

# Algoritmo de Seleção de Classificadores para Reconhecimento de Padrões

Willamos A. S e Silva, Carmelo J. A. Bastos-Filho

<sup>†</sup>*Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil*  
wass@ecomp.poli.br, carmelofilho@ieee.org

**Resumo**—Este artigo apresenta uma proposta de algoritmo para classificação de padrões baseado em comitê de classificadores. A seleção de classificadores é realizada entre as soluções não dominadas geradas por um algoritmo de otimização multi-objetivo. Esta seleção é baseada em uma técnica de exclusão visando minimizar o espaçamento médio das soluções (*spacing*) no espaço de objetivos. Para validação da proposta foram utilizados classificadores do tipo *AdaBoost*. O algoritmo multi-objetivo utilizado para geração das soluções não dominadas foi o *DFMOPSO (Diversity Factor Multi-Objective Particle Swarm Optimization)* e o mecanismo de exclusão de classificadores foi baseado em *spacing*. Os comitês de classificadores gerados apresentaram soluções que dominam as soluções da frente de Pareto gerado a partir do treinamento efetuado. Após aplicar o algoritmo de exclusão de classificadores, obteve-se um resultado melhor do que o comitê com todas as soluções de Pareto inicial. Foram realizados testes num problema de detecção de plaquetas e nesta base observou-se uma melhoria de 28% na taxa de falso negativo (*miss rate*) e 48% na taxa de falso positivo.

**Palavras chave**—otimização multi-objetivo, *AdaBoost*, comitês, seleção de classificadores.

## I. INTRODUÇÃO

Os seres humanos possuem uma grande capacidade de definir padrões para classificar objetos. Esta classificação se dá em nosso cérebro devido à sua alta capacidade de extrair características e, a partir delas, efetuar a classificação [1].

Computacionalmente, o reconhecimento de padrões pode ser realizado de maneira semelhante. A classificação de novos objetos também pode ser feita a partir de características extraídas e, a partir de um treinamento prévio, o dado objeto é classificado como um elemento de uma dada classe. A classe escolhida para o objeto em questão é a que apresentar um menor risco de erro na classificação [2]. Dentre as características que podem ser extraídas com este propósito, encontram-se forma, cor, textura, entre outras.

A seleção de características é importante para a construção do modelo de aprendizado [3]. As características devem ser escolhidas visando minimizar sua quantidade e maximizar o desempenho do classificador [4].

Algumas técnicas são amplamente utilizadas para reconhecimento de padrões, entre elas estão:

- *AdaBoost*: combina classificadores simples (ditos fracos) que juntos formam um classificador forte. Tipicamente são utilizados para classificar objetos usando informação de forma [6] [7];
- Classificadores Bayesianos: fazem uso do conjunto de treinamento para criar um modelo probabilístico e, a partir deste modelo, classificar os objetos apresentados [8];

- Redes Neurais Artificiais (NN): foram inspiradas no cérebro humano e são compostas por unidades de processamento (neurônios) conectadas umas às outras. Para um dado conjunto de entrada realizam uma avaliação ponderada que permite a classificação em uma das classes previstas [9];
- Máquinas de vetor de suporte (SVMs): buscam hiperplanos em espaços de mais alta dimensionalidade que separam os padrões de treinamento [10].

Para melhorar o desempenho de algumas dessas técnicas se faz uso da combinação de alguns classificadores diferentes gerados pelas técnicas para a formação de comitês de classificadores [11] [12]. É fundamental que se escolha bem os classificadores que farão parte do comitê para evitar vieses. Para tal, deve-se ter diversidade nos classificadores do comitê [13]. Uma opção para criar estes classificadores é a partir de algoritmos multiobjetivos, fazendo uso da noção de dominância [14].

