

DESEMPENHO DE MÁQUINAS DE APRENDIZADO EXTREMO COM OPERADORES MORFOLÓGICOS PARA IDENTIFICAÇÃO E CLASSIFICAÇÃO DE LESÕES EM IMAGENS FRONTAIS DE TERMOGRAFIA DE MAMA

MAÍRA A. DE SANTANA, WASHINGTON W. A. DA SILVA, AMANDA L. R. DA SILVA, JESSIANE M. S. PEREIRA, VALTER A. F. BARBOSA, CAMILA A. DINIZ, MARIA B. J. DE ALMEIDA, RITA C. F. DE LIMA, WELLINGTON P. DOS SANTOS

Laboratório de Computação Biomédica, Departamento de Engenharia Biomédica, Universidade Federal Pernambuco
Av. da Arquitetura, Área 2, Departamento de Engenharia Biomédica, sala 17, Cidade Universitária, Recife, Pernambuco, Brasil – CEP: 50670-901
E-mails: mairaasantana@hotmail.com, wellington.santos@ufpe.br

Abstract– Breast cancer is the most common among women, corresponding 28% of new cases per year in Brazil. Breast thermography is recent method for detection and classification of breast lesions. This technique enables the diagnosis in initial stages, besides, thermography has a lower cost than other imaging techniques and do not exposes the patients to ionizing radiations. However, there is a lack of professionals capable to analysis thermography images. Thus, analyzing such images through pattern recognition techniques is a good option to improve the quality of the diagnosis. In this work, we investigate the use of classifiers based in artificial neural network perceptron multilayer (MLP), extreme learning machine (ELM), support vector machine (SVM), trees-based methods and, especially, ELM morphologic to classify frontal images of breast thermography represented by Haralick and Zernike moments as descriptor of texture and shape, respectively. In total, it was used 336 thermography images obtained by Federal University of Pernambuco. The extreme learning machine with morphologic operators obtained good performance in experimental configurations, with 95.19% of accuracy and 0.933 kappa index for the best result.

Keywords– Breast cancer, breast thermography, artificial neural network, extreme learning machine, morphological operators.

Resumo– O câncer de mama é o mais comum entre mulheres, correspondendo a 28% dos novos casos por ano no Brasil. A termografia de mama é um método recente para a detecção e classificação de lesões de mama. Tal técnica possibilita o diagnóstico da doença ainda em estágios iniciais, além de possuir um custo menor do que outras técnicas de imageamento e não expor as pacientes às radiações ionizantes. Contudo, ainda há uma carência de profissionais capacitados para analisar imagens termográficas. Dessa forma, analisar tais imagens através de técnicas de reconhecimento de padrões é uma boa opção para melhorar a qualidade do diagnóstico. Neste trabalho, é investigado o uso de classificadores baseados em redes neurais artificiais perceptron multicamadas (MLP), máquinas de aprendizado extremo (ELM), máquinas de vetor de suporte (SVM), métodos baseados em árvores e, especialmente, o ELM morfológico para classificar imagens frontais de termografia de mama representadas pelos momentos de Haralick e Zernike como descritores de textura e forma, respectivamente. Ao total, foram utilizadas 336 imagens termográficas obtidas pela Universidade Federal de Pernambuco. O ELM com operadores morfológicos obteve um bom desempenho nas configurações experimentadas, apresentando uma acurácia de 95,19% e índice kappa de 0,933 em seu melhor resultado.

Palavras-chave– Câncer de mama, termografia de mama, redes neurais artificiais, máquinas de aprendizado extremo, operadores morfológicos.

1 Introdução

O câncer de mama é o mais comum entre as mulheres em todo o mundo. Só no Brasil ele corresponde a cerca de 28% dos novos casos de câncer por ano, apesar de, no geral, haver um bom prognóstico, esta doença é ainda responsável pelo maior índice de mortalidade por câncer na população feminina (BRASIL, 2015a). De acordo com o Ministério da Saúde (MS), a detecção precoce de tumores, que consiste na identificação do câncer em estágios iniciais, é essencial no sentido de reduzir a mortalidade pela doença (BRASIL, 2015b). Atualmente, os principais métodos de diagnóstico utilizados para identificar o câncer de mama são a mamografia, exame clínico, ultrassonografia e ressonância magnética. Apesar de serem técnicas já consolidadas, elas não vêm sendo suficientes para identificar lesões de mama em seus estágios iniciais principalmente pelo fato de que mulheres mais jovens

geralmente possuem mamas predominantemente densas, tecido esse que dificulta a visualização e diferenciação de lesões mamárias (Etehadtavakol e NG, 2013).

