

OTIMIZAÇÃO GENÉTICA DE CURVAS PRINCIPAIS PARA A CLASSIFICAÇÃO DE CONTATOS

Geancarlo G. M. da Rocha, João B. de O. e Souza Filho, Aline da Rocha Gesualdi

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Unidade Maracanã

Av. Maracanã 229, Maracanã. Rio de Janeiro, RJ, Brasil.

geancarlo.rocha@gmail.com, {jsouza, agesualdi}@cefet-rj.br

José Manoel de Seixas

Universidade Federal do Rio de Janeiro, LPS/COPPE/UFRJ

Centro de Tecnologia, Bloco H, sala 220 e Bloco I, sala 146. Ilha do Fundão

Rio de Janeiro, RJ, Brasil.

seixas@lps.ufrj.br

Resumo – A técnica de curvas principais é um poderoso extrator de características não-lineares que pode ser utilizada para fins de classificação. Classificadores para múltiplas classes podem ser construídos através da modelagem de cada classe por uma curva principal e um critério de decisão simples, tal como atribuir ao evento a classe associada à curva de maior proximidade, permite a construção de classificadores de elevada eficácia. Como, usualmente, as curvas principais são aproximadas por um conjunto de segmentos de reta conectados, para esta proposta é de extrema importância a definição apropriada do número de segmentos que a curva de cada classe possui, as quais possuem relação direta com a eficiência do classificador final. Trata-se, portanto, de um problema de otimização combinatória complexo, em especial quando o número de classes é elevado. Neste artigo é proposta a utilização de algoritmos genéticos para o dimensionamento de curvas principais aplicadas a construção de classificadores baseados em curvas principais. Considerando como aplicação alvo um sistema de classificação automática de contatos, desenvolvido através de uma base de dados que compreende sinais reais provenientes de 8 classes de navios, a técnica proposta resultou em classificadores de maior eficiência média e menor complexidade total, aspecto atrativo para a implementação deste sistema em tempo real.

Palavras-chave – Algoritmos genéticos, Técnicas de otimização, Curvas principais, Classificação, Extração de Características, Sonar Passivo.

Abstract – The principal curves technique is a powerful non-linear feature extractor that may be used for classification purposes. Multiclass classifiers may be built by modeling each class by a principal curve and a simple decision criterion, such as assigning an arbitrary event to the class related to the curve of highest proximity, makes possible the construction of high accuracy classifiers. Since principal curves usually are approximated by a set of connected line segments, for this proposal it is extremely relevant to properly define the number of segments which constitutes the curve of each class, which are directly related to the final classifier's efficiency. Thus, this is a complex combinatorial optimization problem, specially when there are several classes. This paper proposes the use of genetic algorithms to optimize the complexity of principal curves when applied to building classifiers. Considering as target application an automatic contact classification system, developed using a dataset containing signals belonging to 8 classes of ships, the proposed technique resulted in classifiers showing higher mean efficiency and lower overall complexity, which is an attractive aspect to its real-time implementation.

Keywords – Genetic algorithms, Optimization Techniques, Principal curves, Classification, Feature Extraction, Passive Sonar.

1. INTRODUÇÃO

Quando submersos, para a identificação de ameaças, os submarinos utilizam, principalmente, do sistema de sonar passivo. Em contraposição ao sonar ativo, que é baseado na emissão de ruídos específicos e na análise de ecos provenientes desta emissão, o sonar passivo apenas captura o ruído ambiente, evitando que a presença do submarino seja descoberta. Para a identificação de possíveis ameaças, isto é, para a classificação de um contato, o operador de sonar compara o sinal recebido com outras assinaturas conhecidas (padrões), através de sua sonoridade e por informações provenientes de análises complementares, entre elas a LOFAR (*low-frequency analysis and recording*) [1], definindo o contato mais provável.

Neste complexo cenário, sujeito a variadas condições operativas e de estresse, sistemas de classificação automática de contatos (SCAC) são ferramentas úteis. Para a construção deste sistema, a técnica de curvas principais mostra-se bastante atrativa, provendo um poderoso modelo não-linear dos dados [2]. Deste modo, classificadores com número arbitrário de classes podem ser

derivados através da modelagem de cada classe por uma curva representativa, atribuindo-se a um evento arbitrário a classe relativa a curva mais próxima. Trata-se, portanto, de um sistema de classificação altamente modular, com baixo custo computacional na fase de operação e facilmente paralelizável, tanto na fase de extração das curvas quanto de operação.

