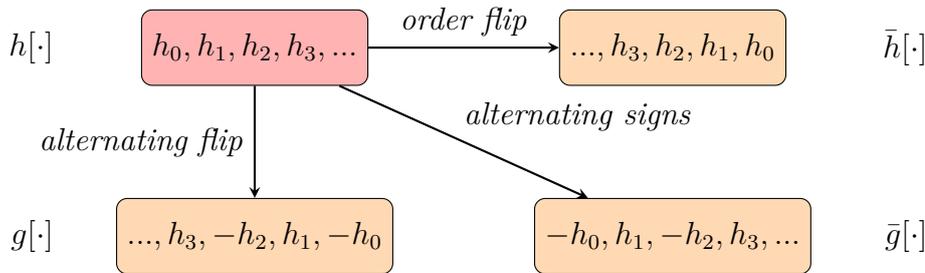


Figura 5 – Relação entre os filtros de análise e síntese (adaptado de (JÚNIOR, 2007))

No presente trabalho, a inversão da DWT não se faz necessária, entretanto, o algoritmo proposto utiliza apenas filtros de reconstrução perfeita, chamados de *Perfect Reconstruction Filter Bank* (PRFB) (WILLIAMS; AMARATUNGA, 1994; AKAY, 1998). Um filtro pode ser considerado PRFB quando  $h[\cdot]$ ,  $g[\cdot]$ ,  $\bar{h}[\cdot]$ , e  $\bar{g}[\cdot]$  mantêm as relações descritas nas Equações 2.11, 2.12 e 2.13 acima, ou seja, as condições de *anti-aliasing* e *no-distortion*, no domínio  $Z$ , representadas nas Equações 2.14 e 2.15, respectivamente, são satisfeitas.

$$\bar{H}[z] = G[-z] \quad , \quad \bar{G}[z] = -H[-z] \quad . \quad (2.14)$$

$$\bar{H}[z]H[z] + \bar{G}[z]G[z] = 2z^{-N+1} \quad . \quad (2.15)$$

### 2.3.1 Cálculo da DWT

Segundo Júnior (2007), para calcular a DWT de um sinal, é possível empregar, com os filtros  $h[\cdot]$  e  $g[\cdot]$ , o algoritmo de Mallat (JENSEN; COUR-HARBO, 2001). O procedimento de cálculo envolve apenas a multiplicação de duas matrizes para cada

nível de transformação. Se  $A[\cdot][\cdot]$  é a matriz de coeficientes dos filtros e  $B[\cdot]$  é o sinal original, então  $C[\cdot] = A[\cdot][\cdot]B[\cdot]$  corresponde ao sinal transformado, sendo a disposição dos coeficientes nas matrizes a seguinte:

$$A[\cdot][\cdot] = \begin{pmatrix} h_0 & h_1 & h_2 & \dots & \dots & \dots & h_{n-1} & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ g_0 & g_1 & g_2 & \dots & \dots & \dots & g_{n-1} & 0 & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & h_0 & h_1 & h_2 & \dots & \dots & h_{n-1} & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ 0 & 0 & g_0 & g_1 & g_2 & \dots & \dots & g_{n-1} & 0 & 0 & 0 & \dots & \dots & 0 & 0 \\ \dots & \dots \\ \vdots & \vdots \\ h_{n-1} & 0 & 0 & 0 & \dots & \dots & 0 & 0 & h_0 & h_1 & \dots & \dots & h_{n-3} & h_{n-2} \\ g_{n-1} & 0 & 0 & 0 & \dots & \dots & 0 & 0 & g_0 & g_1 & \dots & \dots & g_{n-3} & g_{n-2} \end{pmatrix}, \quad (2.16)$$

$$B[\cdot] = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ \vdots \\ \vdots \\ b_{n-2} \\ b_{n-1} \end{pmatrix} \quad e \quad C[\cdot] = \begin{pmatrix} c_0 \\ c_{\frac{n}{2}} \\ c_1 \\ c_{\frac{n}{2}+1} \\ \vdots \\ \vdots \\ c_{n-1} \\ c_{\frac{n}{2}-1} \end{pmatrix}. \quad (2.17)$$

Pode-se notar na matriz  $A[\cdot][\cdot]$  que dois procedimentos estão embutidos no algoritmo de Mallat: *downsampling* e *wrap-around* (JENSEN; COUR-HARBO, 2001), que consiste em fazer com que os coeficientes excedentes dos filtros ocupem as posições iniciais de cada linha, fazendo com que a DWT mantenha sempre o mesmo número de elementos do sinal original.

### 2.3.2 Famílias de Transformadas *Wavelet*

Existem diversas famílias de filtros *wavelet* (JENSEN; COUR-HARBO, 2001), conforme ilustrado nas Figuras 6, 7 e 8 e na Tabela 2. Estas se diferenciam no suporte dos filtros, assim como nas características de resposta em frequência e fase dos mesmos, o que faz com que as funções  $\phi$  e  $\psi$  também sofram reflexo de tais diferenças.

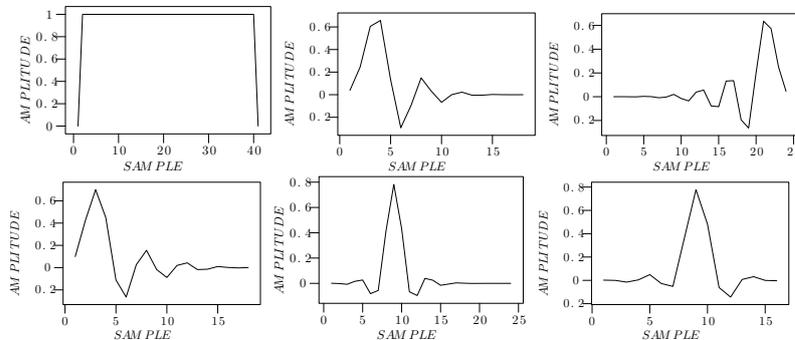


Figura 6 – Formato das respostas ao impulso dos filtros *wavelet* de *Haar*, *Daubechies*, *Vaidyanathan*, *Beylkin*, *Coiflet* e *Symmlet*, respectivamente (extraído de (JÚNIOR, 2007))

Foram utilizadas neste trabalho, para fins de separação dos sinais de áudio em sub-bandas de frequências, as *wavelets* ortogonais de Beylkin, Coiflets, Daubechies, Haar,

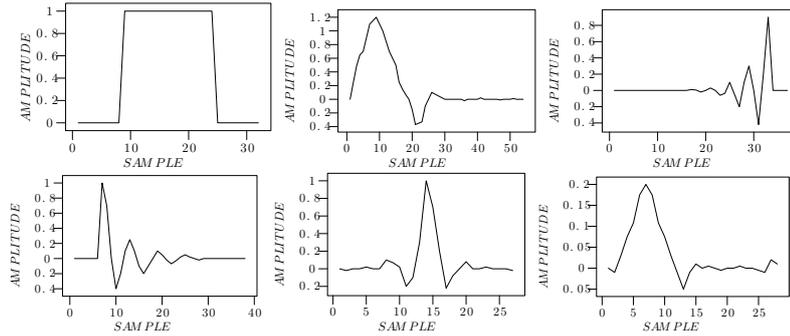


Figura 7 – Formatos das funções *scaling* dos filtros *wavelet* de *Haar*, *Daubechies*, *Vaidyanathan*, *Beylkin*, *Coiflet* e *Symmlet*, respectivamente (extraído de (JÚNIOR, 2007))

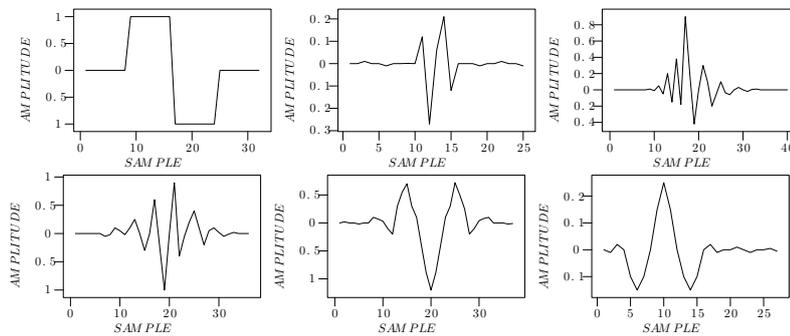


Figura 8 – Formatos das funções *wavelet* dos filtros *wavelet* de *Haar*, *Daubechies*, *Vaidyanathan*, *Beylkin*, *Coiflet* e *Symmlet*, respectivamente (extraído de (JÚNIOR, 2007))

Tabela 2 – Características das famílias de *wavelets* utilizadas no presente trabalho, incluindo a quantidade de momentos da função *wavelet*.

Família	Suporte(n)	Fase	Observação	Momentos
Beylkin	18	não linear	otimizada para áudio em geral	$\frac{n}{2} - 2$
Coiflets	par, múltiplo de 6	quase linear	resposta ao impulso quase simétrica, criada por Ronald Coifman	$\frac{n}{2} - 1$
Daubechies	par, maior que 4	não linear	resposta ao impulso <i>maximally flat</i> , criada por Ingrid Daubechies	$\frac{n}{2}$
Haar	2	linear	é a mais simples das <i>wavelets</i> , criada por Alfred Haar	1
Symmlets	par, múltiplo de 8	não linear	resposta ao impulso mais simétrica	$\frac{n}{2} - 2$
Vaidyanathan	24	não linear	otimizada para voz, criada por P. P. Vaidyanathan	–

Fonte: (JÚNIOR, 2007).

Symmlets e Vaidyanathan com diversos suportes, a exemplo de Júnior (2007) e Souza

(2010).

É importante salientar que todas as famílias aqui empregadas constituem-se de filtros FIR (*Finite Impulse Response*), ou seja, com respostas finitas ao impulso, isto é, limitadas (HAYKIN; VEEN, 2001), sendo que as respostas em frequência se aproximam das ideais à medida que o suporte cresce (SOUZA, 2010).

### 2.3.3 Escala Bark (EB)

Segundo Vieira (2008) e Bosi e Goldberg (2012), a EB permite subdividir a faixa de frequências audíveis pelo ouvido humano em 25 intervalos chamados de bandas críticas. Quando dois ou mais sons, de frequências diferentes mas pertencentes à mesma banda crítica, estão sendo captados simultaneamente, o de maior amplitude mascara o(s) outro(s), portanto, o de menor amplitude não é percebido. A Tabela 3 mostra as bandas críticas, na escala de frequências em Hertz (Hz) e na EB ( $z$ ), que é uma escala adotada para linearizar o intervalo entre as bandas. A conversão entre Hertz e Bark é dada pela Equação 2.18 onde  $f$  representa a frequência em Hertz. Nos casos onde  $z < 2$  então aplica-se a correção  $z \leftarrow z + 0.15(2 - z)$  e onde  $z > 20$  então faz-se  $z \leftarrow z + 0.22(z - 20.1)$ .

$$z = \frac{26.81f}{1960 + f} - 0.53 \quad (2.18)$$

Tabela 3 – As 25 bandas críticas do sistema auditivo humano que são aproveitadas no presente trabalho para obtenção das energias predominantes do sinal.

<i>Bark</i> ( $z$ )	<i>Hertz</i> (Hz)	<i>Bark</i> ( $z$ )	<i>Hertz</i> (Hz)
0	0 - 100	12	1720 - 2000
1	100 - 200	13	2000 - 2320
2	200 - 300	14	2320 - 2700
3	300 - 400	15	2700 - 3150
4	400 - 510	16	3150 - 3700
5	510 - 630	17	3700 - 4400
6	630 - 770	18	4400 - 5300
7	770 - 920	19	5300 - 6400
8	920 - 1080	20	6400 - 7700
9	1080 - 1270	21	7700 - 9500
10	1270 - 1480	22	9500 - 12000
11	1480 - 1720	23	12000 - 15500
		24	15500 - 22050

Fonte: Elaborada pelo autor.

Neste trabalho, a EB é utilizada para simular o funcionamento da audição humana, tendo em vista que, se ela é capaz de identificar os padrões acústicos de interesse, o seu uso no sistema eletrônico desenvolvido deve ser benéfico.



### 3 A ABORDAGEM PROPOSTA

*Como visto, um dos principais problemas dos agricultores é o manejo de insetos-praga. No caso da cultura do café, as cigarras se mostram como um importante entrave, resultando em prejuízos consideráveis todos os anos. O uso de defensivos agrícolas é uma das soluções para o referido manejo, entretanto, antes de aplicá-los, é desejável fazer o monitoramento do problema a fim de estimar a área de incidência e dosar, de forma otimizada, o defensivo ou tomar a melhor decisão em relação ao controle e manejo. Para tal, a EPC é utilizada para selecionar características extraídas no domínio conjunto wavelet-Bark, possibilitando um nível de otimização que permitiu reconhecer os padrões de interesse com base em um classificador SVM clássico.*

#### 3.1 A Estrutura do Sistema

O objetivo do sistema é escolher entre duas possíveis respostas para cada sinal adquirido: “cigarra”, isto é, o sinal contém som de cigarra, ou “ruído”. Um sistema como esse pode ser utilizado para auxiliar o administrador da lavoura na tomada de decisão em relação ao manejo desses insetos-praga, uma vez que permitirá que seja feito um mapeamento automático da área, substituindo o procedimento de monitoramento da lavoura realizado de forma manual com dispêndio de trabalho braçal de humanos.

O sistema implementado foi dividido em quatro módulos distintos, descritos ao final deste capítulo:

- **Primeiro módulo (MPP):** o primeiro módulo é responsável pelo pré-processamento (MPP) do banco de arquivos de áudio coletados em uma lavoura de café por um especialista. Os sinais foram obtidos por meio de gravadores digitais e *smartphones*, amostrados a 44.100 Hz, com 16 bits de quantização. De acordo com Sueur (2002) a espécie *Quesada gigas*, praga-chave da lavoura de café, foco do presente estudo, pode emitir sinais com frequência de aproximadamente 2.4KHz. Os sinais aqui empregados apresentaram uma frequência dominante de aproximadamente 1.6KHz, conforme pode-se observar na Figura 9. Estando os sinais amostrados a 44.100 amostras por segundo os sinais empregados no presente estudo satisfazem o teorema de Nyquist (NYQUIST, 1928);
- **Segundo módulo (MEP):** o segundo, corresponde ao módulo de engenharia paraconsistente (MEP), responsável pelo processamento da EPC. Por meio dele será possível analisar a importância das *features* de entrada antes de testar o classificador proposto;

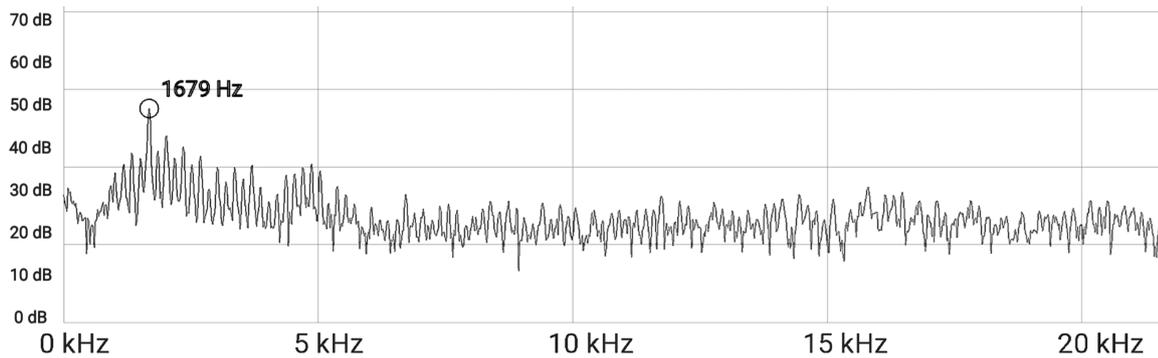


Figura 9 – Espectro de frequências em escala logarítmica de um sinal de *Quesada gigas*.

- **Terceiro módulo (MTT):** O terceiro módulo é responsável por realizar os treinamentos e testes (MTT) a partir dos dados obtidos na etapa de pré-processamento. Assim, uma vez definidas as configurações do sistema, ou seja, o arquivo resultante da etapa de pré-processamento (elegendo assim o filtro, suporte e nível de decomposição desejados), o conjunto de parâmetros do classificador e o limiar percentual de divisão dos arquivos de treinamento e teste, esse módulo realiza um conjunto de treinamentos e testes, percorrendo o conjunto de configurações e o limiar de divisão dos arquivos, aplicando divisões aleatórias entre as amostras em cada etapa, objetivando o melhor resultado possível, imprimindo o correspondente valor de acurácia e sua matriz de confusão;
- **Quarto módulo (MAP):** Por último, têm-se o módulo de aplicação (MAP), que será implantado na lavoura, utilizando o classificador treinado com os melhores parâmetros previamente definidos.

Durante as pesquisas iniciais, uma ideia chamou a atenção: o sistema auditivo humano. Isso porque, se os humanos conseguem facilmente diferenciar os padrões acústicos em questão assim que os ouvem, um algoritmo imitando a estrutura acústico-biológica provavelmente também obtenha sucesso. Assim, atenção foi dedicada para a Escala Bark (EB) (VIEIRA, 2008; BOSI; GOLDBERG, 2012; ISHIZAWA, 2015) e a possibilidade de implementá-la similarmente à cóclea no ouvido humano com base na DWTP, resultando em um método eficiente de mapeamento tempo-frequência (GUIDO, 2017).

Assim, a justificativa da escolha do uso da EB está na investigação da hipótese de imitação do ouvido humano na aquisição do áudio para o presente fim, dividindo as faixas de frequência de cada sinal obtido na lavoura (cigarra ou ruído) como índices de sua escala. Para isso, um passo anterior é o processamento digital dos referidos sinais com a DWTP, ao qual foi escolhida por ser uma técnica amplamente utilizada em pesquisas modernas envolvendo processamento digital de sinais (GUIDO, 2017; GUIDO, 2021; WANG *et al.*, 2020; CAOBIANCO; GUIDO; SILVA, 2021). Além disso, a possibilidade de analisar o

sinal no domínio tempo-frequência se mostrou uma importante hipótese para investigação.

Visando corroborar a escolha da estrutura do sistema proposto, é relevante destacar a importância da seleção das *features* que alimentam o sistema. Apesar dos esforços dispendidos para a melhoria dos algoritmos de aprendizagem de máquina automatizados do tipo “Deep Learning” e “Feature Learning”, que costumam ter como maior desvantagem o alto custo computacional (MASTELINI *et al.*, 2019), os métodos baseados em extração artesanal (*handcrafted extraction*) ainda vigoram como importantes recursos para solucionar problemas de processamento de sinais em inteligência artificial (GUIDO, 2018b). Dessa forma, a escolha cuidadosa dos atributos de entrada é extremamente relevante para uma classificação precisa.

Se, por um lado, é possível desenvolver algoritmos inteligentes que conseguem classificar padrões a partir de um conjunto de entradas de treinamento, sem maiores exigências de rigor em relação às regras de geração de tais dados de entrada, por outro lado, a qualidade dos vetores de entrada pode resultar em algoritmos mais simples e com menor exigência computacional (GUIDO, 2018a).

Como visto, a EPC permite analisar a viabilidade do uso dos conjuntos de entrada em algoritmos mais simples, pois mensura a separabilidade dos dados e suas respectivas classes, mostrando-se importante recurso para ser implementado no presente trabalho, pois, averigua a viabilidade de uso dos 25 atributos de energia provenientes da EB, e ainda, a hipótese de rendimento ótimo mesmo com redução dessa quantidade de entradas no classificador. Além disso, visando uma futura implementação em IoT, pode ser interessante a redução do fluxo de dados transmitidos entre a malha de nós espalhados na lavoura, caso se escolha por transmitir as *features* para processamento em nuvem.

Assim sendo, a justificativa pela escolha da EPC como técnica de seleção de características em diverso a outras opções comuns na literatura, como PCA e ICA, além do crescente interesse do meio científico (FONSECA *et al.*, 2020; WANG *et al.*, 2020; CAOBIANCO; GUIDO; SILVA, 2021) e por se tratar se uma técnica moderna de seleção de características (YAO *et al.*, 2020), está na inexistência, na literatura de trabalho empregando-a para o mesmo fim.

Como classificador do sistema, foi utilizada a Máquina de Vetores de Suporte (*Support Vector Machine* - SVM) (CORTES; VAPNIK, 1995), que constitui um método supervisionado de aprendizagem computacional, muito utilizado em problemas de classificação de padrões (GHOLAMI; FAKHARI, 2017; WEI *et al.*, 2020; JR *et al.*, 2021).

Na Figura 10 está definida a arquitetura da SVM, cuja camada de entrada é composta pelo número de *features* de entrada e que, neste caso, corresponde ao conjunto de 25 energias provenientes de cada uma das bandas da EB. Nota-se, ainda, a camada oculta, onde são ajustados os coeficientes do hiperplano durante o processo de treinamento,

sendo  $X$  correspondente ao número de casos de treinamento, e camada de saída, onde o resultado binário é apresentado de acordo com a classe resultante.

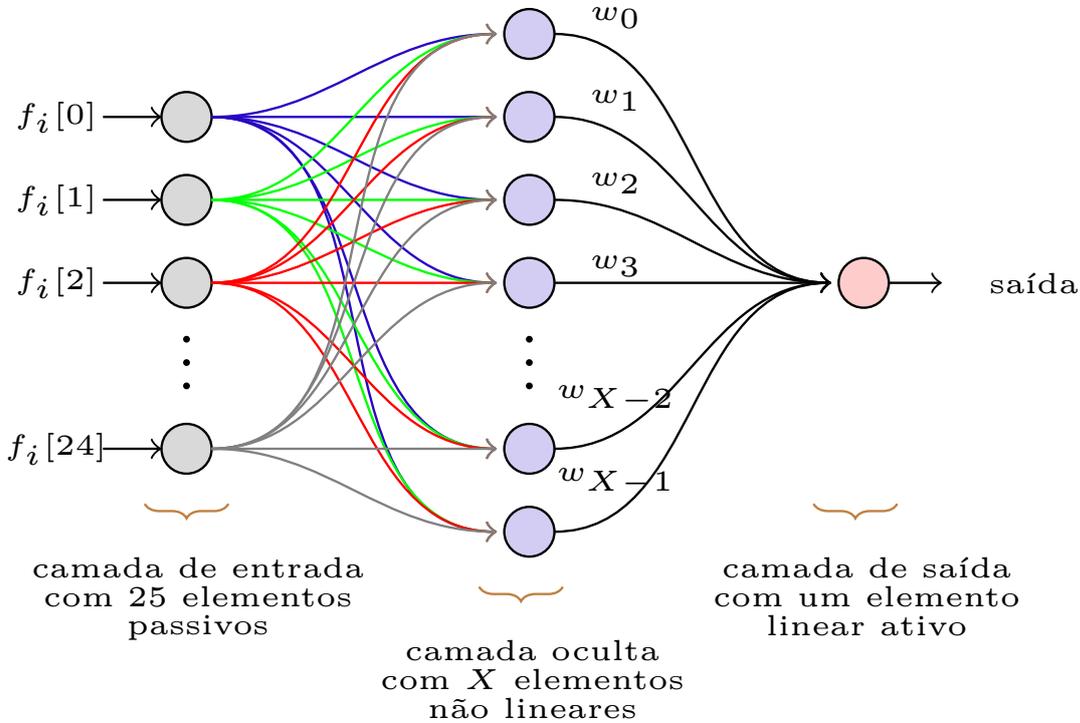


Figura 10 – Estrutura da SVM. Os coeficientes do hiperplano, determinados durante a fase de treinamento supervisionado são  $\{w_0, w_1, \dots, w_{X-1}\}$ . O elemento de saída é resultado da combinação linear das saídas da camada oculta com os referidos coeficientes.

O módulo MTT recebe o conjunto de *features* dos arquivos de áudio capturados na lavoura e o rótulo que identifica cada conjunto de amostras. Dessa forma, esse módulo executa o treinamento da SVM, fixando o valor  $-1$  para a saída de cada arquivo pertencente à classe “cigarra” e o valor  $1$  para a classe “ruído”. A metodologia de treinamento consiste em criar um plano separador, o mais largo possível, como ocorre tradicionalmente.

No caso do presente trabalho, foram testados os *kernels* Linear (Equação 3.1), Gaussiano (Equação 3.2), Polinomial (Equação 3.3) e Sigmoidal (Equação 3.4), sendo *dist* a distância Euclidiana (DENG; O’SHAUGHNESSY, 2018) entre os vetores  $\vec{x}$  e  $\vec{y}$  de teste e de treinamento.

$$\phi(v) = (\vec{x} \cdot \vec{y}), \quad (3.1)$$

$$\phi(v) = \exp(-\gamma \|\vec{x} - \vec{y}\|^2), \quad (3.2)$$

$$\phi(v) = (\gamma \cdot (\vec{x} \cdot \vec{y}) + r)^d, \quad (3.3)$$

$$\phi(v) = \tanh(\gamma \cdot (\vec{x} \cdot \vec{y}) + r), \quad (3.4)$$

Durante o processo de treinamento, pode-se ajustar parâmetros buscando melhorar a acurácia do sistema para a entrada  $\vec{x}$  em relação ao padrão  $\vec{y}$ . O parâmetro de suavização  $C$  é uma constante que permite ajustar uma margem suave (*soft margin*) de separabilidade a fim de evitar possíveis problemas causados por ruídos em amostras, *outliers* ou amostras mal rotuladas durante o processo de separação. Trata-se de uma forma de controle de sobreajuste, sendo as amostras excessivamente próximas à margem são ignoradas.

Os parâmetros  $d$  (*degree*) no *kernel* Polinomial e  $r$  (*deslocamento*), nos *kernels* Polinomial (Equação 3.3) e Sigmoidal (Equação 3.4), permitem ajustar a flexibilidade e o deslocamento da fronteira de decisão, respectivamente, mostrando-se artifícios em busca da melhoria da separabilidade entre as amostras e, conseqüentemente, o aumento da acurácia do sistema.

Outro parâmetro importante que pode ser utilizado nos *kernels* não lineares é o  $\gamma$ , que permite ajustar a flexibilidade da fronteira de decisão. Valores muito pequenos implicam em uma fronteira de decisão quase linear, entretanto valores muito elevados tendem a causar um sobreajuste (*overfitting*) no classificador.

O módulo MAP recebe um sinal de áudio desconhecido e extrai os parâmetros correspondentes, que são os mesmos definidos no módulo MTT. Finalmente, obtém-se o resultado de acordo com o valor de saída apresentado pela SVM, sendo considerado “cigarra” em caso de resposta negativa  $y < 0$ , ou seja, existe canto de cigarra no arquivo de áudio submetido; ou “ruído” em caso de resposta positiva  $y > 0$ , ou seja, não existe canto de cigarra no arquivo submetido. A Figura 11 contém a a estrutura completa do sistema proposto (módulo MAP).

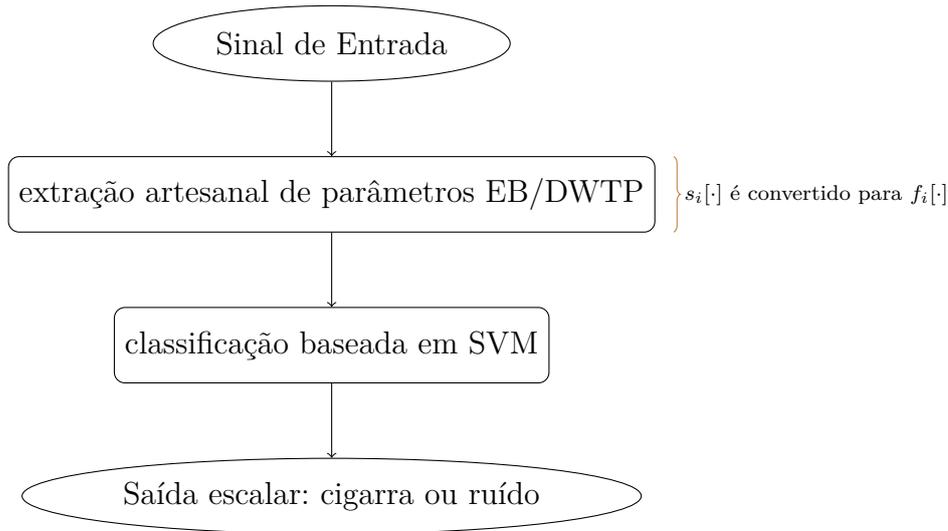


Figura 11 – Estrutura do sistema proposto, no qual cada sinal acústico  $s_i[\cdot]$  é convertido para o seu vetor de características  $f_i[\cdot]$ .

Nas próximas subseções, detalham-se os algoritmos utilizados nos módulos propostos. Os parâmetros mencionados acima, tanto para o treinamento, quanto para teste e aplicação do classificador, serão discutidos no próximo Capítulo.

## 3.2 Implementação do Sistema

Nesta seção, apresentam-se os aspectos técnicos de implementação do sistema, considerando a estrutura detalhada na seção anterior, com o objetivo de corroborar a suposição sustentada no presente trabalho e possibilitar a realização dos referidos testes que serão apresentados no próximo capítulo.

### 3.2.1 Módulo de Pré-Processamento (MPP)

Como já foi mencionado, MPP é o módulo responsável pelo pré-processamento dos dados antes de serem submetidos à etapa de teste. Esse módulo recebe como parâmetro o filtro wavelet (A) e o nível de decomposição desejados (B) e tem por objetivo processar, por meio dos  $X$  arquivos coletados na lavoura de café, os sinais de entrada de áudio extraindo seus dados brutos (Passo 3), efetuar a DWTP do sinal (Passo 4.1), e por fim, calcular as energias do sinal a partir da EB (Passo 4.2).

A seguir, são descritos os passos para implementação do algoritmo proposto para MPP, ilustrados na Figura 12:

- **INÍCIO**
- **MPP (PASSO 1):** (rotina LPE - Leitura dos Parâmetros de Entrada): efetua o processamento dos argumentos de entrada A e B;

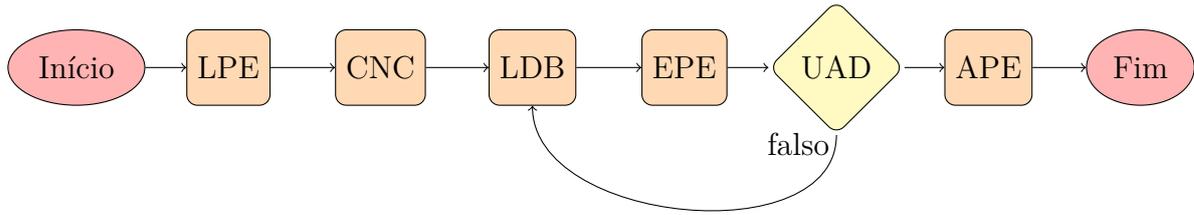


Figura 12 – Estrutura básica proposta para o módulo MPP. As siglas estão explicadas ao longo do texto.

