

THIAGO CASSIN DE CARVALHO OLIVEIRA<sup>1</sup>, JOÃO BAPTISTA DE OLIVEIRA E SOUZA FILHO<sup>1,2</sup>.

<sup>1</sup> Programa de Pós-Graduação em Engenharia Elétrica (PPEL), Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ)  
Avenida Maracanã 229, Maracanã, Rio de Janeiro, RJ, Brasil

<sup>2</sup> Programa de Pós-Graduação em Engenharia Elétrica (PEE/COPPE), Departamento de Engenharia Eletrônica e de Computação (DEL/POLI), Universidade Federal do Rio de Janeiro (UFRJ)  
E-mails: thiago\_cassin@yahoo.com.br, jbfilho@poli.ufrj.br

**Abstract**— Passive sonar contact classification systems require high classification efficiency, and a computational cost that enables its operation in real time systems. In the present work, a hierarchical principal curves classification approach is proposed, associated with optimized strategies for the definition of the curves segmentation and the appropriate depth level of the hierarchical structures. The proposed system was evaluated using real irradiated signals from 28 ships belonging to 8 classes. Using the proposed technique, an average classification efficiency of 96.2% was achieved, as well as the computational cost was reduced by up to 39.5% when compared to the standard technique principal curves.

**Keywords**— Principal curves, contact classification, feature extraction, pattern recognition, passive sonar

**Resumo** — Sistemas de apoio à classificação de sinais de sonar passivo devem possuir uma alta eficiência de classificação, bem como custo computacional compatível a sua operação em sistemas de tempo real. No presente trabalho, é proposta uma abordagem de classificação por curvas principais hierárquicas, associada a estratégias otimizadas para a definição da segmentação das curvas e do nível de profundidade adequado às estruturas hierárquicas. O sistema proposto foi avaliado, utilizando sinais reais irradiados de 28 navios pertencentes a 8 classes. Com o emprego da técnica proposta, uma eficiência média de classificação da ordem de 96,2% foi alcançada, bem como o custo computacional foi reduzido em até 39,5%, quando comparado à técnica de curvas principais tradicionais.

**Palavras-chave**— Curvas principais, classificação de contatos, extração de características, reconhecimento de padrões, sonar passivo

## 1 Introdução

A classificação de possíveis ameaças em submarinos em operação no âmbito militar é conhecida como classificação de contatos. Esta atividade é realizada por meio da aquisição de sinais de sonar passivo, que são processados visando à posterior avaliação pelo operador de sonar. Trata-se de uma tarefa de alta-complexidade e responsabilidade, que envolve vários cenários operativos e risco de vida, sendo realizada, tradicionalmente, por operadores treinados e especializados. Assim, sistemas de apoio à decisão (SAD) são úteis no suporte ao processo de tomada de decisão neste contexto.

Um SAD em operação no ambiente de sonar passivo deve ser capaz de trabalhar dados não lineares, e envolver classificadores com uma boa capacidade de generalização, operando sob múltiplas classes. Para a constituição deste SAD, a técnica de curvas principais (Hastie & Stuetzle, 1989) é indicada, por possuir uma boa capacidade de generalização em conjuntos de dados intrinsecamente não lineares, bem como custo computacional compatível a operação de tais sistemas em tempo real.

Em (Oliveira & Souza Filho, 2016), foi proposta a adoção de curvas principais hierárquicas visando à constituição de SADs para a classificação de sinais de

sonar passivo. Nesta proposta, cada classe foi representada por um banco de curvas, diferentemente das soluções anteriores que adotaram uma única curva, sendo os eventos rotulados de acordo com a sua similaridade (no caso, a distância à curva) aos conjuntos de curvas protótipo de cada classe. Para a obtenção deste conjunto, foi adotada uma estratégia de particionamento hierárquico dos dados de cada classe, também baseada em curvas principais.

Dois desafios inerentes a esta proposta são o dimensionamento das curvas utilizadas para a partição e a classificação dos dados, bem como a determinação do número de níveis hierárquicos adotados. Tais aspectos impactam, diretamente, na eficácia e no custo computacional do classificador, e envolvem um problema combinatório complexo, visto que a sintonia dos parâmetros do classificador de uma classe afeta as demais, sendo a busca exaustiva de uma solução ótima não factível.