Usualmente, no processo de extração das curvas, estas são aproximadas por um conjunto de segmentos de reta conectados. Esta aproximação demanda a definição do quantitativo de segmentos que permita uma adequada representação das diferentes classes. Espera-se, deste modo, que um maior número de segmentos resulte numa melhor representação do conjunto de dados de uma dada classe. Em contra-partida, tem-se o problema do *overfitting*, pelo qual a adoção de curvas de maior complexidade pode vir a prejudicar a capacidade de generalização do classificador. Outro fator a ser considerado é a competição, durante a operação do classificador, das diferentes curvas das classes pelos eventos. Conclui-se, assim, que a otimização das curvas quanto a complexidade deva ser realizada de forma conjunta, isto é, envolvendo todas as curvas de classe envolvidas no sistema de classificação.

Num classificador baseado em curvas, entre opções para a definição da complexidade da curva associada a cada classe, tem-se o teste de todas combinações (força bruta), que possui custo computacional demasiado elevado, frequentemente intratável quando o número de classes é razoável. Índices estatísticos, entre eles a área da ROC, podem ser utilizados, porém, usualmente, tais métodos avaliam cada curva de classe individualmente, não refletindo a competição dos eventos pelas curvas das diferentes classes existente no classificador. Neste cenário, os algoritmos genéticos são uma alternativa interessante, por possuir uma boa capacidade de exploração e de investigação do espaço de soluções, hábeis, portanto, na resolução de problemas de otimização combinacional, tal como o dimensionamento do classificador proposto.

Este artigo propõe a utilização de algoritmos genéticos para a otimização da complexidade das curvas principais integrantes de um sistema de classificação automática de contatos. A estrutura do artigo é a que se segue: inicialmente, são brevemente abordados a técnica de curvas principais e o algoritmo de extração *k-segmentos* utilizado. Após, tem-se a apresentação do SCAC em questão e do algoritmo genético proposto para a aplicação. Por fim, são apresentados a base de dados utilizada, os resultados obtidos e as conclusões.

2. CURVAS PRINCIPAIS

As curvas principais são poderosas ferramentas para a modelagem de conjuntos de dados multidimensionais. Consideradas como uma generalização não-linear da análise de componentes principais, tratam-se de curvas unidimensionais que passam pelo “meio da nuvem formada pelos dados”, promovendo um resumo não-linear destes [2].

Usualmente, as curvas principais são aproximadas por um conjunto de segmentos conectados. Entre variados métodos de extração, destaca-se o algoritmo de *k-segmentos* [3](KSEGS), que é baseado na inserção e conexão de segmentos, definidos por componentes principais locais, em agrupamentos identificados nos dados. A dinâmica do algoritmo KSEGS é ilustrada na Figura 1. Resumidamente, o processo de extração é realizado de forma iterativa. A cada iteração, é procurado um agrupamento de dados para a inserção de um novo segmento. Após o processo de inserção de um segmento, busca-se sua interconexão com a curva, através de um processo baseado em busca gananciosa (*greedy* [4]). Este processo se repete até que um critério de parada seja atingido (o limite de segmentos definido pelo usuário ou o número de pontos do grupo candidato a inserção de segmento seja inferior a 3).