- **MPP (PASSO 2):** (rotina CNC - Consulta Nome da Classe): recupera o nome da classe a partir do nome do diretório atual, onde os arquivos WAV (representados por  $X$ ) estão armazenados:

– **Repita:**

- \* **MPP (PASSO 3):** (rotina LDB - Leitura dos Dados Brutos): extrai os “dados brutos” (amostras digitalizadas) do arquivo atual. Os arquivos de áudio no formato WAV foram capturados com uma taxa de amostragem de 44.100 amostras por segundo e 16 bits de resolução por amostra (BOSI; GOLDBERG, 2012). Esses dados extraídos compõem um vetor  $s_i[\cdot]$ , onde  $(0 \leq i \leq X - 1)$ . Conforme mencionado, uma vez que a DWTP (GUIDO, 2017) é utilizada no próximo passo, é necessário anteriormente executar a técnica *zero-padding*, que consiste em incrementar o tamanho do vetor para que se torne uma potência de dois. Assim, no caso de um arquivo de áudio WAV de 5 segundos, tem-se um vetor de tamanho 220500 que é ajustado para 262144, ou seja,  $44100 \cdot 5 = 220500 \Rightarrow \frac{\log(220500)}{\log(2)} = 17.75$ , sendo o próximo inteiro após 17.75 o número 18, então  $2^{18} = 262144$ ;
- \* **MPP (PASSO 4):** (rotina EPE - Extração dos Parâmetros de Energia): obtém-se  $f_i[\cdot]$  de tamanho 25, onde  $(0 \leq i \leq X - 1)$ ;

· **INÍCIO**

- **PASSO 4.1:** processamento da transformada DWTP do nível de decomposição  $B$ , isto é,  $N_{max} = \frac{\log(262144)}{\log(2)} = 18$  para nível máximo ou  $N_{med} = \frac{N_{max}}{2} = 9$ , para nível médio;
- **PASSO 4.2:** extraem-se as 25 energias de cada banda Bark a partir da Equação 3.5, escolhendo-se as folhas da decomposição que oferecem os limites de frequências referentes à cada banda de interesse;

$$E_z(x) = \sum_{i=a}^b x_i^2, \quad (3.5)$$

onde

$$a = \frac{BS_i(z)}{r}, b = \frac{BS_f(z)}{r} \quad (3.6)$$

e

$$r = \frac{44100}{2^B} \quad (3.7)$$

- **PASSO 4.3:** normaliza-se o conjunto de 25 energias, para evitar variações indevidas às amplitudes, concatenando-as em um vetor único  $f_i[\cdot]$  que representará os parâmetros de entrada da classe atual para o referido classificador.
  - **FIM;**
- **Enquanto** ainda existem arquivos WAV da base de dados ( $X$ ) (rotina UAD - Último Arquivo de Dados);
- **MPP (PASSO 5):** (rotina APE - Armazena Parâmetros de Energia), armazena os parâmetros de energia em um arquivo texto no padrão ASCII “separado por vírgulas”.
  - **FIM.**

### 3.2.2 Módulo de Engenharia Paraconsistente de Características (MEP)

Uma vez que foram obtidos os vetores de entrada, empregou-se a EPC para analisar o comportamento do conjunto de entradas obtido. Nessa técnica, os vetores de entrada são representados como um ponto  $P$  no plano paraconsistente, conforme descrito anteriormente. Quanto mais próximo do canto  $(G_1, G_2) = (1, 0)$  estiver o conjunto de entradas, representadas pelo vértice  $P$ , melhor separabilidade haverá entre as classes (Figura 1).

O presente módulo recebe como parâmetro o arquivo resultante do módulo MPP (A), descrito na subseção anterior, ou seja, o conjunto de vetores de entrada resultantes do processamento da DWTP com o filtro e nível de decomposição definidos, representando todas as amostras  $X$  do sistema proposto. Além desse, também devem ser submetidos como parâmetros o número de classes (B) e o número de atributos por classe (C), tendo por objetivo processar o conjunto de entradas analisando a dissimilaridade interclasses e a similaridade intraclasses, conforme descrito no capítulo anterior, plotando o referido gráfico do plano paraconsistente.

A seguir, são descritos os passos para implementação do algoritmo proposto para MEP, ilustrados na Figura 13:

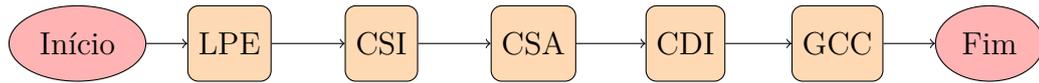


Figura 13 – Estrutura básica proposta para o módulo MEP. As siglas estão explicadas ao longo do texto.

- **INÍCIO**
- **MEP (PASSO 1):** (rotina LPE - Leitura dos Parâmetros de Entrada): efetua o processamento dos argumentos de entrada A, B e C;
- **MEP (PASSO 2):** (rotina CSI - Cálculo de Sobreposições Interclasses): representada por  $R$ , computa o número de sobreposições entre as classes, ou seja, averigua se as classes são linearmente separáveis;
- **MEP (PASSO 3):** (rotina CSA - Cálculo de Similaridade Intraclases): representada por  $\alpha$ , calcula o nível de similaridade entre as classes, conforme detalhado no capítulo anterior;
- **MEP (PASSO 4):** (rotina CDI - Cálculo de Dissimilaridade Interclasses): representada por  $\beta$ , calcula o nível de dissimilaridade entre as classes, conforme detalhado no capítulo anterior;
- **MEP (PASSO 5):** (rotina GCC - Graus de Certeza e Contradição): calcula os graus de certeza e de contradição, representados por  $G_1$  e  $G_2$ , respectivamente, plotando o ponto  $P$  resultante;
- **FIM.**

### 3.2.3 Módulo de Treinamento e Teste (MTT)

Em posse dos vetores de entrada, além de ser possível realizar a análise do conjunto de atributos de entrada por meio da aplicação da EPC, conforme visto na subseção anterior, também é desejável realizar os testes do referido conjunto de entradas em um classificador a fim de verificar sua efetividade para a aplicação proposta.

O presente módulo recebe como parâmetro o arquivo resultante do módulo MPP (A), descrito na subseção 3.2.1, o número de atributos por classe (B) a serem empregados no classificador, e o limiar de treinamento (C), ou seja, o percentual mínimo e máximo de amostras que serão utilizadas em busca do melhor resultado.

Uma vez iniciado o processamento do módulo MTT a partir dos parâmetros iniciais (Passo 1), o sistema inicia um conjunto de iterações, conferindo o limiar de separação (Passo 2), separando as amostras para treinamento e teste, de acordo com o limiar atual (Passo 3). Utilizando a técnica de *Grid Search*, o sistema realiza a escolha dos melhores

hiperparâmetros através de validação cruzada com as amostras de treinamento (Passo 4) e, em seguida o classificador é treinado (Passo 5) com todas as amostras de treinamento e, por fim, é realizado o teste do conjunto de amostras de testes com o classificador previamente treinado (Passo 6) e o resultado é apresentado (Passo 7).

A seguir, descrevem-se os passos para implementação do algoritmo proposto para MTT, ilustrados na Figura 14:

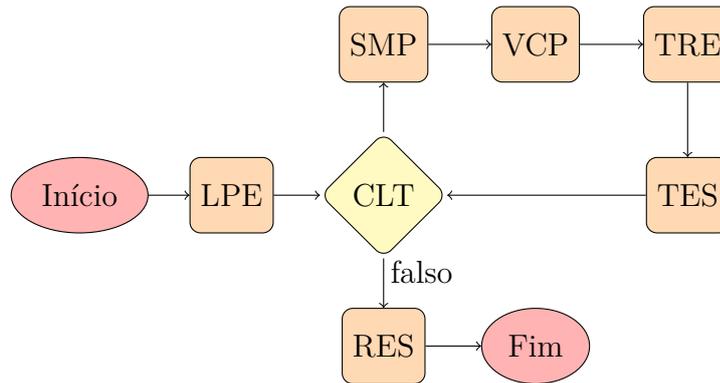


Figura 14 – Estrutura básica proposta para o módulo MTT. As siglas estão explicadas ao longo do texto.

- **INÍCIO**
- **MTT (PASSO 1):** (rotina LPE - Leitura dos Parâmetros de Entrada): efetua o processamento dos argumentos de entrada A, B e C;
- **MTT (PASSO 2):** (rotina CLT - Consulta Limiar de Treinamento): Verifica se a iteração atual está entre o valor mínimo e máximo do limiar de treinamento;
- **MTT (PASSO 3):** (rotina SMP - Separação da Matriz de Parâmetros): Realiza a separação das amostras de treinamento e teste de acordo com o limiar retornado na rotina anterior;
- **MTT (PASSO 4):** (rotina VCP - Verificar Conjunto de Parâmetros): executa a função *GridSearchCV* da biblioteca *scikit-learn*, realizando a pesquisa exaustiva sobre os parâmetros do classificador, conforme Tabela 4, utilizando a técnica *cross-validated grid-search*, elegendo os melhores hiperparâmetros;
- **MTT (PASSO 5):** (rotina TRE - Treinamento): executa o treinamento do classificador com os hiperparâmetros definidos no passo anterior, utilizando o conjunto de treinamento definido no passo 3;
- **MTT (PASSO 6):** (rotina TES - Teste): executa os testes do classificador, ajustado com os melhores hiperparâmetros encontrados no passo anterior, com o conjunto de amostras de teste definidas no passo 3;

- **MTT (PASSO 7):** (rotina RES - Resultado): ao final do processo, apresenta o melhor resultado, limiar correspondente e os hiperparâmetros empregados no melhor resultado;
- **FIM.**

Tabela 4 – Parâmetros utilizados na bateria de testes em MTT

atributos de ajuste	parâmetros
<i>kernel</i>	Linear, Gaussiano, Polinomial, Sigmoidal
<i>C</i>	1, 10, 100, 1000
$\gamma$	1e-3, 1e-4

Fonte: Elaborada pelo autor.

Os parâmetros de ajuste “grau do polinômio” e “deslocamento”, dos *kernels* Polinomial e Sigmoidal (Equações 3.3 e 3.4, respectivamente), empregaram os valores *default* da biblioteca *scikit-learn*, sendo eles:  $d = 3$  e  $r = 0$ .

### 3.2.4 Módulo de Aplicação (MAP)

O presente módulo foi desenvolvido para ser implantado em lavoura de café, após escolha dos hiperparâmetros e treinamento do classificador, ou como sistema final de testes de arquivos de áudio em laboratório.

Inicia-se o processamento tendo como entrada um arquivo de áudio WAV ou o sinal obtido pelo microfone do dispositivo (Passo 1), em seguida extraíndo as energias, conforme detalhado nas seções anteriores (Passo 2). Uma vez que os coeficientes de entrada estão disponíveis, o sistema realiza o teste utilizando a SVM previamente treinada (Passo 3), e apresenta o resultado (Passo 4).

A seguir, são descritos os passos para implementação do algoritmo proposto para MAP, ilustrados na Figura 15:

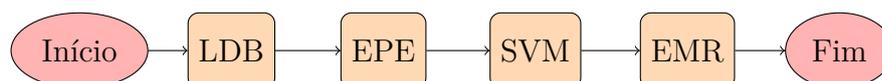


Figura 15 – Estrutura básica proposta para o módulo MAP. As siglas estão explicadas ao longo do texto.

- **INÍCIO**
- **MAP (PASSO 1):** executa a rotina LDB, descrita na subseção 3.2.1, a fim de obter  $s_i[\cdot]$ , ou seja, os dados brutos da amostra a ser classificada;

- **MAP (PASSO 2):** executa a rotina EPE, descrita na subseção 3.2.1, a fim de obter  $f_i[\cdot]$ , ou seja, as energias da amostra a ser classificada;
- **MAP (PASSO 3):** (rotina SVM): injetar os parâmetros extraídos no passo anterior, previamente normalizados, na SVM, obtendo o valor de saída;
- **MAP (PASSO 4):** (rotina EMR - Escolha da Melhor Resposta): a partir do resultado da SVM, o rótulo da classe eleita é apresentado.
- **FIM.**

### 3.3 Detalhes Adicionais

No Capítulo anterior foram apresentados diversos trabalhos encontrados durante o levantamento bibliográfico sobre o tema de monitoramento de cigarras. São estudos envolvendo PDS de áudio, utilização de dispositivos tecnológicos, *softwares* em campo e em laboratório.

Conforme discutido, todos os trabalhos estudados que empregam *software* em laboratório o fazem para análise dos sinais acústicos após terem sido capturados manualmente em campo por meio de equipamentos gravadores e microfones.

Ainda considerando os trabalhos levantados no capítulo anterior, é notável que apenas Aide *et al.* (2013) e Prince *et al.* (2019) utilizam *softwares* em campo, cujo foco está no armazenamento permanente. Em ambos os casos a classificação é realizada em laboratório, sendo que o primeiro envia cada sinal capturado a um servidor remoto e no segundo os sinais são armazenados localmente e coletados manualmente.

Assim, considerando que todos os trabalhos levantados no Capítulo anterior utilizam trabalho manual, o sistema proposto deve se mostrar superior por ser um sistema automatizado de monitoramento, eliminando a necessidade de trabalho manual no processo de captura e classificação dos sinais.

O presente Capítulo descreveu a técnica proposta para o presente trabalho, detalhando os módulos desenvolvidos e as rotinas que os compõem. O próximo Capítulo contém a descrição dos testes e resultados obtidos com o sistema implementado.

## 4 TESTES E RESULTADOS

*Este capítulo está dividido em Materiais e Métodos; Bateria de testes, com quatro subgrupos de testes (GR1, GR2, GR3 e GR4); Protótipo em IoT, onde apresenta-se um protótipo de dispositivo com software embarcado para implantação do sistema proposto e, finalmente, uma breve discussão dos resultados.*

### 4.1 Procedimentos Gerais

Neste capítulo apresentam-se os testes realizados em laboratório objetivando primeiramente:

- 1) Analisar a viabilidade dos atributos de entrada a serem empregados no classificador utilizando a técnica EPC previamente explanada;
- 2) Verificar, ainda via EPC, se um número menor de atributos de entrada pode prover bons resultados, permitindo o uso de um classificador menos robusto (PMC), a fim de simplificar o processo;
- 3) Mensurar a acurácia do classificador adotado no sistema proposto, uma SVM, como classificador binário em um problema de detecção de cigarras;
- 4) Mensurar a acurácia do classificador em modo multiclases, para um problema de estimativa de densidade de cigarras.

Desse modo, conforme mencionado nos capítulos anteriores, os atributos de entrada, também chamados de *features*, são os dados básicos iniciais que são submetidos ao classificador, que após processados resultam em um escalar que é interpretado. No caso deste trabalho, os referidos atributos de entrada são as energias provenientes da EB e o classificador é uma SVM. Para obtenção das referidas energias, a DWTP correspondente é processada via convolução com um filtro *wavelet* em nível médio ou nível máximo. O primeiro provê melhor resolução no tempo e o segundo melhor resolução em frequência, sendo este mais custoso computacionalmente por executar o dobro de iterações em relação ao primeiro.

### 4.2 Materiais e Métodos

Os sinais de áudio utilizados nos testes foram capturados em lavouras de café dos estados de São Paulo e Minas Gerais, em visitas diversas no período de 2016 a 2019. Foram capturados por meio de gravadores digitais e de *smartphones* com 44.100 amostras por segundo em 16 bits de resolução por amostra.

Para realização dos testes foi utilizado um microcomputador com processador Intel(R) Core(TM) i7-4800MQ CPU @ 2.70GHz, 32GB de memória RAM e disco rígido SSHD de 1TB em ambiente Linux.

Os algoritmos foram desenvolvidos em módulos específicos para as diversas tarefas necessárias ao sistema e estão descritas nas subseções abaixo:

#### 4.2.1 Módulo de Pré-Processamento - MPP

Conforme detalhado no capítulo anterior, esse módulo é responsável por receber os arquivos WAV e, a partir de seus dados brutos, executar a DWTP via convolução com um filtro *wavelet* definido. A partir do sinal transformado, as 25 energias provenientes de cada sub-banda da EB são armazenadas em um vetor empregado como recurso de entrada nos testes de EPC e nas SVMs de treinamento e aplicação.

Em um primeiro momento, este módulo foi desenvolvido em linguagem de programação Python (LUNDH, 2001), versão 3.6 em ambiente Linux e implementa as rotinas LDB e EPE detalhadas no Capítulo anterior. Para o desenvolvimento do algoritmo, os seguintes módulos Python foram empregados:

- O módulo “wave”, incorporado à biblioteca padrão da linguagem Python, permite a leitura de arquivos WAV e acesso aos atributos de cabeçalho, como número de canais, resolução e taxa de amostragem;
- A biblioteca “PyWavelets” (LEE *et al.*, 2019) permite a manipulação das diversas transformadas *Wavelet* e procedimentos intrínsecos à pesquisa científica utilizando técnicas de processamento digital de sinais;
- A biblioteca “NumPy” (MCKINNEY, 2012), desenvolvida especificamente para computação científica, permite a manipulação de conjuntos de dados de forma otimizada e disponibiliza importantes recursos como vetores de objetos multidimensionais, transformadas e álgebra linear.
- O módulo “pickle” implementa um algoritmo de serialização de objetos Python. Com ele é possível serializar, ou seja, armazenar dados de uma variável ou objeto em memória permanente. Esse recurso aumenta consideravelmente o desempenho geral do módulo, pois, ao invés de re-executar a DWPT a cada teste, o vetor de dados brutos é recuperado após ser inicialmente processado e armazenado pelo processo. A implantação reduziu o tempo de processamento de um único arquivo WAV de 51.84 segundos para 2.14 segundos.

Em um segundo momento, um algoritmo em C++ foi desenvolvido para implementar o MPP, buscando agilidade do processo, uma vez que a transformada é um dos processos mais penosos do presente trabalho;

Com o desenvolvimento do módulo em C++, os resultados das transformadas foram armazenados em arquivos-texto (padrão ASCII), assim sendo, para cada um dos filtros *wavelet* testados, uma única execução de cada arquivo WAV foi necessária na prática, reduzindo o tempo computacional de maneira considerável

Em relação à serialização de dados, um estudo foi desenvolvido a fim de testar o uso dessa técnica em algoritmos de processamento digital de sinais, buscando analisar a eficiência de sua implementação em algoritmos de processamento de dados, incluindo a DWT, empregando linguagem de programação C. Como resultado, o artigo científico (ESCOLA; DOCUSSE, 2020) foi publicado em revista arbitrada.

#### 4.2.2 Módulo de Engenharia Paraconsistente de Características - MEP

O módulo MEP, desenvolvido em linguagem Python, versão 3.6, tem a função de processar os vetores de entrada, formatados para serem empregados em um possível classificador, a fim de verificar sua viabilidade. Os detalhes em relação a essa técnica foram apresentados nos capítulos anteriores.

Para o desenvolvimento desse módulo, foram seguidas as especificações de (GUIDO, 2018a). No algoritmo desenvolvido, buscando melhor visualização dos resultados, incorporou-se a biblioteca de código aberto Matplotlib (HUNTER, 2007) que permite plotar gráficos em algoritmos desenvolvidos em linguagem Python.

#### 4.2.3 Módulo de Treinamento e Teste - MTT

Para o desenvolvimento deste módulo, a exemplo dos módulos apresentados até o momento, foi utilizada a linguagem Python, versão 3.6. Esses módulos utilizam a biblioteca de código aberto *scikit-learn* (PEDREGOSA *et al.*, 2011), que implementa uma série de ferramentas de análise de dados, classificação, regressão e clusterização, incluindo as SVMs.

As entradas do módulo são executadas a partir de arquivos-texto cujo conteúdo foi resultado do processamento de MPP, cada um com um determinado filtro *wavelet* e nível de decomposição. Os arquivos contém diversos registros compostos pelas 25 energias provenientes da EB com o nome da referida classe.

Uma vez selecionado um determinado arquivo de entrada, MTT faz a separação dos dados de acordo com o perfil do teste escolhido, por exemplo pelo número de amostras de entrada, e executa o referido treinamento da SVM.

A etapa de testes é implementada com variação de diversos atributos, dentre eles o *kernel*, valor de  $\gamma$  (quando aplicável) e os parâmetros de suavização, buscando o melhor resultado possível. O módulo MTT retorna a acurácia resultante e sua matriz de confusão, que fazem parte do relatório de classificação do melhor resultado obtido dentre os parâmetros de ajuste (Tabela 4).

### 4.3 Bateria de Testes

Nesta seção, apresentam-se os testes realizados em laboratório a partir dos sinais capturados em lavoura de café. Os testes foram separados em quatro grupos: GR1, GR2, GR3 e GR4 conforme ilustra a Tabela 5. O primeiro e o segundo grupos dispõem de sinais de áudio em arquivos WAV separados nas classes “cigarra” e “ruído”, pois tem a incumbência de treinar um classificador do tipo “detector de cigarras” para lavoura de café. O terceiro e o quarto grupos estão composto de arquivos separados nas classes “alta”, “média” e “baixa”, representando a densidade de indivíduos em lavoura, pois tem a incumbência de treinar um classificador do tipo “estimador de densidade”. Assim, além de detectar se há cigarras em determinada posição da lavoura, também realiza uma estimativa da densidade populacional naquele ponto. Os arquivos foram separados manualmente por um agrônomo especialista, membro do grupo de pesquisa.

Tabela 5 – Grupos de testes.

Grupo	Classes	Nível DWTP
GR1	cigarras e ruído	médio
GR2	cigarras e ruído	máximo
GR3	alta, média e ruído	médio
GR4	alta, média e ruído	máximo

Fonte: Elaborada pelo autor.

#### 4.3.1 Testes do módulo MPP

Conforme descrito no capítulo anterior, este módulo é responsável pelo pré-processamento dos sinais de entrada, obtidos por meio de dispositivos de captura digitais, em formato de arquivos de áudio WAV. Estes foram separados em diretórios a fim de organizar o processamento e obter as *features* que serão empregadas nos módulos subsequentes.

Ao receber um arquivo WAV, o módulo MPP faz a leitura dos dados brutos e processamento da DWTP com o filtro e nível de decomposição definidos, conforme ilustrado na Figura 16.

Dentre as etapas descritas no capítulo anterior, o módulo MPP é um dos que têm maior custo computacional:  $O(n^2)$  (ESCOLA; DOCUSSE, 2020). Assim sendo, dependendo do equipamento disponível e da linguagem de programação empregada, podendo ser necessários desde poucos segundos, até algumas horas para processar um único sinal de 441.000 amostras (44.100 Hz x 10 segundos). Um conjunto de testes foi realizado com diversas linguagens de programação a fim de analisar o tempo de processamento da DWTP e embasar a escolha da melhor linguagem de programação para ser adotada no presente trabalho. Um artigo foi desenvolvido e submetido para revista científica arbitrada.

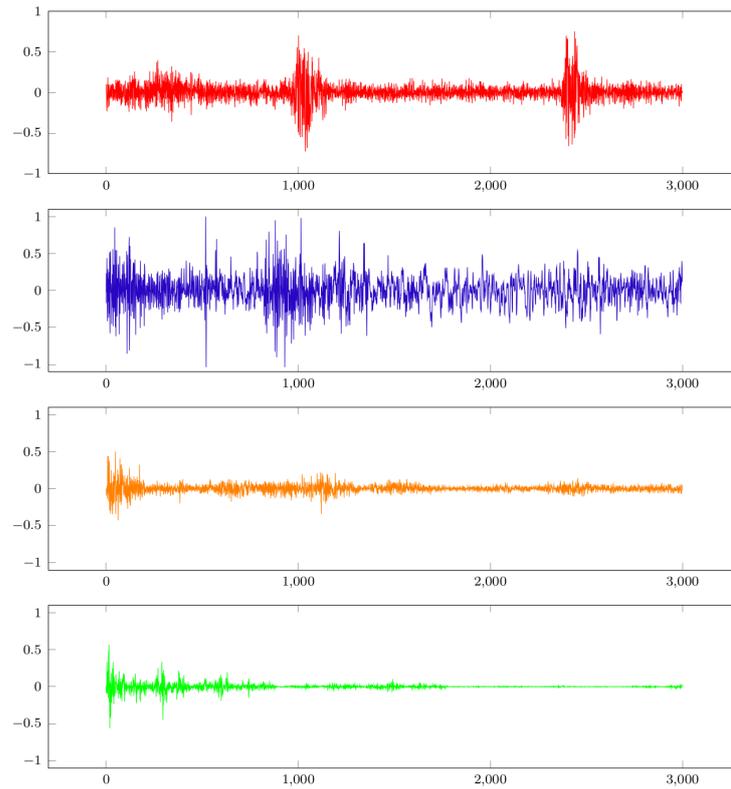


Figura 16 – Exemplos de sinais capturados na lavoura de café: ruído, cigarra, DWTP ruído, DWTP cigarra, respectivamente.

A Figura 17 exibe uma representação das energias dos sinais de cigarra e ruído após processadas com a DWTP, que variam de 0 a 1, partindo do centro do gráfico até sua borda, representando as 25 dimensões de entrada, provenientes da EB, que serão empregadas nos módulos apresentados nas próximas seções. Na Tabela 6 são apresentados os valores dos coeficientes resultantes.

Tabela 6 – Exemplos de coeficientes de energia resultantes do módulo MPP.

<i>Feature</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
cigarra	0,028292	0,029383	0,029362	0,030015	0,031826	0,036751	0,042547	0,043753	0,046845	0,05532	0,061945	0,071075	0,081324
ruído	0,102941	0,086328	0,080748	0,086343	0,091575	0,104249	0,114759	0,112145	0,132017	0,156263	0,171058	0,153671	0,177952
<i>Feature</i>	14	15	16	17	18	19	20	21	22	23	24	25	
cigarra	0,091845	0,108982	0,136941	0,162901	0,205125	0,266861	0,560914	0,594293	0,623496	0,825893	1	0,99842	
ruído	0,212239	0,277644	0,314895	0,372965	0,444860	0,534623	0,760452	1,000000	0,895751	0,760062	0,000000	0,000001	

Fonte: Elaborada pelo autor.

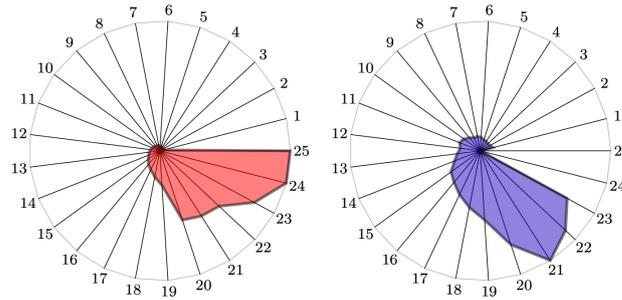


Figura 17 – Exemplos de energias dos sinais de cigarra (esquerda) e ruído (direita).

#### 4.3.2 Testes do módulo MEP

Nesta subseção, apresentam-se os testes realizados com a implementação da EPC. Foram realizados múltiplos testes, variando a quantidade de *features* de entrada (2 a 25 energias da EB) e nível de decomposição, para um conjunto de 47 filtros *wavelet*, e consequentemente seus coeficientes, de acordo com a Tabela 7, empregando os vetores de dados provenientes de MPP, já devidamente processados via DWTP e cujas energias provenientes da EB foram previamente armazenadas.

Tabela 7 – Filtros *wavelet* empregados nos testes do módulo MEP.

Família	Coefficientes
Beylkin	18
Coiflets	6, 12, 18, 24 e 30
Daubechies	4-76 (suporte par)
Haar	2
Symmlets	8 e 16
Vaidyanathan	24

Fonte: Elaborada pelo autor.

Nesse conjunto de testes pode-se analisar o quão importantes são os parâmetros de entrada para a devida classificação proposta. No caso específico deste trabalho, analisaram-se as 25 energias provenientes da EB. Também é possível analisar se a classificação é possível com o emprego de uma quantidade menor de *features*, culminando em uma possível simplificação do sistema e redução da carga de processamento, memória e armazenamento do dispositivo.

O Plano Paraconsistente (PP), apresentado no capítulo anterior, e gráficos de medidas de distâncias, terão o papel de representar o resultado da Engenharia Paraconsistente de Características (EPC) buscando sempre a comparação com o Vértice Ótimo (VO), de coordenadas  $G_1 = 1$  e  $G_2 = 0$ , conforme detalhado no capítulo anterior. Este representa o melhor resultado possível, significando que um classificador PMC é suficiente para classificação das referidas amostras.

### 4.3.2.1 Grupo GR1

Nesse grupo estão alocados os testes relacionados ao sistema de “detecção de cigarras” em que as energias estão divididas nas classes “cigarra” e “ruído”. Foram utilizados  $X = 1526$  arquivos WAV, sendo a metade, ou seja  $\frac{X}{2} = 763$ , contendo sinais de cigarras e o restante compreendendo sinais de ruídos diversos da lavoura, como vento, animais e aves.

Na Figura 18 são apresentados os PPs resultantes nos testes com as 25 energias provenientes da EB. É possível verificar que 100% dos testes apresentaram o ponto P como mais próximo ao vértice que representa “Falsidade”, significando que o classificador alimentado por essas *features* não poderá ser modesto para possibilitar a identificação almejada dos padrões.

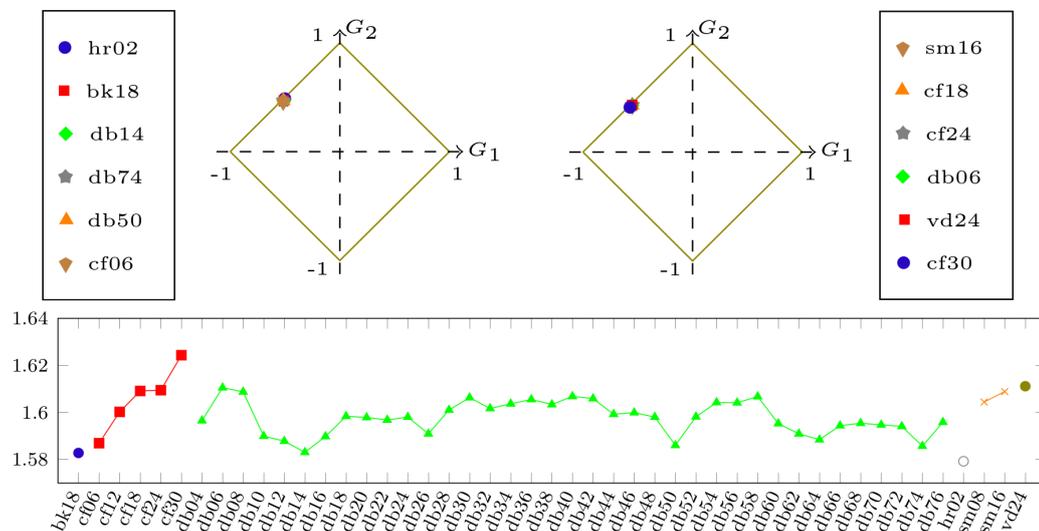


Figura 18 – EPC em GR1 com 25 *features*, com sinais de 10 segundos (441.000 amostras). Acima os planos paraconsistentes, com os melhores resultados (esquerda) e piores resultados (direita), em ordem crescente de distância em relação a VO. Abaixo, as distâncias em relação a VO estão plotadas em diferentes cores de acordo com as famílias *wavelet*.

Conforme previamente elucidado, os arquivos deste grupo foram amostrados a 44.100 Hz, com 16 bits de quantização e contêm 10 segundos, perfazendo um total de  $44100 \cdot 10 = 441.000$  amostras por arquivo, em nível médio de decomposição.

No processo de aquisição das *features*, além de empregar os sinais integrais de 10 segundos, optou-se por analisar a hipótese de utilização de frações menores do sinal, como 5, 2 e 1 segundo, a fim de, futuramente, avaliar a eficiência do sistema no processamento de sinais mais curtos, objetivando redução de custos e expandindo a frequência de resposta do sistema na lavoura.

Na Figura 19 ilustra-se a EPC processada com metade do sinal original, ou seja, 5 segundos. Pode-se verificar que, neste caso, o ponto P se apresenta mais próximo ao vértice que representa “Ambiguidade”, em comparação com os testes com arquivos de

10 segundos. Entretanto, as distâncias para PO diminuíram, permitindo-se considerar o emprego de segmentos de áudio menores para a aplicação do sistema proposto.

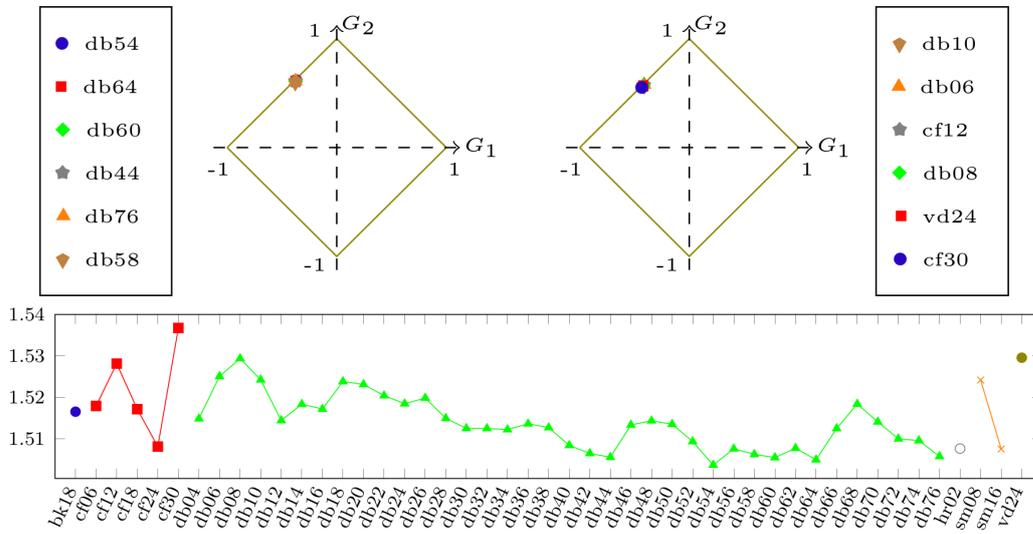


Figura 19 – EPC em GR1 com 25 *features* em sinais de 5 segundos.

