



UNIVERSIDADE FEDERAL DO CEARÁ – CAMPUS SOBRAL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA E DE COMPUTAÇÃO

BORGES, C.D.B

ANÁLISE DOS MODELOS MLP E RBF NA CLASSIFICAÇÃO DE
FORMAS COM DESCRITORES DE CONTORNO BASEADOS EM
REDES NEURAIS RANDOMIZADAS

SOBRAL

2018

LISTA DE ILUSTRAÇÕES

Figura 1 – Região, contorno e esqueleto de uma forma	4
Figura 2 – Exemplos de classes da base de formas <i>Leaf</i>	6
Figura 3 – Curvas de acurácia <i>K-fold</i> para MLPs	11
Figura 4 – Ampliação das curvas de acurácia <i>K-fold</i> para MLPs	12
Figura 5 – Curvas de acurácia <i>K-fold</i> para RBFs	13
Figura 6 – Ampliação das curvas de acurácia <i>K-fold</i> para RBFs	13
Figura 7 – Curvas de acurácia <i>K-fold</i> para as melhores MLPs e RBFs	14
Figura 8 – Ampliação das curvas de acurácia <i>K-fold</i> para as melhores MLPs e RBFs	14
Figura 9 – Exemplos de contornos das classes 25 (folha de lima) e 26 (folha de limão)	17
Figura 10 – Exemplos de contornos das classes 25 (folha de lima) e 27 (<i>P. maliformis</i>)	18

LISTA DE TABELAS

Tabela 1 – Parâmetros fixos do treinamento da MLP no MATLAB	9
Tabela 2 – Taxas médias de acurácia, <i>recall</i> e precisão da MLP sobre cada <i>fold</i> . .	16
Tabela 3 – Taxas médias de acurácia, <i>recall</i> e precisão da MLP sobre todo o <i>dataset</i>	16
Tabela 4 – Taxas médias de acurácia, <i>recall</i> e precisão da RBF sobre cada fold . .	17
Tabela 5 – Taxas médias de acurácia, <i>recall</i> e precisão da RBF sobre todo o <i>dataset</i>	17

SUMÁRIO

1	INTRODUÇÃO	4
2	OBJETIVOS	5
3	MATERIAIS E MÉTODOS	6
3.1	Descritores de Contorno do <i>Dataset Leaf</i>	6
3.2	Validação K-fold	7
3.3	O modelo MLP	9
3.4	O modelo RBF	10
4	EXPERIMENTOS E RESULTADOS	11
4.1	Busca pela melhor arquitetura de MLP	11
4.2	Busca pela melhor arquitetura de RBF	12
4.3	Análise e comparação das melhores arquiteturas de MLP e RBF	14
5	CONCLUSÕES	19
	REFERÊNCIAS	20
	APÊNDICE A – RESUMO DO ARTIGO FONTE DA TÉCNICA DE DESCRIÇÃO DE CONTORNO	22
	APÊNDICE B – MATRIZES DE CONFUSÃO PARA O MELHOR MODELO MLP	27
	APÊNDICE C – MATRIZES DE CONFUSÃO PARA O MELHOR MODELO RBF	37
	APÊNDICE D – CÓDIGOS	47

1 INTRODUÇÃO

A análise de formas é um importante ramo da visão computacional. Sua utilização é motivada pelo extensivo uso da forma no reconhecimento de objetos pela visão humana (1), além de sua estabilidade em relação à iluminação, cor e textura (2). Devido a isso, diversas abordagens para descrição de formas foram desenvolvidas e aplicadas a problemas como reconhecimento de símbolos (3), marcas comerciais (4), detecção de pedestres (5) e auxílio na segmentação de imagens de células (6), de estruturas cardíacas em tomografias computadorizadas (7) e de objetos móveis em planos de fundo complexos (8), entre outros.

O reconhecimento de formas em imagens é normalmente realizado em duas etapas principais: descrição e classificação. A etapa de descrição consiste na extração de atributos discriminadores a partir da representação pictórica da forma. Métodos de descrição de formas podem ser dispostos em três classes principais, sendo baseados em:

1. região: usam todos os *pixels* do objeto na produção dos descritores (9, 10);
2. contorno: utilizam apenas informações dos *pixels* de borda do objeto (11, 12, 13, 14);
3. esqueleto: usam informações topológicas extraídas do eixo medial do objeto (15, 16).

Figura 1 – Região, contorno e esqueleto de uma forma



Fonte: Su et al., 2010

Na etapa de classificação, por sua vez, os vetores de atributos descritores são usados para representar instâncias de formas como pontos em um espaço multidimensional. O objetivo da etapa de classificação é, então, calcular hiperplanos capazes de separar adequadamente as classes de formas. Para isso, diversas técnicas de classificação podem ser empregadas. Neste trabalho, duas abordagens serão analisadas e comparadas: Redes Neurais Multicamadas (MLP, do inglês, *Multi Layer Perceptron*) e Redes de Funções de Base Radial (RBF, do inglês, *Radial Basis Function Network*). Para comparar os dois métodos, o presente trabalho faz uma análise de seus desempenhos de classificação sobre um *dataset* de descritores de contorno de folhas. O conjunto de dados foi produzido pela técnica de descrição apresentada em (17), baseada em redes neurais randomizadas.

2 OBJETIVOS

- (a) Implementar classificação usando MLP e RBF com auxílio do *software* MATLAB;
- (b) Implementar método de validação *K-fold* usando $K = 10$;
- (c) Comparar os resultados obtidos com diferentes arquiteturas de MLP e RBF sobre um *dataset* de descritores de formas de folhas baseados em contorno usando validação 10-*fold* e as métricas acurácia, *recall* e precisão.

3 MATERIAIS E MÉTODOS

Este capítulo apresenta a base de dados utilizada nos experimentos, o método de validação empregado e, enfim, as duas técnicas utilizadas para classificação. Para cada tópico, apresenta-se também trechos de código utilizados para sua implementação. A base de dados do trabalho é composta por descritores de contorno de 30 tipos de folhas. Para classificação, foram testadas diferentes arquiteturas de MLP e RBF. A validação dos modelos foi realizada através do método *K-fold*.