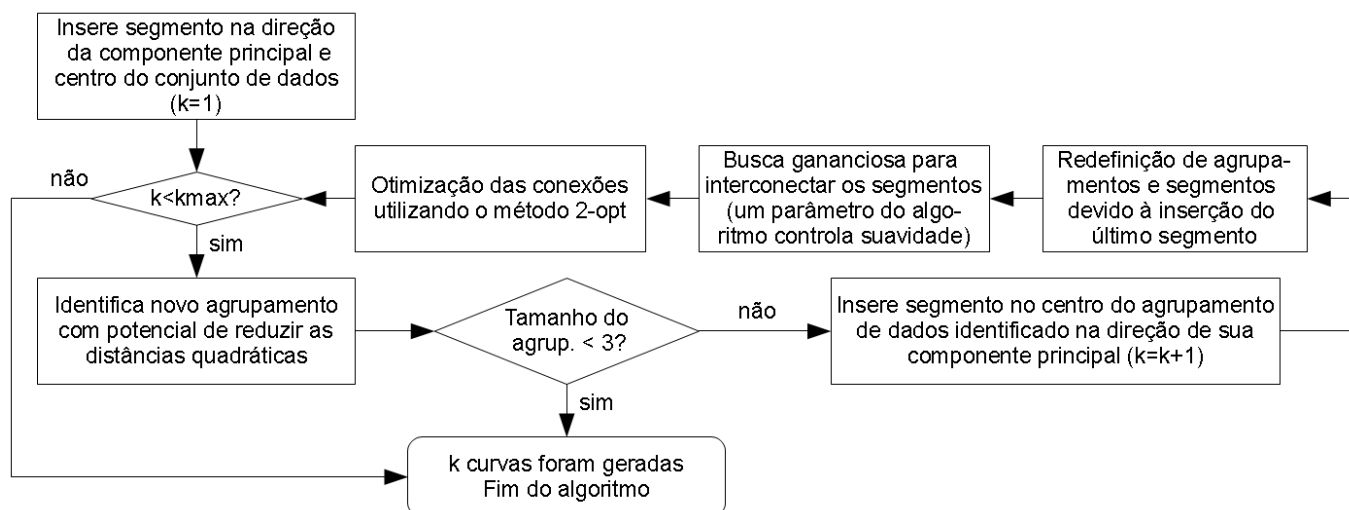


Figura 1: Ilustração do algoritmo de *k-segmentos* de extração de curvas principais.

2.1 SCAC baseado em curvas principais

De posse das curvas representativas de cada classe, a um evento arbitrário (x_i) é possível inferir duas informações relevantes à construção de um SCAC baseado em curvas principais:

- O índice t_{ij} da projeção do evento x_i , que pode ser usado como um indicador da distribuição espacial dos dados ao longo da curva representativa de uma classe arbitrária (j).
- A distância euclidiana do ponto à curva, definida como: $d(x_i, f_j) = \inf_{t \in \mathbb{R}} \|f_j(t) - x_i\|$, simples indicador de similaridade do evento a curva da mesma classe ($f_j(t)$).

O desenvolvimento de um SCAC baseado em curvas principais pressupõe diferentes etapas, conforme ilustrado na Figura 2. Duas fases principais podem ser identificadas: o projeto *offline* e a operação *on-line*. Na primeira, é produzida uma base de assinaturas acústicas, através do pré-processamento e extração de uma curva principal para cada classe de contato, que é realizada pelo algoritmo de *k-segmentos*. Em fase de operação, os sinais são adquiridos pelo sistema de sonar passivo, pré-processados e sua similaridade com relação às curvas representativas de cada classe é inferida, através do critério de distância mais próxima.