Ainda buscando otimização do desempenho do sistema, o próximo teste do grupo GR1 utilizou segmentos de 2 segundos. Pode-se notar na Figura 20 que os resultados se mantiveram próximos aos testes anteriores, significando que há possibilidade de consideração da hipótese da redução dos arquivos originais de 10 para 2 segundos.

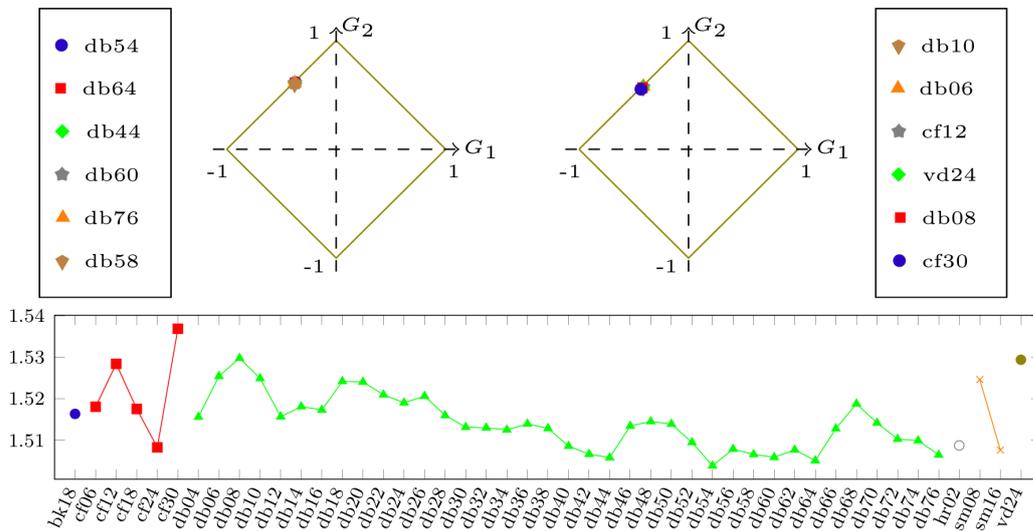


Figura 20 – EPC em GR1 com 25 *features*, em sinais de 2 segundos.

O próximo teste do presente grupo utilizou segmentos de 1 segundo. Na Figura 21 ilustram-se os PPs com os melhores e piores resultados. Pode-se observar que as distâncias para VO se mostraram menores do que todos os testes do presente grupo, apresentando-se, assim, como melhor resultado até o momento, entretanto os pontos P se aproximaram do vértice “Ambiguidade”, reforçando a possível necessidade de um KBC.

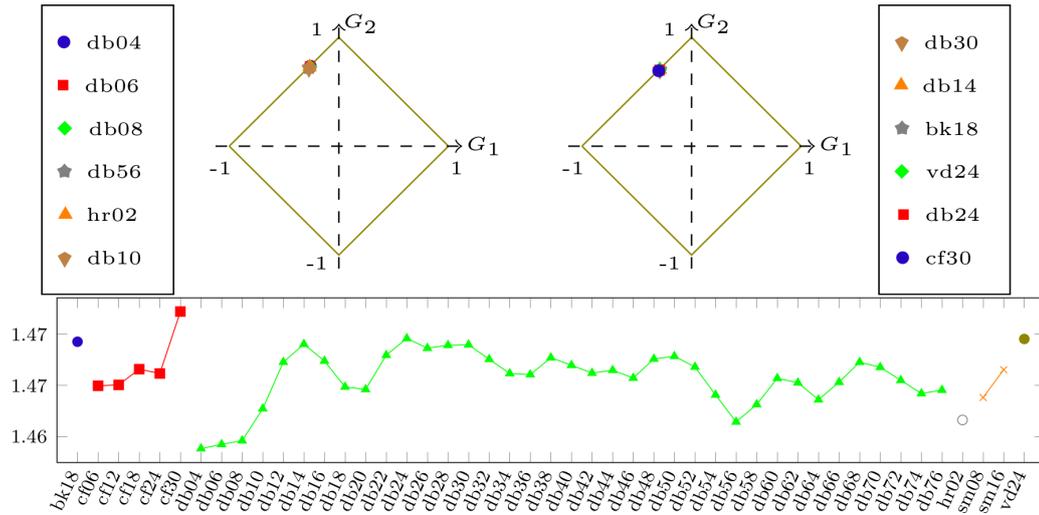


Figura 21 – EPC em GR1 com 25 *features*, em sinais de 1 segundo.

Além da hipótese de redução do comprimento do arquivo de áudio, a fim de reduzir o tempo de captura e, conseqüentemente, os requisitos do sistema, também analisou-se a possibilidade de redução da dimensão das amostras. Na Figura 22 têm-se os PPs de segmentos de 1 segundo (superior esquerdo), 2 segundos (superior direito), 5 segundos (inferior esquerdo) e 10 segundos (inferior direito), testados com quantidades menores de *features*, sendo 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante). Pode-se observar que, em todos os casos, o resultado tem o vértice que representa “Ambigüidade” como mais próximo.

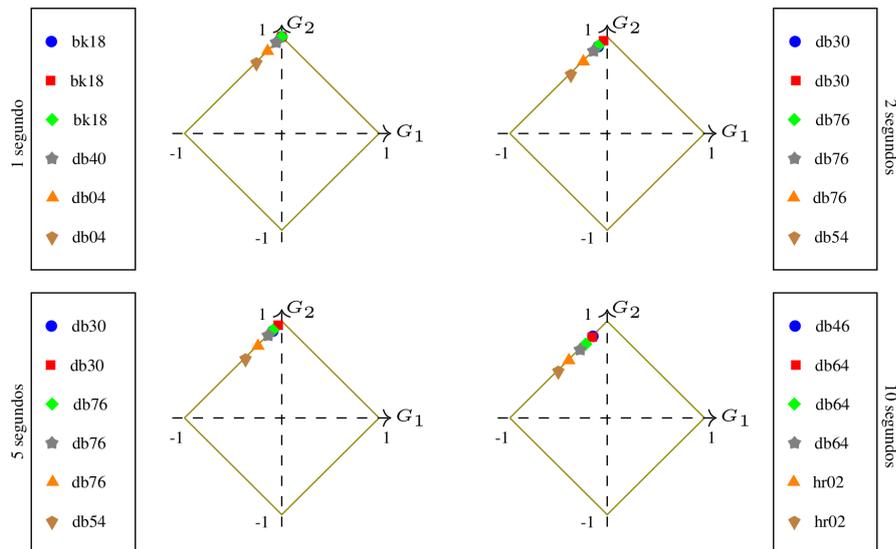


Figura 22 – PPs em GR1 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) *features*.

Na Figura 23 apresenta-se o comparativo de distâncias entre o ponto P e o vértice VO nos diversos testes realizados em GR1, variando-se os filtros *wavelet*, com segmentos

de 1s, 2s, 5s e 10s, além do número de *features*, variando entre 2 e 25.

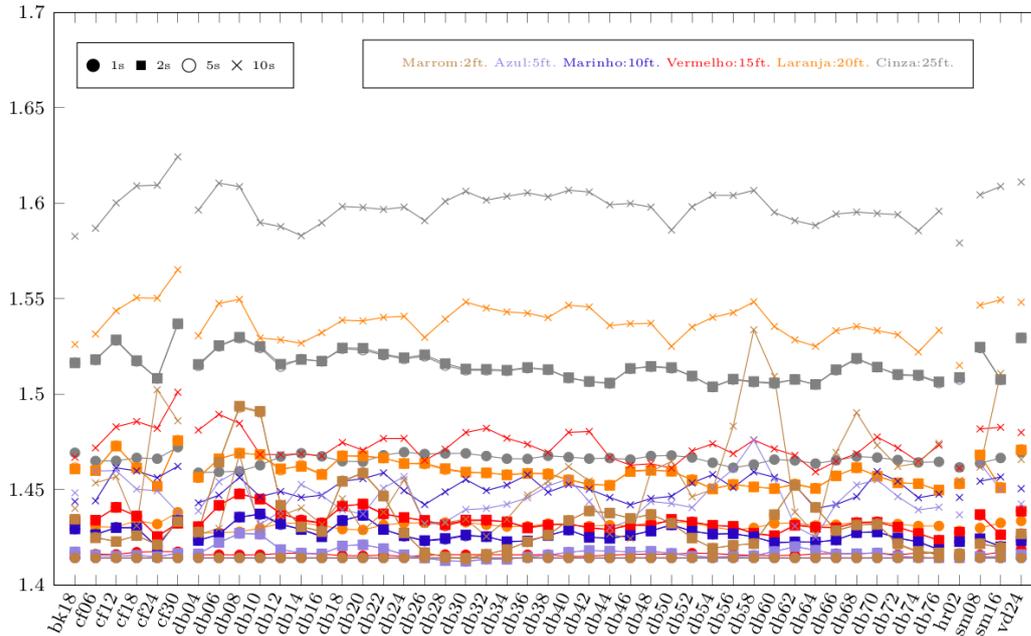


Figura 23 – EPC em GR1: distâncias de VO.

Nos testes realizados até o presente momento, o número  $N$  de características variou entre 2 e 25, utilizando-se sempre as primeiras características do máximo de 25 provenientes da EB. O objetivo fundamental desta etapa é a busca pelas  $N$  “melhores” *features*, dado o limite de  $N$  escolhido, entre 2 e 25.

Apresenta-se, na Figura 24, os melhores resultados da busca pelas melhores *features* testando todas as combinações possíveis (ARFKEN; WEBER, 2005), considerando a distância para VO, conforme Equação 4.1. Na Tabela 8 são representados o número de combinações possíveis  $C(N)$  para cada quantidade  $N$  de *features* testada.

$$C(N) = \frac{25!}{(N! \cdot (25 - N)!)}, \quad (4.1)$$

Os testes utilizaram segmentos de áudio de 1s (esfera), 2s (quadrado), 5s (triângulo) e 10s (círculo). As cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). Verifica-se que as primeiras energias da EB apresentaram os melhores resultados, pois resultam em menores distâncias em relação à distância de VO. As *features* com maior incidência foram as de posição 1 e 2 com 877/9776 e 814/9776 ocorrências, respectivamente, sendo o valor 9776 correspondente ao total de possíveis ocorrências resultante de 47 (filtros) x 52 (2+5+10+15+20) conjuntos de *features* x 4 (tamanhos diferentes de arquivos) = 9776 ocorrências.

Na Figura 25 apresentam-se os PPs com sinais de 1 a 10 segundos, utilizando os  $N$  melhores *features*, sendo  $N=2$  (círculo),  $N=5$  (quadrado),  $N=10$  (losango),  $N=15$

Tabela 8 – Possibilidades na busca das N melhores *features*.

N	cálculo	C(N)
2	$25!/(2! \cdot 1!)$	300
3	$25!/(3! \cdot 2!)$	2300
4	$25!/(4! \cdot 3!)$	12.650
5	$25!/(5! \cdot 4!)$	53.130
6	$25!/(6! \cdot 5!)$	177.100
7	$25!/(7! \cdot 6!)$	480.700
8	$25!/(8! \cdot 7!)$	1.081.575
9	$25!/(9! \cdot 8!)$	2.042.975
10	$25!/(10! \cdot 9!)$	3.268.760
11	$25!/(11! \cdot 10!)$	4.457.400
12	$25!/(12! \cdot 11!)$	5.200.300
13	$25!/(13! \cdot 12!)$	5.200.300
14	$25!/(14! \cdot 13!)$	4.457.400
15	$25!/(15! \cdot 14!)$	3.268.760
16	$25!/(16! \cdot 15!)$	2.042.975
17	$25!/(17! \cdot 16!)$	1.081.575
18	$25!/(18! \cdot 17!)$	480.700
19	$25!/(19! \cdot 18!)$	177.100
20	$25!/(20! \cdot 19!)$	53.130
21	$25!/(21! \cdot 20!)$	12.650
22	$25!/(22! \cdot 21!)$	2.300
23	$25!/(23! \cdot 22!)$	300
24	$25!/(24! \cdot 23!)$	25
25	$25!/(25! \cdot 24!)$	1

Fonte: Elaborada pelo autor.

(estrela), N=20 (triângulo) e N=25 (diamante). Observa-se que os resultados se mantiveram próximos aos obtidos nos primeiros testes, empregando as N primeiras *features*, com os pontos P tendo como vértice mais próximo o vértice “Ambiguidade”.

A Figura 26 contém as distâncias de cada ponto P obtidas nos testes variando os filtros *wavelet* (47 filtros), segmentos de áudio (1s a 10s) e número de *features* (2 a 25 melhores). Verifica-se que os resultados se mantiveram aproximados aos obtidos nos testes utilizando as N primeiras *features*, significando que a busca pelas melhores *features* não permitiu a melhoria do rendimento dos testes em nível médio de decomposição, apesar de possibilitar considerável melhoria de desempenho com a redução do nível de decomposição.

Nesta subseção, foram apresentados os testes relacionados ao sistema de detecção de cigarras em nível médio de decomposição. Na próxima subseção, serão apresentados os testes relacionados ao mesmo sistema, alterando o nível de decomposição para o nível máximo.



Figura 24 – Melhores *features* eleitas por EPC em GR1, com segmentos de 1s (esfera), 2s (quadrado), 5s (triângulo) e 10s (círculo). As cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja).

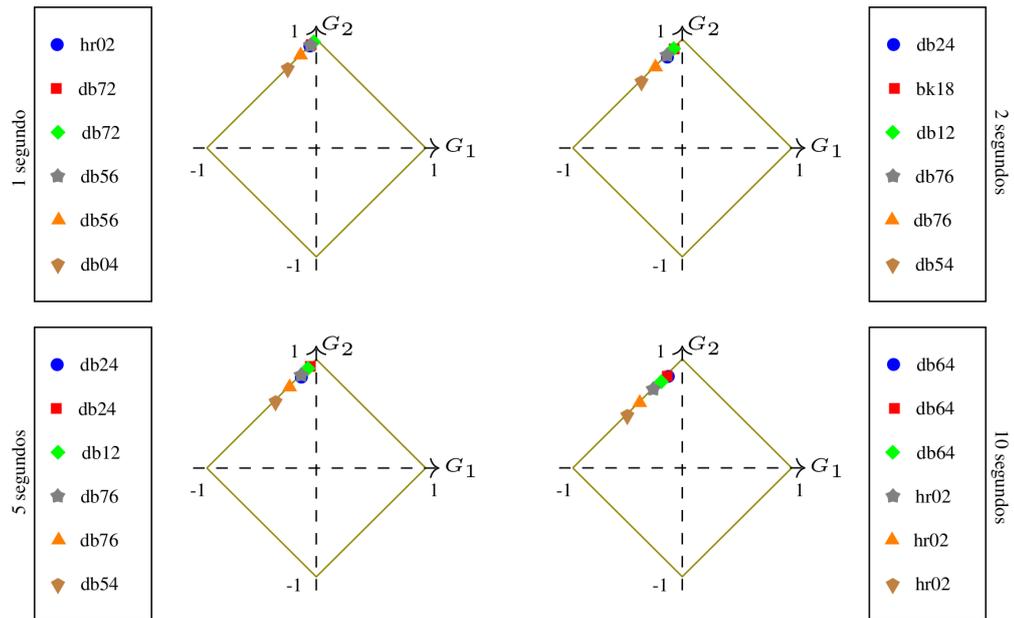


Figura 25 – PPs em GR1 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores *features*.

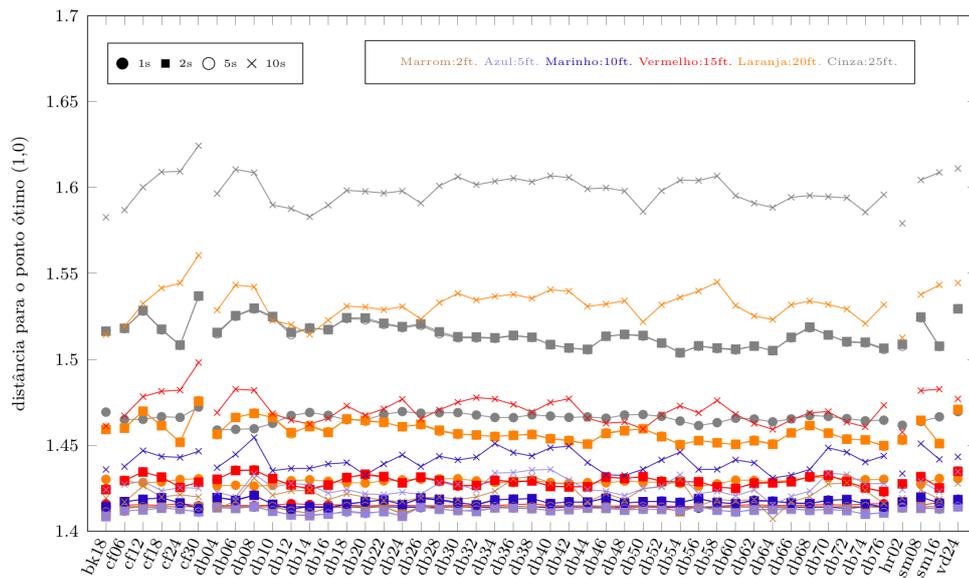


Figura 26 – EPC em GR1: distâncias de VO com as melhores *features*.

#### 4.3.2.2 Grupo GR2

Nesse grupo, a exemplo da subseção anterior, estão alocados os testes relacionados ao sistema de “detecção de cigarras” em que as energias estão divididas nas classes “cigarra” e “ruído”. Também a exemplo da subseção anterior, os arquivos deste grupo foram amostrados a 44.100 Hz, com 16 bits de quantização e contém 10 segundos, perfazendo um total de  $44.100 \cdot 10 = 441.000$  amostras por arquivo, entretanto, agora em nível máximo de decomposição, objetivando maior resolução em frequência.

Na Figura 27 são apresentados os PPs resultantes nos testes com as 25 energias provenientes da EB. Assim como na Figura 18 da subseção anterior, é possível verificar que 100% dos testes apresentaram o ponto P como mais próximo ao vértice que representa “Ambiguidade”, entretanto, agora com menor distância para VO.

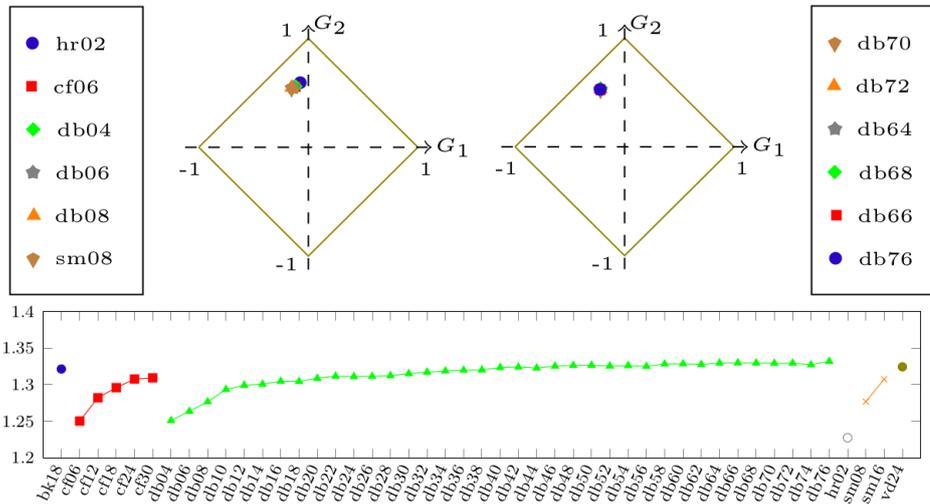


Figura 27 – EPC em GR2 com 25 *features*, em sinais de 10 segundos.

Com o intuito de analisar o desempenho das frações menores dos sinais originalmente definidos para o presente trabalho, na Figura 28 a EPC processada com metade do sinal original é ilustrada, ou seja, 5 segundos. Nota-se que, neste caso, o ponto P se afastando ligeiramente do VO, ainda se apresenta mais próximo ao vértice que representa “Ambiguidade”.

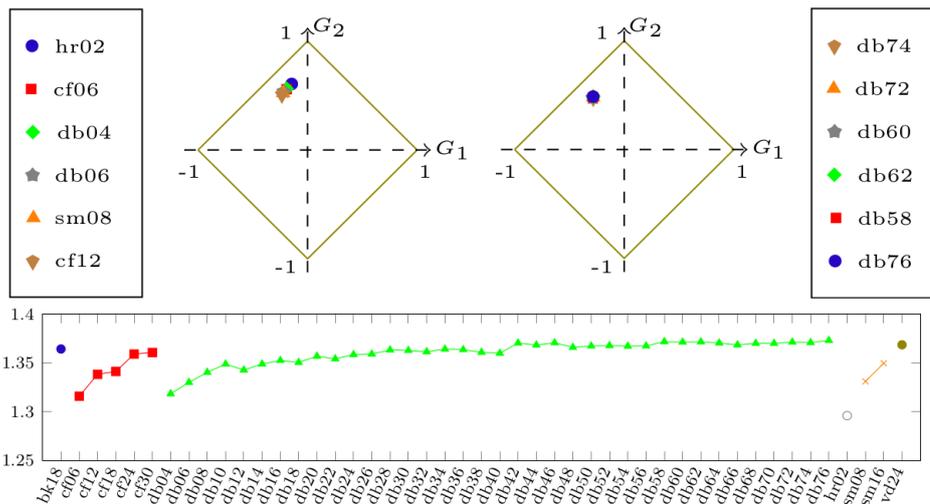


Figura 28 – EPC em GR2 com 25 *features*, em sinais de 5 segundos.

Ainda buscando otimização do desempenho do sistema, o próximo teste do grupo GR2 utilizou segmentos de 2 segundos. Pode-se observar na Figura 29 que os resultados se mantiveram próximos aos testes com 5 segundos, significando que, de acordo com a

EPC, não parece haver diferença considerável de resultados utilizando segmentos de 5 e 2 segundos.

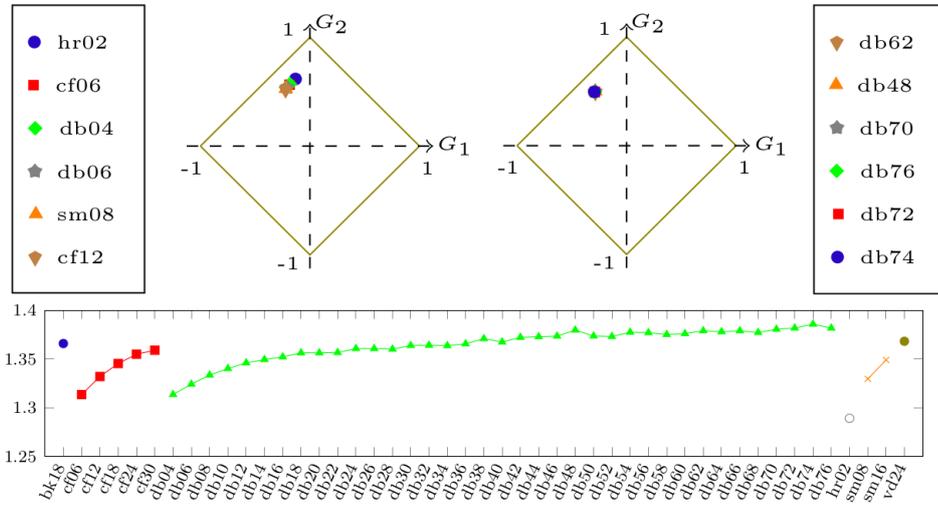


Figura 29 – EPC em GR2 com 25 *features*, em sinais de 2 segundos.

O próximo teste do presente grupo utilizou segmentos de 1 segundo. Na Figura 30 ilustram-se os PPs com os melhores e piores resultados. Comparando todos os gráficos de distância para VO em GR2, nota-se que as distâncias neste caso se mostraram ligeiramente menores do que todos os testes do presente grupo, apresentando-se, assim, como melhor resultado do grupo GR2 até o momento, entretanto os pontos P continuam próximos do vértice “Ambiguidade”, reforçando a possível necessidade de um KBC.

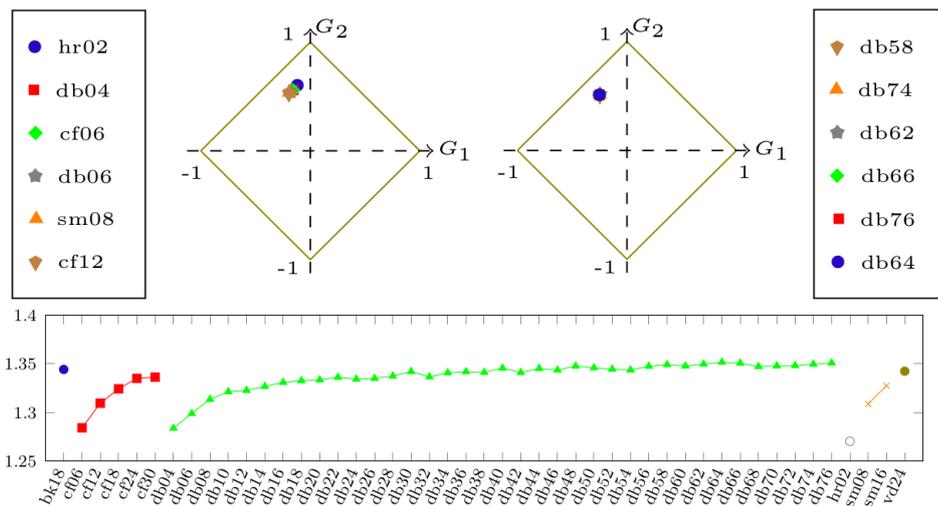


Figura 30 – EPC em GR2 com 25 *features*, em sinais de 1 segundo.

A exemplo do procedimento realizado na seção anterior, também analisou-se a possibilidade de redução da dimensão das amostras. Na Figura 31 são exibidos os PPs de segmentos de 1 segundo (superior esquerdo), 2 segundos (superior direito), 5 segundos (inferior esquerdo) e 10 segundos (inferior direito), testados com quantidades menores de

*features*, sendo 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante). Observa-se que, na maioria dos casos, o resultado tem o vértice que representa “Ambiguidade” como mais próximo, entretanto, o filtro Haar se apresenta mais próximo ao vértice que representa “Falsidade”, representando maior distinção entre os vetores das diferentes classes.

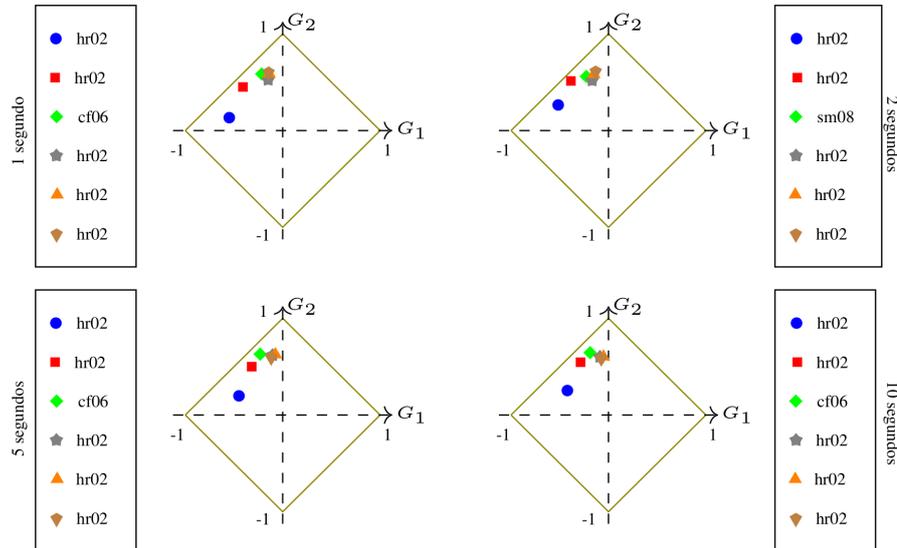


Figura 31 – PPs em GR2 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) *features*.

Na Figura 32 apresenta-se o comparativo de distâncias entre o ponto P e o vértice VO nos diversos testes realizados em GR2, variando-se os filtros *wavelet*, com segmentos de 1s, 2s, 5s e 10s, além do número de *features*, variando entre 2 e 25.

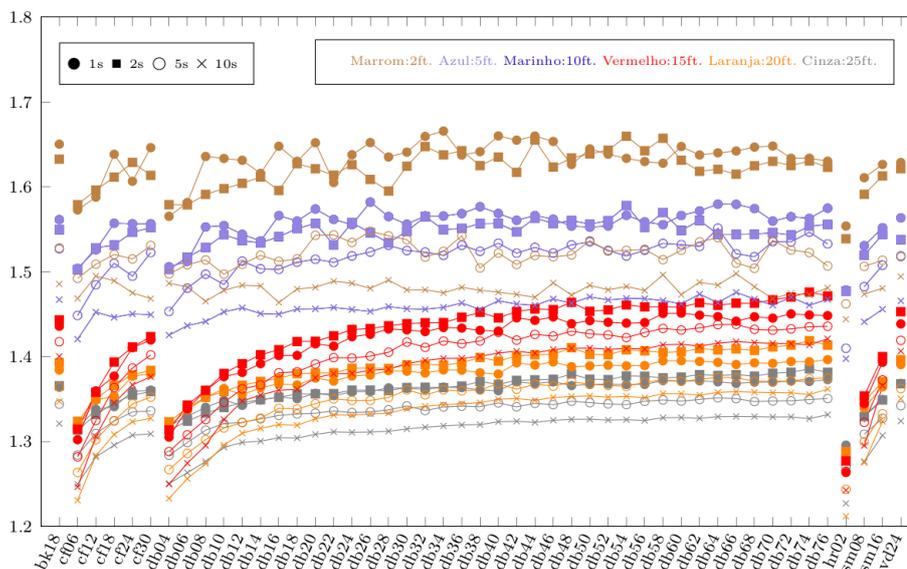


Figura 32 – EPC em GR2: distâncias de VO.

Nesta subseção, os testes realizados até o presente momento variaram o número  $N$  de características entre 2 e 25, utilizando-se sempre as primeiras características do máximo de 25 provenientes da EB. Conforme comentado na subseção anterior, o objetivo fundamental desta etapa é a busca pelas  $N$  “melhores” *features*, dado o limite de  $N$  escolhido, entre 2 e 25.

Apresenta-se, na Figura 33, o resultado da busca pelas melhores *features*, considerando a distância para VO, conforme Equação 4.1 e Tabela 8, apresentadas no grupo de testes anterior. Os testes utilizaram segmentos de áudio de 1s (esfera), 2s (quadrado), 5s (triângulo) e 10s (círculo). As cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). É notável que, desta vez, as primeiras energias da EB não mais apresentaram os melhores resultados, mas sim as posições entre o centro e a região posterior do gráfico. As *features* com maior incidência foram as de posição 16 e 18 com 735/9776 e 732/9776 ocorrências, respectivamente.

Na Figura 34 apresentam-se os PPs com sinais de 1 a 10 segundos, utilizando os  $N$  melhores *features*, sendo  $N=2$  (círculo),  $N=5$  (quadrado),  $N=10$  (losango),  $N=15$  (estrela),  $N=20$  (triângulo) e  $N=25$  (diamante). Pode-se verificar considerável melhoria dos resultados em comparação com a Figura 31, agora utilizando as  $N$  melhores *features* ao invés das  $N$  primeiras. Neste caso, as distâncias para VO foram consideravelmente menores, com o filtro Haar apresentando os melhores resultados.

A Figura 35 confirma as afirmações do parágrafo anterior, apresentando as distâncias de cada ponto P, em relação a VO, obtidas nos testes variando os filtros *wavelet* (47 filtros), segmentos de áudio (1s a 10s) e número de *features* (2 a 25 melhores). É possível verificar que os resultados se mostraram consideravelmente melhores em relação aos realizados com  $N$  primeiras amostras e também em comparação com os testes do grupo anterior.