Com as técnicas de otimização multi-objetivo é gerado como resultado uma frente de Pareto. Uma frente de Pareto é um conjunto de soluções não dominadas, ou seja, que não são comparáveis entre si. Em um problema com dois objetivos, por exemplo, é dito que uma solução A domina a solução B, caso a primeira apresente melhores resultados nos dois objetivos em relação à segunda. Caso a solução A seja melhor que a solução B em um objetivo e pior no outro, diz-se que A e B são incomparáveis entre si. Para este trabalho se tem dois objetivos a serem otimizados pelos algoritmos: diminuir a taxa de falso positivo (quantidade de objetos classificados erroneamente) e minimizar a taxa de falso negativo ou *missrate*, que é a quantidade de elementos da classe que se deseja encontrar e não são reconhecidos. A técnica utilizada terá inspiração em algoritmos de inteligência de enxames, *DFMOPSO* (do inglês *Diversity Factor Multi-Objective Particle Swarm Optimization*) [15] para encontrar um conjunto de soluções, em que cada solução representa uma configuração para um classificador. Entretanto estes algoritmos multi-objetivos podem gerar muitas soluções em uma única região da frente de Pareto, podendo gerar assim um enviesamento do comitê gerado com essas soluções. Neste trabalho propõe-se ainda um critério de exclusão que minimize o desvio padrão do espaçamento médio entre as soluções adjacentes. Este procedimento busca melhor distribuir as soluções que compõem a frente de Pareto e usar este conjunto resultante de soluções para compor um comitê menos enviesado.

Este artigo está organizado da seguinte forma: a Seção II apresenta conceitos básicos dos classificadores utilizados neste trabalho para composição dos comitês, assim como

alguns métodos para criação de comitês. A Seção III apresenta o algoritmo de seleção de classificadores baseado na exclusão de soluções em regiões do espaço de objetivos visando minimizar a inclusão de um viés forte no comitê. A Seção IV apresenta a metodologia e arranjo experimental. As Seções V e VI apresentam os resultados e conclusões, respectivamente.

## II. ESTADO DA ARTE

*Adaboost* (*Adaptive Boosting*) é uma técnica de classificação baseada na forma dos objetos a serem classificados. Esse algoritmo faz uso do método de *Boosting* [16], que forma um classificador forte baseado na combinação de classificadores simples, ditos fracos [17]. Cada estágio da cascata é composto por um classificador fraco e a combinação dos estágios em sequência apresenta um efeito de redução no erro de classificação conforme um objeto é aceito de um estado a outro [7].

Um classificador fraco é composto por uma característica de Haar  $h_i(x)$ , um limiar  $\theta_i$  e uma paridade  $p_i$  [6]. O objetivo do treinamento é encontrar a melhor combinação possível de maneira a associar cada característica  $h_i(x)$  ao seu limiar  $\theta_i$  para evitar uma classificação incorreta. A Equação 1 define um classificador fraco, enquanto a Equação 2 representa um classificador forte. A característica de Haar tem como função detectar formas. Como pode ser visto na Figura 1, a característica mais à esquerda é responsável por detectar bordas verticais, enquanto a seguinte é apropriada para detecção de bordas horizontais. A característica seguinte é responsável por detectar linhas verticais, enquanto a última característica, mais à direita, é apropriada para detecção de bordas diagonais.

$$f(x) = \begin{cases} 1, & \text{se } p_j f_j < p_j \theta_j \\ 0, & \text{caso contrário} \end{cases} \quad (1)$$

$$H(x) = \alpha_1 h_1 + \alpha_2 h_2 + \dots + \alpha_n h_n \quad (2)$$



Figura 1. Características de Haar.

Juntamente com as características de Haar surgiu uma nova forma de representação de informação em imagens. O conceito de imagem integral consiste em representar a informação de um *pixel*  $P$  não mais apenas pela sua intensidade, mas sim pela soma das intensidades de todos os *pixels* localizados acima e a esquerda do ponto  $P$ . A vantagem desta representação é facilitar o cálculo da informação de qualquer retângulo da imagem a partir de uma única iteração do algoritmo [17].

Viola e Jones apresentaram um classificador *AdaBoost* genérico que detecta faces [6]. Posteriormente, Firmo e Bastos-Filho [18] demonstraram que este classificador poderia ser utilizado para detectar ovos de *Schistosoma mansoni* em exames de fezes. Mais recentemente, Nascimento e

Bastos-Filho [19] aplicaram a mesma técnica para detecção de plaquetas sanguíneas. Nascimento e Bastos-Filho também estenderam o reconhecimento para os outros elementos sanguíneos [20]. Mais recentemente, Bastos-Filho e Silva [21] realizaram um estudo comparativo entre técnicas mono-objetivas e multi-objetivas para efetuar o treinamento *AdaBoost* e comparou seus resultados com o mesmo propósito: detectar plaquetas.