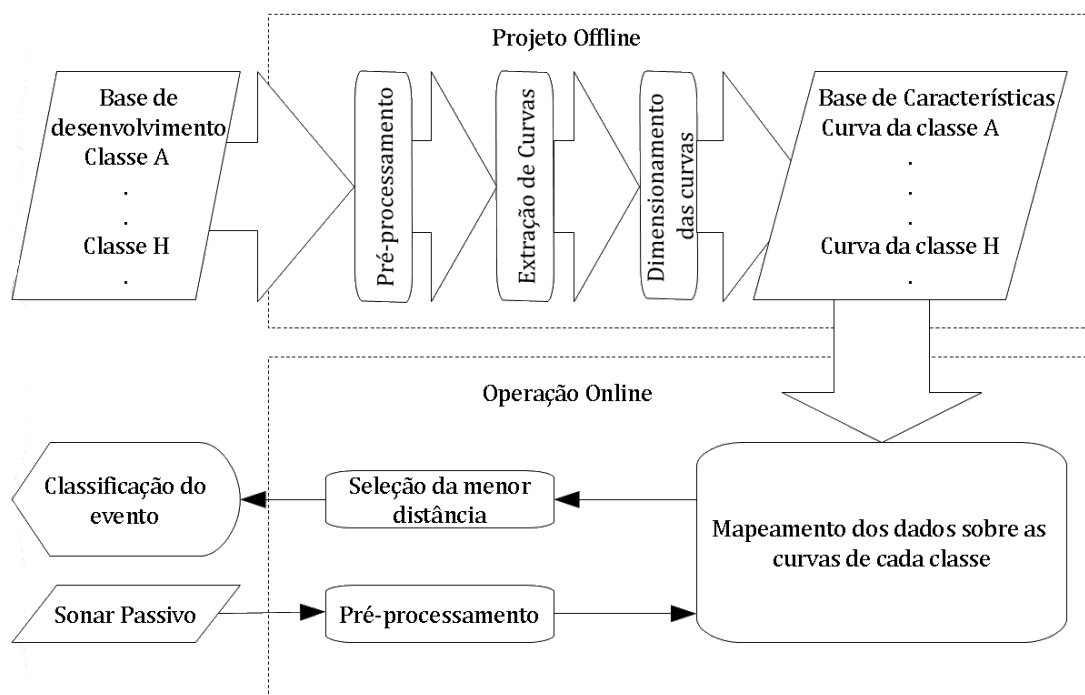


Figura 2: Etapas relativas a produção e operação de um sistema de classificação automática de contatos baseado em Curvas Principais.

Uma importante etapa no projeto do classificador baseado em curvas em discussão consiste na identificação do número de segmentos da curva representativa de cada classe (complexidade de cada curva). Ainda que o algoritmo de extração possa extrair curvas até um limite definido por um parâmetro de entrada, desde que não haja ocorrência das condições de parada, o aumento da complexidade da curva nem sempre resulta numa melhor capacidade de generalização, face a possibilidade de ocorrência de *overfitting* [5].

Uma métrica interessante para a definição destas complexidades é a eficiência de generalização do sistema de classificação. Por esta proposta, para múltiplas classes, um indicador útil é a eficiência média. Assim, na avaliação comparativa de diferentes estratégias para a definição da complexidade, foi realizada uma partição da base de dados disponível em dois conjuntos: projeto e teste; o primeiro utilizado para a extração da curva; e o último, para a avaliação de desempenho.

Alternativamente a avaliação da seleção das complexidades por algoritmos genéticos, também foram consideradas duas alternativas: a seleção pela área da ROC e a construtiva. Na primeira, a curva de cada classe é avaliada isoladamente, sendo determinada uma ROC associada à distribuição das distâncias dos dados da classe para sua curva representativa e das demais classes com respeito a ela. Trata-se, portanto, de um critério que busca curvas cuja distância “média” aos eventos de sua classe é mínima e com relação às demais classes, máxima. Em outras palavras, curvas que possuam máxima discriminabilidade entre os eventos que elas representam e aqueles provenientes de outras classes. Este critério, no entanto, não considera a competição das diferentes curvas de classe pelo eventos de entrada do classificador. Pelo segundo método, é suposto que as curvas possuem uma mesma complexidade, sendo a seleção realizada por validação cruzada, de acordo com o princípio de parcimônia, isto é, buscando-se curvas de complexidade mínima e que resultem numa eficiência máxima. Este critério, no entanto, não considera as especificidades de cada classe, que poderiam resultar em mapeamentos mais ou menos econômicos.

3. ALGORITMOS GENÉTICOS

Algoritmos genéticos (AGs) foram originalmente propostos em [6]. São algoritmos de busca de propósito geral, inspirados em questões genéticas observadas em populações de seres vivos. Por ser um algoritmo de busca heurística, não é possível confirmar se a solução encontrada é ótima, porém os resultados obtidos para uma vasta gama de aplicações são geralmente satisfatórios. O algoritmo genético utilizado neste trabalho pode ser sintetizado como:

1. Gerar população inicial
2. Repetir
 - (i) Avaliar a aptidão dos indivíduos e atualizar a solução;
 - (ii) Verificar os critérios de parada;
 - (iii) Selecionar os ancestrais preservando a elite atual;
 - (iv) Recombinar os ancestrais de acordo com a taxa de recombinação;
 - (v) Mutar os recombinantes e ancestrais restantes, de acordo com a probabilidade de mutação;
 - (vi) Formar uma nova população com a elite atual e os novos descendentes.
 - (vii) Repetir a partir do passo (i) até o limite de gerações

Resumidamente, a seguir serão abordados os principais conceitos desta técnica, com as respectivas escolhas de implementação realizadas.