O presente trabalho tem por objetivo discutir alguns critérios de custo computacional reduzido para a escolha de tais parâmetros, mais especificamente avaliar a aplicabilidade da AUC (Area Under the ROC Curve) (Trees, 2001) como figura de mérito para a sintonia da complexidade das curvas principais adotadas em cada nível de partição e classificação dos dados, bem como a proposição de um critério heurístico para a seleção das curvas principais a integrarem a família de curvas representantes de cada classe. Os critérios

propostos buscam reduzir o total de segmentos das curvas principais envolvidas, portanto resultar em sistemas de classificação com menor custo computacional em fase de operação, no entanto, que possuam uma eficiência de classificação igual ou superior a de sistemas análogos baseados em uma única curva por classe, tal como (Fernandes, et al., 2015) e (Souza Filho, et al., 2011), bem aqueles como baseados em estruturas hierárquicas baseadas em curvas principais (Oliveira & Souza Filho, 2016).

Este artigo é organizado como se segue: na Seção 2, é apresentada uma breve abordagem das técnicas de curvas principais e dos algoritmos hierárquicos de particionamento adotados. São também discutidos alguns aspectos relativos à construção e a avaliação do classificador proposto, bem como referentes à base de dados utilizada e a sua partição para o desenvolvimento e avaliação dos classificadores. Na Seção 3, são apresentados os resultados obtidos. Por fim, tem-se as conclusões e trabalhos futuros na Seção 4.

## 2 Métodos

### 2.1 Curvas Principais

As curvas principais (Hastie & Stuetzle, 1989) denotam curvas suaves que “passam” pelo meio de uma nuvem de dados, oriundas de diferentes realizações de um dado conjunto de variáveis aleatórias. Para uma definição formal das curvas principais, considere um vetor de variáveis aleatórias  $X \in R^d$ , onde o valor da variável  $d$  é arbitrário. A curva principal  $f(\lambda) = [f_1(\lambda), \dots, f_d(\lambda)]$  é uma função do parâmetro  $\lambda$  tal que:

$$f(\lambda) = E[X | \lambda_f(X) = \lambda] \quad (1)$$

Nesta formulação, o valor  $\lambda_f(X)$  é definido de forma a minimizar o valor esperado ( $E$ ) da distância do vetor  $X$  a curva, portanto cada ponto da curva  $f(\lambda)$  é dado pela média dos pontos que nela projetam, propriedade que é conhecida como **autoconsistência**.

Para a obtenção das curvas principais, um algoritmo de destaque é o “**k-segment hard**”, desenvolvido por Verbeek, Vassis e Klöse (Verbeek, et al., 2000), que aproxima a curva principal por um conjunto de segmentos conectados. Segundo este algoritmo, a curva principal é iniciada com um único segmento, sendo inseridos novos segmentos de reta, um a um, até que um número máximo pré-fixado de segmentos ( $k$ ) seja alcançado.

### 2.2 Classificação por curvas principais ordinárias

Segundo a abordagem tradicional de classificação por curvas principais descrita em (Souza Filho, et al., 2011), a atribuição de rótulos aos eventos é realizada tomando-se o dado que se deseja classificar, e calculando a distância deste às curvas principais extraídas

para cada uma das classes, e atribuindo-a ele, o rótulo de classe referente à curva protótipo a qual tenha apresentado a menor distância, portanto, a maior similaridade. Deste modo, a curva protótipo de cada classe é composta por tantos segmentos quantos os necessários para que a referida curva possa representar, adequadamente, a classe em questão, esta curva é extraída com base no conjunto de treinamento. Tal processo é ilustrado na Figura 1.