Pode-se fazer uso de  $n$  classificadores para que tomem a decisão em conjunto, formando um comitê classificador. Trabalhar com a ideia de um comitê é interessante pelo aumento na capacidade de generalização das ferramentas de aprendizado de máquina [11]. Os comitês são simples, e devido a isso, têm sido bastante estudados. Pode-se formar comitês de classificadores compostos por algumas técnicas de aprendizado de máquina, entre elas *Bagging*, *Boosting*, redes neurais, entre outras [12].

No processo de criação do comitê é importante selecionar corretamente os classificadores a serem utilizados no processo de seleção. Não é verdade que um comitê composto por uma quantidade maior de classificadores será necessariamente melhor do que um comitê com menos classificadores, pois comitês maiores podem apresentar sobre-ajuste do ruído no conjunto de treinamento [11]. Logo, é interessante que haja a poda desses elementos sobre-ajustados para que se tenha uma melhor capacidade de generalização.

A capacidade de generalização depende bastante da acurácia e diversidade dos classificadores no comitê [22]. Essa diversidade pode ser observada nos classificadores que formam o Pareto de soluções encontrados pelo algoritmo de treinamento do comitê e deve ser mantida no processo de seleção dos classificadores.

Algumas formas de selecionar os classificadores já foram desenvolvidas, entre elas o *Kappa pruning* e o *concurrency pruning*, que possuem características de busca gulosa [11]. O *Kappa pruning* mede o grau de similaridade entre dois classificadores para fazer avaliação da diversidade do comitê [23]. Os classificadores são comparados dois a dois e, idealmente, os classificadores devem possuir uma boa acurácia individual e serem o menos semelhantes o possível entre si. O *concurrency pruning* baseia-se no acerto do classificador e do comitê para seleção dos classificadores. Um classificador é recompensado por obter uma decisão correta e ainda mais recompensado caso este tome uma decisão correta e o comitê erre na sua decisão. De outra forma, caso o classificador ou o comitê erre, o classificador é penalizado [23]. Outras abordagens buscam a otimização utilizando Algoritmos Genéticos (GA), atribuindo aos classificadores pesos visando buscar a melhor combinação possível dos pesos na tentativa de minimizar o erro durante o treinamento, tornando o algoritmo sensível a ruídos [11]. A escolha da quantidade de classificadores tem impacto na acurácia. A seleção dos classificadores bem como sua quantidade pode ser encarado como um problema de busca [14].

## III. ALGORITMO DE SELEÇÃO DE CLASSIFICADORES NÃO-DOMINADOS

Neste artigo é proposto um algoritmo para seleção dos classificadores que serão utilizados no comitê. O algoritmo leva em consideração a distribuição dos classificadores ao

longo da frente de Pareto, removendo as soluções mais próximas umas das outras, visando a minimizar o desvio da distância média entre as soluções. Para tal é efetuado o cálculo da distância média entre os classificadores do Pareto. O passo seguinte é encontrar o classificador com a menor distância em relação a seus dois vizinhos: anterior, à esquerda, denotada por ( $d_e$ ), e posterior, à direita, denotada por ( $d_d$ ). Se ambas as distâncias forem menores do que a distância média multiplicada por um limiar ( $l_d$ ) o ponto é removido, voltando ao ponto de início do algoritmo. Caso essa condição não seja verdadeira, testa-se ainda se um dos dois vizinhos citados possui distância menor do que um outro limiar ( $l_u$ ). Caso isto aconteça, o classificador em questão deve ser removido. Caso algum classificador seja removido do Pareto, o algoritmo deve ser reiniciado com o conjunto atual de classificadores e a frente de Pareto reavaliada. O pseudocódigo do algoritmo de seleção é apresentado no Algoritmo 1. Os valores de  $l_d$  e  $l_u$  foram escolhidos baseados em testes paramétricos preliminares. Testou-se combinações com seus respectivos valores variando no intervalo [0,2;1,0] sendo incrementados de 0,1. Uma vantagem deste método é a ausência da necessidade de determinação da quantidade de classificadores que devem existir na frente de Pareto final, utilizada para composição do comitê.