A termografia vem sendo explorada como ferramenta auxiliar para diagnóstico de cânceres mamários. Esta técnica se baseia na aquisição de imagens, registradas através de uma câmera de infravermelho, que apresentam a distribuição de temperatura na região. O funcionamento geral da câmera consiste na captura da radiação infravermelha emitida pela superfície de interesse, sem a necessidade de procedimentos invasivos e sem exposição do paciente a radiações ionizantes. A técnica permite a investigação de alterações fisiológicas causadas por doenças a partir da análise do mapeamento das temperaturas na região. Tais alterações tendem a surgir antes de alterações anatômicas, favorecendo, portanto, o diagnóstico em estágios iniciais da doença. No caso da existência de lesões na região de interesse, há um aumento na

atividade metabólica das células locais, muitas vezes acarretando em efeitos como vasodilatação, angiogênese, e, conseqüentemente, no aumento do fluxo sanguíneo na região. Em relação a identificação de lesões mais profundas na mama, a pequena penetrabilidade da técnica não é considerada como uma limitação, pois a aceleração das atividades metabólicas aumentam a temperatura superficial da mama (Etehadtavakol e NG, 2013). Uma possível limitação da técnica é o fato de que as imagens são facilmente influenciadas por fatores externos, variações no ambiente em que a aquisição é realizada, dessa maneira, aspectos como entrada e saída de pessoas, temperatura e umidade da sala devem ser severamente controlados para garantir a validade do exame.

Estudos afirmam que a associação da termografia com a mamografia tendem a aumentar tanto a sensibilidade quanto a especificidade na detecção de lesões mamárias (Bezerra, 2017) e (Kennedy, Lee e Seely, 2009).

Além das limitações intrínsecas aos vários métodos utilizados para identificação de cânceres mamários, a variabilidade dos casos clínicos consiste em um desafio que torna árdua a tarefa de identificação e diferenciação de lesões de mama por parte dos especialistas, independentemente da técnica utilizada. Desta maneira, tem-se investido cada vez mais no desenvolvimento e na otimização de sistemas inteligentes que sejam capazes de realizar a identificação e a classificação automática de tais lesões e, assim, possa auxiliar o trabalho de especialistas no que concerne ao diagnóstico do câncer de mama. Nesse sentido, a maximização de aspectos como acurácia e sensibilidade e minimização do tempo de treino e do custo computacional desses sistemas são fatores almejados, pois podem proporcionar o diagnóstico precoce dessa doença e, portanto, melhorar seu prognóstico, minimizando a taxa de mortalidade associada a ela.

O presente trabalho visa avaliar quantitativa e qualitativamente o desempenho das máquinas de aprendizado extremo do tipo morfológico na identificação da existência de lesão e na classificação do tipo de lesão (quando existente) em imagens termográficas. Ao longo do estudo foram obtidas combinações de parâmetros que possibilitaram até 95% de acerto na classificação.

Este trabalho está organizado como segue: inicialmente são apresentados os materiais e métodos utilizados no decorrer do estudo; em seguida os resultados obtidos através da aplicação do método proposto são exibidos e discutidos; por fim, são apresentadas as conclusões alcançadas.

2 Materiais e Métodos

Atualmente, diversos grupos vêm se dedicando ao estudo da termografia de mama. Aguiar Junior *et al.* (2013) realizaram testes utilizando perceptron multicamadas como classificador para detecção da

existência ou não de lesões em imagens termográficas de mama, obtendo classificações acuradas em cerca de 75%. Resmini *et al.* (2012), por sua vez, obteve resultados próximos a 90% de acerto utilizando outros classificadores (SVM, KNN e Naive Bayes) para também detectar a existência de lesão mamária. No trabalho de Belfort *et al.* (2014) foi introduzido outro método de detecção a partir do qual obtiveram uma acurácia em torno de 66% utilizando SVM como classificador. Em Santana *et al.* (2018) foi obtida uma acurácia em torno de 73% na classificação de lesões de mama em imagens termográficas frontais e laterais utilizando máquinas de aprendizado extremo convencionais.

O diagrama da Figura 1 apresenta uma breve descrição do método utilizado neste trabalho.