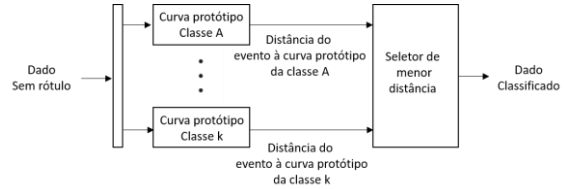


Figura 1: Classificação multiclasse por curvas principais. Adaptado de (Souza Filho, et al., 2011)

### 2.3 Classificador por curvas principais hierárquicas

A principal diferença desta proposta em relação à classificação por curvas principais ordinárias reside em, ao invés de adotar uma curva de alta-complexidade para a representação de cada classe, realiza-se a extração de diversas curvas de baixa complexidade, cada uma baseada em um subconjunto gerado pela partição hierárquica de seus dados constituintes. Mais especificamente, iremos adotar aqui as próprias curvas principais como critério de particionamento dos dados. Este processo é ilustrado na Figura 2.

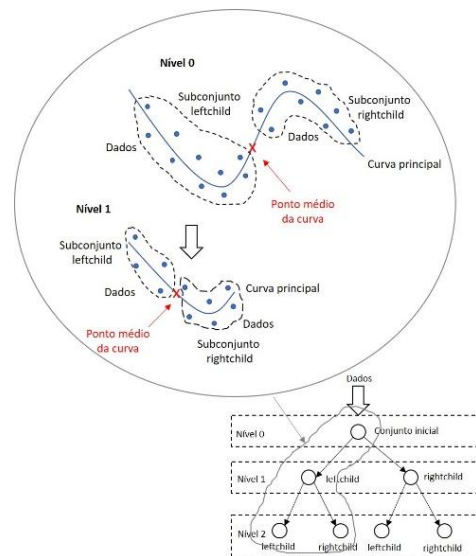


Figura 2: Processo de particionamento hierárquico

Por este processo, os dados, de acordo com o valor de seu índice de projeção à curva principal do nível e nó da árvore correspondente, são divididos em grupos intitulados *ramo esquerdo* e *ramo direito*, isto é, aqueles cujo índice de projeção é menor que a metade

do comprimento da curva definem o “ramo esquerdo”; enquanto os demais, o “ramo direito”. Em sequência, considerando, individualmente, os novos subconjuntos “ramo esquerdo” e “ramo direito”, são produzidas duas novas curvas principais, com as quais é realizado um novo particionamento. Este processo é repetido até que uma profundidade requerida seja atingida.

### 2.3.1 Operação do classificador por curvas principais hierárquicas

Em fase de operação, de forma análoga às curvas ordinárias, tão somente são calculadas as distâncias do dado de interesse às curvas principais representantes de cada uma das classes, e realizada a classificação por meio da identificação da classe cuja uma das curvas representantes apresentou a menor distância dentre todas as demais.

Com respeito a quais curvas produzidas na hierarquia que serão efetivamente empregadas para a classificação, foram consideradas as topologias reportadas por (Silla & Freitas, 2011) como estrutura hierárquica de classificação por níveis e por nós.

No classificador por nós, as curvas referentes todos os nós da árvore competem pelos eventos, conforme ilustrado pela Figura 3.

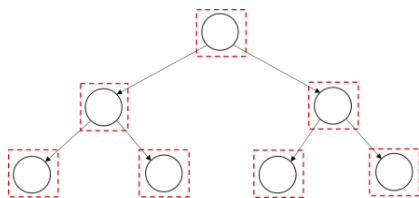


Figura 3: Estrutura hierárquica por nós

No classificador por níveis, somente os nós referentes a um nível de hierarquia selecionado competem pela pertinência destes eventos, conforme mostrado na Figura 4. Claramente, a abordagem por níveis possui um custo computacional significativamente inferior em fase de operação do que aquela por nós, face ao menor quantitativo de curvas envolvidas.

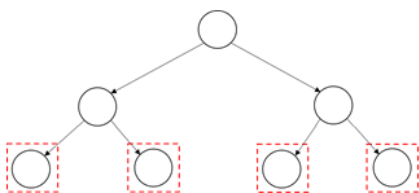


Figura 4: Estrutura hierárquica por níveis

### 2.3 Critérios para a seleção do número de segmentos das curvas.

Foram considerados dois critérios para a definição do número de segmentos utilizados para a construção do banco de curvas representantes de cada

classe: a segmentação pré-definida e a sintonia pela AUC.

Na primeira, é considerado que todos os níveis empregam um mesmo número de segmentos em suas curvas, a serem escolhidas entre uma família de candidatos. Neste trabalho, foram considerados 2, 5 ou 7 segmentos para as curvas principais de cada nível.