---

**Algoritmo 1:** Pseudocódigo do algoritmo de seleção de classificadores .

---

```

1 início
2   └─ Receba o conjunto de classificadores do Pareto;
3 enquanto houver classificadores não checados faça
4   └─ Calcule a distância média ( $d_m$ );
5   └─ para cada classificador  $i$  faça
6     └─ Calcule as distâncias para seus dois vizinhos
7       └─ ( $d_i$ );
8       └─ se  $d_e < d_m \cdot l_d$  e  $d_d < d_m \cdot l_d$  então
9         └─ Remova o classificador  $i$ ;
10      └─ senão se  $d_e < d_m \cdot l_u$  ou  $d_d < d_m \cdot l_u$  então
11        └─ Remova o classificador  $i$ ;
12        └─ se houve remoção então
13          └─ Reinicialize o laço enquanto;
13 Retorne as soluções selecionadas;

```

---

#### IV. METODOLOGIA

Neste trabalho foi utilizado um comitê de classificadores *AdaBoost*. Os parâmetros de treinamento são os mesmos de um classificador fraco: característica de Haar  $h_i(x)$ , que possui largura, altura, posição na imagem (posição nos eixos  $x$  e  $y$ ), um limiar  $\theta_i$  e uma paridade  $p_i$  [6]. Sendo assim, cada classificador fraco da cascata implica em mais seis variáveis no espaço de busca. Foram treinados classificadores que são compostos por cascatas que possuem entre cinco e vinte estágios, totalizando entre trinta e cento e vinte dimensões no espaço de busca. Estes parâmetros são obtidos através de treinamento realizado utilizando a técnica de inteligência computacional DFMOPSO.

O DFMOPSO faz uso de um arquivo externo para armazenar as soluções não dominadas. Nesse algoritmo tem-se a utilização do fator de diversidade (DF), que pode ser observado na equação (3) de cada solução para a escolha de um líder cognitivo a ser alterado caso uma nova solução domine o líder atual. Antes de calcular o DF é necessário calcular a distância média da solução  $i$  para as demais soluções ( $d_i$ ), onde pode-se obter a distância mínima ( $d_{min}$ ) e a distância máxima ( $d_{max}$ ),

$$DF_i = \frac{d_i - d_{min}}{d_{max} - d_{min}}. \quad (3)$$

Para limitar o tamanho do arquivo externo é feito o uso do *crowding distance* (CD) e desprezando as soluções com piores valores nesta métrica. O CD é calculado através da estimativa do maior hiper-cubóide em torno da solução  $i$  sem que qualquer outra solução encontre-se no seu interior.

O treinamento do *DFMOPSO-AdaBoost* utilizou 1000 iterações, 100 partículas na população e tamanho do arquivo externo limitado a 100 partículas. Os valores das constantes de aceleração foram definidas aleatoriamente entre 1,5 e 2,5 e as partículas são inicializadas aleatoriamente. Estes valores foram definidos a partir de experimentos de teste preliminares. O pseudocódigo do algoritmo utilizado no treinamento *AdaBoost* pode ser visto no Algoritmo 2.

---

**Algoritmo 2:** Pseudocódigo de treinamento do *DFMOPSO-AdaBoost* .

---

```

1 Carregue as imagens da base de dados;
2 Inicialize aleatoriamente os indivíduos da população;
3 para cada partícula  $i$  faça
4   └─ Calcule o limiar do AdaBoost;
5   └─ Calcule a polaridade do classificador;
6 Selecione indivíduos para o arquivo externo;
7 enquanto não atinge critério de parada faça
8   └─ para cada partícula  $i$  faça
9     └─ Aplique o operador de turbulência utilizado no
10       └─ MOPSO;
11       └─ Selecione líderes sociais utilizando Torneio
12         └─ Binário com DF;
13         └─ Atualize a velocidade da partícula;
14         └─ Aplique o fator de construção;
15         └─ Atualize a posição da partícula;
16         └─ Avalie a partícula;
17         └─ Atualize o líder cognitivo, caso necessário;
18   └─ Atualize e ajuste o Arquivo Externo utilizando CD;
19   └─ Avalie o DF das soluções dentro do AE;