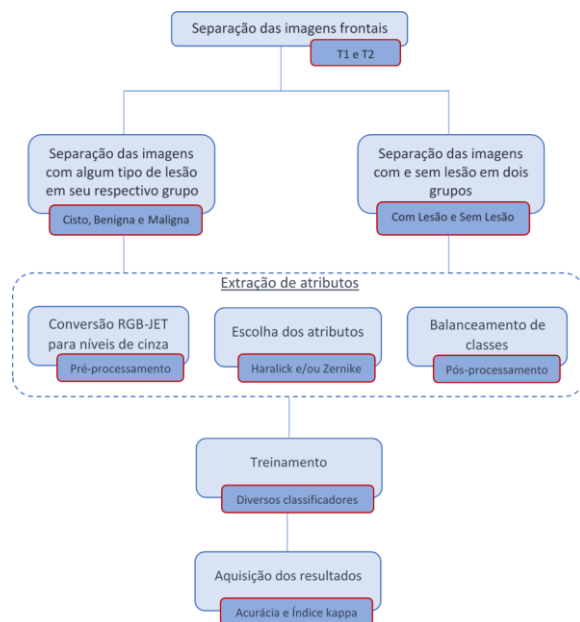


Figura 1. Esquema simplificado do método proposto

A aquisição das imagens termográficas utilizadas no presente estudo foi realizada no Hospital das Clínicas de Pernambuco, Recife, pelo grupo de pesquisa em termografia de mama do Departamento de Engenharia Mecânica da Universidade Federal de Pernambuco. Tais imagens foram adquiridas no formato JPG utilizando uma câmera de infravermelho FLIR S45 e com o auxílio do aparato mecânico apresentado na Figura 2, o qual foi proposto e desenvolvido por Oliveira (2012b), com o intuito de padronizar fatores como ângulo e distância entre câmera e paciente, evitando variações significativas no posicionamento das pacientes.

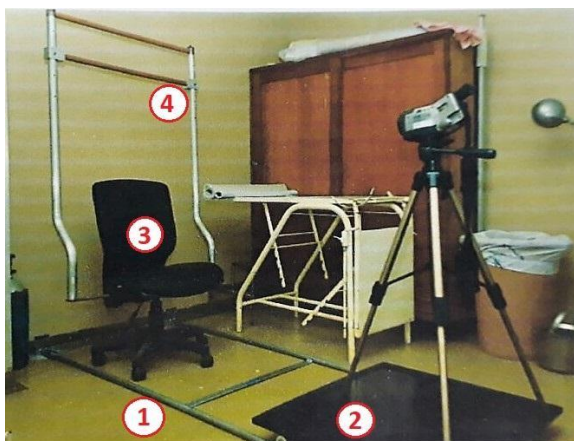


Figura 2. Aparato mecânico para aquisição das imagens termográficas. Em (1) encontram-se os trilhos utilizados para mover o carro de suporte da câmera; (2) placa móvel de suporte do tripé da câmera; (3) cadeira giratória onde a paciente é colocada; (4) suporte para os braços, que consiste em uma barra horizontal para posicionar os braços da paciente para cima durante a aquisição, essa barra se move verticalmente para ajuste de acordo com a altura da paciente.

Além disso, devido ao efeito de troca de calor entre paciente e ambiente, com o intuito de minimizar influências externas nas imagens proporcionando o equilíbrio térmico, as pacientes foram submetidas a um período de aclimatização de 10 minutos, no qual deveriam ficar devidamente posicionadas e sem contato com as mamas (Oliveira, 2012b).

Para cada paciente foram adquiridas imagens da região da mama em oito posições distintas: duas imagens frontais de ambas as mamas (T1 e T2) e três imagens de cada mama isoladamente, direita e esquerda, em ângulos distintos, sendo eles frontal (MD e ME), lateral externa (LEMD e LEME) e lateral interna (LIMD e LIME) (Oliveira, 2012b). Exemplos dessas posições podem ser visualizados na Figura 3.

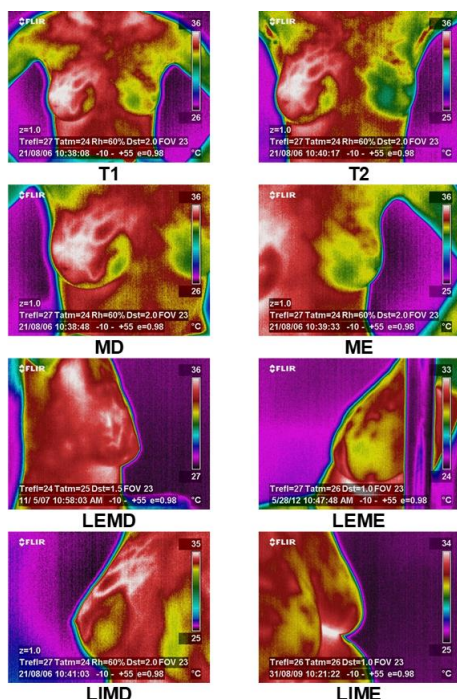


Figura 3: Exemplo das posições de aquisição das imagens por paciente: T1 e T2 são imagens frontais de ambas as mamas com as

mãos na cintura e mãos pra cima, respectivamente; MD e ME correspondem a aquisição frontal do centro para direita e do centro para esquerda, respectivamente; LEMD e LEME são imagens da lateral externa das mamas direita e esquerda, nessa ordem; LIMD e LIME correspondem às visões das laterais interna das mamas