3.1 População e Cromossomos

Um cromossomo representa um indivíduo numa população do algoritmo genético. A informação contida neste vetor é utilizada como parâmetro da função de aptidão, devendo ser codificada de acordo com os operadores de recombinação e mutação utilizados. Para este problema de definição da complexidade das curvas, optou-se pela codificação do cromossomo em 8 genes, cada qual contendo a complexidade de uma classe.

3.2 Função de Aptidão

A função de aptidão infere a qualidade da solução carregada pelo indivíduo da população. Quanto maior sua aptidão, melhor é a solução que seu cromossomo carrega e também maior é a chance de que ele se reproduza. A função de aptidão adotada consiste no valor da eficiência média de classificação para um conjunto de curvas de classe com complexidade definida por dado indivíduo, inferida com base no conjunto de teste.

3.3 Elitismo

O elitismo consiste simplesmente na conservação de alguns indivíduos de maior aptidão, com a finalidade de garantir que sua carga genética seja explorada e possivelmente recombinada com um novo indivíduo de maior aptidão.

3.4 Seleção de Indivíduos

Novos indivíduos são gerados através de recombinação e mutação a cada geração. O processo para a seleção de quais indivíduos reproduzirão para gerar descendentes é feito através de um processo estocástico que leva em consideração a aptidão *a priori* dos indivíduos. O método de seleção utilizado é a *seleção estocástica universal* [7].

Para se formar uma nova geração, deverão ser escolhidos dois ancestrais para cada descendente recombinante, e um ancestral para preencher o resto da população, já levando em conta o número de indivíduos da elite. Desta forma, o total de ancestrais selecionados é calculado como:

$$n_{Recombinantes} = \text{round}(taxa_{Recombinantes} * (tamanhoPopulacao - n_{Elites})) \quad (1)$$

$$n_{Mutantes} = tamanhoPopulacao - n_{Recombinantes} - n_{Elites} \quad (2)$$

$$n_{Ancestrais} = 2 * n_{Recombinantes} + n_{Mutantes} \quad (3)$$

3.5 Operador de Recombinação

A recombinação entre os indivíduos tem por objetivo explorar as soluções encontradas, aproveitando-se do fato de que a maioria delas possuem boa aptidão. O operador baseado no *swap* de sequências de gene foi criado pensando-se no *fenótipo* dos indivíduos, ou seja, naquilo que os indivíduos representam, um conjunto de complexidades de curvas principais.

Para cada par de ancestrais, primeiramente, calcula-se o tamanho da sequência a ser recombinada segundo uma função de probabilidade que atribui uma maior chance às sequências curtas. A função proposta é dada por:

$$p[n] = \begin{cases} \left(\sum_{n=1}^{N-1} n \right)^{-1} & \text{se } n = N - 1 \\ (N - n) \times p[N - 1] & \text{se } 1 \leq n \leq N - 2 \\ 0 & \text{caso contrário} \end{cases} \quad (4)$$

onde N é o número de genes do cromossomo (número de classes, no nosso caso). Para $N = 8$, tem-se:

$$p[1..8] = [0.2500, 0.2143, 0.1786, 0.1429, 0.1071, 0.0714, 0.0357, 0]$$

Determinado o tamanho n da sequência, o índice i do primeiro gene desta sequência é escolhido utilizando-se uma função de probabilidade uniforme. Por fim, um dos dois ancestrais é selecionado para formar a base do cromossomo e receber a sequência do outro ancestral.

3.6 Operador de Mutação

A mutação objetiva a exploração do espaço de soluções, levando à descoberta de novas soluções e a redescoberta de informações eventualmente perdidas. Desta forma, trata-se de uma estratégia para a fuga de soluções ótimas locais.

O operador definido para este fim primeiramente calcula o número de genes a serem mutados (n), segundo uma função de probabilidade uniforme. Selecionado o número de genes, seus índices são sorteados sem reposição. Cada um dos genes escolhidos recebe então um novo valor, pertencente a faixa de valores possíveis, ou seja, entre 1 e a maior complexidade admissível a curva respectiva de dada classe.