Segundo a sintonia AUC, a cada nível é escolhido o melhor quantitativo de segmentos entre uma família de candidatos, utilizando a AUC como figura de mérito. Neste caso, são considerados a distância dos eventos da classe correspondente relativamente a curva associada a dado nível e profundidade, versus as distâncias dos eventos remanescentes à esta curva. Quanto maior for a AUC, maior é o potencial de discriminação daquela curva, considerando a distância como estatística suficiente para a classificação.

### 2.4 Estratégia de Poda

As árvores de classificação frequentemente atingem grandes tamanhos e complexidades, podendo tornar-se imprecisas e incompreensíveis (Patil & Bichkar, 2012). Neste contexto, surgem as estratégias de poda, que são divididas, tradicionalmente, como: pré-poda e pós-poda (Breiman, et al., 1984). Como nossa proposta envolve problemas análogos aos envolvidos nas árvores de classificação tradicionais, iremos explorar alguns conceitos e idéias desta estratégia de classificação.

Para o presente trabalho, foi desenvolvida uma estratégia de poda denominada como pós-poda heurística (PPH). A PPH “visita” as árvores de classificação referentes à cada classe, poda estas árvores com base em uma figura de mérito, que considera o compromisso entre eficiência e complexidade, e as recompõe, visando a otimização da estrutura hierárquica multiclasse que compõe o classificador final. A Figura 5 apresenta o diagrama de blocos da PPH.



Figura 5: Diagrama de Blocos da Pós Poda Heurística

O passo inicial deste processo é a indução das árvores de classificação para cada classe de forma “top-down” (Sun & Lim, 2002), até o nível de profundidade máximo permitido (aqui foi adotado um limite de 5 níveis), e para vários valores distintos de complexidade de curvas.

O algoritmo da PPH é então implementado em cinco etapas: seleção, corte, ordenação, combinação e avaliação de desempenho. Na etapa de seleção, é realizada uma varredura em todas as árvores geradas e representativas das classes em busca de ramos que

apresentem uma maior eficiência igual ou superior associada ao emprego de menos segmentos totais em suas curvas quando comparados com classificadores baseados em curvas ordinárias. Caso não sejam encontrados ramos que atendam a estes requisitos, é considerada uma curva ordinária para representar a respectiva classe.

Na etapa de corte, as árvores são podadas na altura dos ramos selecionados na etapa de seleção. Em sequência na etapa de ordenação, primeiramente, as novas árvores podadas têm sua relação de eficiência versus custo computacional quantificada pelo cálculo do índice “S” representado na Equação 2, onde “E” representa a eficiência de classificação atingida na classe correspondente e “C” representa a complexidade (número total de segmentos empregados nas curvas).

$$S = \frac{E}{C} \quad (2)$$

As árvores podadas e quantificadas são organizadas segundo seus índices “S” em ordem decrescente, de forma que quanto maior for o índice “S”, melhor é a relação entre o custo computacional e a eficiência. Após esta ordenação, as árvores de cada classe que apresentaram o maior valor de índice “S” são combinadas, compondo o classificador multiclasse otimizado. Por fim, o sistema de classificação definido pela PPH tem seu desempenho avaliado, utilizando o conjunto de dados de teste.

## 2.5 Base de dados

Para o desenvolvimento do presente trabalho, foi utilizada uma base de dados fornecida pelo Instituto de Pesquisa da Marinha Brasileira (IPqM), que consiste em um conjunto de ruídos irradiados por 28 navios, distribuídos entre 8 classes de embarcações.

Tais sinais foram pré-processados segundo uma cadeia descrita detalhadamente em (Souza Filho, et al., 2011), resultando em janelas de 557 dimensões, onde cada classe possui de 2143 a 7075 espectros.

Para o desenvolvimento e a avaliação da abordagem proposta, este conjunto foi particionado, aleatoriamente, nos seguintes subconjuntos: treino, validação e teste, segundo as seguintes proporções: 70%, 10% e 20%. O primeiro conjunto foi utilizado para a extração das curvas; o segundo, para o ajuste dos parâmetros das curvas; enquanto o último, para a avaliação do sistema proposto.