O banco de imagens foi também organizado de acordo com os tipos de diagnósticos, os quais foram previamente estabelecidos baseados em métodos já consolidados, tais como exames clínicos, biópsias, mamografias e ultrassonografias de mama (Dourado Neto, 2014). Os possíveis diagnósticos são: cisto, lesão benigna, lesão maligna e sem lesão.

No estudo em questão as imagens foram divididas de duas maneiras distintas. Em um primeiro momento foram utilizadas apenas as imagens de casos nos quais foram identificados cistos, lesões benignas e lesões malignas, com o intuito de verificar a eficácia do sistema no sentido de classificar as imagens em seus respectivos grupos de lesões. Em seguida, as imagens dos grupos dos diferentes tipos de lesões foram alocadas para um único grupo, denominado “com lesão”; foram então realizados testes utilizando apenas os grupos “com lesão” e “sem lesão”, nesse caso, para observar a capacidade do sistema em detectar a existência de algum tipo de lesão de mama, ou seja, saber diferenciar um tecido normal de um lesionado. Além disso, os experimentos foram realizados apenas com as imagens frontais T1 e T2, pois considera-se que essas condições favorecem a identificação da região de interesse. Exemplos de imagens das 4 diferentes classes podem ser visualizados na Figura 4.

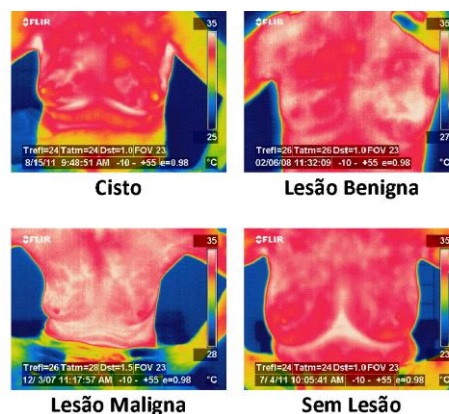


Figura 4: Exemplos de imagens de cada grupo de diagnóstico.

Ao todo, 336 imagens foram usadas nos testes, sendo: 73 de cistos, 121 de lesões benignas e 76 de lesões malignas (totalizando 270 imagens com lesão) e 66 sem lesão. O banco de dados foi composto por pacientes com suspeitas de lesões mamárias e idade acima dos 35 anos, idade a partir da qual é permitida a realização de mamografias no Brasil.

Inicialmente, as imagens foram submetidas a uma etapa de pré-processamento de conversão de RGB-JET para níveis de cinza, com os tons mais claros indicando temperaturas mais altas.

O método propõe ainda a utilização de momentos de Haralick e momentos de Zernike como extratores dos atributos para o treinamento do conjunto de dados. Os momentos de Haralick fornecem informações

relacionadas à textura, a partir da avaliação da probabilidade de ocorrência das combinações entre os níveis de cinza da imagem (Oliveira, 2012a). Já os de Zernike foram utilizados no reconhecimento de padrões relacionados à forma (Felipe *et al.*, 2015).

Como exposto anteriormente, a quantidade de imagens no banco de dados utilizado é diferente para cada classe, fato esse que pode acarretar em um resultado tendencioso durante o treinamento, as lesões poderiam ser mais comumente classificadas como da classe que possui mais representantes; para evitar este problema, foi realizado o balanceamento das classes a partir da inserção de instâncias sintéticas através da combinação linear de vetores de atributos da mesma classe (Lima, Silva-Filho e Santos, 2016).

Após a configuração dos parâmetros supracitados, os atributos foram gerados de três maneiras distintas: utilizando apenas momentos de Haralick, apenas momentos de Zernike e ambos ao mesmo tempo. Tal procedimento foi realizado tanto para a etapa de classificação dos diferentes tipos de lesão (3 classes) como na etapa de detecção da existência de lesão (2 classes).