Nesta subseção, foram apresentados os testes relacionados ao sistema de detecção de cigarras em nível máximo de decomposição. Na próxima seção, serão apresentados os testes relacionados ao sistema de estimativa de densidade populacional de cigarras na lavoura em nível médio de decomposição da DWTP.

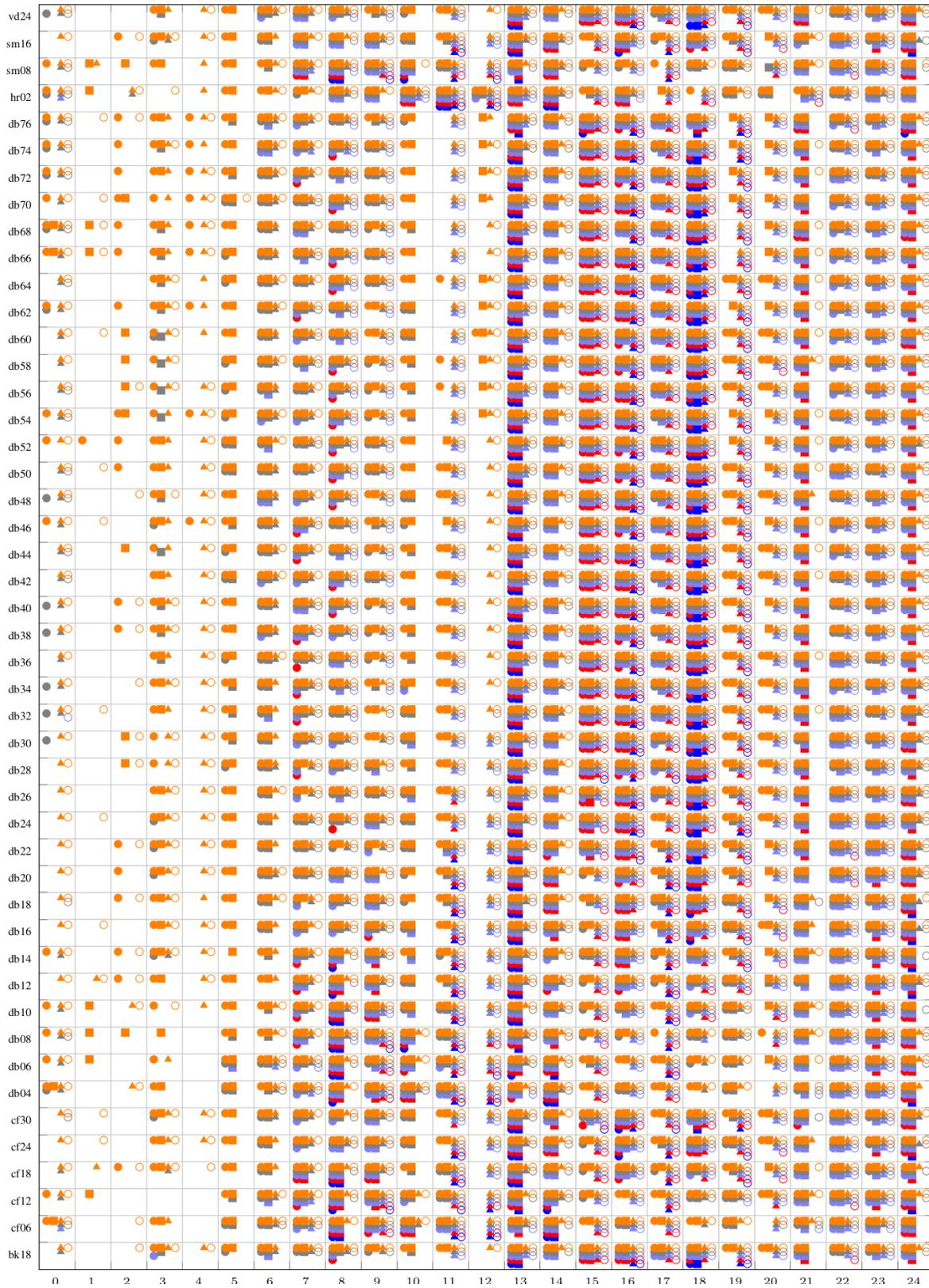


Figura 33 – Melhores *features* eleitas por EPC em GR2, com segmentos de 1s (esfera), 2s (quadrado), 5s (triângulo) e 10s (círculo). As cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja).

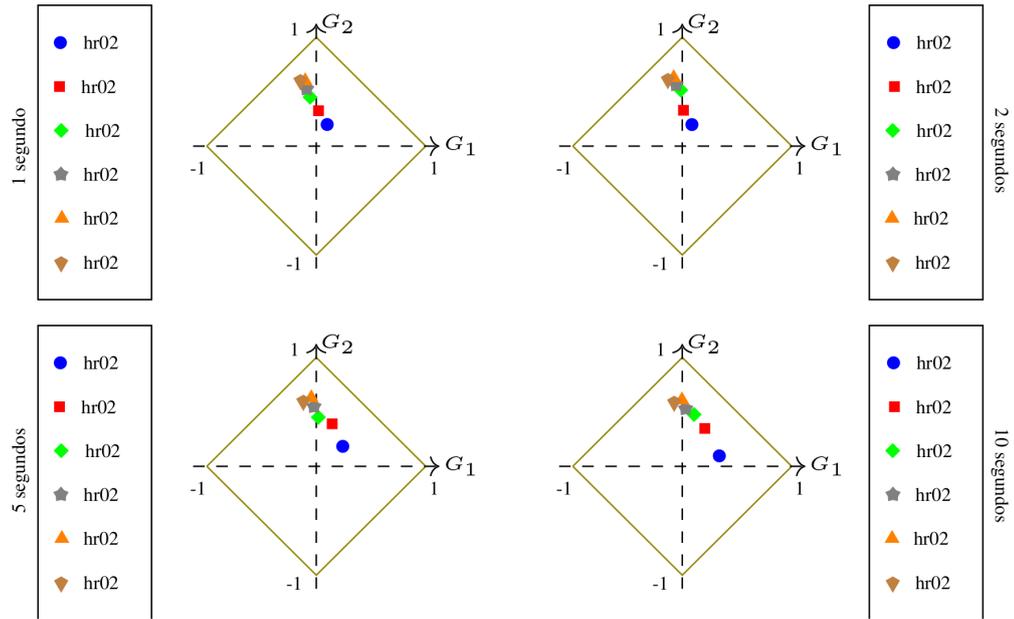


Figura 34 – PPs em GR2 testados com 1s, 2s, 5s e 10s, respectivamente. Utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores *features*.

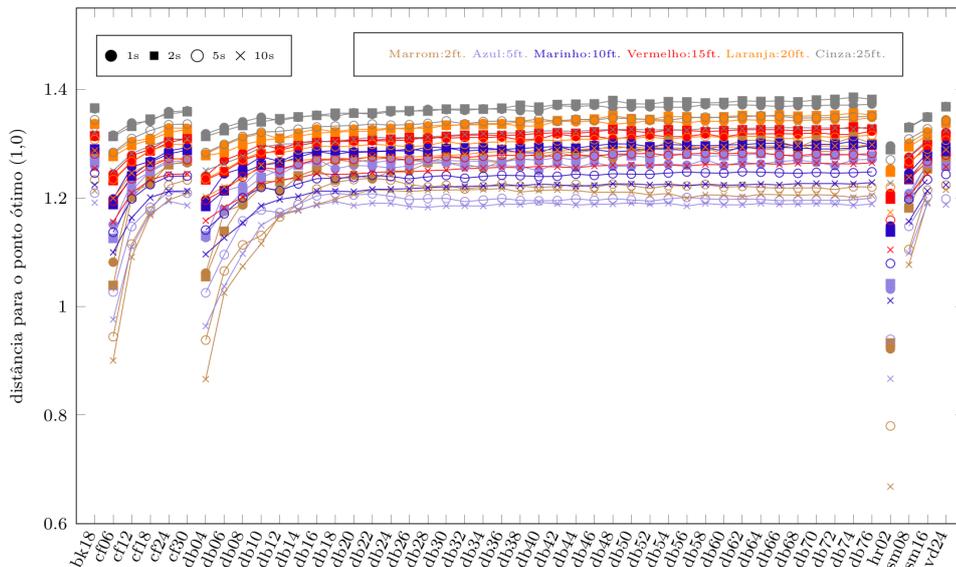


Figura 35 – EPC em GR2: distâncias de VO com as melhores *features*.

#### 4.3.2.3 Grupo GR3

Nesse grupo estão os testes relacionados ao sistema de “densidade de cigarras”, em que as energias estão divididas nas classes “alta”, “baixa” e “ruído”, representando as possíveis densidades de indivíduos estimadas a partir do som capturado em determinada posição da lavoura.

Foram utilizados  $X = 315$  arquivos WAV no total, divididos em três classes de 105 arquivos cada. A primeira fração contém sinais de duas ou mais cigarras, a segunda fração

contendo sinais de uma única cigarra e o restante compreendendo sinais de ruídos diversos da lavoura, como vento, animais e aves.

A separação dos arquivos de áudio, entre as três classes definidas, foi realizada por um engenheiro agrônomo especialista em entomologia, pertencente ao grupo de pesquisa do qual faz parte o presente trabalho.

Assim como nos dois últimos grupos de testes, os arquivos deste grupo foram amostrados a 44.100 Hz, com 16 bits de quantização, entretanto, neste caso contém 1000 milissegundos, perfazendo um total de 44.100 amostras por arquivo, em nível médio de decomposição, objetivando maior resolução no tempo.

Na Figura 36 são apresentados os PPs resultantes nos testes com as 25 energias provenientes da EB. É possível verificar que 100% dos testes apresentaram o ponto P como mais próximo ao vértice que representa “Ambiguidade”, significando a possível necessidade de um KBC.

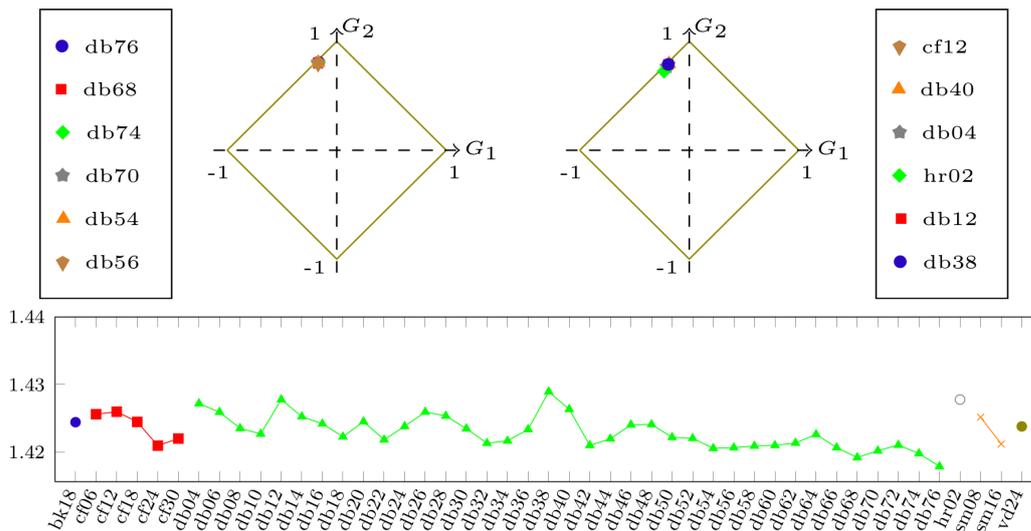


Figura 36 – EPC em GR3 com 25 *features*, com sinais de 1 segundo.

A exemplo do procedimento realizado no grupo anterior, também analisou-se a possibilidade de redução da dimensão das amostras. Na Figura 37 têm-se o PP testado com quantidades menores de *features*, sendo 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e, para efeito de comparação, 25 (diamante). É possível verificar que o resultado tem o vértice que representa “Ambiguidade” como mais próximo, mesmo com a variação da quantidade de *features*.

Na Figura 38 apresenta-se o comparativo de distâncias entre o ponto P e o vértice VO nos diversos testes realizados em GR3, alternando-se os filtros *wavelet*, com número de *features* variando entre 2 e 25.

Conforme realizado nas dois primeiros grupos da presente subseção, os testes iniciais de cada grupo objetivam variar o número N de características entre 2 e 25, utilizando-se

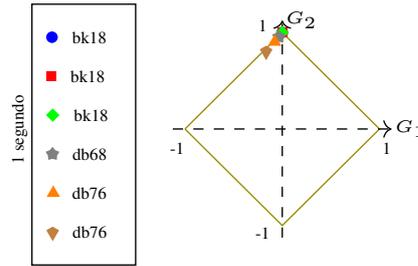


Figura 37 – PP em GR3 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) *features*.

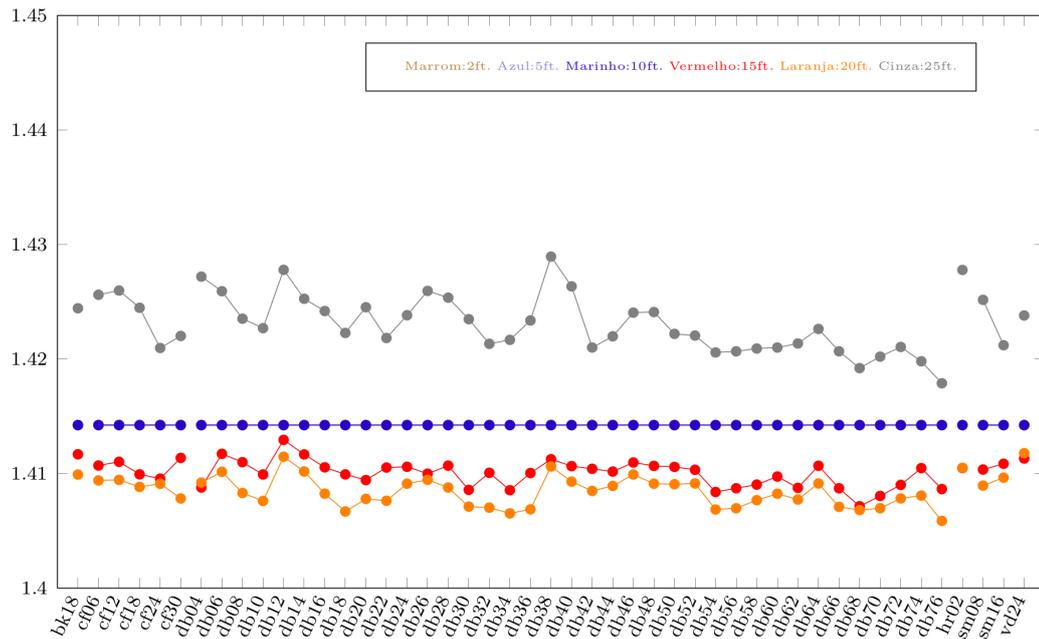


Figura 38 – EPC em GR3: distâncias de VO.

sempre as primeiras características do máximo de 25 provenientes da EB. Como o objetivo fundamental desta etapa é a busca pelas  $N$  “melhores” *features*, dado o limite de  $N$  escolhido, entre 2 e 25, na Figura 39 apresenta-se o resultado da busca pelas melhores *features*, considerando a distância para VO, conforme detalhado anteriormente.

Conforme mencionado, os testes utilizaram somente segmentos de áudio de 1 segundo e as cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). As *features* com maior incidência foram as de posições 0 e 1 com 199/2444 e 179/2444 ocorrências, respectivamente, sendo o total correspondente ao produto entre a quantidade de filtros (47) e o número de *features* testado 52 (2+5+10+15+20).

Na Figura 40 apresenta-se o PP com comprimento de 1 segundo, utilizando as  $N$  melhores *features*, sendo  $N=2$  (círculo),  $N=5$  (quadrado),  $N=10$  (losango),  $N=15$  (estrela),  $N=20$  (triângulo) e  $N=25$  (diamante). Verifica-se que não houve considerável melhora nos resultados em comparação com a Figura 37, onde o primeiro utiliza as  $N$  melhores *features*

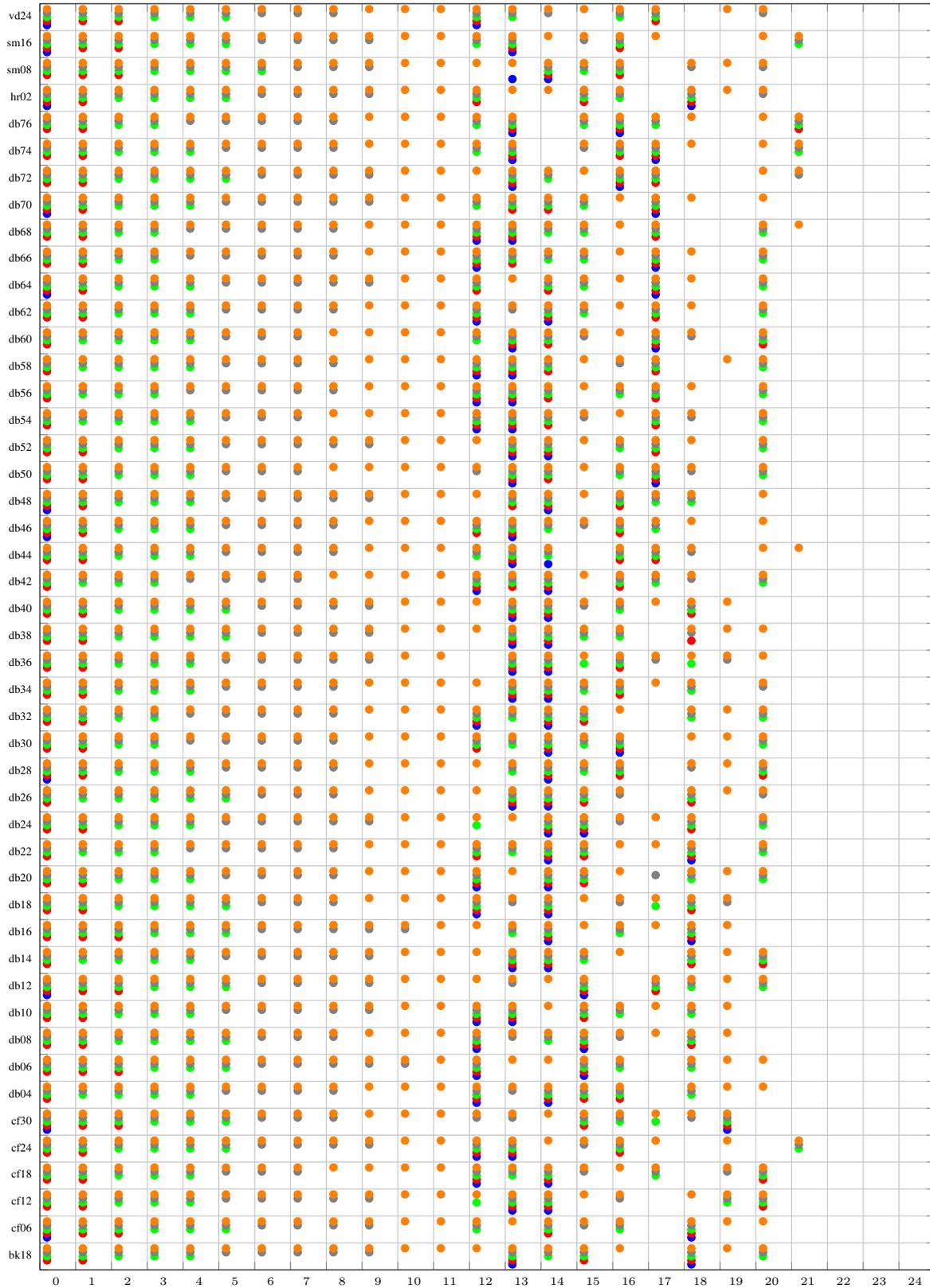


Figura 39 – Melhores *features* eleitas por EPC em GR3. As cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja).

e o segundo emprega as  $N$  primeiras.

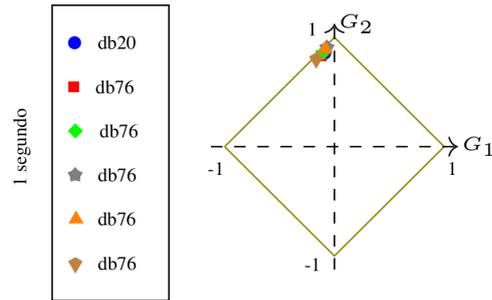


Figura 40 – PPs em GR3 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores *features*.

A Figura 41 confirma as afirmações do parágrafo anterior, apresentando as distâncias de cada ponto P, em relação a VO, obtidas nos testes variando os filtros *wavelet* (47 filtros) e o número de *features* (2 a 25 melhores). É notável que os resultados se mostraram ligeiramente melhores em relação aos realizados com N primeiras amostras.

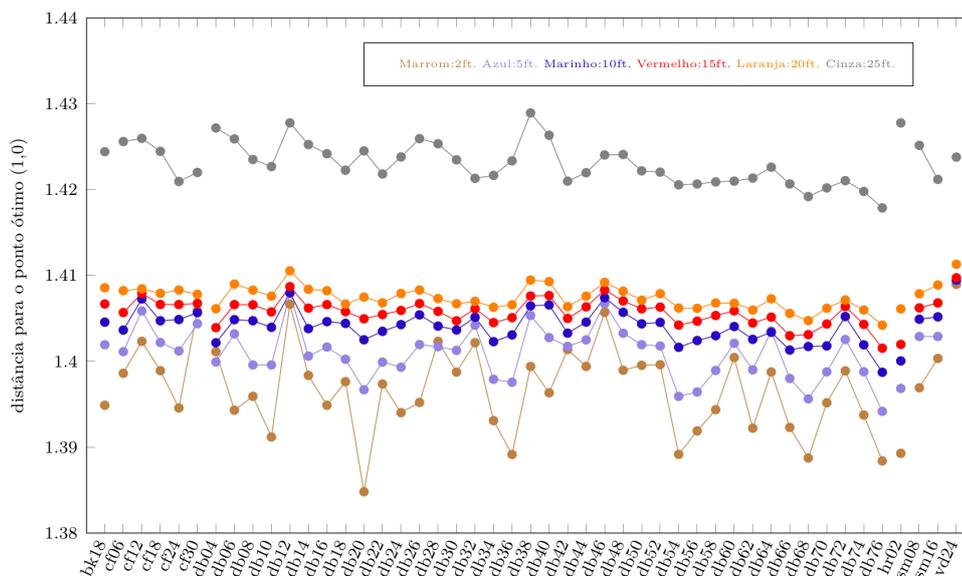


Figura 41 – EPC em GR3: distâncias de VO com as melhores *features*.

Nesta subseção, foram apresentados os testes relacionados ao sistema de densidade de cigarras em nível médio de decomposição. Na próxima subseção, serão apresentados os testes relacionados ao mesmo sistema em nível máximo de decomposição da DWTP.

#### 4.3.2.4 Grupo GR4

A exemplo do grupo anterior, nesse estão apresentados os testes relacionados ao sistema de “densidade de cigarras”, em que as energias estão divididas nas classes “alta”, “baixa” e “ruído”, representando as possíveis densidades de indivíduos estimadas a partir do som capturado em determinada posição da lavoura.

Igualmente ao grupo anterior, os arquivos deste grupo foram amostrados a 44.100 Hz, com 16 bits de quantização, entretanto, neste caso foram processados em nível máximo de decomposição, objetivando maior resolução em frequência.

Na Figura 42 são apresentados os PPs resultantes nos testes com as 25 energias provenientes da EB. É possível verificar que 100% dos testes apresentaram o ponto P como mais próximo ao vértice que representa “Ambiguidade”, entretanto, comparando com a Figura 36 da subseção anterior, vê-se que os sinais transformados em nível máximo apresentam menores distâncias em relação a VO do que os sinais em nível médio.

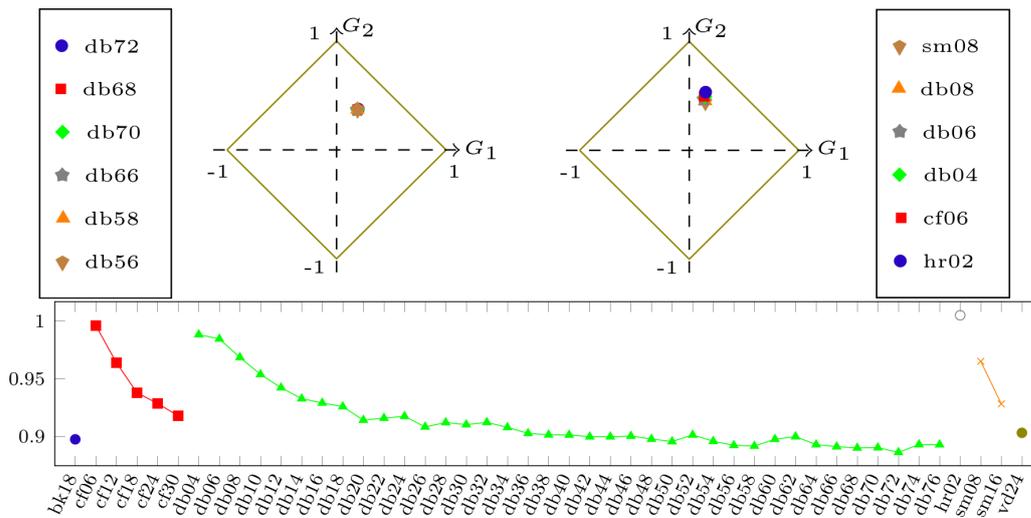


Figura 42 – EPC em GR4 com 25 *features*, em sinais de 1 segundo.

A exemplo do procedimento realizado no grupo anterior, também analisou-se a possibilidade de redução da dimensão das amostras. Na Figura 43 têm-se o PP testado com quantidades menores de *features*, sendo 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e, para efeito de comparação, 25 (diamante). Observa-se que os pontos P estão dispostos mais ao centro do gráfico. Esses resultados são melhores, em comparação com os resultados apresentados na subseção anterior, pois a distância para VO é menor em todos casos de testes em nível máximo.

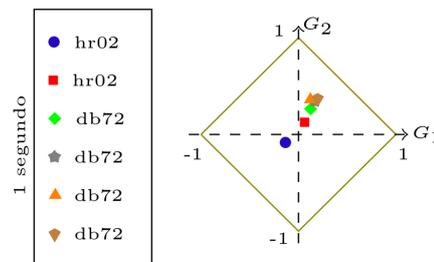


Figura 43 – PP em GR4 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) *features*.

Na Figura 44 apresenta-se o comparativo de distâncias entre o ponto P e o vértice

VO nos diversos testes realizados em GR4, alternando-se os filtros *wavelet*, com número de *features* variando entre 2 e 25, confirmando menor distância para VO em relação ao mesmo gráfico na subseção anterior, na Figura 38.

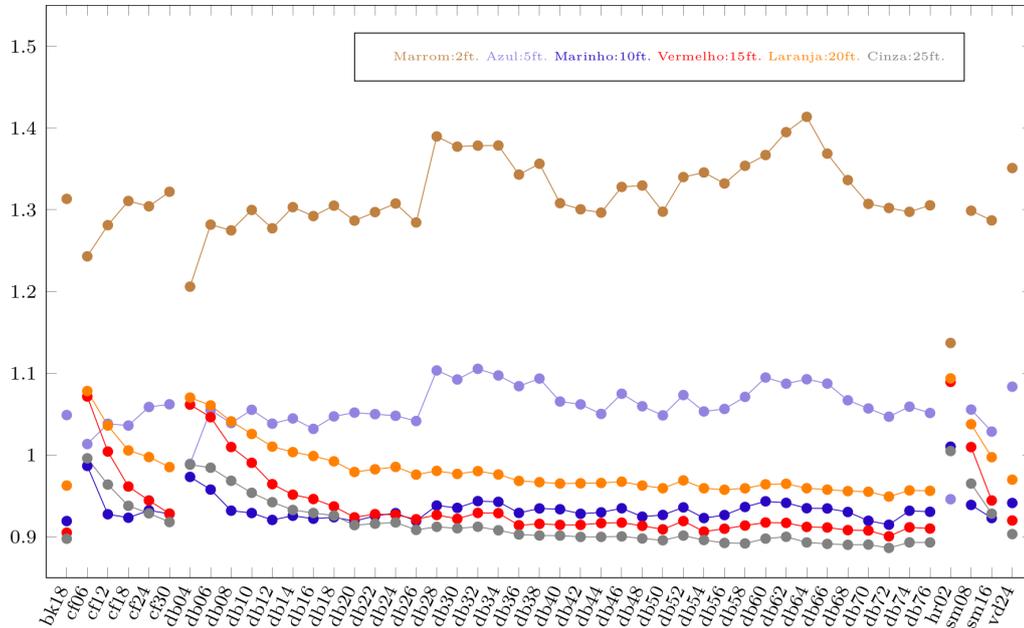


Figura 44 – EPC em GR4: distâncias de VO.

Conforme realizado nos dois primeiros grupos de testes da presente seção, os testes iniciais do grupo variaram o número  $N$  de características entre 2 e 25, utilizando-se sempre as primeiras características do máximo de 25 provenientes da EB. Como o objetivo fundamental desta etapa é a busca pelas  $N$  “melhores” *features*, dado o limite de  $N$  escolhido entre 2 e 25, apresenta-se na Figura 45 o resultado da busca pelas melhores *features*, considerando a distância para VO, conforme detalhado anteriormente.

Os testes utilizaram segmentos de áudio de 1 segundo e as cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja). As *features* com maior incidência foram as de posições 1, 20, 21 e 24 com 221, 221, 200 e 191 ocorrências, respectivamente, de um total de 2.444, ou seja, o produto entre a quantidade de filtros (47) e o número de *features* testado 52 (2+5+10+15+20).

Na Figura 46 apresenta-se o PP utilizando as  $N$  melhores *features*, sendo  $N=2$  (círculo),  $N=5$  (quadrado),  $N=10$  (losango),  $N=15$  (estrela),  $N=20$  (triângulo) e  $N=25$  (diamante). É verificável que houve considerável melhora nos resultados em comparação com a Figura 43, onde as três primeiras ocorrências, ou seja, os pontos P que utilizam as 2, 5 e 10 melhores *features* estão mais próximas ao VO em comparação com os quatro vértices do PP. Uma linha pontilhada transversal à origem do gráfico dividindo o primeiro quadrante ao meio, em cinza escuro, foi incluída no gráfico para permitir melhor visualização.