```

---

As imagens de treinamento foram obtidas através do banco de imagens sanguíneas da Sociedade Americana de Hematologia (em inglês, ASH) [24]. A Figura 2 apresenta exemplo de imagens de plaquetas utilizadas, enquanto a Figura 3 apresentam exemplos negativos. Foram utilizadas apenas imagens de plaquetas, porém, posteriormente, outros tipos de imagens serão testados para validar o algoritmo de geração de comitês de classificadores. Deste banco foram recortadas 147 imagens de plaquetas e 556 não-plaquetas. Estas imagens foram utilizadas sob o esquema de *8-folds*

selecionados aleatoriamente. Desses oito grupos, seis foram escolhidos aleatoriamente e utilizados para a etapa de treinamento do *AdaBoost*. Após a criação dos classificadores, os outros 2 grupos são selecionados para teste. Este processo de teste é repetido por trinta vezes. Foram realizados treinamentos com classificadores cuja quantidade de estágios varia entre 5 e 20 estágios. Para as constantes  $l_d$  e  $l_u$  foram atribuídos os valores de 0,9 e 0,6, respectivamente.

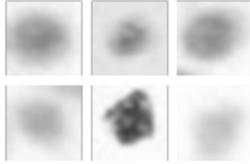


Figura 2. Exemplo de imagens de plaquetas.

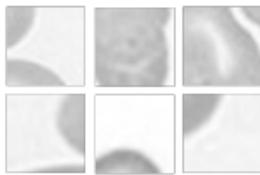


Figura 3. Exemplos de imagens que não representam plaquetas.

## V. RESULTADOS

Em teoria, classificadores com uma menor quantidade de estágios tendem a ser menos restritivos, apresentando uma menor taxa de falso negativo, porém com uma maior taxa de falso positivos. Por outro lado, classificadores com maiores quantidades de estágios tendem a apresentar o inverso: um comportamento mais restritivo com uma maior taxa de falso negativo e menor taxa de falso positivo. Foram geradas soluções com cascatas com número específico de classificadores fracos. Depois as soluções foram agrupadas e foi formado uma frente de Pareto com os melhores classificadores dentre todos os tamanhos de cascata testados. Foi avaliado o desempenho do comitê composto por todas estas soluções em que o processo de decisão foi feito através de votação majoritária. Pode-se observar na Figura 4 as soluções da frente de Pareto, bem como o resultado apresentado pelo comitê. Pode-se observar que o resultado do comitê domina as soluções da frente de Pareto, mas quando se considera o desvio padrão das 30 execuções ainda existe uma sobreposição forte entre o resultado do comitê e da frente de Pareto.

Após aplicar o processo de exclusão apresentado na Seção III, pode-se observar na Figura 5 uma redução considerável na quantidade de pontos na frente de Pareto, bem como a melhora no resultado gerado pelo comitê. A Tabela I apresenta a comparação entre os resultados obtidos pelos comitês formados pelas frentes de Pareto antes e depois de ser submetido a um processo de seleção, apresentado nas Figuras 4 e 5. Nesta tabela encontram-se as taxas de falso negativo (FN) e seu desvio padrão ( $\sigma$  FN), bem como a taxa de falso positivo (FP) e seu desvio padrão ( $\sigma$  FP). Foram realizados testes não-paramétricos de Wilcoxon para cada um dos objetivos (taxa de falso positivo e falso negativo)

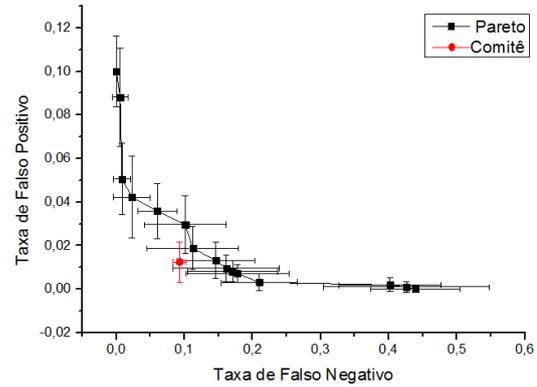


Figura 4. Pareto resultante utilizando diferentes cascatas após o treinamento realizado pelo *DFMOPSO-AdaBoost*.

e através deles foi constatado que o resultado do comitê obtido sem o processo de seleção alcançou valores maiores, ou seja, um resultado pior para ambos os objetivos para uma significância de 97%. A redução obtida foi de 28% na taxa de falso negativo e 48% na taxa de falso positivo. As barras ao redor de cada classificador mostram seu desvio padrão  $\sigma$  para cada eixo e estão apresentadas nas frentes de Pareto mostradas nas Figuras 4 e 5.