Logo em seguida, os conjuntos de dados foram submetidos a etapa de treinamento utilizando três configurações distintas do classificador Máquina de Aprendizagem Extremo (ELM, do inglês *Extreme Learning Machine*), foram elas: ELM clássico, ELM morfológico com *kernel* dilatação e ELM morfológico com *kernel* erosão. Apesar de ser uma ferramenta ainda não tão difundida quanto outras redes neurais, tais como SVM ou MLP, estudos apontam algumas vantagens no uso da ELM por ser uma rede versátil, permitir classificações multiclasse com uma menor taxa de erro e menor complexidade computacional para realizar o treinamento (Huang *et al.*, 2012).

O ELM morfológico (mELM), proposto por Azevedo *et al.* (2017), consiste na aplicação de *kernels* baseados nos operadores morfológicos não lineares de dilatação e erosão na camada escondida da rede. Tais operadores são básicos na morfologia matemática, por isso, são bastante utilizados no processamento de imagens digitais e tendem a gerar modificações de forma, mas preserva o conteúdo presente na imagem. O treinamento com mELM foi realizado através do ambiente GNU Octave, versão 4.0.3 (Eaton *et al.* 2018).

Por fim, foram realizados treinamentos com o uso de alguns classificadores mais comumente utilizados em trabalhos previamente estudados na literatura, foram eles: *Bayes Net*, *Naive Bayes*, *Support Vector Machine* (SVM/SMO), *Multilayer Perceptron* (MLP), *J48*, *Random Forest* e *Random Tree*. A ferramenta utilizada para a aplicação desses diferentes algoritmos na base de dados foi o *software* WEKA (*Waikato Environment for Knowledge Analysis*), versão 3.8.1 (Ferreira e Wayne, 2012).

A Tabela 1 mostra os parâmetros modificados para cada classificador.

Tabela 1: Parâmetros escolhidos na configuração dos classificadores.

Classificador	Parâmetros
BayesNet	-
NaiveBayes	-
J48	-
SVM	<i>kernel</i> linear
MLP	Camadas escondidas: a * Taxa de aprendizagem: 0.3 <i>Momentum</i> : 0.2 Iterações: 500
Random Forest	Árvores: 100
Random Tree	-
ELM	Neurônios na camada escondida: 400 <i>kernel</i> : sigmoide
mELM	Neurônios na camada escondida: 400 <i>kernel</i> : dilatação e erosão

* 'a' = (attribs + classes) / 2 = 85 hidden layers

Para todos os classificadores foram realizados treinamentos utilizando dois formatos de teste: validação cruzada com 10 *folds* (método *k-fold*) e divisão percentual com 75% da base utilizada para treino e 25% para teste.

Após a aquisição dos resultados, foi realizada a análise dos mesmos levando em consideração alguns fatores extraídos: acurácia e índice kappa. A acurácia consiste na taxa de acerto da rede, ou seja, o quanto ela classificou os casos corretamente. O índice kappa, por sua vez, fornece informações sobre o nível de concordância entre os resultados obtidos, possui valor máximo 1, indicando total concordância, e valores próximos ou abaixo de 0 (zero) indicando nenhuma concordância.

3 Resultados

Durante os testes realizados foram extraídos os fatores de interesse citados anteriormente: taxa de acerto e índice kappa. Tais fatores são essenciais para a análise da qualidade da classificação realizada pelas redes em todos os casos testados, permitindo assim, a determinação de qual entre os métodos possui melhor desempenho frente ao problema proposto.

A Tabela 2 apresenta os resultados obtidos para a classificação em três classes, ou seja, classificação dos tipos de lesões, utilizando diferentes formas de extração de atributos. Nessa tabela são apresentados os resultados utilizando a validação cruzada (10 *folds*) como formato de teste.

Tabela 2: Resultados de acurácia e índice kappa obtidos para classificação dos tipos de lesões utilizando validação cruzada (10 *folds*) para realizar os testes.

HARALICK		
Classificador	Acurácia	Índice kappa
BayesNet	57,44%	0,3616
NaiveBayes	53,31%	0,2996
MLP	67,36%	0,5103
SMO / SVM	61,29%	0,4194
J48	64,33%	0,4649
RandomForest	70,25%	0,5537
RandomTree	57,02%	0,3554
ELM sigmoide	70,10%	0,6636
mELM dilatação	83,73%	0,815
mELM erosão	84,02%	0,8191
ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	58,95%	0,3843
NaiveBayes	59,37%	0,3905
MLP	82,92%	0,7438
SMO / SVM	78,79%	0,6818
J48	58,26%	0,374
RandomForest	77,27%	0,6591
RandomTree	54,96%	0,3244
ELM sigmoide	77,55%	0,749
mELM dilatação	78,37%	0,7561
mELM erosão	79,47%	0,7677
HARALICK E ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	58,26%	0,374
NaiveBayes	55,37%	0,3306
MLP	88,43%	0,8264
SMO / SVM	82,78%	0,7417
J48	61,71%	0,4256
RandomForest	72,73%	0,5909
RandomTree	58,40%	0,376
ELM sigmoide	79,90%	0,7739
mELM dilatação	91,60%	0,905
mELM erosão	89,93%	0,8871

Para esse mesmo conjunto de dados, com três classes possíveis, também foram obtidos resultados utilizando divisão percentual (75%) para realizar o treinamento da rede. A Tabela 3 apresenta os resultados referentes a essa configuração.