A Figura 47 confirma as afirmações do parágrafo anterior, apresentando as distâncias

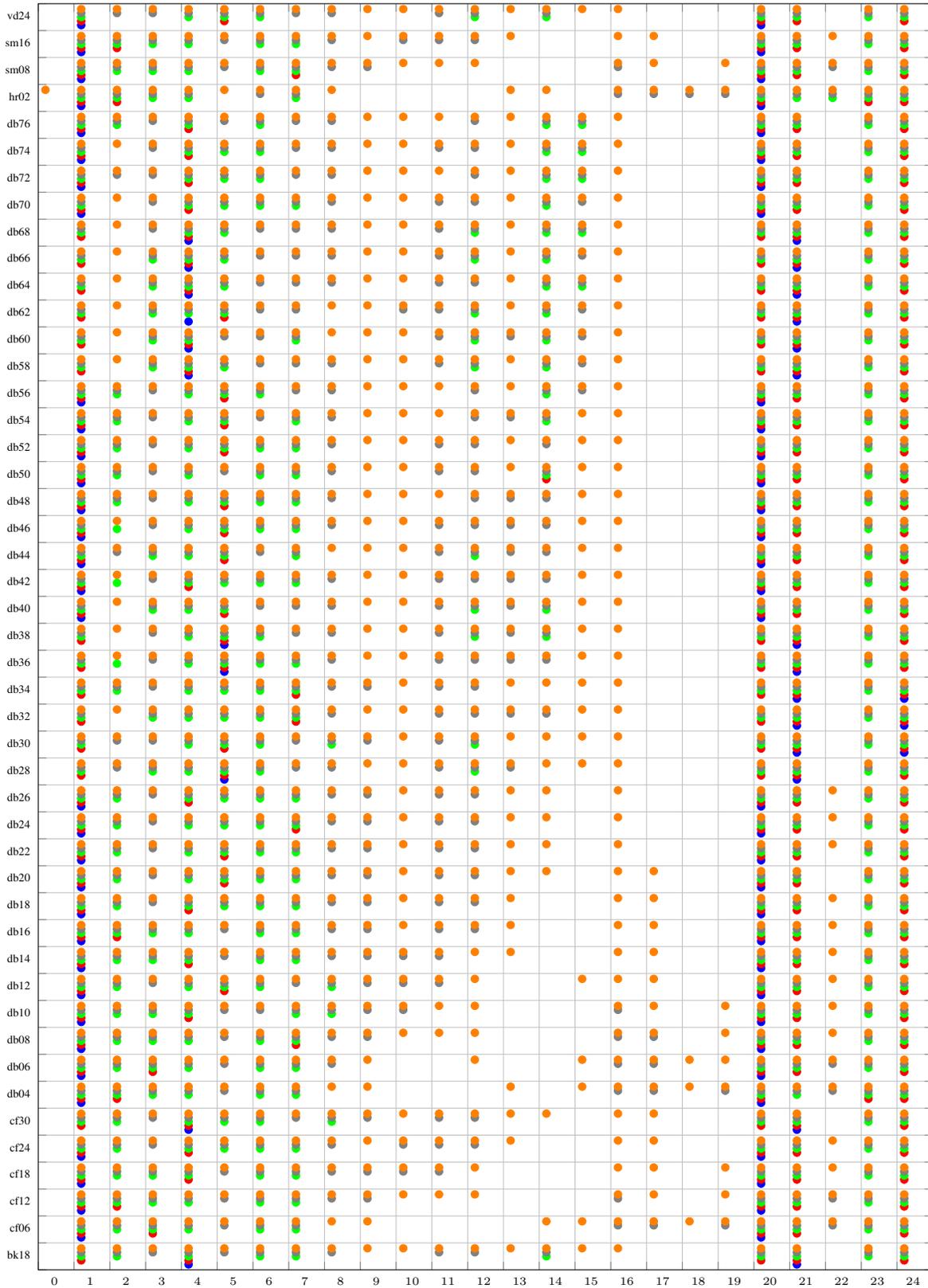


Figura 45 – Melhores *features* eleitas por EPC em GR4. As cores representam o número de *features* utilizadas: 2 (azul), 5 (vermelha), 10 (verde), 15 (cinza) e 20 (laranja).

de cada ponto P, em relação a VO, obtidas nos testes variando os filtros *wavelet* (47 filtros) e o número de *features* (2 a 25 melhores). Observa-se que as distâncias em relação a VO

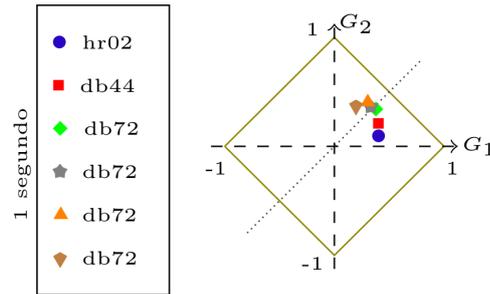


Figura 46 – PPs em GR4 utilizando 2 (círculo), 5 (quadrado), 10 (losango), 15 (estrela), 20 (triângulo) e 25 (diamante) melhores *features*.

foram reduzidas em comparação com os testes realizados com  $N$  primeiras amostras, sendo que os filtros Haar e Daubechies-4, ambos empregando duas *features*, apresentaram as menores distâncias em relação a VO.

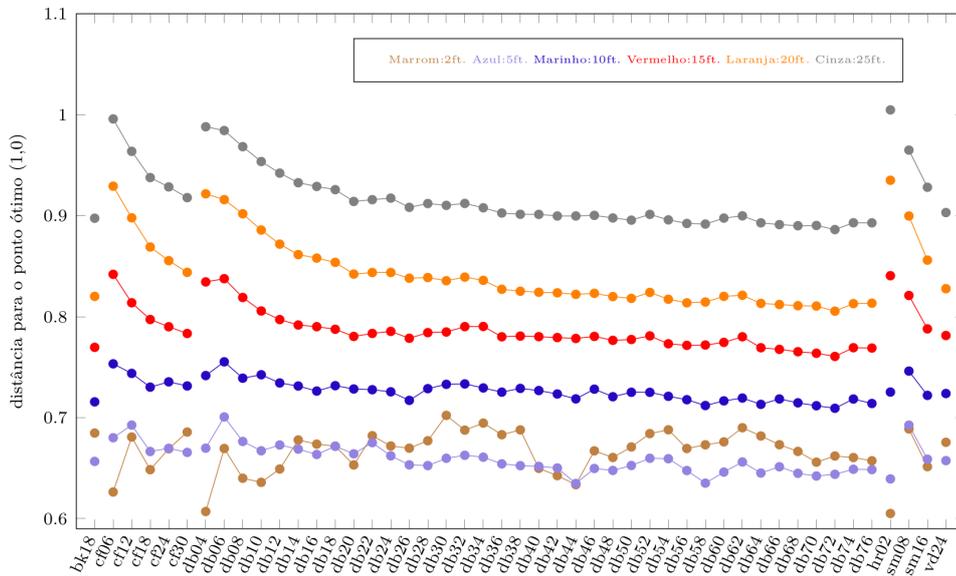


Figura 47 – EPC em GR4: distâncias de VO com as melhores *features*.

Nesta seção, foram apresentados os testes relacionados ao Módulo de Engenharia Paraconsistente de Características (MEP). Na próxima seção, serão apresentados os testes relacionados ao Módulo de Treinamento e Teste (MTT).

### 4.3.3 Testes do módulo MTT

Conforme apresentado na seção anterior, o presente módulo utiliza o classificador SVM para treinar e testar o conjunto de amostras de entrada recebidos como parâmetro. Além dessas, o MTT também recebe como parâmetro o número de *features* a serem empregadas no processo, conforme detalhado no capítulo anterior.

Os parâmetros de ajuste, disponíveis na tabela 4, apresentada no capítulo anterior, são alternados a fim de buscar o melhor rendimento possível ao final do processamento.

O limiar de treinamento, rotina CLT apresentada no capítulo anterior, foi definido entre 20 e 60, assim, o número de amostras de entrada deve variar entre  $\frac{X*20}{100}$  e  $\frac{X*60}{100}$  no processo de treinamento do classificador, totalizando 40 testes para cada filtro.

Buscando ilustrar a análise dos resultados obtidos nos testes do classificador, foram utilizados recursos gráficos amplamente empregados na literatura, como diagrama de caixa, matriz de confusão (MC) e curva ROC, descritos a seguir.

- **Diagrama de Caixa (Box Plot):** é um tipo de gráfico que fornece informações detalhadas sobre a distribuição das amostras de conjuntos numéricos através de seis medidas descritivas, sendo elas: *outliers*, máximo, primeiro quartil, segundo quartil (mediana), terceiro quartil e mínimo (WILLIAMSON; PARKER; KENDRICK, 1989).
- **Matriz de Confusão (Confusion Matrix)** é um recurso que permite comparar os resultados de algoritmos de classificação de padrões, composto por uma tabela onde as linhas representam as instâncias das classes e as colunas representam as classes de resposta do algoritmo (PAULUS; HORNEGGER, 2003; GUIDO *et al.*, 2021). As cores representam o índice atingido, assim, tons mais escuros representam valores mais altos e as cores mais claras representam valores mais baixos. O ideal é que a diagonal principal ( $d_1$ ) tenha cor azul escura e a diagonal secundária ( $d_2$ ) seja branca.
- **Curva ROC (Receiver Operating Characteristic):** é um gráfico que permite analisar o desempenho de um conjunto de testes, provendo importante informação acerca do desempenho do sistema por meio da comparação entre a taxa de verdadeiros-positivos (eixo vertical) e de falsos-positivos (eixo horizontal), além do índice AUC (*Area Under the Curve*) por meio do cálculo da área sob a referida curva (FAN; UPADHYE; WORSTER, 2006).

#### 4.3.3.1 Grupo GR1

No tocante aos testes do grupo GR1, que processaram os sinais em nível médio de decomposição e que consideram as classes “cigarra” e “ruído”, com o objetivo de validar o sistema de “detecção de cigarras”, apresentam-se na Figura 48 os diagramas de caixa

com as acurácias dos 40 testes com sinais de 10, 5, 2 e 1 segundos, utilizando 47 diferentes filtros *wavelet*, empregando 25 *features*. Pode-se contemplar resultados abaixo do esperado, sendo que, em todos os casos, a acurácia média apresentou-se inferior aos 54%.

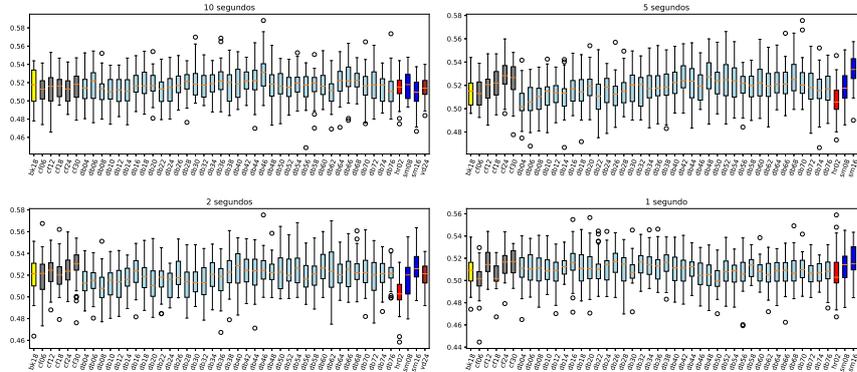


Figura 48 – Acurácia resultante dos testes utilizando 25 *features*.

Nas matrizes de confusão ilustradas na Figura 49, apresentam-se os melhores resultados dos testes empregando 25 *features* em diferentes comprimentos de sinal. O cálculo da acurácia ( $Ac$ ) é realizado por meio da Equação 4.2, onde  $d_1$  representa a diagonal principal e  $d_2$  representa a diagonal secundária.

$$Ac = \frac{d_1}{(d_1 + d_2)}, \quad (4.2)$$

Os resultados abaixo do esperado foram confirmados, com testes alcançando o percentual de 63.4% de falsos-positivos, no caso dos arquivos de 1 segundo, que utilizou filtro Haar e obteve acurácia de 55.9% com *kernel* linear e parâmetro de suavização  $C=10$ , utilizando 55% de arquivos de treinamento, conforme apresentado na Tabela 9.

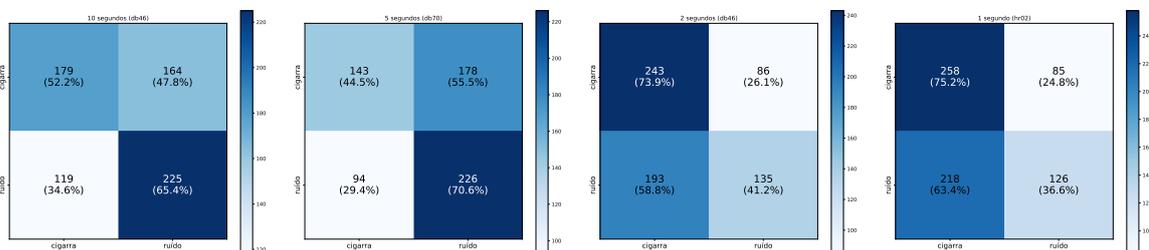


Figura 49 – Matrizes de confusão resultantes dos testes utilizando 25 *features*.

A Figura 50 apresenta as curvas ROC com o melhor resultado (azul) e o pior resultado (vermelho) para sinais de 10 segundos, 5 segundos, 2 segundos e 1 segundo empregando 25 *features*. Nos melhores resultados, a área (AUC) resultante foi inferior a 0.6 em todos os casos, reforçando a necessidade de busca de melhores *features* na investigação da solução do problema.

Tabela 9 – Melhores resultados de testes em GR1 com 25 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db46	0.5881	0.55	1000	-	linear
5 segundos	db70	0.5757	0.58	1000	-	linear
2 segundos	db46	0.5753	0.57	100	-	linear
1 segundo	hr02	0.559	0.55	10	-	linear

Fonte: Elaborada pelo autor.

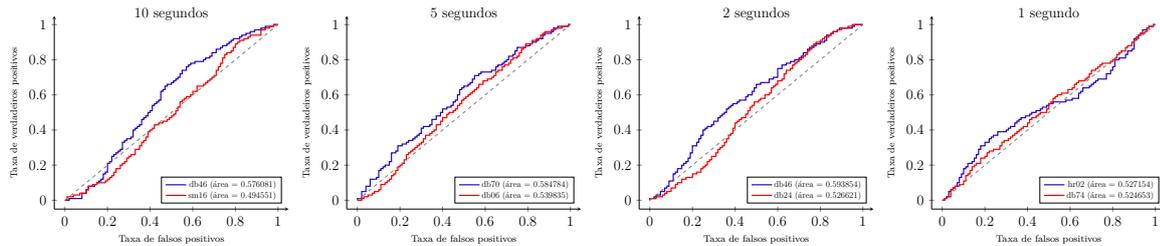


Figura 50 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 25 *features*.

Além dos testes utilizando 25 *features*, também foram analisados os resultados utilizando quantidades menores de *features*. Para isso, empregaram-se as melhores posições eleitas nos testes EPC da subseção anterior, entretanto, os testes com as 20 melhores *features* não apresentaram considerável melhoria no rendimento, conforme ilustra a Figura 51, onde a média apresentada ficou em torno de 52%.

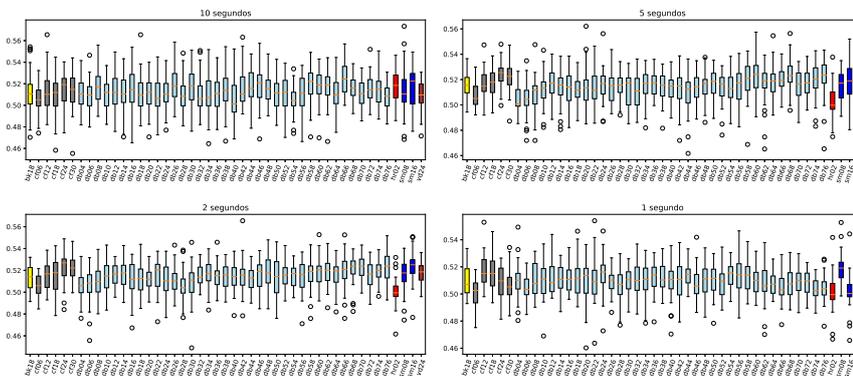


Figura 51 – Acurácia resultante dos testes utilizando 20 *features*.

Na Tabela 10 apresentam-se os melhores resultados dos testes do presente grupo empregando 20 *features*. Nota-se que a melhor acurácia foi obtida pelo filtro Symlet-8 com segmentos de 10 segundos, utilizando 46% das amostras para treinamento, *Kernel* linear e parâmetro de suavização  $C=10$ .

Na Figura 52 as MCs dos melhores resultados obtidos nos testes com 20 *features*. A MC do melhor resultado com segmentos de 1 segundo mostrou taxa de falsos-positivos de 50.8%. Além desse, chegou-se a uma taxa de falsos-negativos de 69% para os sinais de áudio de comprimento de 10 segundos.

Tabela 10 – Melhores resultados de testes em GR1 com 20 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	sm08	0.5733	0.46	10	-	linear
5 segundos	db20	0.5621	0.33	1000	-	linear
2 segundos	db42	0.5655	0.48	100	-	linear
1 segundo	db22	0.554	0.43	10	-	linear

Fonte: Elaborada pelo autor.

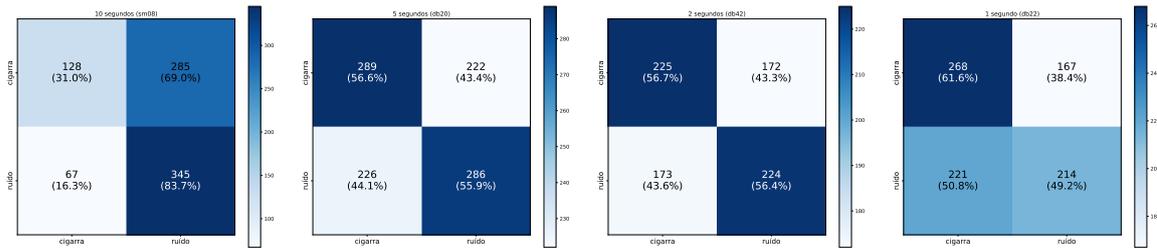


Figura 52 – Matriz de confusão resultante dos testes utilizando 20 *features*.

Os valores mostrados na Figura 53 confirmam os resultados obtidos, tendo nos melhores resultados a AUC variando entre 0.54 e 0.55.

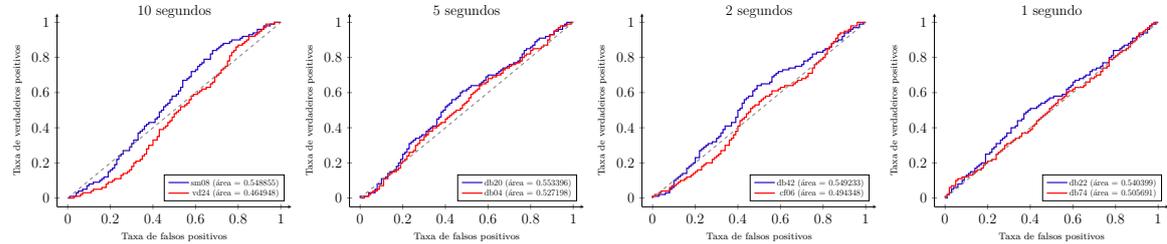


Figura 53 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 20 *features*.

Os resultados médios dos testes com 15 *features* ficaram nos mesmos patamares dos testes anteriores do presente grupo, não havendo considerável melhoria do rendimento no classificador (Figura 54).

Na Tabela 11 estão organizados os resultados obtidos nos testes com 15 *features* do presente grupo, onde pode-se comprovar que não houve melhoria dos resultados com a redução do número de *features* para o caso de GR1.

Na Figura 55 apresentam-se as MCs resultantes dos testes com 15 *features*, ainda com altas taxas de falsos-positivos e falsos-negativos, cuja maior taxa de falsos-positivos foi apresentada nos testes com arquivos de 2 segundos (57%).

Analisando as curvas ROC da Figura 56, nos testes com 15 *features*, tem-se como melhor resultado a taxa AUC de aproximadamente 0.59 obtida com filtro Beylkin-18 com arquivos de 10 segundos.

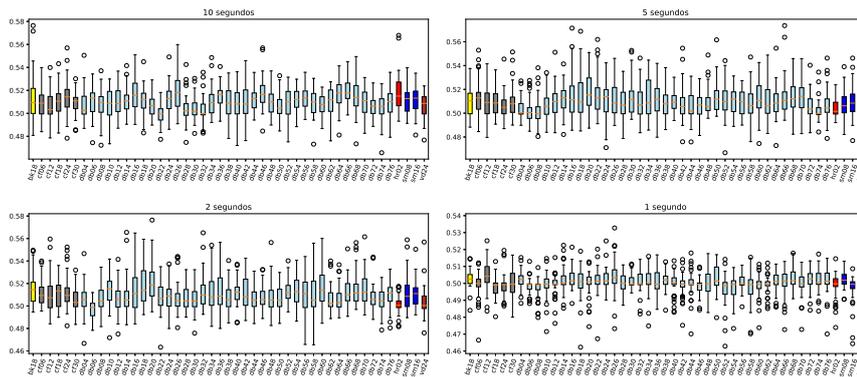


Figura 54 – Acurácia resultante dos testes utilizando 15 *features*.

Tabela 11 – Melhores resultados de testes em GR1 com 15 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	bk18	0.5764	0.49	1000	-	linear
5 segundos	db66	0.5736	0.43	10	-	linear
2 segundos	db20	0.5764	0.55	100	-	linear
1 segundo	vd24	0.5374	0.23	1	0.0001	poly

Fonte: Elaborada pelo autor.

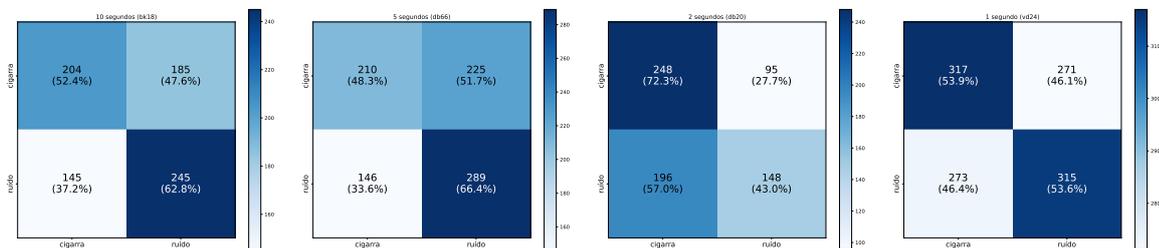


Figura 55 – Matriz confusão resultante dos testes utilizando 15 *features*.

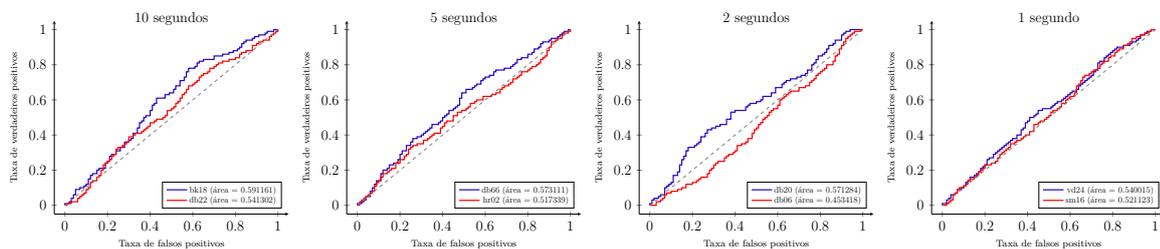


Figura 56 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 15 *features*.

Os resultados com 10 *features* se mantiveram nos mesmos patamares dos testes com quantidades maiores de *features*, onde as acurácias médias estiveram abaixo de 52% em todos os conjuntos de testes (Figura 57).

Na Tabela 12 estão organizados os resultados obtidos nos testes com 10 *features* do presente grupo. Com a acurácia de 57.1% o conjunto de testes com segmentos de 10 segundos apresentou o melhor resultado de GR1 utilizando 10 *features*.

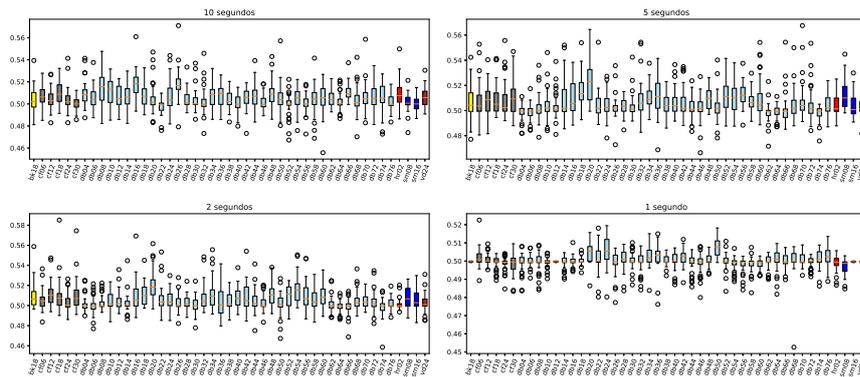


Figura 57 – Acurácia resultante dos testes utilizando 10 *features*.

Tabela 12 – Melhores resultados de testes em GR1 com 10 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db26	0.571	0.53	1000	-	linear
5 segundos	db70	0.5675	0.52	1000	-	linear
2 segundos	cf18	0.5851	0.43	1000	-	linear
1 segundo	cf06	0.5226	0.55	10	-	linear

Fonte: Elaborada pelo autor.

Ainda em relação aos testes com 10 *features*, as MCs (Figura 58) confirmam o baixo rendimento obtido nos testes, com alta taxa de falsos-positivos: 63.8%, 34.7% e 31.7% para sinais de 10, 5 e 2 segundos, respectivamente. Nos testes com sinais de 1 segundo a taxa de falsos-positivos foi de 12.8%, entretanto, a taxa de falsos-negativos foi elevadíssima, passando dos 82%.

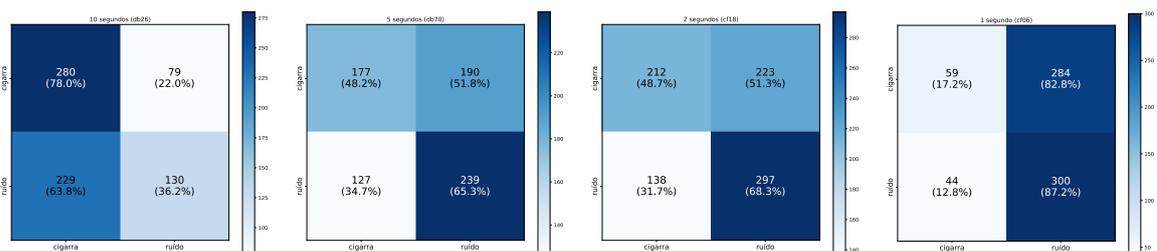


Figura 58 – Matriz confusão resultante dos testes utilizando 10 *features*.

Confirmando os baixos resultados, a Figura 59 apresenta as curvas ROC resultantes dos testes com 10 *features*, com AUCs abaixo de 0.58.

A Figura 60 apresenta as acurácias obtidas nos testes empregando 5 *features*. Os resultados médios continuam abaixo dos 52%, com outliers passando dos 58% no caso dos segmentos de 10 segundos.

Na Tabela 13 estão organizados os resultados obtidos nos testes com 5 *features* do presente grupo, onde as acurácias máximas obtidas variam de 51.93%, com filtro

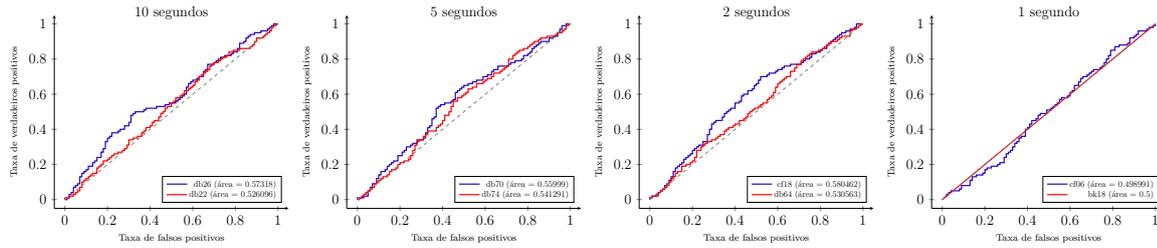


Figura 59 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 10 *features*.

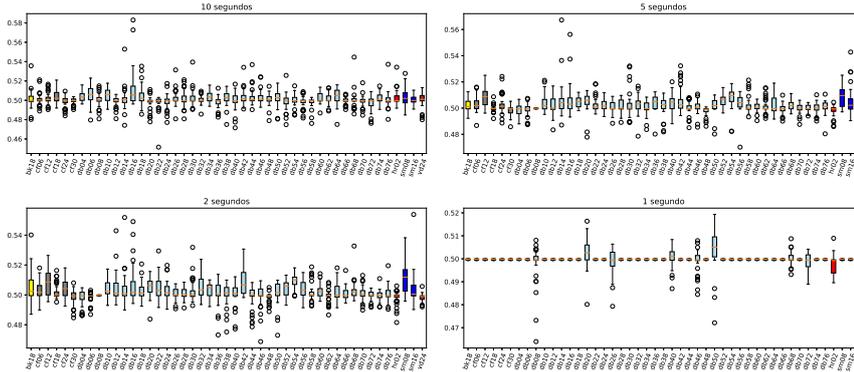


Figura 60 – Acurácia resultante dos testes utilizando 5 *features*.

Daubechies-50 a 58.3% com filtro Daubechies-16.

Tabela 13 – Melhores resultados de testes em GR1 com 5 *features*

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db16	0.583	0.46	1000	-	linear
5 segundos	db14	0.5673	0.44	100	-	linear
2 segundos	sm16	0.5539	0.52	100	-	linear
1 segundo	db50	0.5193	0.44	1000	-	linear

Fonte: Elaborada pelo autor.

Os gráficos da Figura 61 permitem analisar melhor os resultados da figura anterior empregando 5 *features*, pois apresenta um percentual de falsos-negativos acima de 92% com segmentos de 1 segundo e alta taxa de falsos-positivos, com 43.6% nos testes com segmentos de 2 segundos.

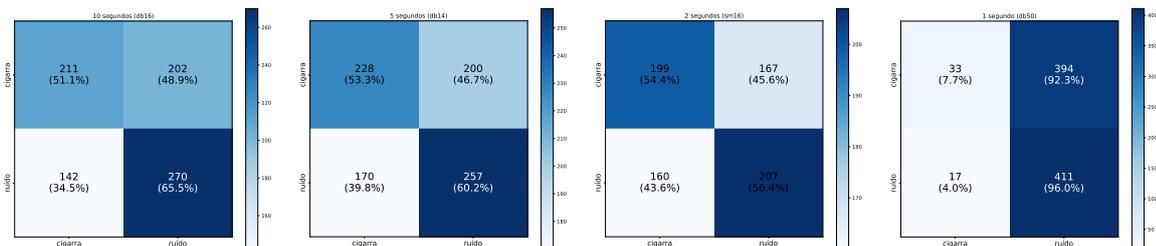


Figura 61 – Matriz confusão resultante dos testes utilizando 5 *features*.

As curvas ROC ilustradas na Figura 62 confirmam os resultados abaixo do esperado, com AUC variando, nos melhores casos, de 0.48 em segmentos de 1 segundo a 0.59 em segmentos de 10 segundos.

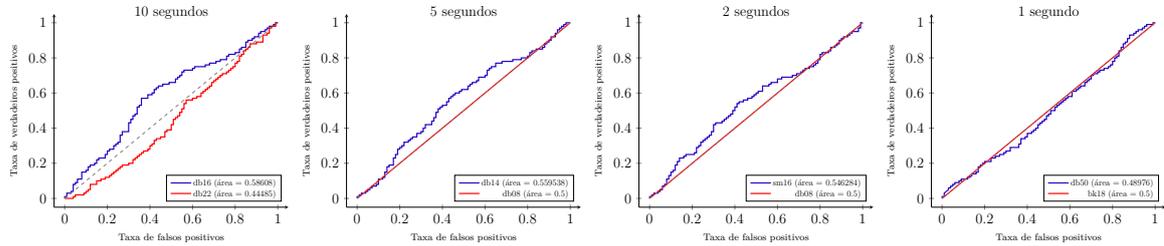


Figura 62 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 5 *features*.

Na Figura 63 apresentam-se as acurácias resultantes dos testes empregando as 2 melhores *features*. É possível verificar que, também nesse caso, os resultados ficaram longe dos ideais, apresentando acurácia média inferior a 52%, significando que os parâmetros testados no presente grupo não provêm o rendimento almejado.

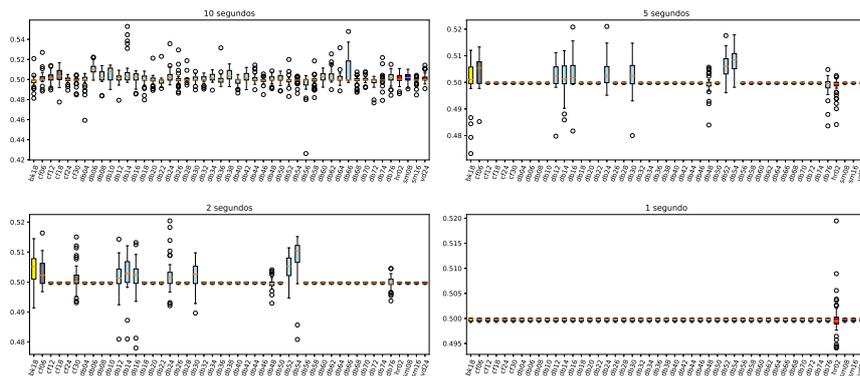


Figura 63 – Acurácia resultante dos testes utilizando 2 *features*.