3.1 Descritores de Contorno do *Dataset Leaf*

A base de formas *Leaf* é composta por 600 amostras, totalizando 30 classes de folhas com 20 representantes cada. Exemplos das classes são exibidos na Figura 2. Os descritores das folhas são extraídos a partir de seu contorno através do método das redes neurais randomizadas. Tal método é explanado em detalhes no Apêndice A, que apresenta um resumo do trabalho desenvolvido por (17). Os descritores utilizados no presente trabalho foram gerados utilizando a abordagem da junção $\vec{\Omega}$ (Seção II-C-3 do Apêndice A), com número de neurônios $Q = 4$, parâmetros de vizinhança $K = \{2, 4, 6, 8\}$ e parâmetros de porção do contorno $P = \{5, 10, 15, 20\}$. A expressão que representa o descritor é $\vec{\Omega}(4)_{5,10,15,20}^{2,4,6,8}$. Ainda de acordo com (17), os vetores descritivos produzidos por essa abordagem possuem 40 dimensões.

Figura 2 – Exemplos de classes da base de formas *Leaf*.



Fonte: (17)

No MATLAB, os descritores das folhas são armazenados em uma matrix X de dimensões 40×600 , em que cada coluna representa um dos vetores descritivos e cada linha representa uma de suas dimensões. Os rótulos estão inicialmente dispostos numericamente indicando as categorias de cada amostra. No *script* desenvolvido, esses valores são transformados em uma matriz Y , com dimensões 30×600 , que representa a classe das amostras de X , em cada coluna, usando a codificação *one-hot*. O trecho de código a seguir, dos *scripts* *mlp10fold.m* e *rbf10fold.m*, mostra a extração de valores importantes do *dataset* e a transformação do vetor de categorias na matriz de rótulos Y .

```

7 % Variáveis calculadas a partir dos dados de entrada e rótulos
8 % Respectivamente, tamanho da camada de entrada, tamanho da camada
9 % de saída, número total de amostras, valor K usado no K-fold e
10 % tamanho de cada fold.
11 inputsize = size(Xraw,1);
12 outputsize = size(unique(Yraw),1);
13 samples = size(Xraw,2);
14 K = 10;
15 foldsize = samples/K;
16
17 % Os valores brutos dos dados de entrada já estão ajustados.
18 X = Xraw;
19
20 % Os valores brutos dos rótulos estão numerados de 0 a 29.
21 % Este trecho de código primeiramente numera de 1 a 30.
22 % Depois altera a representação da classe para a codificação one-hot.
23 Yraw = Yraw + 1;
24 Yraw = Yraw';
25 Y = zeros(outputsize, samples);
26 for i = 1:samples
27     Y(Yraw(1,i),i) = 1;
28 end

```

3.2 Validação K-fold

O *K-fold* é um método de validação cruzada cujo objetivo principal é verificar a capacidade de generalização do modelo treinado através do teste com amostras que não foram utilizadas durante o treinamento. Os passos para execução do *K-fold* são:

1. dividir a base de dados em K grupos G_1, G_2, \dots, G_K ;
2. selecionar um grupo G_i ainda não usado para teste e usar todos os outros grupos G_j com $i \neq j$ para treinamento do modelo;
3. testar o modelo treinado sobre as amostras do grupo G_i ;
4. registrar a taxa de acerto do modelo sobre G_i ;
5. voltar ao passo 2 até que todos os grupos tenham sido usados para teste.

Nos experimentos deste trabalho, utilizou-se $K = 10$. O número de amostras do *dataset* é 600, então cada grupo acomoda 60 amostras. Ademais, foram tomadas medidas para garantir que cada grupo contenha o mesmo número de amostras de cada classe. Considerando que há 30 classes e 20 amostras para cada, o algoritmo de divisão do *K-fold* implementado neste trabalho garante que todos os grupos possuem exatamente 2 amostras de cada classe. A seleção de amostras e seus grupos é feita de maneira aleatória no início do experimento usando a função `crossvalind`, disponível no MATLAB. O trecho de código a seguir mostra como o *K-fold* foi implementado nos arquivos *mlp10fold.m* e *rbf10fold.m*.


```

30 % Variáveis para armazenamento dos K folds.
31 % KX armazena os dados de entrada separados em K folds.
32 % KY armazena os dados de saída separados em K folds.
33 % h armazena a quantidade de itens em cada fold.
34 KX = zeros(K,inputsize,foldsize);
35 KY = zeros(K,outputsize,foldsize);
36 h = zeros(1,K);
37
38 % Este trecho de código utiliza a função crossvalind para selecionar
39 % randomicamente um fold para cada amostra. Essa seleção ocorre
40 % separadamente para cada classe, garantindo assim que cada fold
41 % contenha o mesmo número de amostras de cada classe.
42 indices = [];
43 for i = 1:outputsize
44     indices = [indices; crossvalind('Kfold', samples/outputsize, K)];
45 end
46
47 % Assinala cada amostra ao seu fold selecionado previamente.
48 for i = 1:samples
49     h(indices(i)) = h(indices(i)) + 1;
50     KX(indices(i),:,h(indices(i))) = X(:,i);
51     KY(indices(i),:,h(indices(i))) = Y(:,i);
52 end
53
54 success = zeros(K,1); % armazena taxas de acerto de cada um dos K folds.
55 total = zeros(K,1); % armazena taxas de acerto sobre todo o conjunto.
56
57 % Realiza K rodadas de treinamento e teste, uma rodada para cada fold;
58 for i = 1:K
59     % Separa e ajusta as dimensões do conjunto de teste
60     Xtest = KX(i, :, :);
61     Ytest = KY(i, :, :);
62     Xtest = reshape(Xtest,inputsize,foldsize);
63     Ytest = reshape(Ytest,outputsize,foldsize);
64
65     % Separa e ajusta as dimensões do conjunto de treinamento
66     Xtrain = [];
67     Ytrain = [];
68     for j = 1:K
69         if j ~= i
70             xx = reshape(KX(j, :, :),inputsize,foldsize);
71             yy = reshape(KY(j, :, :),outputsize,foldsize);
72             Xtrain = [Xtrain xx];
73             Ytrain = [Ytrain yy];
74         end
75     end

```

3.3 O modelo MLP

De forma geral, as MLPs usadas neste experimento possuem camada de entrada de comprimento 40, correspondente à dimensionalidade dos descritores da base de dados; e camada de saída com 30 neurônios, cada um representante de uma classe de folha da base. No MATLAB, a criação das redes MLP foi realizada com a função `feedforwardnet`. O algoritmo de treinamento utilizado foi *Scaled Conjugate Gradient Backpropagation*, representado pelo identificador `trainscg`. A Tabela 1 exibe os parâmetros fixos usados para criação e treinamento das MLPs.