Este operador é aplicado ao conjunto de indivíduos formados pelos ancestrais não recombinados e aos descendentes recombinantes. Os indivíduos resultantes desta operação são concatenados a elite, formando a população da próxima geração.

4. RESULTADOS

Este trabalho utilizou um conjunto de dados constituído pelo ruído irradiado por 25 navios, pertencentes a 8 diferentes classes, durante 263 corridas realizadas na raia acústica da Marinha do Brasil. Em cada corrida, um navio percorre a raia, mantendo constantes as condições de maquinaria e operação. Diferentes condições operativas são experimentadas. Este ruído é capturado por um hidrofone junto ao fundo da raia e processado pela cadeia ilustrada na Figura 3.

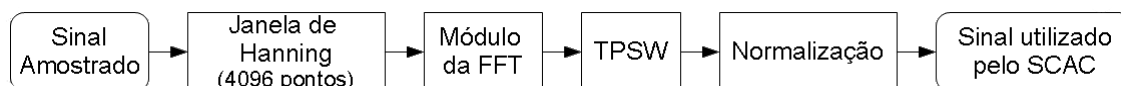


Figura 3: Cadeia de pré-processamento aplicada aos dados.

Inicialmente, o sinal do hidrofone é amostrado a 22,05 kHz e quantizado em 16 bits de resolução. O sinal resultante é submetido a um janelamento de Hanning, que visa minimizar os efeitos de recorte do sinal, e em seguida é calculada sua transformada de Fourier. São considerados apenas os 557 primeiros pontos (em módulo) de cada janela, que responde pela faixa de frequências de 0 a ≈ 3 kHz, pois concentra informações relacionadas à maquinaria dos contatos [8]. Em seguida, é realizado o algoritmo TPSW [9], que visa diminuir o ruído de fundo do sinal, o qual é normalizado para energia unitária. A Tabela 1 resume o número de corridas e janelas espectrais derivadas pela sequência de processamento apresentada.

Tabela 1: Resumo do conjunto de dados empregado

	Classe A	Classe B	Classe C	Classe D	Classe E	Classe F	Classe G	Classe H	# Total
# Corridas	22	26	43	29	66	27	19	31	263
# Espectros	2432	3432	4797	3072	7075	2934	2143	3392	29277

A proposta de definição da complexidade da curva por algoritmos genéticos foi comparada com dois outros métodos: a seleção por área da ROC e a incremental. Para cada método foram avaliadas as eficiências de classificação obtidas, juntamente com a complexidade das curvas. Para a estimação da eficiência dos classificadores, foi utilizado o método de *10-fold cross-validation* [10]. A extração da curvas principais representativas de cada classe foi baseada no algoritmo KSEGS, considerando um máximo de 30 segmentos (valor de $kmax$), assim como os seguintes parâmetros: $kmax = 30$, $sigma = 0$ e $lambda = 1$ [3]. Os parâmetros do algoritmo genético, ajustados através de alguns ensaios, foram: *tamanho da população* igual a 1000, *preservação de 2 indivíduos de elite*, *taxa de recombinação* igual a 80% e *chance de mutação* igual a 60%.

Na Tabela 2 são apresentados o valor médio e o desvio padrão ($\hat{\mu} \pm \hat{\sigma}$) associado aos valores de eficiência para cada classe de contato para os diferentes métodos de seleção. Em negrito são exibidas as maiores eficiências obtidas por classe. É possível observar que a seleção por algoritmos genéticos resulta sempre em maiores eficiências por classe, exceto para a classe C.

Tabela 2: Valores de eficiência por classe para cada método de seleção de complexidade de curva (média e desvio padrão - maior é melhor) - valores percentuais.

Método	A	B	C	D	E	F	G	H
INC	81,0 ± 11,4	64,6 ± 16,7	58,8 ± 22,3	89,5 ± 13,6	79,9 ± 15,4	90,3 ± 7,0	89,7 ± 8,6	73,2 ± 18,8
aROC	80,7 ± 13,2	69,0 ± 19,6	65,3 ± 19,1	87,4 ± 17,0	83,0 ± 12,3	88,6 ± 8,9	85,7 ± 9,4	70,1 ± 22,4
AG	84,3 ± 11,3	73,5 ± 18,6	61,6 ± 23,9	92,7 ± 7,8	89,4 ± 9,1	95,0 ± 3,7	92,6 ± 4,7	78,5 ± 12,9