Tabela 3: Resultados de acurácia e índice kappa obtidos para classificação dos tipos de lesões utilizando divisão percentual (75%) para realizar os testes.

HARALICK		
Classificador	Acurácia	Índice kappa
BayesNet	53,04%	0,3047
NaiveBayes	50,28%	0,2693
MLP	64,09%	0,4549
SMO / SVM	60,22%	0,4125
J48	53,59%	0,3043
RandomForest	61,88%	0,4307
RandomTree	54,14%	0,3131
ELM sigmoide	63,74%	0,6007

mELM dilatação	82,43%	0,8018
mELM erosão	82,97%	0,8069
ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	66,85%	0,5041
NaiveBayes	62,98%	0,4471
MLP	78,45%	0,6769
SMO / SVM	79,01%	0,6864
J48	58,56%	0,3748
RandomForest	76,80%	0,6522
RandomTree	52,49%	0,2846
ELM sigmoide	71,98%	0,681
mELM dilatação	75,82%	0,7256
mELM erosão	80,77%	0,7868
HARALICK E ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	59,12%	0,3956
NaiveBayes	54,70%	0,3352
MLP	90,06%	0,8507
SMO / SVM	81,22%	0,7192
J48	62,43%	0,4338
RandomForest	71,27%	0,5704
RandomTree	58,01%	0,3706
ELM sigmoide	75,82%	0,732
mELM dilatação	90,66%	0,8907
mELM erosão	91,76%	0,9049

Como dito anteriormente, foram também realizados experimentos com o banco de dados sendo dividido em duas classes, para verificar a capacidade do sistema de detectar a existência de algum tipo de lesão e, portanto, dividir o conjunto de dados nas classes “com lesão” e “sem lesão”.

A Tabela 4, a seguir, apresenta os resultados de acurácia e kappa obtidos quando utilizada validação cruzada como formato de teste.

Tabela 4: Resultados de acurácia e índice kappa obtidos para etapa de detecção da existência de lesões utilizando validação cruzada (10 *folds*) para realizar os testes.

HARALICK		
Classificador	Acurácia	Índice kappa
BayesNet	77,69%	0,5537
NaiveBayes	72,22%	0,4444
MLP	87,69%	0,7537
SMO / SVM	77,96%	0,5593
J48	80,37%	0,6074
RandomForest	85,83%	0,7167
RandomTree	79,81%	0,5963
ELM sigmoide	87,59%	0,8243
mELM dilatação	92,22%	0,8899
mELM erosão	91,76%	0,8856
ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	77,41%	0,5481
NaiveBayes	76,94%	0,5389
MLP	92,96%	0,8593
SMO / SVM	84,26%	0,6852

J48	76,76%	0,5352
RandomForest	87,59%	0,7519
RandomTree	73,33%	0,4667
ELM sigmoide	87,78%	0,8277
mELM dilatação	84,26%	0,7814
mELM erosão	84,72%	0,7873
HARALICK E ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	76,02%	0,5204
NaiveBayes	71,67%	0,4333
MLP	95,09%	0,919
SMO / SVM	88,33%	0,7667
J48	80,65%	0,613
RandomForest	89,26%	0,7852
RandomTree	74,44%	0,4889
ELM sigmoide	88,33%	0,8322
mELM dilatação	94,44%	0,9233
mELM erosão	94,35%	0,922

Os resultados utilizando divisão percentual para realizar os testes de classificação em duas classes (detecção da existência de lesão) se encontram na Tabela 5.

Tabela 5: Resultados de acurácia e índice kappa obtidos para etapa de detecção da existência de lesões utilizando divisão percentual (75%) para realizar os testes.