Tabela 1 – Parâmetros fixos do treinamento da MLP no MATLAB

<code>net.trainFcn</code>	<code>trainscg</code>
<code>net.trainParam.epochs</code>	1000
<code>net.trainParam.goal</code>	10^{-3}
<code>net.trainParam.time</code>	inf
<code>net.trainParam.min_grad</code>	0

Como função de ativação das camadas ocultas foi utilizada a tangente hiperbólica sigmoide (`tansig`), enquanto na camada de saída utilizou-se a função linear. Ambas as definições de função de ativação são as opções padrão do MATLAB. Segue o código utilizado para definir, treinar e testar a MLP no arquivo *mlp10fold.m*.

```

77 % Configura a MLP
78 net = feedforwardnet(layers);
79 net.divideParam.trainRatio = 1;
80 net.divideParam.valRatio = 0;
81 net.divideParam.testRatio = 0;
82 net.trainFcn = 'trainscg';
83 net.trainParam.goal = 0.001;
84 net.trainParam.epochs = 1000;
85 net.trainParam.min_grad = 0;
86 net.trainParam.showWindow = false;
87
88 % Treina a MLP com os folds de treinamento
89 net = train(net, Xtrain, Ytrain);
90
91 % Testa a MLP com o fold de teste. Em seguida, seleciona os neurônios
92 % da camada de saída que tiveram a maior ativação e verifica se estes
93 % correspondem à classe das amostras. Caso positivo, registra acerto.
94 y = net(Xtest);
95 [m1,r1] = max(y);
96 [m2,r2] = max(Ytest);
97 result = (r1==r2);
98 success(i) = sum(result)/foldsize;

```

3.4 O modelo RBF

Assim como na MLP, a camada de entrada possui 40 elementos, enquanto a camada de saída possui 30 neurônios, respectivamente devido ao número de atributos dos descritores e ao número de classes. A função `newrb` foi responsável pela criação e treinamento da RBF. De acordo com a documentação da função `newrb`, durante sua execução, o número de neurônios na camada oculta é iterativamente incrementado até que o erro seja menor que a variável `goal` ou que o número máximo de neurônios, `maxneurons`, seja atingido. Foi verificado experimentalmente que o valor de erro dificilmente atinge zero. Dessa forma, para garantir que a RBF utilize o número especificado de neurônios, a variável `goal` é definida como 0 e `maxneurons` é configurado para ser igual ao número desejado de neurônios. Segue o código utilizado para definir, treinar e testar a RBF no arquivo *rbf10fold.m*.

```
77 % Configura e treina a RBF
78 % Utiliza-se a função evalc para capturar o output de texto da
79 % função newrb na variável T e assim desativar a impressão de
80 % texto no terminal MATLAB.
81 [T, net] = evalc('newrb(Xtrain, Ytrain, 0, 1.0, maxneurons, ...
    maxneurons+10)');
82
83 % Testa a RBF com o fold de teste. Em seguida, seleciona os neurônios
84 % da camada de saída que tiveram a maior ativação e verifica se estes
85 % correspondem à classe das amostras. Caso positivo, registra acerto.
86 y = net(Xtest);
87 [m1,r1] = max(y);
88 [m2,r2] = max(Ytest);
89 result = (r1==r2);
90 success(i) = sum(result)/foldsize;
```