## 3 Resultados

Primeiramente, realizamos uma avaliação comparativa entre os sistemas de classificação baseados em curvas ordinárias, curvas hierárquicas por nós e curvas hierárquicas por níveis. Em sequência, avaliamos a adoção do critério AUC para o dimensionamento das curvas por níveis e o critério de pós-poda

heurística para a otimização da estrutura de classificação adotada.

Neste processo, foram construídos os seguintes sistemas:

- **Sistema 1:** baseou-se em uma curva ordinária para cada classe, adotando-se um número máximo de segmentos igual a 30, valor que foi estabelecido conforme resultados de (Fernandes, et al., 2015) e (Souza Filho, et al., 2011)

- **Sistema 2:** trata-se de um classificador hierárquico por níveis, considerando curvas principais compostas por 2, 5 ou 7 segmentos.

- **Sistema 3:** é um classificador hierárquico por nós, formado por curvas principais de 2, 5 ou 7 segmentos.

Em todas as análises, as figuras de mérito foram a quantidade total de segmentos envolvidos e a eficiência de classificação final de cada sistema.

### 3.1 Classificadores por curvas ordinárias, hierárquicas por nós e hierárquicas por níveis.

A Figura 6, resume os resultados relativos às diferentes modalidades de classificadores por curvas principais, onde “L” representa o quantitativo de níveis e “k” o número de segmentos por curva.

É possível observar que em termos de eficiência de classificação, os classificadores hierárquicos por nós e níveis apresentaram um desempenho muito similar. A maior diferença entre estas abordagens foi observada para a configuração com “L=2” e “k=5” (0,8%). Foi também verificada uma tendência de aumento das eficiências de classificação com a adoção de curvas mais complexas, bem como uma maior profundidade da estrutura de classificação. Observa-se, também, que para “k=7”, a partir de “L=3”, as abordagens hierárquicas atingem uma eficiência média de classificação superior a abordagem ordinária de até 1 ponto percentual.

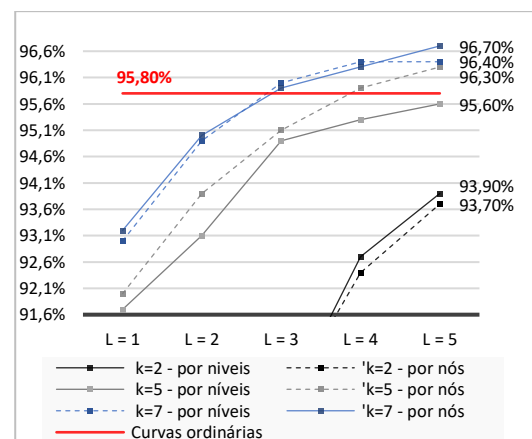


Figura 6: Eficiência dos classificadores por nós e níveis para diferentes valores de profundidade e de complexidade das curvas.



Na Figura 7 é exibido o quantitativo de segmentos empregado em cada uma das abordagens de classificação consideradas na Figura 6.

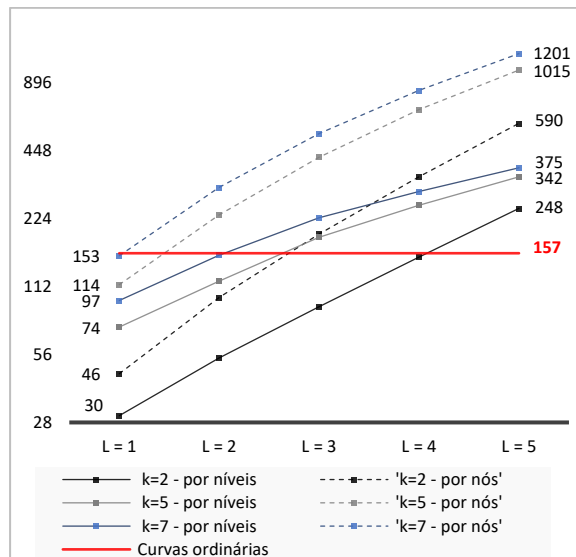


Figura 7: Total de segmentos empregados entre os classificadores por nós e níveis para diferentes valores de profundidade e segmentação das curvas.