HARALICK		
Classificador	Acurácia	Índice kappa
BayesNet	79,63%	0,5894
NaiveBayes	76,67%	0,5287
MLP	86,67%	0,7327
SMO / SVM	83,33%	0,6643
J48	76,67%	0,5333
RandomForest	86,30%	0,7255
RandomTree	76,30%	0,5255
ELM sigmoide	84,44%	0,7803
mELM dilatação	95,19%	0,9289
mELM erosão	89,26%	0,8618
ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	76,67%	0,531
NaiveBayes	74,07%	0,4782
MLP	88,15%	0,7617
SMO / SVM	83,33%	0,6633
J48	73,70%	0,4744
RandomForest	82,96%	0,6573
RandomTree	70,37%	0,4105
ELM sigmoide	87,04%	0,8226
mELM dilatação	83,70%	0,7739
mELM erosão	84,44%	0,7793
HARALICK E ZERNIKE		
Classificador	Acurácia	Índice kappa
BayesNet	77,78	0,5517
NaiveBayes	74,44%	0,4835
MLP	93,70%	0,8738
SMO / SVM	89,26%	0,7835
J48	77,04%	0,5421
RandomForest	89,26%	0,7848

RandomTree	78,15%	0,5622
ELM sigmoide	85,56	0,7847
mELM dilatação	95,19%	0,933
mELM erosão	92,59%	0,894

4. Discussão

A partir da análise dos resultados, é possível observar que a qualidade da classificação tende a melhorar quando os extratores de Haralick e Zernike são associados.

Quanto a comparação entre os classificadores, o destaque do método ELM morfológico é notável na grande maioria das configurações experimentadas. Tanto o *kernel* de dilatação quanto o de erosão se mostraram bastante eficientes para lidar com os problemas propostos na base de dados utilizada nos testes.

No geral, os melhores resultados durante a etapa de classificação dos tipos de lesões foram 91,60% de acurácia e kappa de 0,905 utilizando validação cruzada como formato de teste e acurácia de 91,76% e 0,9049 de índice kappa utilizando divisão percentual. Tais resultados foram obtidos utilizando o ELM morfológico, com *kernel* dilatação no primeiro caso e erosão no segundo.

Na etapa de detecção da existência ou não de lesões os resultados foram ainda mais satisfatórios. Foi obtida uma taxa de acerto de 95,09% com índice kappa de 0,919 para testes com validação cruzada utilizando a rede MLP, enquanto que realizando experimentos com divisão percentual obteve-se acurácia de 95,19% e kappa de 0,933 com ELM morfológico com *kernel* de dilatação. Embora a rede MLP tenha tido um desempenho melhor do que a morfológico no primeiro caso, a diferença entre a acurácia desses classificadores foi menor que 1%, ou seja, uma diferença pouco relevante considerando o problema aqui abordado.

Analisando a matriz de confusão dos melhores resultados, pode-se afirmar que a sensibilidade do método proposto foi de aproximadamente 92% para classificação dos tipos de lesões e 99% para detecção da existência de lesão. Já a especificidade ficou em torno de 96% para o primeiro grupo e 92% para o segundo.

5 Conclusão

A melhora no desempenho das classificações quando utilizada a combinação dos extratores de Haralick e Zernike demonstra que ambos os aspectos de textura e de geometria são importantes tanto na detecção da existência de algum tipo de lesão quanto na diferenciação entre os tipos de lesões, quando se trata de imagens termográficas.

Ainda que o ELM morfológico não tenha sido o melhor classificador em todas as configurações testadas, seu desempenho ficou próximo do melhor

obtido em todas as situações. Além disso, as Máquinas de Aprendizado Extremo se destacam, no geral, por suas vantagens em relação ao tempo de treino, custo computacional e desvio padrão, fatores que tendem a ser menores, quando comparados aos demais classificadores testados (Azevedo *et al.*, 2017a) e (Azevedo *et al.*, 2017b).

No geral, utilizando ELM morfológico, a eficiência do método na classificação em diferentes tipos de lesões ficou em torno de 94%. Na etapa de detecção da existência de lesão a eficiência do método ficou próxima de 96%. Em ambos os casos a eficiência ficou próxima do valor máximo (100%), indicando que a utilização de ELM morfológico como método de classificação de massas em termografias de mama se mostrou bastante eficiente. A utilização deste classificador é ainda indicada pois minimiza o tempo de treinamento e, consequentemente, o custo computacional envolvido.

Agradecimentos

À professora Rita de Cássia do Departamento de Engenharia Mecânica da UFPE pela disponibilização das imagens térmicas utilizadas para confecção da base de dados utilizada no presente trabalho.

Ao CNPq e à FACEPE pelo suporte financeiro parcial.