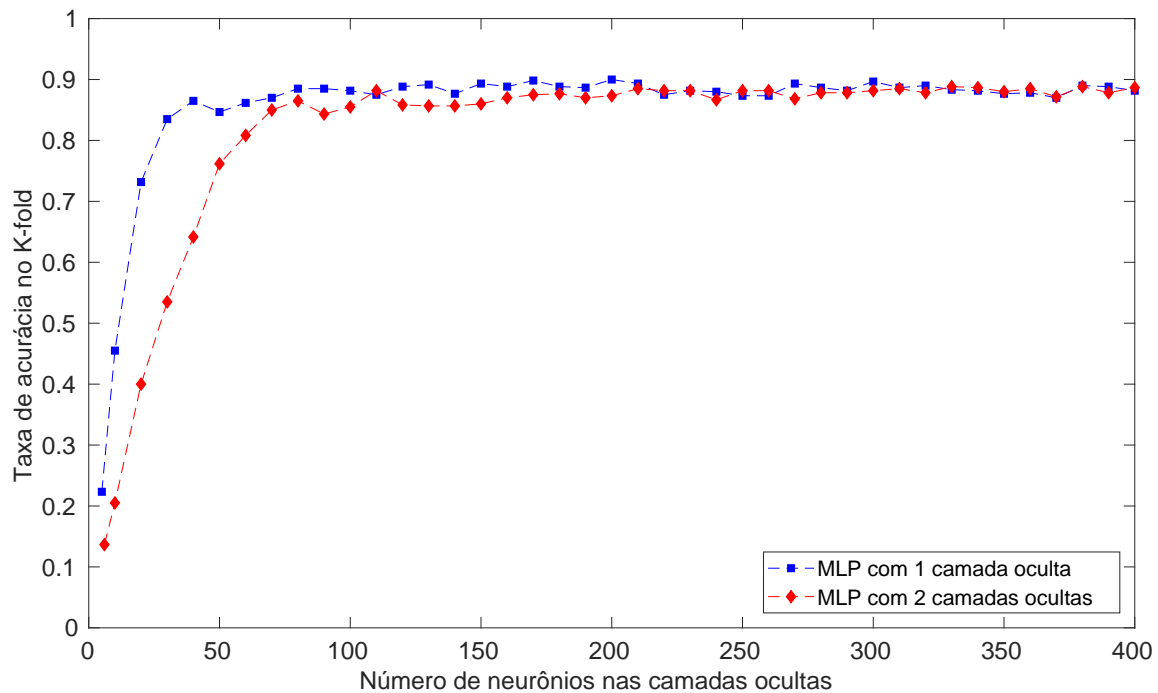
4 EXPERIMENTOS E RESULTADOS

Esta seção descreve os experimentos realizados e, paralelamente, apresenta os resultados obtidos. Dois experimentos são realizados. Primeiramente, uma busca por arquiteturas de MLP e RBF capazes de obter boa generalização no problema de classificação de descritores de contorno. O segundo experimento analisa com mais detalhes as arquiteturas com maior taxa de acurácia obtidas no primeiro experimento. Os códigos utilizados para ambos os experimentos podem ser encontrados nos arquivos *mlp.m* e *rbf.m*.

4.1 Busca pela melhor arquitetura de MLP

O presente trabalho explora arquiteturas de MLP com 1 e 2 camadas ocultas através de variações nos números de neurônios n contidos nessas camadas. Os valores avaliados foram $n = 5, 10, 20, 30, 40, \dots, 400$. Para as arquiteturas contendo duas camadas, foram definidas camadas ocultas com os mesmos números de neurônios $n = 5, 10, 20, 30, 40, \dots, 400$, mas dessa vez divididos igualmente entre as duas camadas. Uma exceção foi feita para o caso $n = 5$, que tornou-se $n = 6$, para acomodar igualmente 3 neurônios em cada camada. No gráfico da Figura 3 são exibidas as taxas médias de acurácia obtidas pela validação K -fold em função do número de neurônios nas camadas ocultas. O código de geração desse gráfico encontra-se no arquivo *mlpplot.m*.

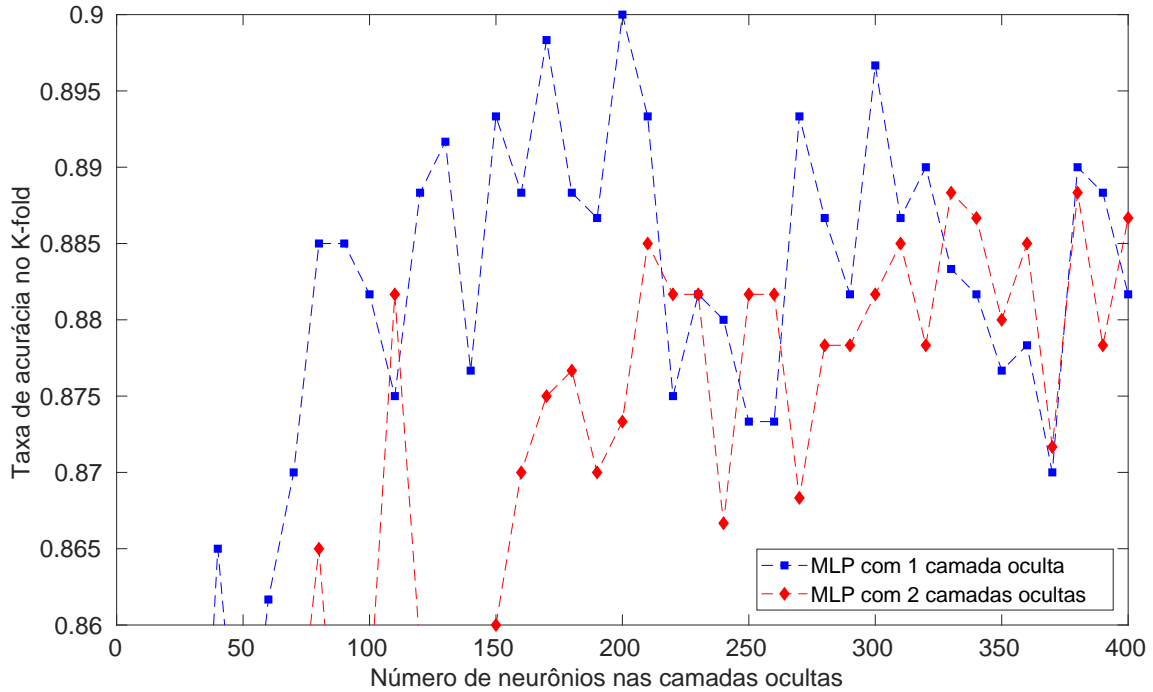
Figura 3 – Curvas de acurácia K -fold para MLPs



Fonte: Autoria própria

Observa-se na Figura 3 que ambas as arquiteturas de MLP, com 1 e 2 camadas ocultas, apresentam um rápido crescimento das taxas de acurácia. No entanto, conforme o número de neurônios nas camadas ocultas aumenta, após a marca dos 150 neurônios, os incrementos já não contribuem de forma significativa para o crescimento da acurácia.