Na Tabela 14 estão organizados os resultados obtidos nos testes com 2 *features* do presente grupo. O melhor resultado apresentou percentual de acertos de aproximadamente 55% nos testes com sinais de 10 segundos.

Tabela 14 – Melhores resultados de testes em GR1 com 2 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db14	0.5529	0.43	1000	-	linear
5 segundos	db24	0.521	0.33	1000	-	linear
2 segundos	db24	0.5204	0.23	1000	-	linear
1 segundo	hr02	0.5195	0.26	1	0.001	gaussiano

Fonte: Elaborada pelo autor.

No conjunto de gráficos de MC apresentados na Figura 64 tem-se um considerável montante de erros de classificação que reforçam a necessidade de novos parâmetros para o

classificador, tendo como pior resultado o índice de 81.8% de falsos-positivos nos testes com segmentos de 5 segundos.

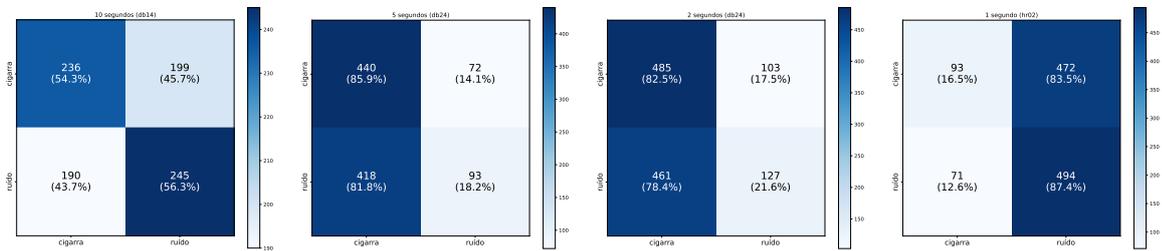


Figura 64 – MC resultante dos testes utilizando 2 *features*.

Mostrando-se abaixo de 0.57, as AUCs apresentadas na Figura 65 confirmam a inviabilidade da adoção desses parâmetros para o problema proposto e reforçam a importância de análise dos resultados empregando novas configurações de entrada no classificador.

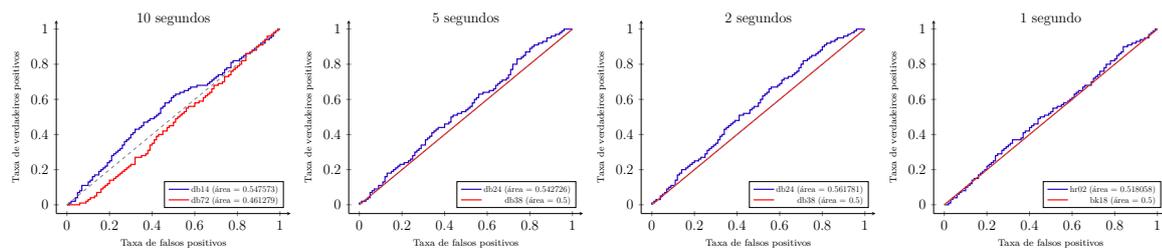


Figura 65 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 2 *features*.

Nesta seção foram apresentados os resultados dos testes do sistema de detecção de cigarras em nível médio de decomposição da DWTP. A próxima seção apresentará os resultados do mesmo sistema utilizando nível máximo de decomposição da DWTP.

#### 4.3.3.2 Grupo GR2

No tocante aos testes do grupo GR2, que processaram os sinais em nível máximo de decomposição e que consideram as classes “cigarra” e “ruído”, com o objetivo de validar o sistema de “detecção de cigarras”, apresentam-se na Figura 66 os diagramas de caixa com as acurácias dos testes com sinais de 10, 5, 2 e 1 segundos, utilizando 47 diferentes filtros *wavelet*, empregando 25 *features*. É aceitável verificar que, em todos os casos, a acurácia apresentou-se acima dos 90%, resultado superior aos obtidos nos testes em nível médio de decomposição (grupo anterior).

Na Tabela 15 estão organizados os resultados obtidos nos testes com 25 *features* do presente grupo. Pode-se notar que os melhores resultados apresentados estiveram acima dos 99%, significando que os parâmetros estão adequados para implantação em lavoura do sistema proposto.

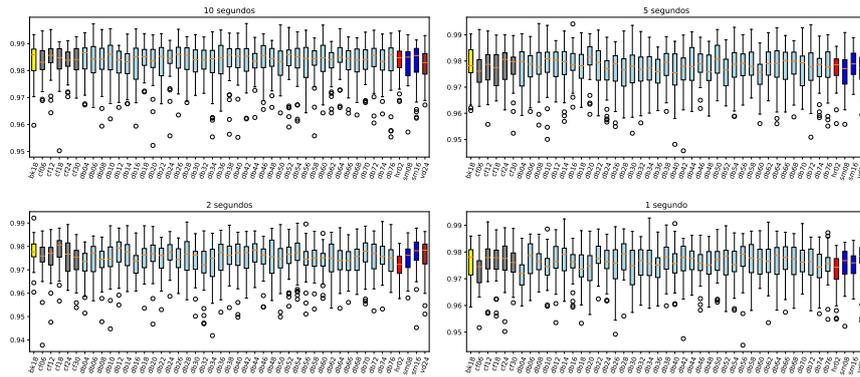


Figura 66 – Acurácia resultante dos testes utilizando 25 *features*.

Tabela 15 – Melhores resultados de testes em GR2 com 25 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db42	0.9974	0.49	1000	-	linear
5 segundos	db08	0.9943	0.54	1000	-	linear
2 segundos	bk18	0.9921	0.5	1000	-	linear
1 segundo	db34	0.9927	0.55	1000	-	linear

Fonte: Elaborada pelo autor.

Nas MCs ilustradas na Figura 67, apresentam-se os melhores resultados dos testes empregando 25 *features* em diferentes comprimentos de sinal. No caso dos testes utilizando arquivos de 10 segundos, pode-se ver que apenas 2 arquivos definidos como “cigarra” foram classificados como “ruído”. Destaca-se também a MC utilizando segmentos de 5 segundos, com filtro Daubechies-8, apresentando somente 4 falsos-negativos.

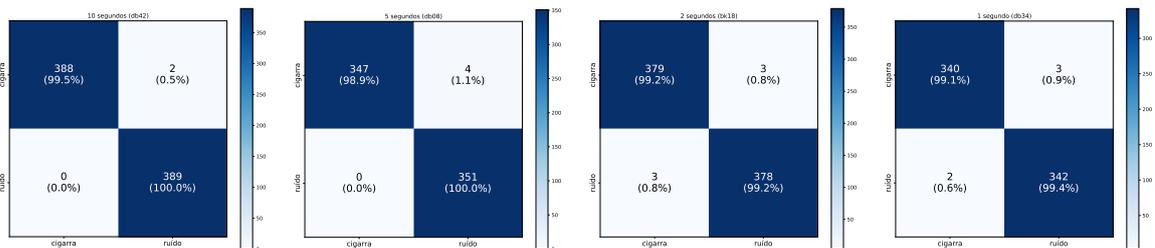


Figura 67 – Matrizes confusão resultantes dos testes utilizando 25 *features*.

A Figura 68 apresenta as curvas ROC com o melhor resultado (azul) e o pior resultado (vermelho) para sinais de 10 segundos, 5 segundos, 2 segundos e 1 segundo empregando 25 *features*. A área (AUC) resultante foi superior a 0.98 em todos os casos testados (incluindo os não apresentados nos gráficos), significando excelente desempenho na classificação das amostras.

Seguindo o padrão de testes do presente trabalho, testou-se o uso de uma quantidade de 20 *features*, utilizando sempre as melhores posições eleitas nos testes de EPC, onde

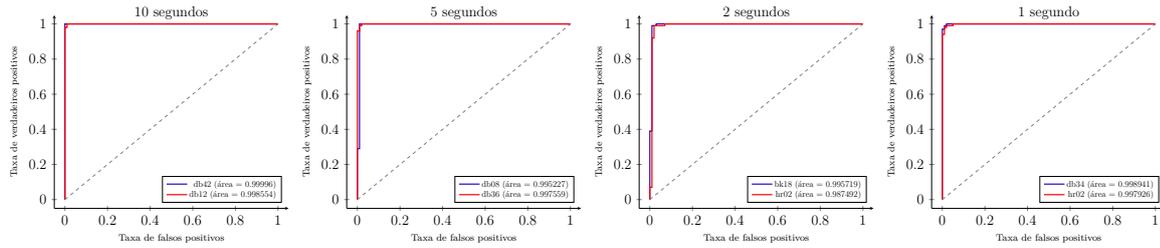


Figura 68 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 25 *features*.

também obteve-se altas taxas de acurácia conforme ilustra a Figura 69.

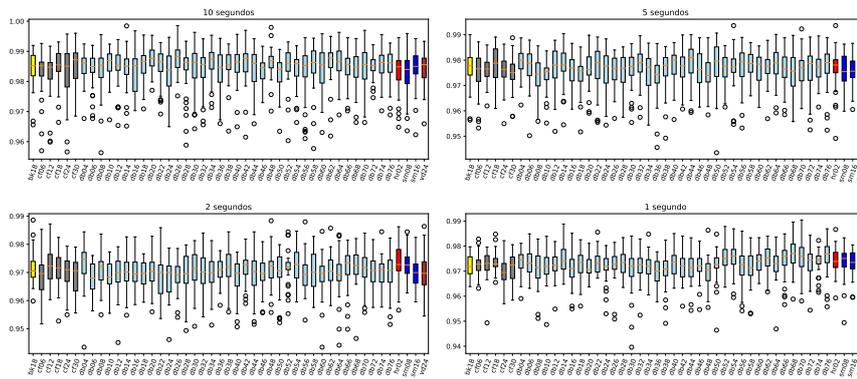


Figura 69 – Acurácia resultante dos testes utilizando 20 *features*.

Na Tabela 16 estão organizados os resultados obtidos nos testes com 20 *features* do presente grupo. Constata-se que a taxa de acurácia se manteve alta, mesmo extraindo 5 unidades de *features* do conjunto de treinamento e testes.

Tabela 16 – Melhores resultados de testes em GR2 com 20 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db26	0.9985	0.55	1000	-	linear
5 segundos	hr02	0.9936	0.59	1000	-	linear
2 segundos	bk18	0.9886	0.54	1000	-	linear
1 segundo	db70	0.9904	0.59	1000	-	linear

Fonte: Elaborada pelo autor.

O conjunto de MC dos melhores resultados obtidos nos testes com 20 *features*, apresentadas na Figura 70 mostram baixa taxa de falsos-negativos, com quantidade máxima de 8 amostras nos testes de 2 segundos, além de uma taxa de falsos-positivos quase nula, com uma única amostra de ruído classificada como cigarra, no caso dos sinais de 10 segundos.

As curvas ROC da Figura 71 confirmam o bom resultado nesta etapa de testes com 20 *features*, com taxa AUC mínima de 0.99 em todos os resultados, incluindo os piores casos.

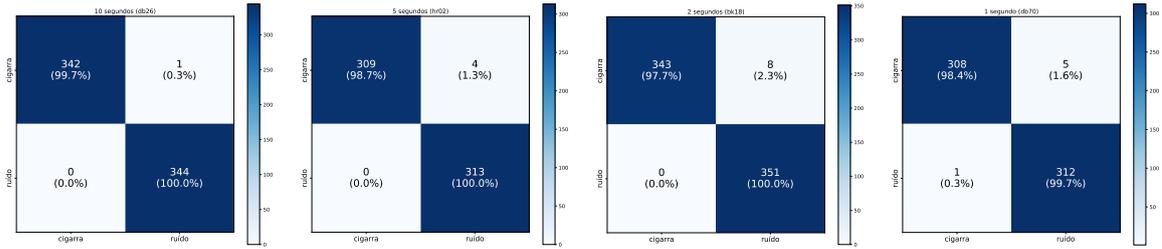


Figura 70 – MC resultante dos testes utilizando 20 *features*.

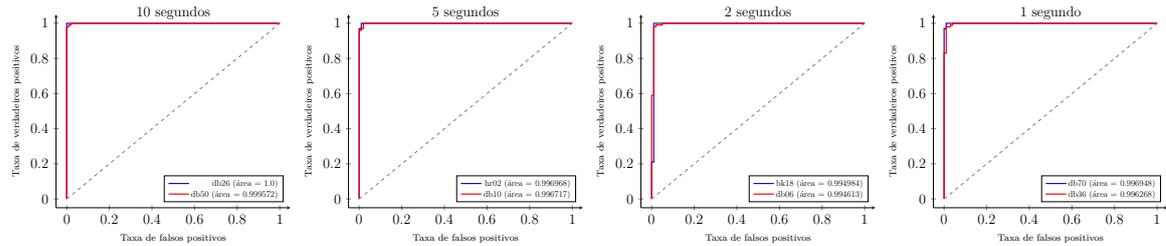


Figura 71 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 20 *features*.

A Figura 72 apresenta os gráficos boxplot dos resultados com as 15 primeiras *features*. Nota-se que, apesar de menores, os índices de acurácia continuam consideravelmente altos, apresentando valores médios entre 95% e 99%.

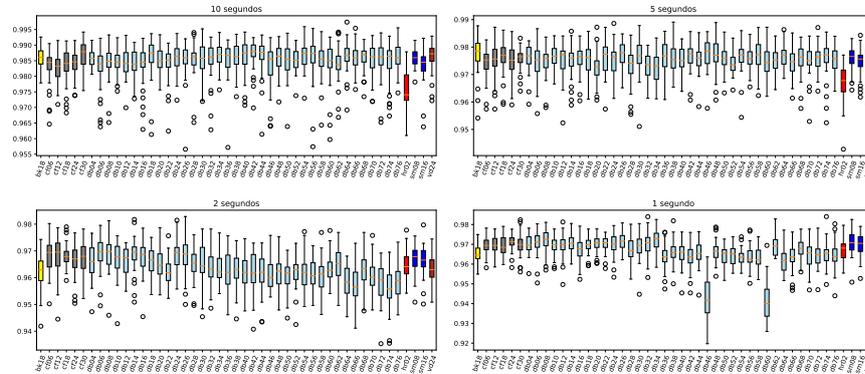


Figura 72 – Acurácia resultante dos testes utilizando 15 *features*.

Na Tabela 17 estão organizados os melhores resultados obtidos nos testes com 15 *features* do presente grupo, onde a acurácia mínima apresentou-se com índice de 98%, sendo superior no conjunto de testes com sinais de 10 segundos.

A fim de confirmar os bons resultados do conjunto de testes atual, apresenta-se o grupo de MCs com os melhores resultados utilizando as 15 melhores *features* (Figura 73), com índices de falsos-positivos nulos em metade dos casos, sendo que, na outra metade os índices foram considerados extremamente baixos, com apenas 2 arquivos sendo erroneamente positivamente classificados nos testes com arquivos de 5 segundos e, no máximo, 10 arquivos erroneamente classificados como falsos em todos os casos.

Tabela 17 – Melhores resultados de testes em GR2 com 15 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db64	0.9974	0.5	1000	-	linear
5 segundos	db38	0.9891	0.58	1000	-	linear
2 segundos	db26	0.9828	0.58	1000	-	linear
1 segundo	db74	0.984	0.59	1000	-	linear

Fonte: Elaborada pelo autor.

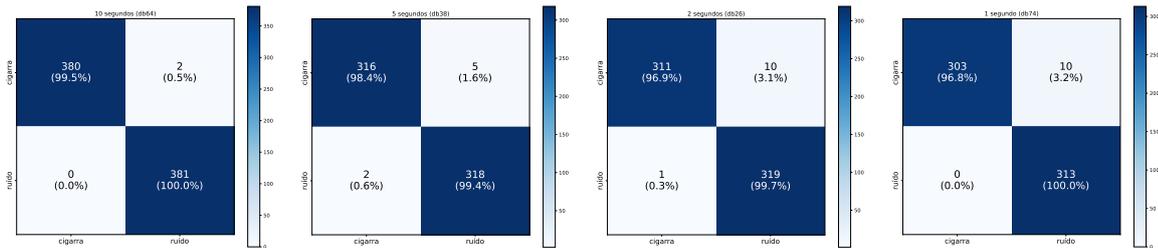


Figura 73 – Matriz confusão resultante dos testes utilizando 15 *features*.

Comparando-se esses resultados com os da Figura 74, confirmamos o bom resultado empregando 15 *features* com taxas de AUC acima de 0.99, mesmo nos piores casos apresentados.

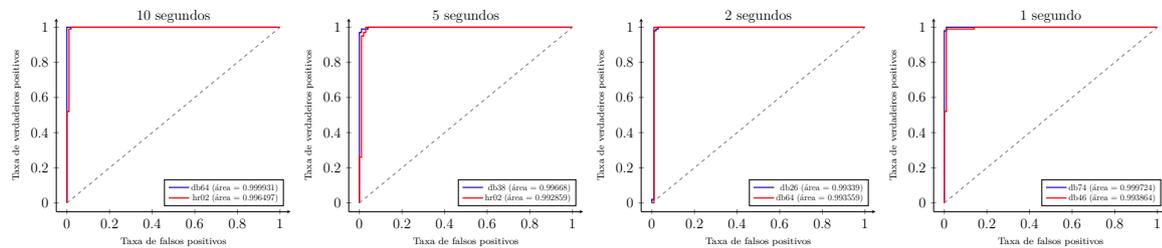


Figura 74 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 15 *features*.

Os resultados com 10 *features* se mostraram promissores, onde alguns filtros apresentaram resultado médio acima de 98% nos segmentos de 10 segundos, 97% nos segmentos de 5 segundos, 95% nos segmentos de 2 segundos e de 1 segundo (Figura 75).

Na Tabela 18 estão organizados os resultados obtidos nos testes com 10 *features* do presente grupo, predominando o uso do *kernel* linear com parâmetro de suavização  $C=1000$  do classificador.

Ainda em relação aos testes com 10 *features*, metade das MCs não apresentou falsos-positivos e a outra metade apresentou valores considerados baixos, com 0.4% e 1.2% nos segmentos de 2 e 1 segundo, respectivamente.

Também incorporando importantes resultados, a Figura 77 apresenta AUCs acima de 99% em todos os melhores resultados, com índice AUC geral mínimo de 96%.

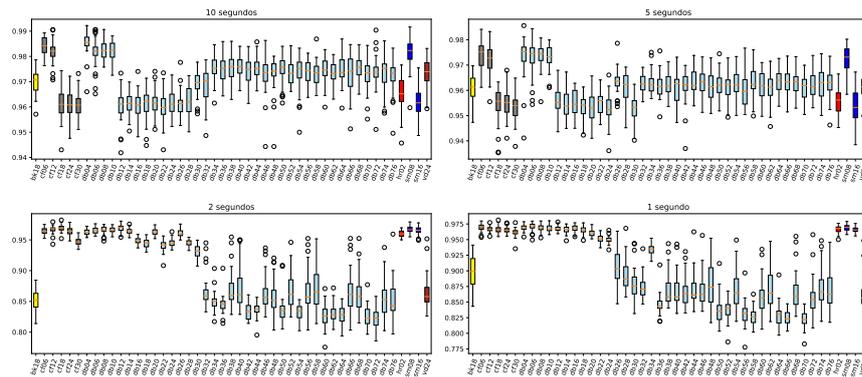


Figura 75 – Acurácia resultante dos testes utilizando 10 *features*.

Tabela 18 – Melhores resultados de testes em GR2 com 10 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	db04	0.9922	0.58	1000	-	linear
5 segundos	db04	0.9856	0.59	1000	-	linear
2 segundos	cf18	0.9822	0.41	1000	-	linear
1 segundo	cf18	0.9813	0.58	1000	-	linear

Fonte: Elaborada pelo autor.

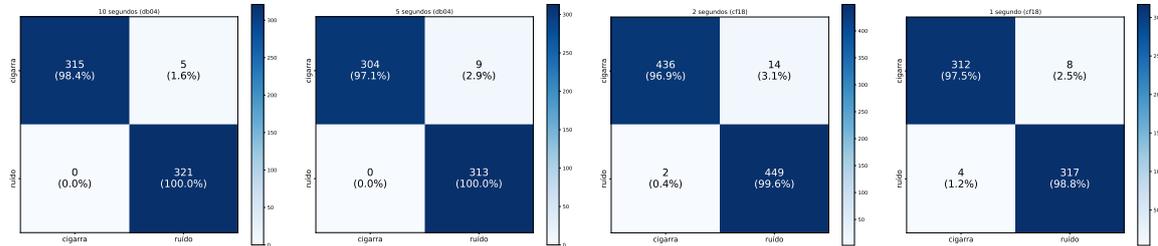


Figura 76 – Matriz confusão resultante dos testes utilizando 10 *features*.

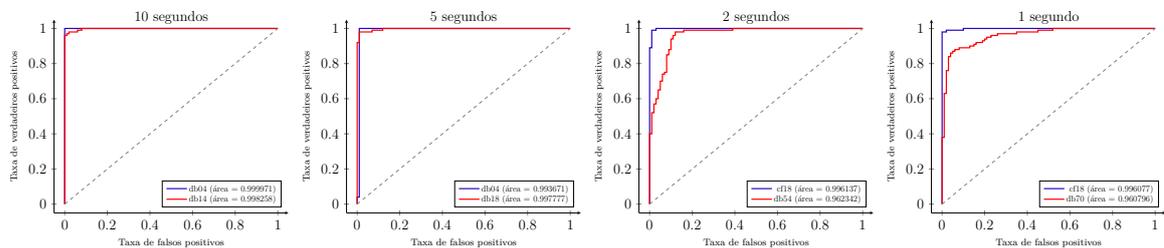


Figura 77 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 10 *features*.

A Figura 78 apresenta as acurácias obtidas nos testes empregando 5 *features*. Os resultados médios, a exemplo dos testes com 10 *features*, ficaram acima de 95% em muitos casos de testes utilizando filtros *wavelet* com menores valores de suporte.

Na Tabela 19 estão organizados os melhores resultados obtidos nos testes com 5 *features* do presente grupo, apresentando a predominância do uso do *kernel* linear com

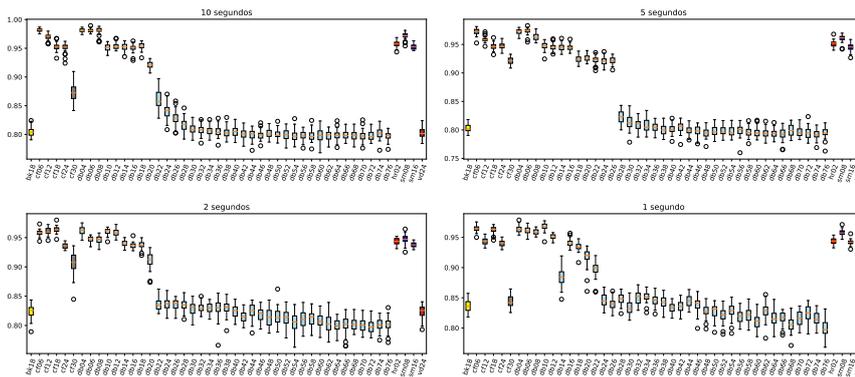


Figura 78 – Acurácia resultante dos testes utilizando 5 *features*.

$C=1000$ .

Tabela 19 – Melhores resultados de testes em GR2 com 5 *features*.

Sinal	Filtro	Acurácia	Treinamento	$C$	$\gamma$	Kernel
10 segundos	db06	0.9894	0.38	1000	-	linear
5 segundos	db06	0.983	0.46	1000	-	linear
2 segundos	cf18	0.9799	0.51	1000	-	linear
1 segundo	db04	0.9786	0.54	1000	-	linear

Fonte: Elaborada pelo autor.

No gráficos da Figura 79 nota-se que, na maioria dos casos, o número de falsos-positivos foi nulo, apresentando discrepância somente no melhor resultado utilizando segmentos de 5 segundos, onde uma amostra de ruído foi erroneamente classificada como cigarra (falso-positivo). O maior índice de erros ocorreu com os falsos-negativos, chegando a mais de 4% com filtro Daubechies-4 em segmentos de 1 segundo.

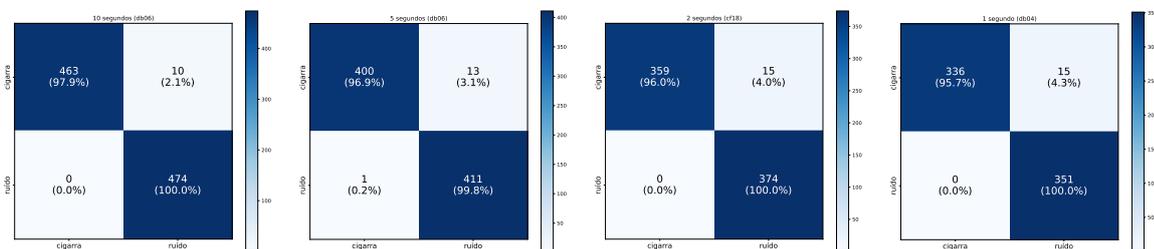


Figura 79 – Matriz confusão resultante dos testes utilizando 5 *features*.

As curvas ROC ilustradas na Figura 80 apresentam os melhores resultados utilizando 5 *features*. É possível verificar que os melhores resultados apresentaram AUC acima de 0.99 e o pior resultado atingido foi de 0.89, com segmentos de 5 segundos.

Para finalizar o presente grupo de testes, apresentamos os gráficos boxplot com os testes realizados com apenas as duas melhores *features*. Pode-se verificar na Figura 81, que

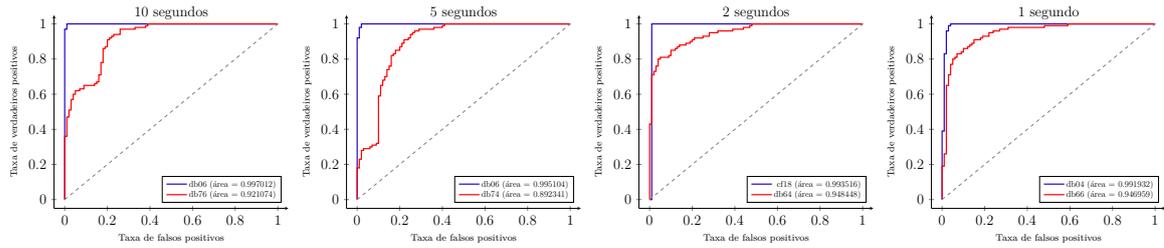


Figura 80 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 5 *features*.

houveram casos onde os resultados médios atingiram patamares acima dos 90%, mesmo empregando somente 2 *features*.

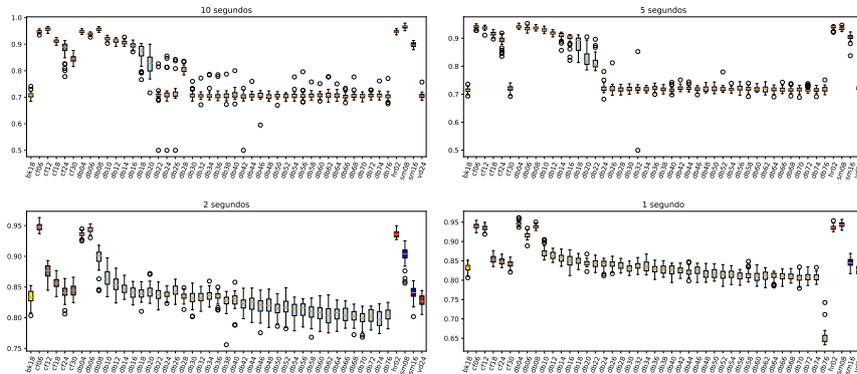


Figura 81 – Acurácia resultante dos testes utilizando 2 *features*.

Na Tabela 20 estão organizados os resultados obtidos nos testes com 2 *features* do presente grupo. A acurácia máxima foi obtida nos testes com segmentos de 10 segundos (aproximadamente 98%), entretanto, mesmo no sinais de 1 segundo atingiu-se um alto valor de acurácia (aproximadamente 96%).

Tabela 20 – Melhores resultados de testes em GR2 com 2 *features*.

Sinal	Filtro	Acurácia	Treinamento	C	$\gamma$	Kernel
10 segundos	sm08	0.9792	0.59	1000	-	linear
5 segundos	db04	0.9537	0.59	100	-	linear
2 segundos	cf06	0.963	0.54	1000	-	linear
1 segundo	db04	0.9617	0.59	1000	-	linear

Fonte: Elaborada pelo autor.

Seguem no conjunto de gráficos de MCs apresentados na Figura 82 os melhores resultados obtidos nos testes com 2 *features*, com o filtro Symlet-8, no conjunto de 10 segundos, apresentando somente 13 arquivos de áudio de cigarra reconhecidos como ruído, com 300 unidades sendo corretamente reconhecidos nessa classe.

Mostrando-se acima de 0.98, as AUCs apresentadas na Figura 83 apresentam resultado condizente ao esperado, mostrando ser possível a implantação do sistema em

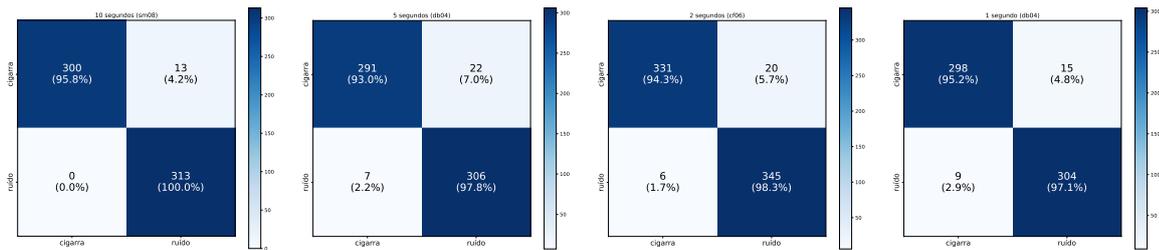


Figura 82 – Matriz de confusão resultante dos testes utilizando 2 *features*.

lavouira utilizando somente duas energias, provendo simplificação do sistema com diminuição de consumo dos recursos de energia, processamento, memória e principalmente transmissão de dados entre os dispositivos espalhados na lavouira.

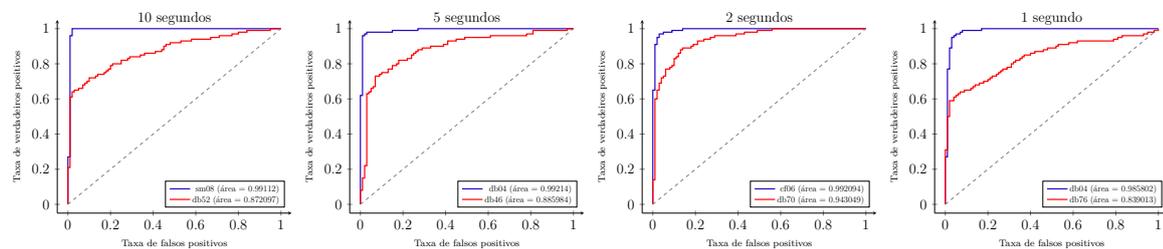


Figura 83 – Curvas ROC: melhor (azul) e pior (vermelho) resultado utilizando 2 *features*.

Nesta subseção foram apresentados os resultados dos testes do sistema de detecção de cigarras em nível máximo de decomposição da DWTP. As próximas duas subseções encerram os testes práticos do presente capítulo com os testes do sistema de densidade de cigarras. A primeira apresenta os resultados em nível médio de decomposição (GR3) e a segunda em nível máximo de decomposição da DWTP (GR4).

#### 4.3.3.3 Grupo GR3

Esta subseção apresenta os resultados obtidos nos testes do grupo GR3, que consideram as classes “alta”, “baixa” e “ruído”, com o objetivo de validar o sistema de “densidade de cigarras”, processando os sinais de 1 segundo (44.100 amostras) em nível médio de decomposição.

Como é sabido, a SVM é um classificador binário em sua essência. Na biblioteca *scikit-learn*, ao submeter à SVM um vetor com mais de duas classes, considera o problema como *multiclass*, empregando a técnica ‘um-contra-resto’ por padrão em sua função de decisão. Nesse caso, a classe eleita é a que retorna a maior pontuação no conjunto de comparações (PEDREGOSA *et al.*, 2011).