Em seguida, os métodos foram comparados em termos da complexidade das curvas por eles selecionadas. Na Tabela 3 são exibidos o valor médio e o desvio padrão das complexidades das curvas obtidas. Pode-se observar que a seleção genética resultou em curvas menos complexas (em média) para metade das classes (A,C,F e H). O método incremental, no entanto, resultou em curvas "menores" para as classes B e E, enquanto resultado similar foi obtido pela área da ROC para as classes D e G. Interessante observar que para as classes B, D, E e G, a seleção genética resultou em valores intermediários de complexidade com relação aos demais métodos. Outro fato é que as maiores eficiências de classificação por classe nunca estiveram relacionadas as maiores complexidades médias, o que evidencia que curvas de maior complexidade tendem, de fato, a comprometer a capacidade de generalização do classificador.

Tabela 3: Complexidades das curvas por classes (média e desvio padrão - menor é melhor).

Método	A	B	C	D	E	F	G	H
INC	9,9 ± 4,8	14,1 ± 8,0	14,1 ± 8,0	5,1 ± 1,8	14,1 ± 8,0	13,5 ± 7,3	8,3 ± 3,9	13,7 ± 7,5
aROC	7,0 ± 4,1	23,6 ± 8,3	22,7 ± 6,5	2,0 ± 1,9	19,3 ± 9,6	14,9 ± 6,0	4,9 ± 4,0	8,5 ± 6,5
AG	4,4 ± 4,4	16,5 ± 7,8	11,6 ± 10,2	2,1 ± 1,4	16,5 ± 10,4	8,8 ± 7,4	5,5 ± 3,1	5,5 ± 4,6

Na Tabela 4 são sintetizados os valores de eficiência média e de complexidade total para os diferentes métodos de seleção. A seleção genética resulta num classificador de maior eficiência e com menor complexidade média. Na Figura 4 são apresentados os *boxplots* relativos as eficiências médias e complexidades totais. Pode-se observar que o seleção genética, quanto a eficiência, possui valores interquartil (25% e 75%) superiores aos demais métodos e que este intervalo é menor que os demais, sinalizando uma menor flutuação das eficiências, ou seja, uma menor susceptibilidade desta técnica de seleção quanto a escolha do conjunto de extração das curvas, aspecto que é atrativo. Por outro lado, em termos de complexidade, a distância interquartil para os algoritmos genéticos apresenta um valor intermediário entre os demais métodos.

Tabela 4: Valores de eficiência média e da complexidade total para os diferentes métodos de seleção.

Método	Eficiência Média	# Total de Segmentos
INC	78,4 ± 4,7	177,6 ± 96,3
aROC	78,7 ± 5,2	197,8 ± 31,6
AG	83,4 ± 4,6	133,8 ± 42,1

5. CONCLUSÃO

Este artigo discutiu a produção de sistemas de classificação baseados em curvas principais, para os quais a seleção da complexidade de cada curva representativa de classe é realizada por algoritmos genéticos.

Considerando o problema da classificação automática de contatos de sonar passivo, este sistema de classificação possui como principais atrativos a modularidade, eficácia e robustez, além de baixo custo computacional em fase de operação, o que é de extrema valia para sua necessária operação em tempo real.

A técnica genética se mostrou mais eficaz que outros métodos em termos da eficiência média e da complexidade total do classificador. Foi também observada uma menor susceptibilidade quanto à eficiência média em respeito à escolha do conjunto de projeto do classificador. Através desta proposta, foi possível obter um classificador cuja mínima eficiência obtida por classe foi de 61,6 ± 23,9% (classe C); a máxima, de 95,0 ± 3,7% (classe F), totalizando uma eficiência média de 83,4 ± 4,6. Em termos de complexidades, a classe D foi a que demandou curvas de menor complexidade (2,1 ± 1,4); e a classe E, maior (16,5 ± 10,4).

Como trabalhos futuros, pretende-se investigar outros critérios de classificação alternativos à distância a curva mais próxima. É também previsto que outros métodos de integração do conhecimento extraído dos dados pelas curvas sejam explorados.