Referências Bibliográficas

- Aguiar Junior, P. S., Belfort, C. N. S., Silva, A. C., Diniz, P. H. B., Lima, R. C. F., Conci, A., Paiva, A. C. (2013). Detecção de Regiões Suspeitas de Lesão na Mama em Imagens Térmicas Utilizando *Spatioqram* e Redes Neurais. São Luís.
- Azevedo, W. W., Lima, S. M. L., Fernandes, I. M. M., Rocha, A. D. D., Cordeiro, F. R., Silva-Filho, A. G., Santos, W. P. (2017a). *Fuzzy Morphological Extreme Learning Machines to Detect and Classify Masses in Mammograms*. Recife, Pernambuco.
- Azevedo, W. W., Lima, S. M. L., Fernandes, I. M. M., Rocha, A. D. D., Cordeiro, F. R., Silva-Filho, A. G., Santos, W. P. (2017b). *Morphological Extreme Learning Machines Applied to Detect and Classify Masses in Mammograms*. Recife, Pernambuco.
- Belfort, C. N. S., Motta, S. A. C. S., Silva, A. C. (2015). Detecção de lesões decorrentes do câncer de mama em imagens termográficas utilizando funções geostatísticas e SVM.
- Bezerra, L. A., (2007). Uso de imagens termográficas em tumores mamários para validação de simulação computacional [dissertação]. Recife: Universidade Federal de Pernambuco; 2007.
- BRASIL, (2015a). Ministério da Saúde, Instituto Nacional de Câncer. Diretrizes para a Detecção Precoce do Câncer de Mama no Brasil. Rio de Janeiro.
- BRASIL, (2015b). Ministério da Saúde, Instituto Nacional de Câncer. Estimativa 2016: Incidência de Câncer no Brasil. Rio de Janeiro.
- Dourado Neto, H. M. (2014). Segmentação e análise automática de termogramas: um método auxiliar na detecção do câncer de mama [dissertação]. Recife: Universidade Federal de Pernambuco.
- Eaton, J. W., Bateman, D., Hauberg, S., Wehbring, R. (2018). *GNU Octave: A high-level interactive language for numerical computations*. Ed. 4. Maryland, USA.
- Etehadtavakol M. e NG E. Y. K., (2013). *Breast Thermography as a Potential Non-contact Method in the Early Detection of Cancer: a review*. Em: Journal of Mechanics in Medicine and Biology. Vol. 13, No. 2. doi: 10.1142/S0219519413300019.
- Felipe, J. C., Olioti, J. B., Traina, A. J. M. (2005). Discriminação de Aspectos Malignos em Massas Tumerais de Mamografias Usando Características de Forma das Imagens. São Paulo.
- Ferreira, T., Wayne, R. (2012). *ImageJ User Guide*. Maryland, USA.
- Huang, G., Zhou, H., Ding, X., Zhang, R. (2012). *Extreme Learning Machine for Regression and Multiclass Classification*. Em: IEEE Transactions on Systems, man, and cybernetics – part B, Vol. 42.
- Kennedy, D. A., Lee, T., e Seely, D. (2009). *A Comparative Review of Thermography as a Breast Cancer Screening Technique*. Em: Integrative Cancer Therapies / Vol. 8, No. 1.
- Lima, S. M., Silva-Filho, A. G., e Santos, W. P. (2016). *Detection and classification of masses in mammographic images in a multiKernel approach*. Em: Computer Methods and Programs in Biomedicine.
- Oliveira L. F., Narloch, A. L. M., Kist, D. M., Soares Filho, M. P., Maneghello, G. E., Cavalheiro, G. G. H., Tillmann, M. A. A. (2012a). Extração de Características de Forma utilizando matriz de co-ocorrência e Atributos de Haralick. Curitiba, Paraná.
- Oliveira, M. M. (2012b). Desenvolvimento de protocolo e construção de um aparato mecânico para padronização da aquisição de imagens termográficas de mama [dissertação]. Recife: Universidade Federal de Pernambuco.
- Resmini, R., Conci, A., Borchardt, T. B., Lima, R. C. F., Montenegro, A. A., Pantaleão, C. A. (2012). Diagnóstico Precoce de Doenças Mamárias Usando Imagens Térmicas e Aprendizado de Máquina. Rio de Janeiro.
- Santana, M. A., Pereira, J. M. S., Silva, F. L., Lima, N. M., Sousa F. N., Arruda, G. M. S., Lima, R. C. F., Silva, W. W. A., Santos, W. P. (2018). *Breast cancer diagnosis based on mammary thermography and extreme learning machines*. Em: Research on Biomedical Engineering. doi: <http://dx.doi.org/10.1590/2446-4740.05217>.