É importante salientar a política de desempate empregada na *scikit-learn*, biblioteca utilizada no presente trabalho. Atuando como classificador binário, resultados positivos elegem a segunda classe definida e, no caso de múltiplas classes, a classe com maior

pontuação é a eleita, de acordo com a ordem definida no vetor inicial. Assim sendo, no caso de empate, por padrão, a primeira classe do conjunto é considerada vencedora (binária) ou a classe de menor índice no vetor (multiclasse).

Apesar de não ter sido a opção escolhida no presente trabalho, no modo *multiclass*, a biblioteca *scikit-learn* possibilita definir uma política de desempate no classificador, empregando os valores de confiança da função de decisão para quebrar empates, caso necessário.

Na Figura 84 apresentam-se os diagramas de caixa com as acurácias dos testes utilizando 47 diferentes filtros *wavelet* e empregando as melhores *features* obtidas nos testes de EPC realizados nas primeiras seções do presente capítulo.

Pode-se contemplar resultados abaixo do esperado para um sistema de classificação, sendo que, na maioria dos casos, a acurácia média apresentou-se, em geral, abaixo dos 35%, sendo esse resultado considerado insatisfatório.

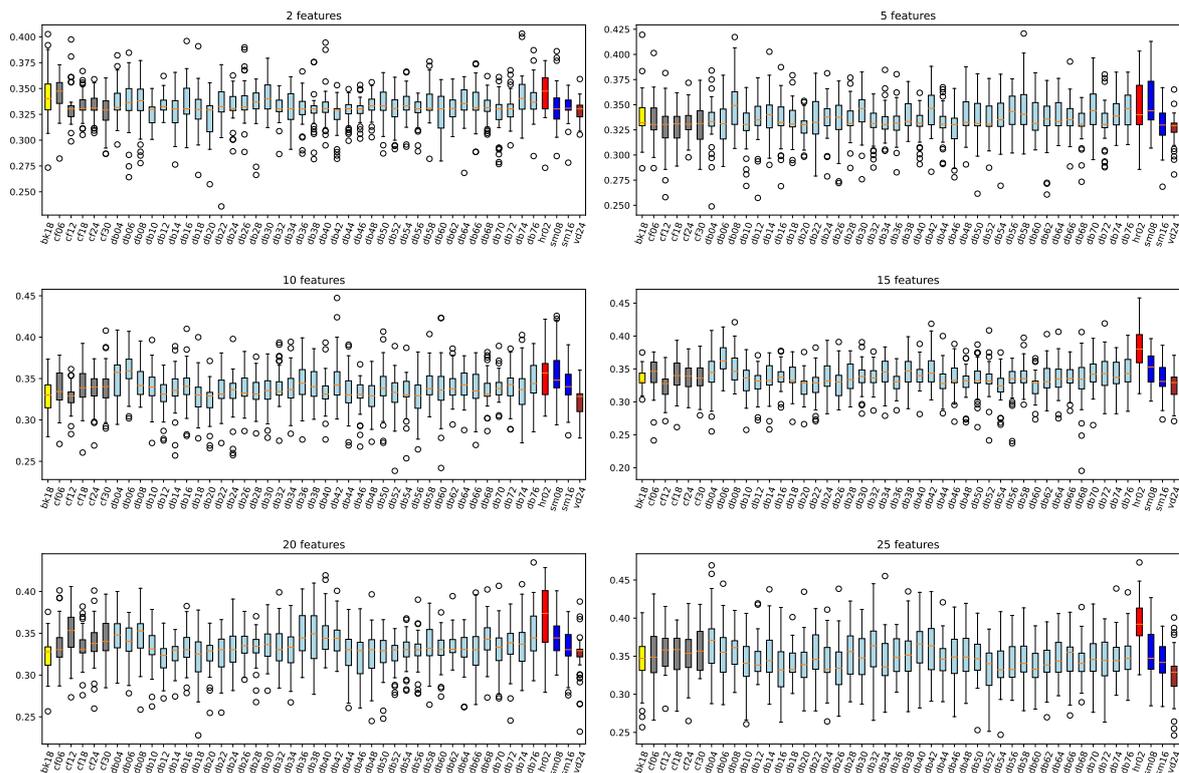


Figura 84 – Acurácia resultante em GR3

Na Tabela 21 estão apresentados os melhores resultados obtidos nos testes do presente grupo, sendo que, o maior percentual de acurácia (47.31%) foi apresentado pelo filtro Haar no conjunto de testes com 25 *features*.

Nas MCs ilustradas na Figura 85, apresentam-se os melhores resultados obtidos nos testes do grupo atual. É notável a confirmação do resultado abaixo do esperado, incluindo o melhor caso de teste utilizando 5 *features* que alcançou o percentual de 65.6%

Tabela 21 – Melhores resultados de testes em GR3.

<i>Features</i>	Filtro	Acurácia	Treinamento	C	$\gamma$	<i>Kernel</i>
2	db74	0.4032	0.41	1000	-	linear
5	db58	0.4208	0.42	1000	-	linear
10	db42	0.4474	0.52	1000	-	linear
15	hr02	0.4578	0.21	1000	-	linear
20	db76	0.4348	0.49	100	-	linear
25	hr02	0.4731	0.47	10	-	linear

Fonte: Elaborada pelo autor.

de falsos-positivos no piores casos. Destaca-se também o baixo índice de acertos da classe “alta” utilizando 2 *features*, com apenas 3.2%.

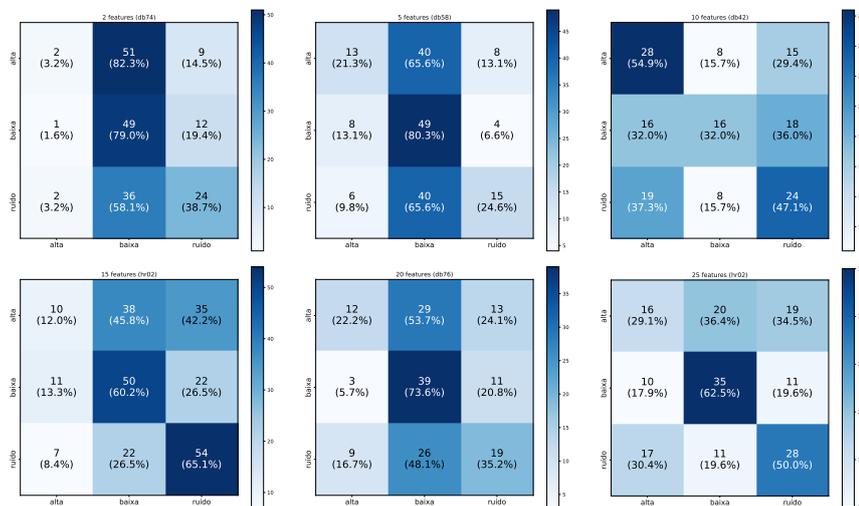


Figura 85 – Matrizes confusão resultantes em GR3.

A Figura 86 apresenta as curvas ROC com o melhor resultado (azul) e o pior resultado (vermelho), empregando as melhores *features*. Nos melhores resultados, a área (AUC) resultante foi inferior a 0.65, confirmando o baixo rendimento do classificador utilizando sinais transformados pela DWTP em nível médio.

Nesta subseção foram apresentados os resultados dos testes do sistema de densidade de cigarras em nível médio de decomposição da DWTP. A próxima seção apresentará os resultados do mesmo sistema utilizando nível máximo de decomposição da DWTP.

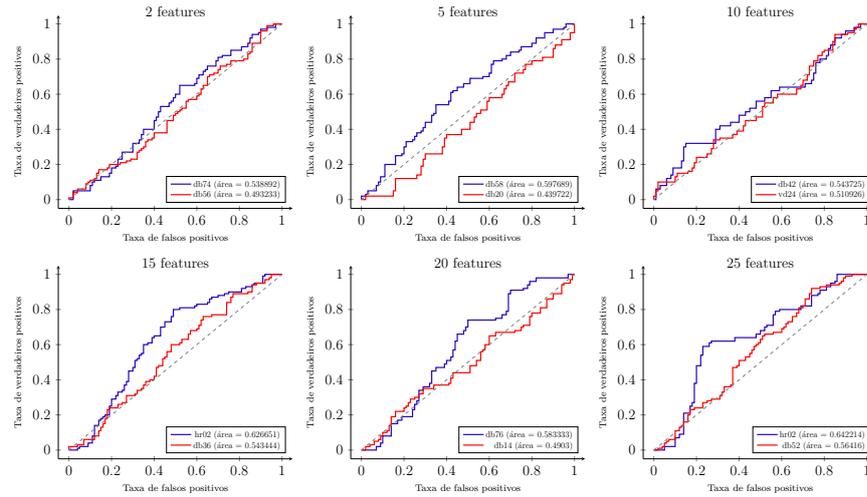


Figura 86 – Curvas ROC resultantes em GR3: melhor resultado (azul) e pior resultado (vermelho).

#### 4.3.3.4 Grupo GR4

Esta subseção apresenta os resultados obtidos nos testes do grupo GR4, que consideram as classes “alta”, “baixa” e “ruído”, com o objetivo de validar o sistema de “densidade de cigarras”, processando os sinais de 1 segundo (44.100 amostras), agora em nível máximo de decomposição.

Na Figura 87 apresentam-se, a exemplo da seção anterior, os diagramas de caixa com as acurácias dos testes utilizando 47 diferentes filtros *wavelet* e empregando as melhores *features* obtidas nos testes de EPC realizados nas primeiras seções do presente capítulo.

Pode-se contemplar os excelentes resultados obtidos onde, na maioria dos casos, a acurácia média apresentou-se acima dos 92%, com exceção de alguns filtros no caso de 2 *features*, sendo esse resultado considerado condizente com as expectativas iniciais.

Na Tabela 22 estão apresentados os melhores resultados obtidos nos testes do presente grupo. São notáveis os excelentes índices de acurácia em todos os casos, com valores acima de 97%.

Tabela 22 – Melhores resultados de testes em GR4

<i>Features</i>	Filtro	Acurácia	Treinamento	C	$\gamma$	<i>Kernel</i>
2	db72	0.9706	0.57	1000	-	linear
5	db72	0.993	0.55	1000	-	linear
10	db64	1.0	0.54	1000	-	linear
15	cf12	0.9933	0.53	1000	-	linear
20	db72	1.0	0.47	100	-	linear
25	db74	1.0	0.5	100	-	linear

Fonte: Elaborada pelo autor.

Nas MCs ilustradas na Figura 88, apresentam-se os melhores resultados obtidos

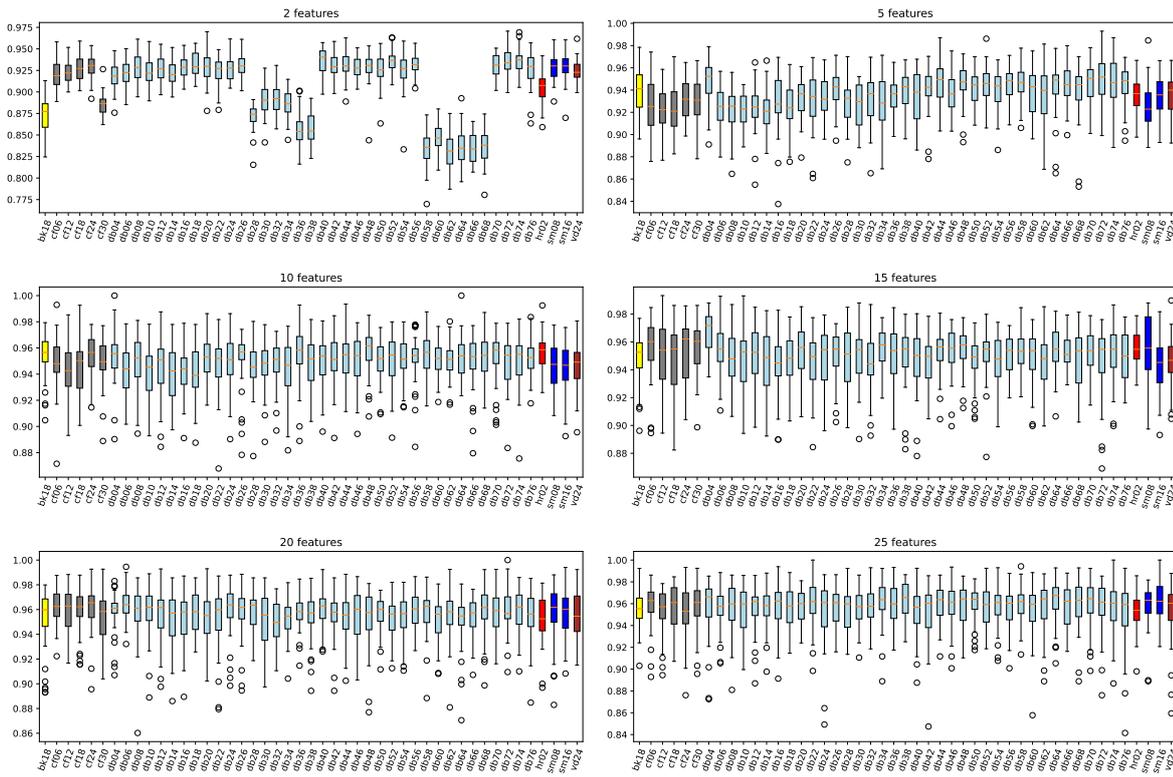


Figura 87 – Acurácia resultante em GR4.

nos testes do grupo atual. É notável a confirmação dos bons resultados, onde, na maioria dos casos o número de falsos-positivos e falsos-negativos foi nulo. Nos demais casos esse percentual ficou abaixo de 2.2%, excetuando-se o caso utilizando 2 *features* em que a incidência de falsos-positivos apresentou-se em 6.5%.

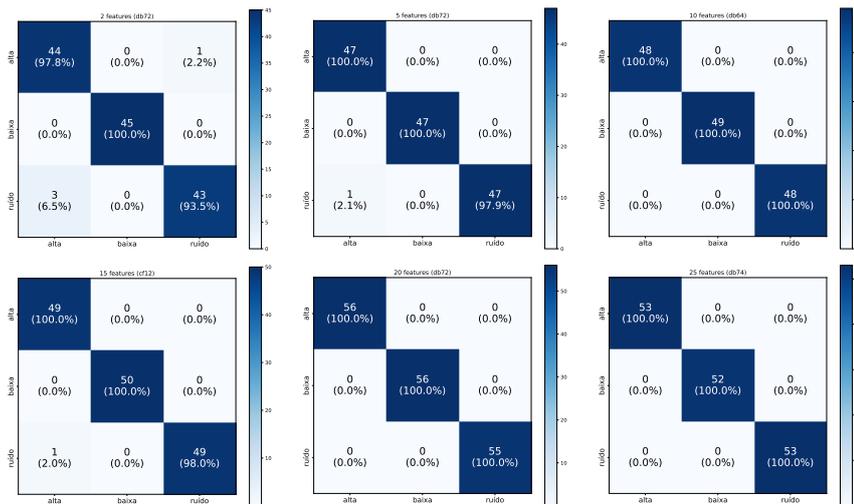


Figura 88 – Matrizes confusão resultantes em GR4.

A Figura 89 apresenta as curvas ROC com o melhor resultado (azul) e o pior resultado (vermelho), empregando as melhores *features*. Pode-se verificar que a taxa, em todos os casos, foi superior a 0.99%.

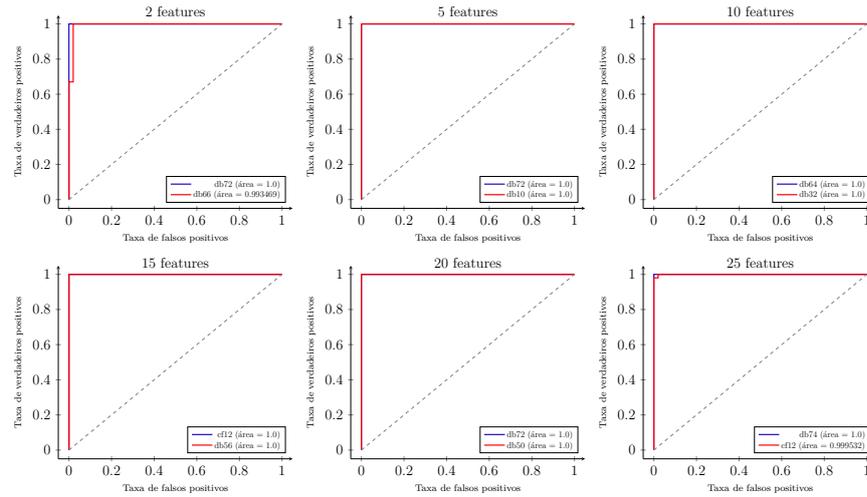


Figura 89 – Curvas ROC resultantes em GR4: melhor resultado (azul) e pior resultado (vermelho).

Nesta subseção foram apresentados os resultados dos testes do sistema de densidade de cigarras em nível máximo de decomposição da DWTP. A próxima seção, última do presente capítulo, apresentará o protótipo em IoT desenvolvido para implantação do presente sistema em lavoura de café.

#### 4.4 Protótipo em IoT

Entre as principais soluções possíveis com IoT está a implantação de aplicações de baixo custo, tamanho reduzido e baixo consumo de energia, nas mais diversas áreas (ATZORI; IERA; MORABITO, 2010). Assim, desenvolveu-se um protótipo, ilustrado na Figura 90, que contém um algoritmo que processa os módulos aqui apresentados com o objetivo de capturar sinais na lavoura e enviar esses dados para um servidor em nuvem. Com base nesses atributos, um relatório é apresentado ao agricultor por meio de uma página *Web*.

O dispositivo, nomeado INS em homenagem ao professor Ivan Nunes da Silva, utiliza o conceito de *Edge Computing* (AI; PENG; ZHANG, 2018) que possibilita integração de dados coletados em dispositivos IoT com servidores em nuvem. Um conjunto de dispositivos forma uma rede *Mesh* (NANDIRAJU *et al.*, 2007) na qual os dados são trocados entre os dispositivos até que um *Gateway* seja encontrado e o conjunto de dados seja enviado e armazenado na Internet.

O processamento da DWTP e a obtenção das energias é feito no próprio microcontrolador, resultando em uma *String* que contém as 25 energias da Escala Bark (reduzidas a 8 energias após a EPC), obtidas a partir do módulo de som MAX9814 (Figura 92-J) com segmentos de aproximadamente 1000 milissegundos amostrados a uma taxa de, aproximadamente, 13 kHz, dispendendo, em média, 56522 milissegundos por iteração. Essas



Figura 90 – Dispositivo INS acondicionado em uma caixa de plástico ABS.

energias são processadas no servidor, que faz a devida classificação e armazena o resultado para consulta posterior. No caso de algumas famílias de filtros *wavelet*, como Daubechies 34 e Daubechies 52, a acurácia chega a 100% com máxima resolução em frequência.

Conforme descrito anteriormente, a referida taxa de 13 kHz é suficiente para satisfazer o teorema de Nyquist, visto que os sinais aqui empregados apresentaram uma frequência máxima de aproximadamente 1.6 kHz, entretanto, o ideal nesse caso seria incluir um filtro analógico, limitando a frequência máxima de entrada no sistema.

Como primeiro protótipo, nomeado de “Guido” em homenagem ao Prof. Dr. Rodrigo Capobianco Guido, mini-PC *Orange Pi* (HAUKA *et al.*, 2021) foi empregado no processamento dos sinais (Figura 91), incorporando tanto a DWTP quanto a SVM, enviando para o *Gateway* a resposta final do classificador (DEZOTTI *et al.*, 2017; DEZOTTI *et al.*, 2019; ESCOLA *et al.*, 2020). Mais recentemente, buscando economia de recursos, um dispositivo empregando ESP32 foi desenvolvido (ESCOLA *et al.*, 2021).

Com a substituição do *Orange Pi* pelo microcontrolador ESP32, os recursos de *hardware* foram consideravelmente reduzidos. Para analisar a viabilidade de adoção desse dispositivo, um estudo foi realizado a fim de conhecer sua capacidade de processar a DWTP em campo. Os testes e análises de desempenho mostraram que o dispositivo é capaz de processar pouco mais de 1000 milissegundos de áudio por iteração, com taxa de amostragem de, aproximadamente, 13 kHz. Um artigo científico foi produzido e submetido a uma revista internacional arbitrada.

A análise de desempenho geral do INS empregando o microcomputador ESP32 é um dos objetivos do presente trabalho. Um conjunto de experimentos foi realizado

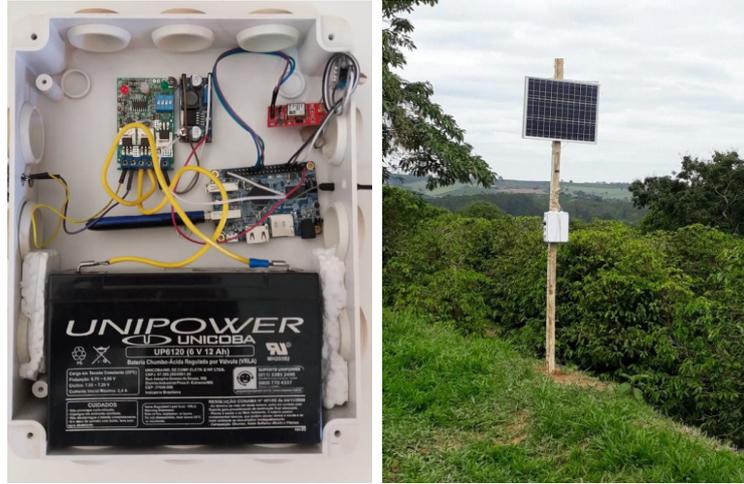


Figura 91 – Dispositivo “Guido”: disposição dos componentes internos (esquerda) e implantação na lavoura (direita).

em laboratório, variando-se os filtros *wavelet*, atingindo valores importantes de acurácia. Um artigo científico está em desenvolvimento e será submetido para revista científica internacional arbitrada.

Um conjunto de sensores integrados ao microcontrolador ESP32 (Figura 92-A), permite capturar dados que são incorporados à *String*, como temperatura/umidade (Figura 92-I), luminosidade (Figura 92-D), nível de chuva (Figura 92-E) e localização geográfica (Figura 92-B), agregando mais valor ao processo com a inclusão de informações adicionais que podem ser importantes na tomada de decisão do agricultor. O vetor resultante é transmitido pelo módulo de comunicação sem fio HC12 (Figura 92-C) para o próximo dispositivo INS ao alcance ou ao próprio *Gateway*. O gerenciamento de energia é realizado empregando energia solar por meio de uma placa fotovoltaica (Figura 92-G) e um controlador (Figura 92-F), armazenando-a em uma bateria de chumbo-ácido (Figura 92-H).

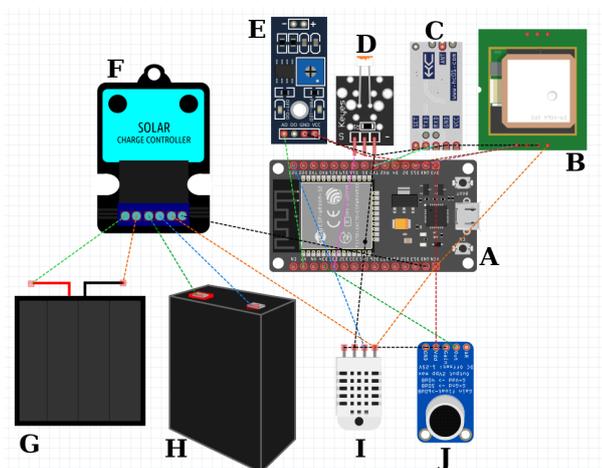


Figura 92 – Esquema de integração dos módulos físicos do INS.

Conhecer o alcance real de cada um dos dispositivos que serão espalhados na lavoura é um dos objetivos do presente trabalho, entretanto é notório que a implantação deverá ser realizada por amostragem, visto que as lavouras de café podem abranger áreas de milhares de hectares (Figura 93), assim, antes mesmo da implantação é necessário conhecer o alcance de cada dispositivo a fim de estimar o custo de implantação por hectare.

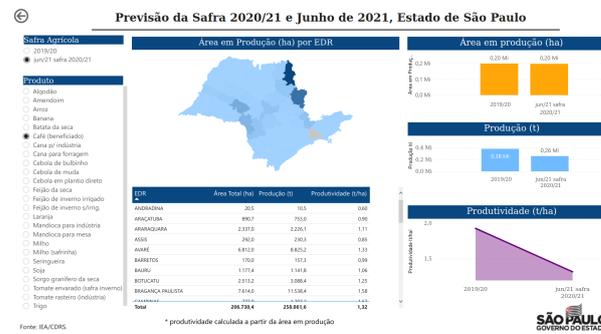


Figura 93 – Área de plantio de café no estado de São Paulo em 2021 (Fonte: (IEA... , 2021)).

O alcance do dispositivo INS na lavoura foi medido em experimentos práticos realizados em laboratório e publicados em (ESCOLA *et al.*, 2021), onde um *smartphone* emitindo som de cigarras foi posicionado a uma distância variando de 2m a 20m do dispositivo INS. A acurácia resultante foi superior a 84%.

O próximo passo no estudo da estimativa de alcance/ha será a implantação em área aberta, que deverá ser inicialmente realizada na unidade agrícola do IFSP, em Barretos-SP. Na Figura 94 apresentam-se dois protótipos de suporte para posicionamento do INS durante os experimentos agendados para medição do alcance do módulo de comunicação e, conseqüentemente, da eficiência da rede *mesh*, responsável por permitir a troca de dados entre os dispositivos espalhados na lavoura até serem encaminhados à borda (*Edge Computing*).

Para viabilizar a comunicação em rede *mesh* e gerenciar o envio/recebimento de dados pelos dispositivos INS, foi necessário o desenvolvimento de uma rotina personalizada, uma vez que as bibliotecas disponíveis na plataforma Arduino não permitiam esse recurso específico. Assim, foi desenvolvido um protocolo para implementação de redes *mesh* para aplicações de PDS em *Smart Farm*, sendo desenvolvido um artigo científico e submetido a uma revista arbitrada.

Com um conjunto de dispositivos espalhados pela lavoura, o processo de atualização do software integrado ao microcontrolador pode ser um processo penoso. Atrasos de resposta ocasionados por falhas de *hardware* ou *software* também podem resultar em transtornos, tornando necessário o deslocamento até o ponto de instalação. Normalmente a atualização de código-fonte em projetos com microcontroladores como o Arduino e o ESP32 é feita por meio de um cabo USB. Uma possibilidade para tornar o processo mais prático é a utilização



Figura 94 – Protótipos de suportes para posicionamento experimental de dispositivos INS.

da tecnologia *Over-The-Air* (OTA), que permite atualização do microcomputador por meio de sua interface WiFi.

Tendo sido a solução OTA implementada com sucesso no código-fonte do INS, iniciou-se o desenvolvimento de um sistema para monitoramento sem fio, a fim de registrar o tempo de atividade do dispositivo e conhecer, por exemplo, a autonomia da bateria ou falhas de recepção das *Strings* de dados. O sistema provê uma interface *online* e registra periodicamente na nuvem o *status* do dispositivo. Assim, um artigo científico foi escrito e enviado para revista arbitrada.

#### 4.5 Comentários Gerais Sobre os Resultados

A partir dos resultados obtidos nos experimentos com a EPC e o classificador SVM, pode-se considerar a viabilidade do sistema proposto diante do rendimento apurado em nível máximo de decomposição.

Durante os experimentos, os 47 diferentes filtros *wavelet* foram analisados, além da variação do tempo do sinal entre 1s e 10s e do nível de decomposição da DWTP com um conjunto de 2 a 25 energias provenientes da EB.

Nos grupos de testes GR1 e GR2, para o sistema de detecção de cigarras, foram empregados 1526 arquivos WAV de 10 segundos, variando para 5, 2 e 1 segundos nos testes. Já em GR3 e GR4, no sistema de densidade de cigarras, utilizou-se 315 arquivos de 1 segundo. Nesse segundo caso, a quantidade menor de sinais de entrada de comprimento mínimo se justifica pela dificuldade em separar os sinais capturados na lavoura, que majoritariamente são desbalanceados, na maioria de ruídos e de muitas cigarras.

Nos testes de EPC, que visam analisar a importância dos parâmetros de entrada antes da implementação de qualquer tipo de classificador, além das  $N$  primeiras *features*,

foram testadas as  $N$  melhores *features*. Essas foram eleitas, após a execução de um conjunto de testes que comparou as distâncias de VO com todas as combinações possíveis para cada conjunto de  $N$  entradas, variando entre 2 (300 combinações possíveis) e 25 (1 combinação possível), com máximo de combinações possível nos casos de 12 e 13 *features* (mais de 5 milhões de combinações).

O processo de cálculo das combinações demandou alto esforço computacional, sendo que, no caso de  $N = 12$  e  $N = 13$ , o número de possibilidades é  $> 5 \cdot 10^6$ , sendo necessários vários dias de processamento do algoritmo desenvolvido em Python. Para a solução do problema, a plataforma de programação paralela Cuda (SOYATA, 2018; ODEN, 2020) foi empregada e o tempo de processamento significativamente reduzido.

Com a EPC foi possível previamente analisar as *features* de entrada, obtidas a partir da EB, prevendo como resultado, que os sinais em nível médio de decomposição exigiriam maior esforço do classificador em relação aos sinais em nível máximo. Além disso, foi possível prever que os filtros com suportes mais baixos seriam mais propensos a bons resultados em detrimento aos de suportes mais altos. Em diversos casos o filtro Haar, de suporte 2, apresentou a menor distância para VO.

Em relação ao classificador SVM foram realizados diversos testes, alternando os parâmetros de treinamento: *kernel*, suavização e  $\gamma$ , apresentando resultados condizentes com as expectativas iniciais em nível máximo de decomposição, com acurácia acima de 90%. No caso do nível médio de decomposição, não foi possível encontrar configurações adequadas para o bom desempenho do classificador.

O desenvolvimento de um protótipo em IoT mostra que se buscou uma real implantação em lavoura, ou seja, uma solução de baixo custo que possa ser efetivamente utilizada no campo. Os mesmos algoritmos utilizados nos testes em laboratório foram portados para a linguagem de programação específica do dispositivo, atingindo acurácia de 100% com alguns filtros *wavelet*, em nível máximo de decomposição.

No próximo Capítulo apresentam-se as conclusões deste trabalho, discorrendo-se ainda sobre os trabalhos futuros.

## 5 CONCLUSÕES E TRABALHOS FUTUROS

Após a apresentação dos objetivos deste trabalho, análise de trabalhos correlatos e a devida revisão dos conceitos, foi apresentado um sistema de detecção e monitoramento de densidade de cigarras para implantação em lavouras de café, o qual auxilia no manejo de cigarras por permitir o mapeamento da incidência dos insetos em uma determinada área. Para viabilizá-lo, a EPC foi utilizada visando selecionar as mais relevantes *features* extraídas no domínio conjunto *wavelet*-Bark e oriundas de sinais acústicos coletados em campo nos estados de São Paulo e Minas Gerais.

O processamento inicial dos sinais acústicos foi realizado por meio da DWTP, utilizando as energias provenientes da EB, que provê um total de 25 escalares, mas que, de acordo com os testes, não são, em sua totalidade, necessários para um ótimo desempenho do classificador utilizado, isto é a SVM. No melhor caso, apenas 2 dos 25 atributos de entrada fornecidos pela EB foram necessários para uma acurácia superior a 90%, demonstrando a relevância da EPC para a tarefa em questão.

Adicionalmente, apresentou-se um protótipo físico que utiliza IoT para capturar o som em determinado ponto da lavoura e envia dados a um servidor remoto, que é acessado pelo administrador da lavoura, visando acessar informações como temperatura, umidade, posição geográfica, incidência de chuva, luminosidade, além da densidade de cigarras presentes no momento da leitura.