É notório que do ponto de vista do custo computacional, a abordagem hierárquica por nós não é uma opção atrativa, visto que o quantitativo de segmentos envolvido é muito superior ao das abordagens hierárquica por níveis e ordinária, o qual não se refletiu num ganho de eficiência de classificação significativo.

### 3.2 Sintonia das curvas por nível $x$ uso de curvas com complexidade fixa

Esta avaliação objetiva verificar se a sintonia das complexidades das curvas a cada nível, utilizando a AUC como figura de mérito, é mais efetiva que a adoção de um quantitativo de segmentos fixo para todos os níveis, pré-definido em 2, 5 ou 7 segmentos, conforme realizado em (Oliveira & Souza Filho, 2016). Esta avaliação se restringiu ao classificador por níveis, visto os resultados da análise anterior. Os resultados são resumidos na Figura 8.

Através da Figura 8, é possível verificar que o emprego do indicador AUC resultou em valores de eficiência superiores às curvas principais ordinárias a partir do nível de profundidade " $L=3$ ". Uma queda brusca na taxa de variação da eficiência de classificação também pode ser percebida a partir deste nível de profundidade.

Com relação ao total de segmentos empregados para a construção das curvas, observamos que o nível  $L=2$  de profundidade continua a representar uma profundidade hierárquica limítrofe com respeito a ser atrativo o emprego de curvas hierárquicas a este problema.

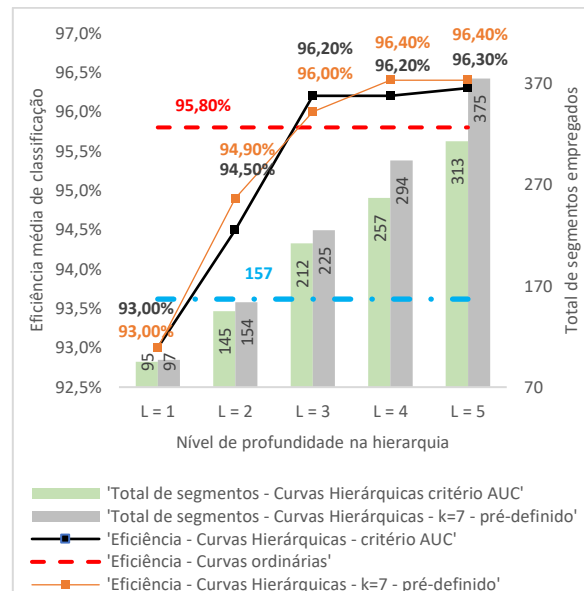


Figura 8: Eficiências e total de segmentos empregados pelo uso do critério da AUC.

Relativamente à seleção pré-definida das complexidades das curvas, o critério AUC empregou um quantitativo menor de segmentos para todas as profundidades avaliadas. Quanto à eficiência de classificação, embora ambas abordagens tenham apresentado um desempenho similar, o critério AUC alcançou maiores valores de eficiência média para a profundidade " $L=3$ ".

### 3.3 Pós-poda heurística

Nesta avaliação foi considerada a aplicação do critério PPH à estrutura de classificação por níveis, considerando os dois critérios anteriormente descritos para o dimensionamento das curvas: fixo e por nível. O quantitativo de níveis e segmentos selecionados para cada uma das abordagens, juntamente com as eficiências médias de classificação obtidas, são resumidos na Tabela 1.

Da Tabela 1, é possível observar que o algoritmo de PPH, segundo o seu índice de ordenação, selecionou apenas as profundidades  $L=\{1,2\}$  para o particionamento dos dados e geração das curvas principais correspondentes (que possuem até cinco níveis de profundidade). Este fato indica que uma partição dos dados em níveis hierárquicos mais profundos não resultou em relações interessantes entre o custo computacional e a eficiência de classificação resultante, parâmetros que definem o índice " $S$ ".

Adicionalmente, verifica-se que, no caso do critério AUC, a seleção pela PPH selecionou apenas curvas com 7 segmentos, que corresponde ao valor de complexidade máximo definida a priori. Este fato indica que a adoção de curvas com menos do que 7 segmentos tendeu a produzir uma menor discriminabilidade local.