Figura 4 – Ampliação das curvas de acurácia K -fold para MLPs

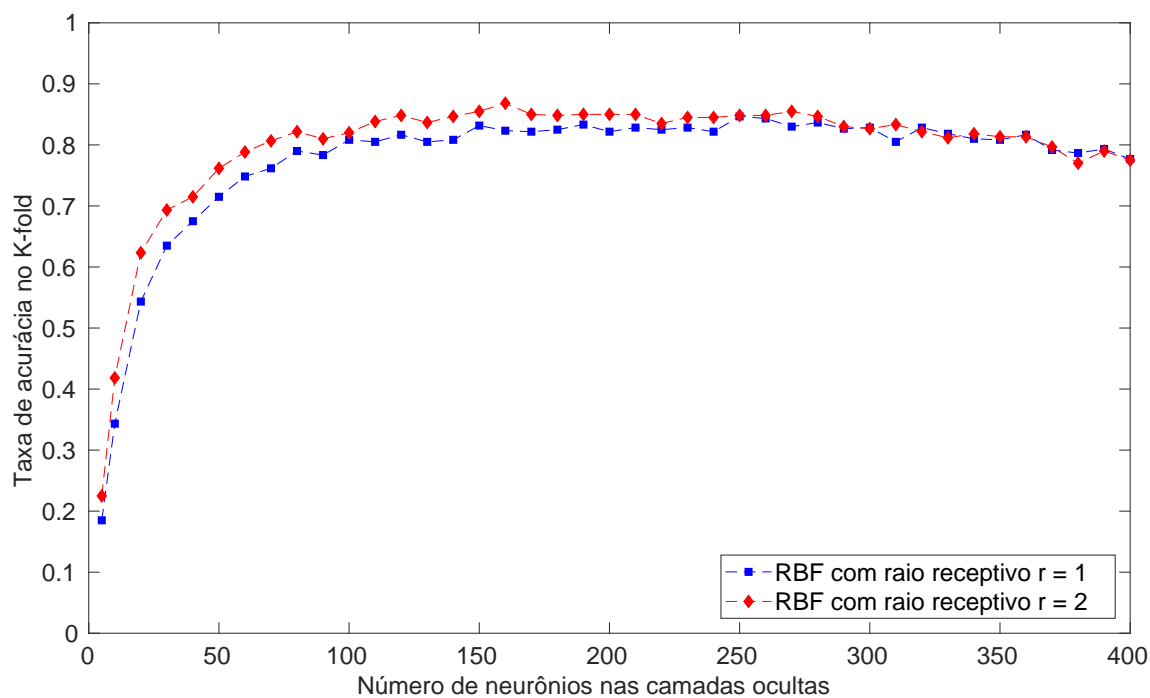


Fonte: Autoria própria

A Figura 4 é uma ampliação do gráfico da Figura 3 para exibir apenas o intervalo do eixo y entre 0.86 e 0.90. Vê-se que, após $n = 150$, os valores das taxas de acerto de ambas as arquiteturas estabilizam-se nesse intervalo. O valor de pico ocorre com apenas uma camada oculta com $n = 200$ neurônios, em que a taxa de sucesso obtida pela validação K -fold atinge 90%. Essa arquitetura foi escolhida para uma análise detalhada na seção 4.3.

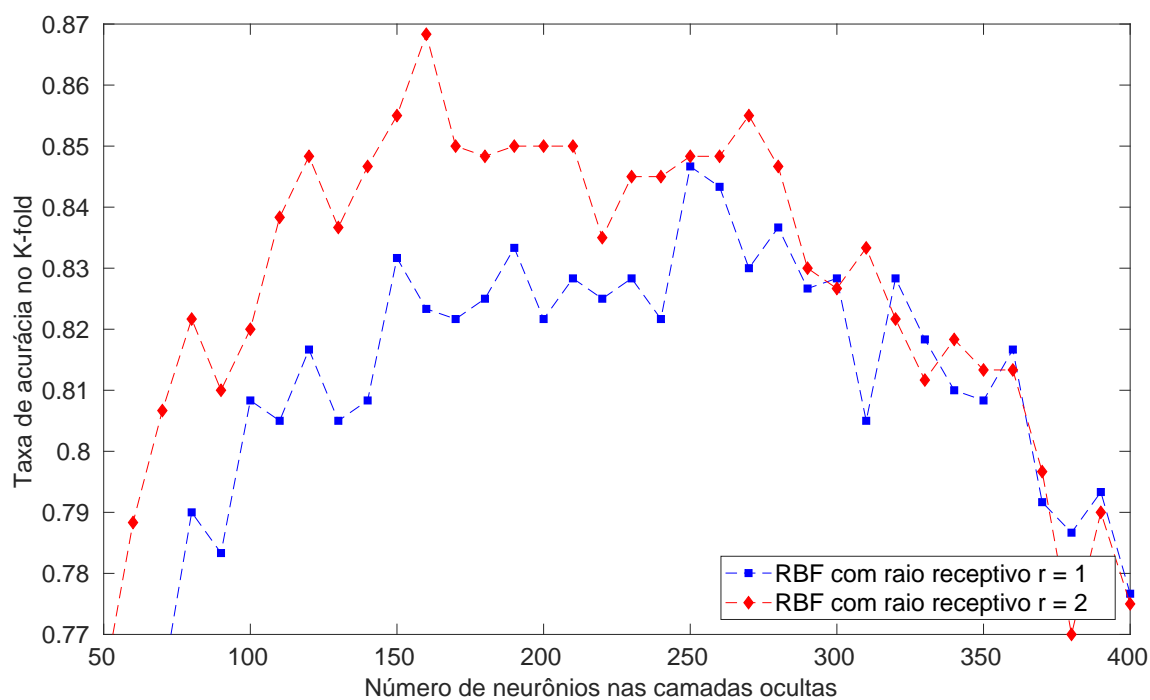
4.2 Busca pela melhor arquitetura de RBF

Neste trabalho, o número de neurônios na camada oculta da RBF foi explorado para obtenção de um modelo adequado ao problema. Os números de neurônios n avaliados foram $n = 5, 10, 20, 30, 40, \dots, 400$. Adicionalmente, o tamanho do raio receptivo dos neurônios de base radial também foi explorado através de dois valores, $r_1 = 1$ e $r_2 = 2$. Novamente, o critério utilizado foi a taxa de acurácia média obtida pela validação K -fold. Na Figura 5 são exibidas as curvas obtidas. O gráfico foi gerado pelo arquivo *rbfplot.m*. Observa-se também para a RBF um rápido crescimento das taxas de acurácia quando o valor de n é pequeno, menor que 50. No entanto, depois disso, as curvas tendem a se estabilizar e, ao final, dão sinais de decréscimo.