Uma possibilidade para trabalhos futuros é a inclusão de uma rotina de pré-processamento do sinal no INS a fim de testar um limiar mínimo de energia ou entropia do sinal antes de transmitir o resultado, buscando reduzir falsos positivos e economizando recursos. Entretanto, especializando o dispositivo IoT apresentado, essa abordagem reduziria a flexibilidade existente atualmente no sistema, que permite a substituição do classificador na nuvem, podendo assim, ser usado para outras aplicações como monitoramento de pássaros ou de caça predatória. O desempenho do algoritmo desenvolvido pode ser ainda investigado com base na associação dele com outros algoritmos tais como *Principal Component Analysis* (ZOU; XUE, 2018; ABONIZIO *et al.*, 2020), *Independent Component Analysis* (SOMPAIRAC *et al.*, 2019) e *Random Forest Importance* (SANTANA *et al.*, 2018), mostrando-se excelente oportunidade de pesquisa futura.

Outra possibilidade para trabalhos futuros é o desenvolvimento do dispositivo de monitoramento INS 2.0, substituindo o microcontrolador por um *smartphone*. Visto que esse dispositivo possui conectividade WiFi e *Bluetooth*, GPS, microfone, bateria embutidos, normalmente com alta capacidade de memória e processamento, a substituição permitiria economia de recursos utilizando lixo eletrônico (E-Waste) (FORTI *et al.*, 2020),

possibilitando testes de implementação da DWTP, possivelmente com sinais mais longos.

Deseja-se também, em trabalhos futuros, ampliar o número de sinais de áudio de entrada para reanalisar a resposta do sistema, incluindo o uso de outras técnicas como as descritas em Guido (2019) e aplicando as técnicas para monitoramento de *Quesada gigas* em outros ambientes como Cerrado Stricto-Sensu e Mata de Galeria, no estado de Goiás, onde as espécies emitem sinais em frequências mais altas. Também deseja-se fazer a devida implantação de um conjunto de dispositivos de captura em uma lavoura de café no campus Muzambinho do Instituto Federal do Sul de Minas Gerais para analisar a viabilidade do sistema em campo.

Por fim, mas não menos importante, destaca-se que a contribuição deste trabalho é original: não há registro na literatura científica do uso da EPC para a análise de sinais acústicos representados no domínio conjuntowavelet-Bark, para a finalidade em questão. Assim, acredita-se que este trabalho, o qual encontra-se publicado em periódico especializado e com alto fato de impacto, tenha efetivamente contribuído com as pesquisas da área.

## REFERÊNCIAS

- ABONIZIO, H. Q. *et al.* Language-independent fake news detection: English, portuguese, and spanish mutual features. **Future Internet**, Multidisciplinary Digital Publishing Institute, v. 12, n. 5, p. 87, 2020.
- ACEMOGLU, D.; RESTREPO, P. **8. Artificial Intelligence, Automation, and Work**. [S.l.: s.n.]: University of Chicago Press, 2019.
- ADDISON, P. S. **The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance**. [S.l.: s.n.]: CRC press, 2017.
- AI, Y.; PENG, M.; ZHANG, K. Edge computing technologies for internet of things: a primer. **Digital Communications and Networks**, Elsevier, v. 4, n. 2, p. 77–86, 2018.
- AIDE, T. M. *et al.* Real-time bioacoustics monitoring and automated species identification. **PeerJ**, PeerJ Inc., v. 1, p. e103, 2013.
- AKAY, M. **Time Frequency and Wavelets in Biomedical Signal Processing**. [S.l.: s.n.]: IEEE press series in Biomedical Engineering, 1998.
- ALMEIDA, J. E. M. de. Controle biológico de cigarras-do-cafeeiro. **X REUNIÃO ITINERANTE DE FITOSSANIDADE DO INSTITUTO BIOLÓGICO CAFÉ**, p. 101, 2004.
- ALONSO, R. S. *et al.* An intelligent edge-iot platform for monitoring livestock and crops in a dairy farming scenario. **Ad Hoc Networks**, Elsevier, v. 98, p. 102047, 2020.
- ALVES, L. *et al.* A mesh sensor network based on bluetooth: Comparing topologies to crop monitoring. *In*: SBC. **Anais Estendidos do IX Simpósio Brasileiro de Engenharia de Sistemas Computacionais**. [S.l.: s.n.], 2019. p. 125–130.
- ANDRADE, S. d. C. Aspectos bioecológicos de *quesada gigas* (olivier, 1790)(hemiptera: Cicadidae) associados à cultura do café. Universidade Estadual Paulista (UNESP), 2018.
- ANITESCU, C. *et al.* Artificial neural network methods for the solution of second order boundary value problems. **Computers, Materials and Continua**, Tech Science Press, v. 59, n. 1, p. 345–359, 2019.
- ARFKEN, G. B.; WEBER, H. J. **Mathematical Methods for Physicists**. 6th ed. ed. [S.l.: s.n.]: Elsevier, 2005.
- ARTUZO, F. D.; FOGUESATTO, C. R.; SILVA, L. X. da. Agricultura de precisão: inovação para a produção mundial de alimentos e otimização de insumos agrícolas. **Revista Tecnologia e Sociedade**, Universidade Tecnológica Federal do Paraná, v. 13, n. 29, p. 146–161, 2017.
- ATZORI, L.; IERA, A.; MORABITO, G. The internet of things: A survey. **Computer networks**, Elsevier, v. 54, n. 15, p. 2787–2805, 2010.

BASSOI, L. H. *et al.* Agricultura de precisão e agricultura digital. **Embrapa Pecuária Sudeste-Artigo em periódico indexado (ALICE)**, TECCOGS, n. 20, jul./dez., 2019., 2019.

BLUMER, A. *et al.* Occam's razor. **Information processing letters**, Elsevier, v. 24, n. 6, p. 377–380, 1987.

BOSI, M.; GOLDBERG, R. E. **Introduction to digital audio coding and standards**. [S.l.: s.n.]: Springer Science & Business Media, 2012. v. 721.

CAOBIANCO, L. G.; GUIDO, R. C.; SILVA, I. N. da. Wavelet-based features selected with paraconsistent feature engineering successfully classify events in low-voltage grids. **Measurement**, Elsevier, v. 170, p. 108711, 2021.

CARDOSO, J. R. F. Desenvolvimento de estrutura robótica para aquisição e classificação de imagens (eraci) de lavoura de cana-de-açúcar. Universidade Estadual Paulista (UNESP), 2020.

CARDOSO, J. R. F. *et al.* Development of a robotic structure for acquisition and classification of images (eraci) in sugarcane crops. **Revista Ciência Agronômica**, v. 51, p. 1–15, 2020.

CARDOSO, J. R. F. *et al.* Otimização de recursos de hardware de estrutura robótica para aquisição e classificação de imagens (eraci) por meio da distribuição de processos em rede. RESI Uemg Frutal, 2020.

CARNIELLI, W. A.; CONIGLIO, M. E. **Paraconsistent logic: Consistency, contradiction and negation**. [S.l.: s.n.]: Springer, 2016.

CHEN, C.-H.; SHIAO, S.-F. Two new species of the genus euterpnosia matsumura (hemiptera: Cicadidae) from taiwan. **The Pan-Pacific Entomologist**, BioOne, v. 84, n. 2, p. 81–91, 2008.

CHEN, S.-H.; CHEN, S.-H.; GUIDO, R. C. Music genre classification algorithm based on dynamic frame analysis and support vector machine. *In*: IEEE. **Multimedia (ISM), 2010 IEEE International Symposium on**. [S.l.: s.n.], 2010. p. 357–361.

CHIAVACCI, S. J.; BEDNARZ, J. C.; MCKAY, T. The emergence densities of annual cicadas (hemiptera: Cicadidae) increase with sapling density and are greater near edges in a bottomland hardwood forest. **Environmental Entomology**, v. 44, p. 859–867, 2015. ISSN 1938-2936.

CLARIDGE, M. Acoustic signals in the homoptera: behavior, taxonomy, and evolution. **Annual review of entomology**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 30, n. 1, p. 297–317, 1985.

COLE, J. A. A new cryptic species of cicada resembling tibicen dorsatus revealed by calling song (hemiptera: Auchenorrhyncha: Cicadidae). **Annals of the Entomological Society of America**, BioOne, v. 101, n. 5, p. 815–823, 2008.

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, Springer, v. 20, n. 3, p. 273–297, 1995.

DARDAR, M. A.; BELAL, H. M.; BASHEER, A. M. The occurrence of the cicada *cicadatra persica* on apple trees, *malus domestica*, in erneh, syria. **Journal of Insect Science**, v. 13, n. 1, p. 42, 2013. Disponível em: +<http://dx.doi.org/10.1673/031.013.4201>.

DENG, L.; O'SHAUGHNESSY, D. **Speech processing: a dynamic and optimization-oriented approach**. [*S.l.: s.n.*]: CRC Press, 2018.

DEPRAETERE, M. *et al.* Monitoring animal diversity using acoustic indices: implementation in a temperate woodland. **Ecological Indicators**, Elsevier, v. 13, n. 1, p. 46–54, 2012.

DEZOTTI, A. K. *et al.* Dispositivo de monitoramento de densidade populacional de insetos a partir de sinais acústicos emitidos. **Brazilian Journal of Animal and Environmental Research**, v. 2, n. 5, p. 1781–1785, 2019.

DEZOTTI, A. K. *et al.* Software capaz de reconhecer cigarras no cafeeiro. *In*: CALIXTO, E. S.; SILINGARDI, H. M. T. (ed.). **Temas Atuais em Ecologia Comportamental e Interações**. Uberlândia, MG: Editora Composer, 2017. cap. 1, p. 657–674.

D'UTRA, G. Cicadas in coffee plantations (in portuguese). **Boletim da Agricultura**, v. 9, n. 5, 1908.

ESCOLA, J. P. L. **Reconhecimento de assinaturas baseado em seus ruídos caligráficos**. 2014. Dissertação (Mestrado) — Instituto de Física de São Carlos, University of São Paulo, São Carlos-SP, 2014.

ESCOLA, J. P. L.; DOCUSSE, T. A. Serialização de dados em processamento digital de sinais: um estudo de caso. **Revista Tecnologia**, v. 41, n. 2, 2020.

ESCOLA, J. P. L. *et al.* Automated acoustic detection of a cicadid pest in coffee plantations. **Computers and Electronics in Agriculture**, v. 169, p. 105215, 2020. ISSN 0168-1699. Disponível em: <https://www.sciencedirect.com/science/article/pii/S016816991930081X>.

ESCOLA, J. P. L. *et al.* Análise de área de cobertura de dispositivo iot para monitoramento em smart farm/iot device coverage analysis for smart farm monitoring. **RISTI (Revista Iberica de Sistemas e Tecnologias de Informacao)**, AISTI (Iberian Association for Information Systems and Technologies), n. 42, p. 1–12, 2021.

FAN, J.; UPADHYE, S.; WORSTER, A. Understanding receiver operating characteristic (roc) curves. **Canadian Journal of Emergency Medicine**, Cambridge University Press, v. 8, n. 1, p. 19–20, 2006.

FARINA, A.; GAGE, S. H. **Ecoacoustics: The ecological role of sounds**. [*S.l.: s.n.*]: John Wiley & Sons, 2017.

FEURER, M. *et al.* Auto-sklearn: efficient and robust automated machine learning. *In*: **Automated Machine Learning**. [*S.l.: s.n.*]: Springer, Cham, 2019. p. 113–134.

FONSECA, E. S. *et al.* Acoustic investigation of speech pathologies based on the discriminative paraconsistent machine (dpm). **Biomedical Signal Processing and Control**, Elsevier, v. 55, p. 101615, 2020.

FONSECA, E. S. *et al.* Wavelet time-frequency analysis and least squares support vector machines for the identification of voice disorders. **Computers in Biology and Medicine**, Elsevier, v. 37, n. 4, p. 571–578, 2007.

FORNAZIER, M. J. Controle da cigarra do cafeeiro em regiões declivosas no estado do espírito santo<sup>1</sup>. 2000.

FORTI, V. *et al.* The global e-waste monitor 2020: Quantities, flows and the circular economy potential. United Nations University/United Nations Institute for Training and Research . . . , 2020.

GHOLAMI, R.; FAKHARI, N. Support vector machine: principles, parameters, and applications. *In: Handbook of Neural Computation*. [S.l.: s.n.]: Elsevier, 2017. p. 515–535.

GHOSH, A.; PAL, A.; DAS, N. R. An approach to design photonic crystal gas sensor using machine learning. **Optik**, Elsevier, v. 208, p. 163997, 2020.

GONÇALVES, W.; FARIA, A. M. Inseticidas sistêmicos granulados no controle das ninfas móveis das cigarras e seus efeitos na produtividade de cafeeiros. Instituto Agrônômico (IAC), 1989.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.: s.n.]: MIT press, 2016.

GUIDO, R. C. Effectively interpreting discrete wavelet transformed signals. **IEEE Signal Processing Magazine**, IEEE, v. 34, n. 3, p. 89–100, 2017.

GUIDO, R. C. Paraconsistent feature engineering [lecture notes]. **IEEE Signal Processing Magazine**, IEEE, v. 36, n. 1, p. 154–158, 2018.

GUIDO, R. C. A tutorial review on entropy-based handcrafted feature extraction for information fusion. **Information Fusion**, Elsevier, v. 41, p. 161–175, 2018.

GUIDO, R. C. Enhancing teager energy operator based on a novel and appealing concept: Signal mass. **Journal of the Franklin Institute**, Elsevier, v. 356, n. 4, p. 2346–2352, 2019.

GUIDO, R. C. Nearly symmetric orthogonal wavelets for time-frequency-shape joint analysis: Introducing the discrete shapelet transform's third generation (dst-iii) for nonlinear signal analysis. **Communications in Nonlinear Science and Numerical Simulation**, v. 97, p. 105685, 2021. ISSN 1007-5704. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1007570420305153>.

GUIDO, R. C.; ADDISON, P. S.; WALKER, J. Introducing wavelets and time-frequency analysis [introduction to the special issue]. **IEEE Engineering in Medicine and Biology Magazine**, IEEE, v. 28, n. 5, p. 13–13, 2009.

GUIDO, R. C. *et al.* Introducing the discrete path transform (dpt) and its applications in signal analysis, artefact removal, and spoken word recognition. **Digital Signal Processing**, Elsevier, p. 103158, 2021.

GUIDO, R. C. *et al.* A neural-wavelet architecture for voice conversion. **Neurocomputing**, Elsevier, v. 71, n. 1, p. 174–180, 2007.

GUPTA, M. *et al.* Security and privacy in smart farming: Challenges and opportunities. **IEEE Access**, IEEE, v. 8, p. 34564–34584, 2020.

GUPTA, S. *et al.* Residual neural network precisely quantifies dysarthria severity-level based on short-duration speech segments. **Neural Networks**, v. 139, p. 105–117, 2021. ISSN 0893-6080. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0893608021000502>.

HAENLEIN, M.; KAPLAN, A. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. **California management review**, SAGE Publications Sage CA: Los Angeles, CA, v. 61, n. 4, p. 5–14, 2019.

HAUKA, M. *et al.* Automated sensory monitoring system prototype for continuous monitoring of material and structure state. *In*: **ENVIRONMENT. TECHNOLOGIES. RESOURCES. Proceedings of the International Scientific and Practical Conference**. [*S.l.: s.n.*], 2021. v. 2, p. 58–61.

HAYKIN, S. S. *et al.* **Neural networks and learning machines/Simon Haykin**. [*S.l.: s.n.*]: New York: Prentice Hall,, 2009.

HAYKIN, S. S.; VEEN, B. V. **Sinais e sistemas**. [*S.l.: s.n.*]: Bookman, 2001.

HILL, K. B.; MARSHALL, D. C. The song, morphology, habitat, and distribution of the elusive north american cicada okanagana viridis (auchenorrhyncha: Cicadidae). **Annals of the Entomological Society of America**, BioOne, v. 106, n. 5, p. 598–603, 2013.

HOU, Z. *et al.* Ecology of meimuna mongolica (hemiptera: Cicadidae) nymphs: Instars, morphological variation, vertical distribution and population density, host-plant selection, and emergence phenology. **Journal of Insect Science**, v. 15, n. 1, p. 42, 2015. Disponível em: [+http://dx.doi.org/10.1093/jisesa/iev031](http://dx.doi.org/10.1093/jisesa/iev031).

HUNTER, J. D. Matplotlib: A 2d graphics environment. **Computing in science & engineering**, IEEE Computer Society, v. 9, n. 3, p. 90, 2007.

IEA Instituto de Economia Agrícola. São Paulo: [*S.l.: s.n.*], 2021. Disponível em: <http://iea.agricultura.sp.gov.br>. Acesso em: 17 set. 2021.

ISHIZAWA, W. H. **Análise das concentrações energéticas no limiar entre fonemas vozeados e não-vozeados e suas implicações para fins de reconhecimento de locutores dependente do discurso**. 2015. Tese (Doutorado) — Universidade de São Paulo, 2015.

JENSEN, A.; COUR-HARBO, A. la. **Ripples in mathematics: the discrete wavelet transform**. [*S.l.: s.n.*]: Springer Science & Business Media, 2001.

JESUS, L. F. de; SILVA, V. B. da; ROCHA, F. da G. Uso de software para detecção de doenças na cultura da soja com o auxílio de um drone autônomo. **Anais do Computer on the Beach**, p. 552–553, 2015.

JIANG, X. *et al.* Hybrid low-power wide-area mesh network for iot applications. **IEEE Internet of Things Journal**, IEEE, v. 8, n. 2, p. 901–915, 2020.

- JR, G. C. de A. *et al.* Corneal tomography multivariate index (ctmvi) effectively distinguishes healthy corneas from those susceptible to ectasia. **Biomedical Signal Processing and Control**, Elsevier, v. 70, p. 102995, 2021.
- JÚNIOR, S. B. **Dynamic Time Warping baseado na transformada wavelet**. 2007. Tese (Doutorado) — Universidade de São Paulo, 2007.
- JUNIOR, S. B. *et al.* Improved dynamic time warping based on the discrete wavelet transform. *In: IEEE. Multimedia Workshops, 2007. ISMW'07. Ninth IEEE International Symposium on*. [S.l.: s.n.], 2007. p. 256–263.
- JUNIOR, S. T. D. *et al.* Oviposition of *quesada gigas* (hemiptera: Cicadidae) in coffee plants. **Revista Colombiana de Entomologia**, Sociedad Colombiana de Entomología, v. 38, n. 1, p. 1–5, 2012.
- KODALI, R. K.; YERROJU, S.; SAHU, S. Smart farm monitoring using lora enabled iot. *In: IEEE. 2018 second international conference on green computing and internet of things (ICGCIoT)*. [S.l.: s.n.], 2018. p. 391–394.
- LAMPSON, B. *et al.* Automatic detection and identification of brown stink bug, *euschistus servus*, and southern green stink bug, *nezara viridula*, (heteroptera: Pentatomidae) using intraspecific substrate-borne vibrational signals. **Computers and electronics in agriculture**, Elsevier, v. 91, p. 154–159, 2013.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015.
- LEE, G. *et al.* Pywavelets: A python package for wavelet analysis. **Journal of Open Source Software**, v. 4, n. 36, p. 1237, 2019.
- LEE, Y. J. Revised synonymic list of cicadidae (insecta: Hemiptera) from the korean peninsula, with the description of a new species and some taxonomic remarks. **Proceedings of the Biological Society of Washington**, BioOne, v. 121, n. 4, p. 445–467, 2008.
- LIAKOS, K. G. *et al.* Machine learning in agriculture: A review. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 18, n. 8, p. 2674, 2018.
- LIPSCHUTZ, S.; ABELLANAS, L.; ONTALBA, C. M. **Álgebra lineal**. [S.l.: s.n.]: McGraw-Hill, 1992. v. 366.
- LUNDH, F. **Python standard library**. [S.l.: s.n.]: "O'Reilly Media, Inc.", 2001.
- LUNZ, A. M. *et al.* Método para monitoramento de ninfas de cigarras e controle com inseticidas em reflorestamentos com paricá. **Pesquisa Agropecuária Brasileira**, scielo, v. 45, p. 631 – 637, 07 2010. ISSN 0100-204X. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-204X2010000700001&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-204X2010000700001&nrm=iso).
- MACCAGNAN, D. H. B. **Cigarra (Hemiptera: Cicadidae): emergência, comportamento acústico e desenvolvimento de armadilha sonora**. 2008. Tese (Doutorado) — PhD Thesis, Faculdade de Filosofia, Ciências e Letras da Universidade de São ..., 2008.

- MACCAGNAN, D. H. B.; MATUO, T. K.; MARTINELLI, N. M. Oviposition of *quesada gigas* in weed no hostess: implication in pest management. **Ciência Rural**, SciELO Brasil, v. 47, 2017.
- MACCAGNAN, D. H. B. *et al.* Primeiro registro de cigarra em reflorestamentos com paricá no estado de mato grosso, brasil. **Revista de Ciências Agrárias Amazonian Journal of Agricultural and Environmental Sciences**, v. 57, n. 4, p. 451–454, 2014.
- MALLAT, S. G. Multiresolution approximations and wavelet orthonormal bases of  $l^2(\mathbb{R})$ . **Transactions of the American mathematical society**, v. 315, n. 1, p. 69–87, 1989.
- MANKIN, R. *et al.* Perspective and promise: a century of insect acoustic detection and monitoring. **American Entomologist**, The Oxford University Press, v. 57, n. 1, p. 30–44, 2011.
- MARTINELLI, N. M. *et al.* Modo de aplicação e eficiência de inseticidas granulados sistêmicos para o controle de cigarras (hemiptera: Cicadidae) do cafeeiro. **Anais da Sociedade Entomológica do Brasil**, Sociedade Entomológica do Brasil, p. 133–140, 1998.
- MARTINELLI, N. M.; ZUCCHI, R. A. Cigarras (hemiptera: Cicadidae: Tibiciniidae) associadas ao cafeeiro: distribuição, hospedeiros e chave para as espécies. **Anais da Sociedade Entomológica do Brasil**, SciELO Brasil, p. 133–143, 1997.
- MASTELINI, S. M. *et al.* Multi-output tree chaining: An interpretative modelling and lightweight multi-target approach. **Journal of Signal Processing Systems**, Springer, v. 91, n. 2, p. 191–215, 2019.
- MCKINNEY, W. **Python for data analysis: Data wrangling with Pandas, NumPy, and IPython**. [S.l.: s.n.]: "O'Reilly Media, Inc.", 2012.
- MICHAELIS. **Dicionário Brasileiro da Língua Portuguesa**. [S.l.: s.n.], 2015. Disponível em: <https://michaelis.uol.com.br/>. Acesso em: 20 set. 2021.
- MOLIN, J. P. Agricultura de precisão: números do mercado brasileiro. **Boletim Técnico**, v. 3, 2017.
- MOLNAR, C. **Interpretable machine learning**. [S.l.: s.n.]: Lulu.com, 2020.
- MONTGOMERY, G. A. *et al.* Standards and best practices for monitoring and benchmarking insects. **Frontiers in Ecology and Evolution**, Frontiers, v. 8, p. 513, 2021.
- NANDIRAJU, N. *et al.* Wireless mesh networks: Current challenges and future directions of web-in-the-sky. **IEEE Wireless Communications**, IEEE, v. 14, n. 4, p. 79–89, 2007.
- NILSSON, N. J. **Principles of artificial intelligence**. [S.l.: s.n.]: Morgan Kaufmann, 2014.
- NYQUIST, H. Certain topics in telegraph transmission theory. **Transactions of the American Institute of Electrical Engineers**, IEEE, v. 47, n. 2, p. 617–644, 1928.
- ODEN, L. Lessons learned from comparing c-cuda and python-numba for gpu-computing. *In*: IEEE. **2020 28th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)**. [S.l.: s.n.], 2020. p. 216–223.

OLIVEIRA, S. de. **Internet das coisas com ESP8266, Arduino e Raspberry PI**. [S.l.: s.n.]: Novatec Editora, 2017.

PAULUS, D.; HORNEGGER, J. **Applied pattern recognition: algorithms and implementation in C++**. [S.l.: s.n.]: Springer Science & Business Media, 2003.

PEDREGOSA, F. *et al.* Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEREIRA, N. A. Distribuição espacial de posturas e ninfas de *quesada gigas* (olivier, 1790)(hemiptera: cicadidae) na cultura do cafeeiro. Universidade Estadual Paulista (UNESP), 2017.

PRINCE, P. *et al.* Deploying acoustic detection algorithms on low-cost, open-source acoustic sensors for environmental monitoring. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 19, n. 3, p. 553, 2019.

PRINGLE, J. A physiological analysis of cicada song. **Journal of Experimental Biology**, The Company of Biologists Ltd, v. 31, n. 4, p. 525–560, 1954.

REIS, P.; SOUZA, J. d. Cigarras-do-cafeeiro, dano e controle. **Lavras: EPAMIG/CRSM, 5p.(Circular Técnica)**, 1991.

RIBEIRO, J. G.; MARINHO, D. Y.; ESPINOSA, J. W. M. Agricultura 4.0: desafios à produção de alimentos e inovações tecnológicas. *In*: **SIMPÓSIO DE ENGENHARIA DE PRODUÇÃO**. [S.l.: s.n.], 2018. v. 2, p. 1–7.

RICHARDSON, M.; WALLACE, S. **Getting started with raspberry PI**. [S.l.: s.n.]: "O'Reilly Media, Inc.", 2012.

SANBORN, A. F.; HEATH, J. E.; HEATH, M. S. Long-range sound distribution and the calling song of the cicada *beameria venosa* (uhler)(hemiptera: Cicadidae). **The Southwestern Naturalist**, BioOne, v. 54, n. 1, p. 24–30, 2009.

SANTANA, E. J. *et al.* Predicting poultry meat characteristics using an enhanced multi-target regression method. **Biosystems Engineering**, Elsevier, v. 171, p. 193–204, 2018.

SANTOS, R. S. Ocorrência de *fidicinoides pauliensis* boulard & martinelli, 1996 (hemiptera: Cicadidae) em cafeeiro em tapiratiba, sp. **Revista de Agricultura**, v. 82, p. 311–314, 2007.

SANTOS, T. C. dos; ESPERIDIÃO, T. L.; AMARANTE, M. dos S. Agricultura 4.0. **Revista Pesquisa E Ação**, v. 5, n. 4, p. 122–131, 2019.

SATO, Y.; SATO, S. Spring temperature predicts the long-term molting phenology of two cicadas, *cryptotympana facialis* and *graptopsaltria nigrofuscata* (hemiptera: Cicadidae). **Annals of the Entomological Society of America**, v. 108, n. 4, p. 494, 2015. Disponível em: [+http://dx.doi.org/10.1093/aesa/sav036](http://dx.doi.org/10.1093/aesa/sav036).

SAZIMA, I. Insect cornucopia: various bird types prey on the season's first giant cicadas in an urban park in southeastern brazil. **Biota Neotropica**, SciELO Brasil, v. 9, n. 1, p. 0–0, 2009.

- SEABRA, S. G.; QUARTAU, J. A.; BRUFORD, M. W. Spatio-temporal genetic variation in sympatric and allopatric mediterranean cicada species (hemiptera, cicadidae). **Biological Journal of the Linnean Society**, Wiley Online Library, v. 96, n. 2, p. 249–265, 2009.
- SILVA, I. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. **São Paulo: Artliber**, p. 33–111, 2010.
- SOARES, V. C. *et al.* Análise espacial da distribuição de cigarras (quesada gigas oliver) em povoados de paricá (schizobium amazonicum huber ex ducke) na região de dom eliseu, pa. **Rev. Árvore**, v. 32, p. 251–258, 2008. ISSN 1806-9088.
- SOMPAIRAC, N. *et al.* Independent component analysis for unraveling the complexity of cancer omics datasets. **International Journal of molecular sciences**, Multidisciplinary Digital Publishing Institute, v. 20, n. 18, p. 4414, 2019.
- SOUZA, J.; REIS, P. R.; SILVA, R. A. **Cigarras-do-cafeeiro em Minas Gerais: histórico, reconhecimento, biologia, prejuízos e controle**. 2nd. ed. Belo Horizonte, MG: EPAMIG, 2007. ISSN 0101-062X.
- SOUZA, L. M. d. **Detecção inteligente de patologias na laringe baseada em máquinas de vetores de suporte e na transformada Wavelet**. 2010. Tese (Doutorado) — Universidade de São Paulo, 2010.
- SOUZA, U. B. de; ESCOLA, J. P. L.; BRITO, L. da C. A survey on hilbert-huang transform: Evolution, challenges and solutions. **Digital Signal Processing**, v. 120, p. 103292, 2022. ISSN 1051-2004. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1051200421003316>.
- SOYATA, T. **GPU parallel program development using CUDA**. [*S.l.: s.n.*]: CRC Press, 2018.
- STRANG, G.; NGUYEN, T. **Wavelets and filter banks**. [*S.l.: s.n.*]: SIAM, 1996.
- SUEUR, J. Audiospectrographical analysis of cicada sound production: a catalogue (hemiptera, cicadidae). **Deutsche Entomologische Zeitschrift**, Wiley Online Library, v. 48, n. 1, p. 33–51, 2001.
- SUEUR, J. Cicada acoustic communication: potential sound partitioning in a multispecies community from mexico (hemiptera: Cicadomorpha: Cicadidae). **Biological Journal of the Linnean Society**, Oxford University Press, v. 75, n. 3, p. 379–394, 2002.
- SUEUR, J.; FARINA, A. Ecoacoustics: the ecological investigation and interpretation of environmental sound. **Biosemiotics**, Springer, v. 8, n. 3, p. 493–502, 2015.
- SUEUR, J. *et al.* Rapid acoustic survey for biodiversity appraisal. **PloS one**, Public Library of Science, v. 3, n. 12, p. e4065, 2008.
- SUNDERLAND, T. *et al.* Sdg2: Zero hunger: Challenging the hegemony of monoculture agriculture for forests and people. **Katila P, Colfer C, de Jong W, Galloway G, Pacheco P, Winkel G, Sustainable Development Goals: Their Impacts on Forests and People**. Cambridge University Press.[Google Scholar], p. 48–71, 2019.

TRIPLEHORN, C. A.; JOHNSON, N. F. Estudo dos insetos. **São Paulo: Cengage Learning**, 2011.

ULLOA, J. S. *et al.* Estimating animal acoustic diversity in tropical environments using unsupervised multiresolution analysis. **Ecological Indicators**, Elsevier, v. 90, p. 346–355, 2018.

VALLEE, M. The science of listening in bioacoustics research: Sensing the animals' sounds. **Theory, Culture & Society**, SAGE Publications Sage UK: London, England, v. 35, n. 2, p. 47–65, 2018.

VIEIRA, L. S. **Conversão de voz baseada na transformada wavelet**. 2008. Tese (Doutorado) — Universidade de São Paulo, 2008.

VILLET, M. Sound pressure levels of some african cicadas (homoptera: Cicadoidea). **Journal of the Entomological Society of Southern Africa**, Entomological Society of South Africa (ESSA), v. 50, n. 2, p. 269–273, 1987.

WALKER, J. S. **A primer on wavelets and their scientific applications**. [*S.l.: s.n.*]: CRC press, 2002.

WANG, K. *et al.* Wavelet packet analysis for speaker-independent emotion recognition. **Neurocomputing**, Elsevier, v. 398, p. 257–264, 2020.

WANG, L. *et al.* Denoising speech based on deep learning and wavelet decomposition. **Scientific Programming**, Hindawi, v. 2021, 2021.

WEI, P. *et al.* Research on sound classification based on svm. **Neural Computing and Applications**, Springer, v. 32, n. 6, p. 1593–1607, 2020.

WILLIAMS, J. R.; AMARATUNGA, K. Introduction to wavelets in engineering. **International journal for numerical methods in engineering**, Wiley Online Library, v. 37, n. 14, p. 2365–2388, 1994.

WILLIAMSON, D. F.; PARKER, R. A.; KENDRICK, J. S. The box plot: a simple visual method to interpret data. **Annals of internal medicine**, American College of Physicians, v. 110, n. 11, p. 916–921, 1989.

YAO, X. *et al.* Exploration of glottal characteristics and the vocal folds behavior for the speech under emotion. **Neurocomputing**, Elsevier, v. 410, p. 328–341, 2020.

ZHAO, Q. *et al.* Long-term coffee monoculture alters soil chemical properties and microbial communities. **Scientific reports**, Nature Publishing Group, v. 8, n. 1, p. 1–11, 2018.

ZOU, H.; XUE, L. A selective overview of sparse principal component analysis. **Proceedings of the IEEE**, IEEE, v. 106, n. 8, p. 1311–1320, 2018.