Classe	Quantitativo Pré-definido		Quantitativo por AUC	
	Profundidade (L)	"k" por curva (máx)	Profundidade (L)	"k" por curva (máx)
A	Curva ordinária		Curva ordinária	
B	1	7	1	7
C	2	7	2	7
D	2	7	Curva ordinária	
E	1	5	1	7
F	1	7	1	7
G	3	5	3	7
H	Curva ordinária		Curva ordinária	
Total de segmentos empregados	107		93,40%	
Eficiência média de classificação	110		94,10%	

Tabela 1: Composição e resultados do classificador hierárquico por níveis, em seus dois modos de definição da complexidade das curvas principais adotadas, após a aplicação do critério de PPH.

Observa-se, também, que a PPH selecionou as curvas ordinárias para as classes A e H (quantitativo pré-definido) e A, D e H (critério AUC), o que se deve ao fato de que nenhuma das árvores podadas referentes a estas classes apresentaram um compromisso entre custo computacional inferior e eficiência superior à abordagem ordinária.

### 3.4 Abordagem hierárquica versus abordagem ordinária

Para possibilitar uma melhor análise do comportamento das abordagens propostas frente a técnica de classificação por curvas principais ordinárias, foram selecionadas algumas configurações de classificadores hierárquicos reportadas na Tabela 2.

Configuração	Definição do número de segmentos por curva	Profundidade (L)	Segmentos "k" por curva (máx.)
1	Pré-definido	1	7
2	Pré-definido	2	5
3	Pré-definido	2	7
4	Pré-definido	3	7
5	Sintonia pela AUC	1	7
6	Sintonia pela AUC	2	7
7	Sintonia pela AUC	3	7
8	Pré-definido + PPH	2	5 ou 7
9	Sintonia pela AUC + PPH	2	7

Tabela 2: Configurações de classificadores utilizados no comparativo global.

As configurações 1, 2, 3 e 4 retratam classificadores hierárquicos por níveis com complexidades pré-definidas e diferentes níveis de profundidade "L". As configurações 5, 6 e 7 consistem em classificadores hierárquicos por níveis, com sintonia da complexidade de suas curvas pela AUC, para diferentes níveis de profundidade. As configurações 8 e 9 foram concebidas de forma a combinar as duas diferentes estratégias de segmentação propostas com a PPH.

Na Figura 9 é apresentada uma comparação da perda ou ganho percentual entre todas as configurações de classificadores hierárquicos apresentados na Tabela 2, relativamente à eficiência média de classificação atingida, e ao total de segmentos adotado, em comparação com a solução baseada em curvas ordinárias.

É possível verificar que o emprego de estruturas hierárquicas até o nível dois de profundidade, de maneira geral, impactou em uma redução expressiva do quantitativo de segmentos empregados nas curvas. É interessante pontuar que foram observados significativos ganhos relativos à redução do total de segmentos empregados, da ordem de quase 40 pontos percentuais (configuração 5), enquanto que relativamente à eficiência média de classificação, as maiores perdas (configurações 1 e 5) são da ordem de 3 pontos percentuais.

As configurações 4 e 7 apresentaram ganhos de eficiência de classificação de 0,2 e 0,4 pontos percentuais aproximadamente, ao custo de um um forte incremento no total de segmentos empregados (43,31% e 35,03%, respectivamente).

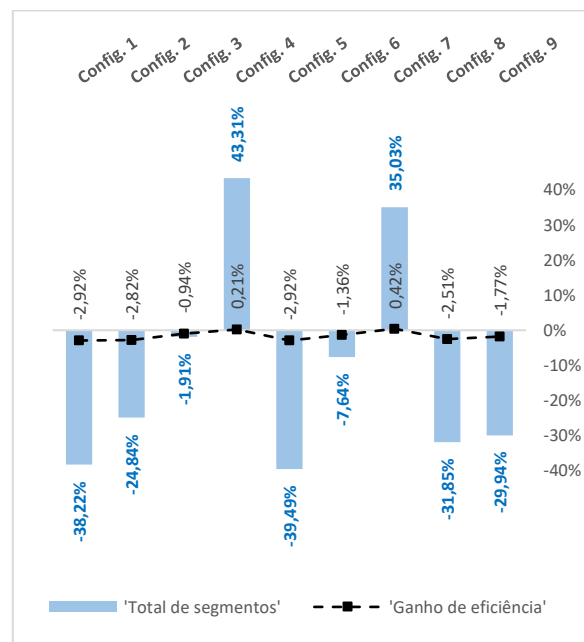


Figura 9: Valores de ganho e perda quanto a eficiência média e custo computacional para as diferentes implementações avaliadas (vide texto).