Figura 5 – Curvas de acurácia K -fold para RBFs

Fonte: Autoria própria

O gráfico da Figura 6 exibe com maior detalhe o comportamento das curvas de acurácia do modelo RBF. O intervalo de concentração das curvas é entre 0.77 e 0.87, sendo o valor de pico obtido com $n = 160$ e $r = 2$. Esta arquitetura obteve uma taxa de acurácia na validação K -fold igual a 86,83%, superando todas as outras RBFs analisadas e, por isso, foi escolhida para análise detalhada na seção 4.3.

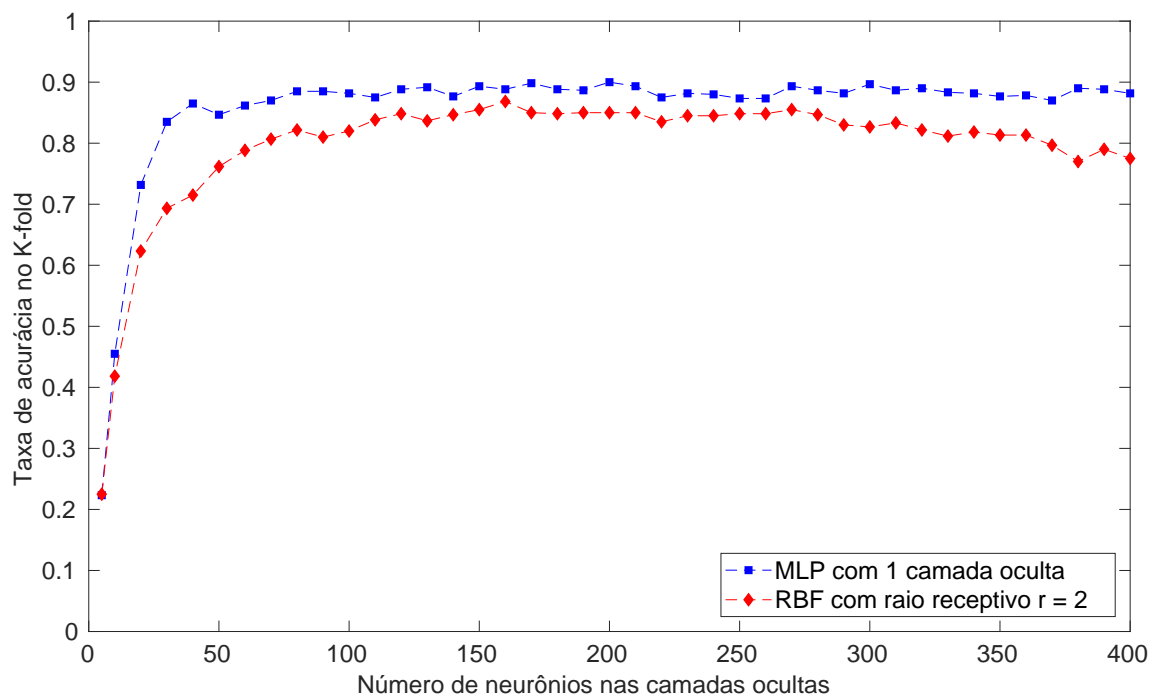
Figura 6 – Ampliação das curvas de acurácia K -fold para RBFs

Fonte: Autoria própria

4.3 Análise e comparação das melhores arquiteturas de MLP e RBF

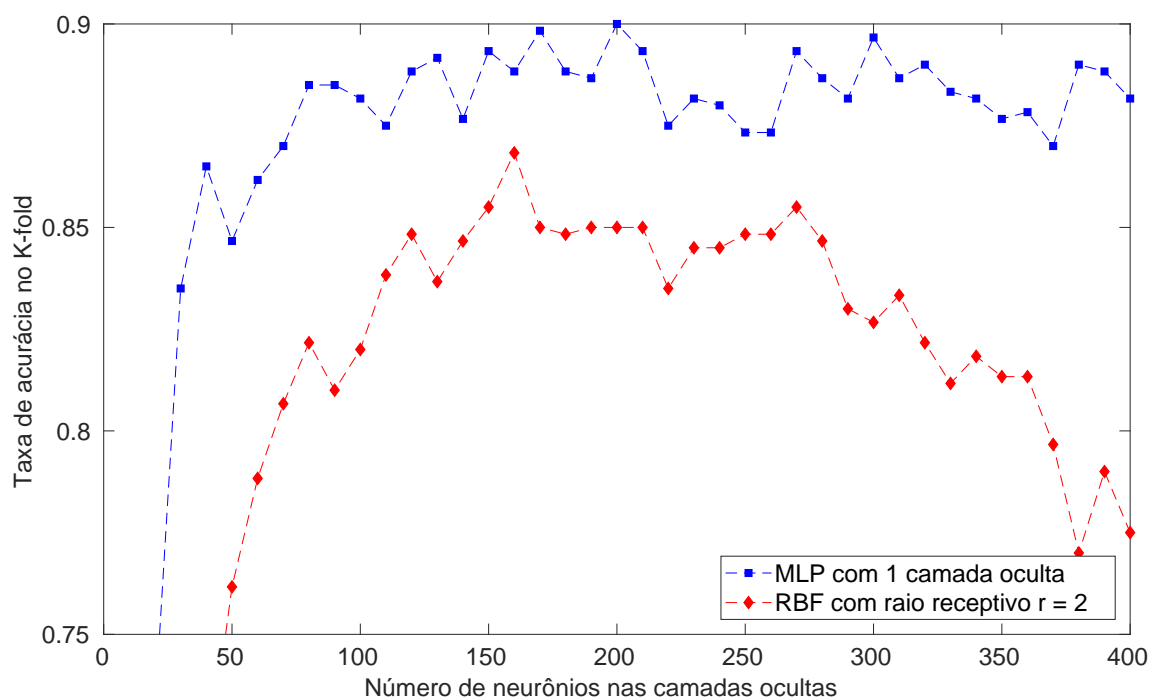
As curvas de acurácia dos modelos MLP com 1 camada oculta e RBF com neurônios de raio receptivo $r = 2$, que obtiveram maior sucesso nos experimentos anteriores, são comparadas nas Figuras 7 e 8 (geradas pelo script *mlprbfplot.m*). Através delas, pode-se notar que a acurácia da MLP se mantém superior à da RBF em todo o intervalo avaliado.