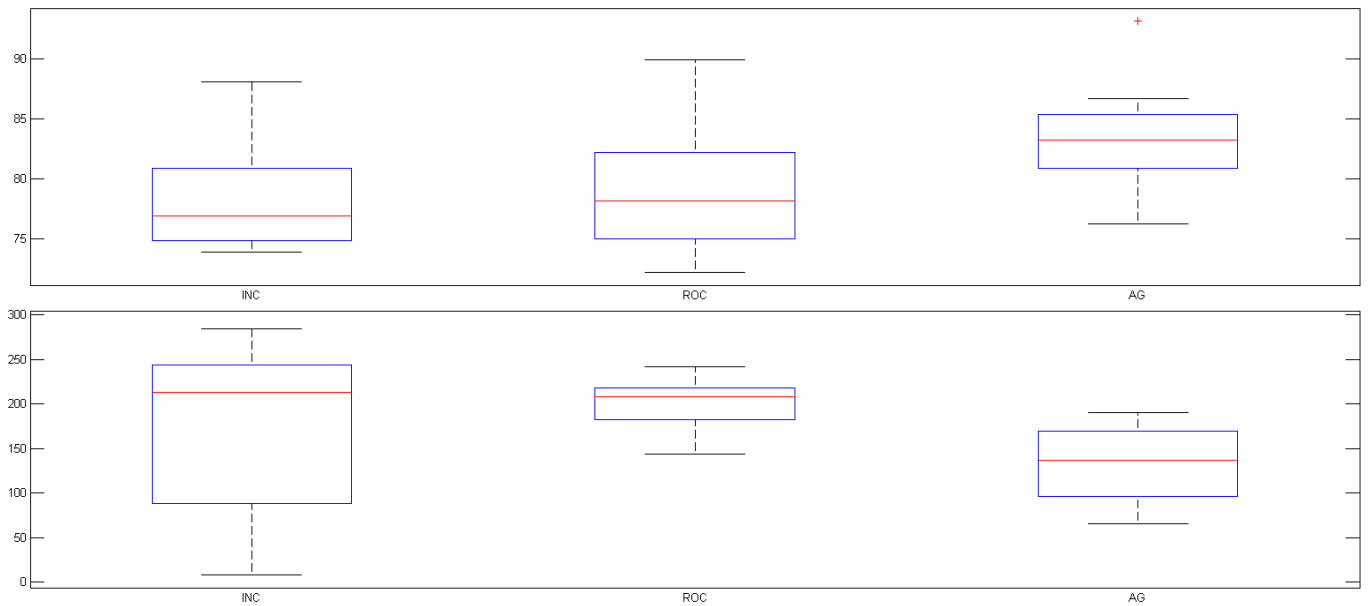


Figura 4: Eficiências médias(acima) e complexidades totais(abaixo) obtidas para cada método de seleção.

REFERÊNCIAS

- [1] G. R. Arrabito, B. E. Cooke and S. M. McFadden. “Recommendations for enhancing the role of the auditory modality for processing sonar data”. *Applied Acoustics*, vol. 66, pp. 986–1005, 2005.
- [2] T. Hastie and W. Stuetzle. “Principal Curves”. *Journal of the American Statistical Association*, vol. 84, no. 406, pp. 502–516, 1989.
- [3] J. Verbeek, N. Vlassis and B. Kröse. “A k-segments algorithm for finding principal curves”. *Pattern Recognition Letters*, vol. 23, no. 8, pp. 1009–1017, 2002.
- [4] S. Dasgupta, C. H. Papadimitriou and U. Vazirani. *Algorithms*. McGraw-Hill, Julho 2006.
- [5] R. DUDA, P. HART and D. STORK. *Pattern Classification*. John Willey & Sons, 2001.
- [6] J. H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975.
- [7] J. E. Baker. “Reducing bias and inefficiency in the selection algorithm”. In *Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application*, pp. 14–21, Hillsdale, NJ, USA, 1987. L. Erlbaum Associates Inc.
- [8] W. S. Filho. “Classificação do Ruído Irradiado por Navios usando Redes Neurais”. Ph.D. thesis, COPPE/UFRJ, 2001.
- [9] R. O. Nielsen. *Sonar Signal Processing*. Artech House, 1991.
- [10] R. Kohavi. “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection”. pp. 1137–1143. Morgan Kaufmann, 1995.