Quanto ao melhor compromisso entre a eficiência média de classificação e o custo computacional envolvido, destacamos a configuração 5, que considerou a sintonia da complexidade das curvas pelo critério AUC. A configuração menos atrativa foi a número 3, com uma redução do total de segmentos em torno de 2%, e perda de quase 1 ponto percentual de eficiência média.

## 5 Conclusão

Neste trabalho foi avaliado o potencial da construção de sistemas de classificação baseados em curvas principais geradas sobre partições hierárquicas dos dados para a classificação de contatos em sinais de sonar passivo.

Como resultado, verificou-se que a estrutura hierárquica por níveis é mais atrativa do que a por nós, visto o custo computacional excessivo da última, que não é acompanhado de um ganho expressivo quanto à eficiência média de classificação. Adicionalmente, por meio do critério proposto de dimensionamento das curvas por AUC e do critério de PPH, foi possível obter melhores compromissos entre a complexidade das curvas adotadas e a eficiência do classificador resultante, obtendo-se uma estrutura para a constituição do classificador com uma redução do quantitativo de segmentos da ordem de 40%, associada a uma perda da eficiência média de aproximadamente, 3 pontos percentuais.

Como trabalhos futuros, visamos explorar algoritmos meta-heurísticos de busca para uma definição dos níveis de profundidade e segmentação conjunta das curvas, de forma a, não somente, reduzir o custo computacional envolvido, mas também resultar numa maior eficiência média de classificação das curvas hierárquicas.

## Agradecimentos

Agradecemos ao Instituto de Pesquisas da Marinha (IPqM) pela disponibilização do banco de dados e pela cooperação técnica, fundamentais para o desenvolvimento deste trabalho.

## Referências Bibliográficas

- Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A., 1984. *Classification and Regression Trees*. Florida: Chapman & Hall.
- Fernandes, T. S. ..., Souza Filho, J. B. D. O. & Alves, E. S., 2015. *Identificação de Navios em Sinais de Sonar Passivo por Curvas Principais*. Curitiba, Associação Brasileira de Inteligência Computacional.
- Hastie, T. & Stuetzle, W., 1989. Principal Curves. *Journal of the American Statistical Association*, Volume vol. 84, pp. p. 502-516.
- Oliveira, T. C. d. C. & Souza Filho, J. B. D. O., 2016. *Uso de Curvas Principais Hierárquicas na Classificação de Sonar Passivo*. Espírito Santo, Sociedade Brasileira de Automática.
- Patil, D. V. & Bichkar, R. S., 2012. Issues in Optimization of Decision Tree learning: A Survey. *International Journal of Applied Information Systems*, Volume vol 3.
- Silla, C. & Freitas, A. A., 2011. A Survey of Hierarchical Classification Across Different Application Domains. *Data Mining and Knowledge Discovery*, Volume vol 22, pp. pp 31-72.
- Souza Filho, J. B. D. O., Gesualdi, A. R., Seixas, J. M. & Guedes, R. M., 2011. *Classificadores Multi-Classe de Contatos por Curvas Principais*. São João Del Rei, Sociedade Brasileira de Automação Inteligente.
- Sun, A. & Lim, E. P., 2002. *Hierarchical Text Classification and Evaluation*. California, s.n., pp. p.p 521-528.
- Trees, V., 2001. Detection, Estimation and Modulation Theory, part I and III. Em: s.l.:Jonh Wiley Sons.
- Verbeek, J., Vlassis, N. & Krose, B., 2000. A K-segments Algorithm for Finding Principal Curves. *Pattern Recognition Letters*, Volume vol. 23, pp. pp. 1009-1017.