Figura 7 – Curvas de acurácia K -fold para as melhores MLPs e RBFs



Fonte: Autoria própria

Figura 8 – Ampliação das curvas de acurácia K -fold para as melhores MLPs e RBFs



Fonte: Autoria própria

Ademais, a presente seção seleciona os dois melhores modelos de MLP e RBF e os compara através das métricas acurácia, *recall* e precisão, obtidas a partir de classificações sobre o conjunto de teste do 10-*fold* e também sobre o *dataset* completo. Essas métricas são mecanismos clássicos para verificação da qualidade de modelos de classificação:

- Acurácia: mede a porcentagem de amostras cuja classe foi corretamente predita. A equação para cálculo da acurácia A_i de uma classe i é

$$A_i = TP/NC \quad (4.1)$$

em que TP é o número de amostras corretamente preditas da classe i e NC é o total de amostras da classe i .

- *Recall*: mede a probabilidade de o modelo determinar corretamente que uma amostra pertence à classe i , dado que a amostra realmente pertence à classe i . A equação para cálculo do *recall* R_i de uma classe i é

$$R_i = TP/(TP + FN) \quad (4.2)$$

onde TP é o número de amostras corretamente preditas da classe i e FN é o número de amostras da classe i que o modelo erroneamente atribuiu a outra classe qualquer. *Recall* também pode ser calculado através da matriz de confusão M do modelo:

$$R_i = M_{ii} / \sum_j M_{ij} \quad (4.3)$$

- Precisão: dado que o modelo prediz que uma amostra pertence à classe i , mede a probabilidade de o modelo ter acertado. O cálculo da precisão P_i para uma classe i é dado por

$$P_i = TP/(TP + FP) \quad (4.4)$$

em que TP é o número de amostras corretamente preditas da classe i e FP é o número de amostras de outra classe qualquer que o modelo erroneamente atribuiu à classe i . Precisão também pode ser calculada através da matriz de confusão M :

$$P_i = M_{ii} / \sum_j M_{ji} \quad (4.5)$$

A arquitetura de MLP escolhida para análise neste experimento possui camada de entrada de tamanho 40, camada oculta de 200 neurônios com função de ativação tangente hiperbólica sigmoide e camada de saída de 30 neurônios com função de ativação linear. O experimento deu-se através da execução da validação K -*fold*, com $K = 10$, e análise das taxas de acurácia, *recall* e precisão para cada uma das 10 etapas de treinamento e teste.

A Tabela 2 exibe os resultados obtidos de acurácia, *recall* e precisão sobre o grupo de teste para cada uma das 10 etapas do 10-*fold*. Além disso, são mostrados também as médias μ e o desvio padrão σ de cada métrica. Esses valores foram calculados a partir das

matrizes de confusão do modelo sobre os conjuntos de teste; essas matrizes foram omitidas do presente relatório, mas encontram-se presentes nos arquivos do projeto. Para facilitar a visualização, mostram-se as médias de *recall* e precisão considerando todas as 30 classes do *dataset*. Através da análise da Tabela 2, verifica-se que há consistência entre acurácia, *recall* e precisão em todas as etapas do 10-*fold*.

Ainda de acordo com as estatísticas da validação 10-*fold* exibidas na Tabela 2, a quarta etapa resultou na melhor capacidade de generalização, considerando as médias de 93% de acurácia e *recall* e 95% de precisão.

Tabela 2 – Taxas médias de acurácia, *recall* e precisão da MLP sobre cada *fold*

	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	μ	σ
acurácia	0.88	0.92	0.88	0.93	0.90	0.90	0.90	0.88	0.90	0.90	0.90	0.01
<i>recall</i>	0.88	0.92	0.88	0.93	0.90	0.90	0.90	0.88	0.90	0.90	0.90	0.01
precisão	0.89	0.94	0.90	0.95	0.93	0.90	0.91	0.93	0.92	0.94	0.92	0.02

Adicionalmente, outro experimento foi realizado para medir a capacidade do modelo de realizar predições sobre todo o *dataset* disponível, não apenas sobre o conjunto de teste especificado no 10-*fold*. Nessa análise, para cada etapa do 10-*fold*, o modelo foi treinado com 9 dos grupos, mas testado sobre o *dataset* completo. Os resultados de acurácia, *recall* e precisão, e suas médias e desvios, são exibidos na Tabela 3. Os valores foram calculados com base nas matrizes de confusão de cada etapa do 10-*fold*; as 10 matrizes podem ser visualizadas no Apêndice B. Verifica-se que as taxas alcançadas no teste com o *dataset* completo superam com larga vantagem os valores obtidos no experimento realizado apenas com os grupos de teste.

A partir da análise de ambas as tabelas, pode-se confirmar que o treinamento realizado foi capaz de ajustar a MLP para classificar com alta acurácia, *recall* e precisão tanto as amostras que foram apresentadas durante o treino quanto as que não haviam sido vistas anteriormente. O valor de pico atingido foi 98% na sexta etapa do 10-*fold*.

Tabela 3 – Taxas médias de acurácia, *recall* e precisão da MLP sobre todo o *dataset*

	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	μ	σ
acurácia	0.97	0.97	0.97	0.97	0.97	0.98	0.97	0.97	0.97	0.96	0.97	0.003
<i>recall</i>	0.97	0.97	0.97	0.97	0.97	0.98	0.97	0.97	0.97	0.96	0.97	0.003
precisão	0.98	0.97	0.97	0.97	0.97	0.98	0.97	0.98	0.97	0.97	0.97	0.003

A arquitetura de RBF selecionada para análise possui uma camada de entrada de tamanho 40, camada oculta com 160 neurônios de base radial com raio $r = 2$ e uma camada de saída de 30 neurônios com função de ativação linear. Como feito anteriormente, o foco deste experimento é a análise dos valores de acurácia, *recall* e precisão nos testes realizados durante a validação 10-*fold*. Essas medidas foram determinadas tanto para as classificações feitas sobre o conjunto de teste (Tabela 4) quanto para avaliações realizadas

sobre o *dataset* completo (Tabela 5). Os valores foram calculados com base nas matrizes de confusão de cada etapa do 10-*fold*. As matrizes obtidas sobre o conjunto de teste foram omitidas, mas encontram-se nos arquivos do projeto. As matrizes calculadas sobre o *dataset* completo relativos à RBF podem ser visualizadas no Apêndice C.