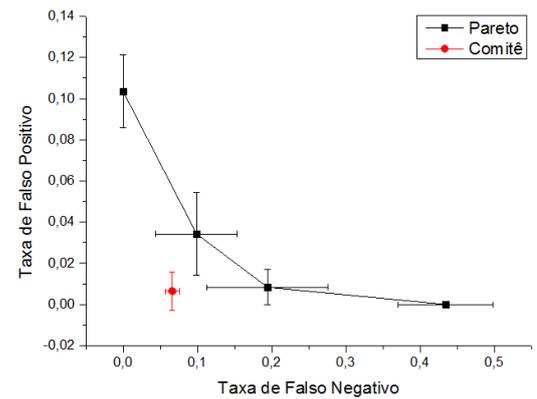


Figura 5. Pareto resultante aplicando o algoritmo de seleção após o treinamento realizado pelo *DFMOPSO-AdaBoost*.

Tabela I  
COMPARAÇÃO ENTRE AS MÉDIAS E A QUANTIDADE DE CLASSIFICADORES UTILIZADOS ENTRE AS FRENTES DE PARETO FORMADOS POR CLASSIFICADORES SELECIONADOS E NÃO SELECIONADOS.

Processo de Seleção?	FN	$\sigma$ FN	FP	$\sigma$ FP
Sim	0,06637	0,00932	0,00645	0,00932
Não	0,09305	0,00922	0,01242	0,00922

## VI. CONCLUSÕES

Este artigo apresentou uma metodologia para projeto de um comitê de classificadores baseado em otimização multi-objetiva e mecanismos de exclusão baseado em *spacing*. Foi testado com um algoritmo baseado em inteligência de enxames (*DFMOPSO*) e classificadores do tipo *AdaBoost* para detecção de plaquetas em imagens sanguíneas.

Com a utilização de um comitê composto por soluções de classificadores gerados pelo algoritmo de otimização multi-objetivo, acaba-se por obter uma solução que domina a

frente de Pareto compostas pelas soluções dos classificadores individualmente. Este comitê ainda pôde ser melhorado utilizando um processo de exclusão de classificadores em regiões específicas da frente de Pareto, evitando assim um forte viés. Ao se efetuar este processo de exclusão, tem-se um menor número de classificadores no comitê, diminuindo o tempo computacional necessário para a classificação e melhorando o desempenho do comitê.

Outro aspecto importante obtido nos resultados é que o desvio padrão dos resultados considerando o método de exclusão diminui quando comparado ao caso onde não é utilizado o método de exclusão. Isto implica em um aumento de robustez após a aplicação do método de exclusão.

Como trabalhos futuros podem ser listados a comparação utilizando outras bases, aplicação de outros algoritmos de otimização multi-objetivos e o desenvolvimento de outros mecanismos de exclusão.

#### REFERÊNCIAS

- [1] C. Wang and D. J. Hill, "Deterministic learning and rapid dynamical pattern recognition," *IFAC Proceedings Volumes (IFAC PapersOnline)*, vol. 16, no. 3, pp. 180–185, 2005.
- [2] W. S.-j. C. Xu, "A new Model of Pattern Recognition Theory and Its Application," *Proceedings of the International Joint Conference on Neural Networks*, vol. 3, pp. 2258–2262, 2003.
- [3] S. N. Ghazavi and T. W. Liao, "Medical data mining by fuzzy modeling with selected features," *Artificial Intelligence in Medicine*, vol. 43, pp. 195–206, 2008.
- [4] S. M. Vieira, L. F. Mendonça, G. J. Farinha, and J. a. M. C. Sousa, "Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients," *Applied Soft Computing Journal*, vol. 13, pp. 3494–3504, 2013.
- [5] A. M. Ahmed and A. Alkilany, "Pattern Discovery Using K-Means Algorithm," in *World congress on Computer Applications and Information Systems*, (Hammamet), pp. 1–4, IEEE, 2014.
- [6] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [7] R. E. Schapire, Y. Freund, P. Bartlett, and W. S. Lee, "Boosting the margin: A new explanation for the effectiveness of voting methods," *Annals of Statistics*, vol. 26, no. April, pp. 1651–1686, 1998.
- [8] A. Almonayyes, "Categorizing Fanatic Texts by Integrating Explanation Patterns with Na&#239;ve Bayes Classifier," in *2005 International Conference on Neural Networks and Brain*, vol. 2, (Beijing), pp. 1279–1283, IEEE, 2005.
- [9] J. Kuruvilla and K. Gunavathi, "Lung cancer classification using neural networks for CT images," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 202–209, 2014.
- [10] J. Qing, L. Ting, Y. Yan, and C. Jinhui, "Study of Rolling Bearing SVM Pattern Recognition Based on Correlation Dimension of IMF," in *International Conference on Intelligent System Design and Engineering Application (ISDEA)*, no. 4, (Sanya, Hainan), pp. 1132–1135, IEEE, 2012.
- [11] H. Chen, P. Tino, and X. Yao, "Predictive ensemble pruning by expectation propagation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 7, pp. 999–1013, 2009.
- [12] T. P. F. D. Lima and T. B. Ludermir, "Optimizing Dynamic Ensemble Selection Procedure by Evolutionary Extreme Learning Machines and a Noise Reduction Filter," in *International Conference on Tools with Artificial Intelligence (ICTAI)*, (Recife), pp. 546–552, 2013.
- [13] L. L. Minku, A. P. White, and X. Yao, "The impact of diversity on online ensemble learning in the presence of concept drift," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 5, pp. 730–742, 2010.
- [14] A. A. F. A. M. P. C. Neto, "Meta-Learning and Multi-Objective Optimization to Design Ensemble of Classifiers," in *Brazilian COnference on Intelligent Systems (BRACIS)*, (São Paulo), pp. 91–96, IEEE, 2014.
- [15] D. R. C. Silva and C. J. a. Bastos-filho, "A Speed-Constrained Multi-Objective Particle Swarm Optimizer Guided by a Diversity Metrics," in *Latin American Congress on Computational Intelligence*, (San Carlos de Bariloche - Argentina), pp. 1–6, 2014.
- [16] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Computational learning theory*, vol. 55, pp. 119–139, Aug. 1995.
- [17] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 1, no. C, pp. I-511–I-518, 2001.
- [18] A. C. A. Firmo, *Classificação Adaboost com Treinamento por Enxame de Partículas para Diagnóstico da Esquistossomose Mansônica no Litoral de Pernambuco*. PhD thesis, Universidade de Pernambuco, Recife, 2010.
- [19] D. N. O. Nascimento, *Classificação AdaBoost para detecção e contagem automática de plaquetas*. Trabalho de conclusão de curso, 2011.
- [20] D. N. O. Nascimento, *Classificação AdaBoost para Reconhecimento de Elementos Figurados do sangue*. Dissertação de mestrado, UPE, Recife, 2014.
- [21] C. J. A. Bastos-Filho, W. A. S. Silva, and L. R. M. Lira, "Comparing Meta-heuristics for AdaBoost Training Applied to Platelets Detection," *Latin America Transactions*, vol. 12, no. 5, pp. 942–950, 2014.
- [22] H. Chen and X. Y. X. Yao, "Evolutionary Multiobjective Ensemble Learning Based on Bayesian Feature Selection," *2006 IEEE International Conference on Evolutionary Computation*, pp. 267–274, 2006.
- [23] R. E. Banfield, L. O. Hall, K. W. Bowyer, and W. P. Kegelmeyer, "Ensemble diversity measures and their application to thinning," *Information Fusion*, vol. 6, no. 1, pp. 49–62, 2005.
- [24] ASH, "American Society of Hematology," 2012.