Tabela 4 – Taxas médias de acurácia, *recall* e precisão da RBF sobre cada fold

	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	μ	σ
acurácia	0.85	0.82	0.80	0.88	0.80	0.92	0.93	0.92	0.90	0.87	0.87	0.05
<i>recall</i>	0.85	0.82	0.80	0.88	0.80	0.92	0.93	0.92	0.90	0.87	0.87	0.05
precisão	0.89	0.84	0.85	0.92	0.83	0.92	0.95	0.94	0.92	0.90	0.90	0.04

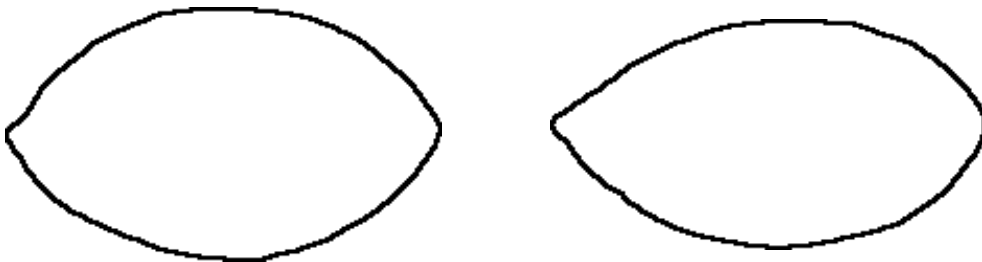
Tabela 5 – Taxas médias de acurácia, *recall* e precisão da RBF sobre todo o *dataset*

	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	μ	σ
acurácia	0.95	0.94	0.94	0.95	0.95	0.95	0.95	0.96	0.96	0.95	0.95	0.005
<i>recall</i>	0.95	0.94	0.94	0.95	0.95	0.95	0.95	0.96	0.96	0.95	0.95	0.005
precisão	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.96	0.96	0.95	0.95	0.005

Pela comparação entre as Tabelas 2 e 4, observa-se que, durante a validação 10-*fold* com o conjunto de teste, a MLP apresentou maiores médias μ de acurácia, *recall* e precisão. Além disso, pode-se afirmar que os treinamentos realizados com a MLP produziram resultados mais estáveis, visto que os desvios σ para acurácia, *recall* e precisão foram bem menores do que os obtidos pela RBF: enquanto os desvios da MLP giraram em torno de 0.01 e 0.02, para a RBF os mesmos valores atingiram 0.05 e 0.04. As Tabelas 3 e 5, referentes ao teste dos modelos sobre o *dataset* completo, refletem as mesmas características observadas na validação *K-fold*: a MLP supera a RBF em acurácia, *recall* e precisão e também apresenta maior estabilidade nas 10 etapas de treinamento e teste.

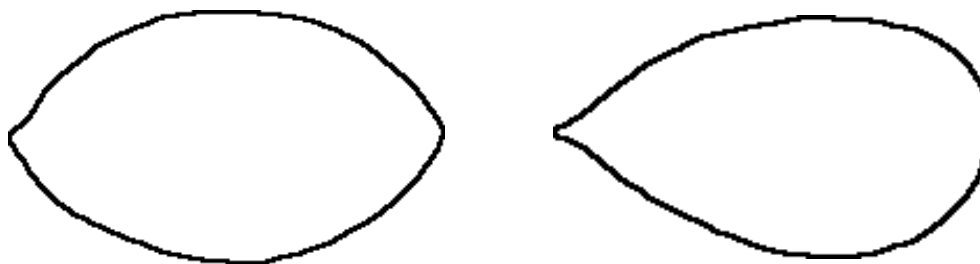
Uma informação interessante que pode ser extraída das matrizes de confusão dos Apêndice B e C é a existência de classes de difícil classificação pela MLP e pela RBF. São muito recorrentes os erros de classificação entre as classes 25 (folha de lima) e 26 (folha de limão), especialmente nas rodadas 3, 7 e 10 da MLP e 2, 3, 4, 5, 6, 7, 9 e 10 da RBF. Na Figura 9 são exibidos os contornos referentes a essas duas classes; pode-se ver que seus aspectos são bastante semelhantes, o que justifica a confusão dos modelos.

Figura 9 – Exemplos de contornos das classes 25 (folha de lima) e 26 (folha de limão)



Um outro erro recorrente, observado principalmente nas matrizes de confusão associadas à RBF, é a classificação de elementos da classe 25 (folha de lima) como pertencentes à classe 27 (folha da espécie *P. maliformis*), cujas semelhanças de contorno podem ser visualizadas na Figura 10. Através dessas observações, pode-se apontar a classe 25 (folha de lima) como a mais difícil de categorizar corretamente, na base de dados utilizada, por ambos os modelos MLP e RBF.

Figura 10 – Exemplos de contornos das classes 25 (folha de lima) e 27 (*P. maliformis*)



5 CONCLUSÕES

O presente trabalho fez uma avaliação dos modelos MLP e RBF para a tarefa de classificação de um *dataset* de descritores de contorno de folhas. Para alcançar esse objetivo, os modelos MLP e RBF foram implementados através de funções do *software* MATLAB. Diversas arquiteturas foram analisados por validação 10-*fold* e as mais bem sucedidas no critério acurácia foram selecionadas para uma análise mais detalhada. As melhores arquiteturas de MLP e RBF foram submetidas a um exame de acurácia, *recall* e precisão dos modelos treinados em cada etapa do 10-*fold* e testados, tanto sobre o conjunto de teste especificado pelas regras da validação, quanto para a base de dados completa.

Os resultados apontaram a superioridade da MLP para tratamento do problema especificado. A arquitetura de MLP com uma única camada oculta superou, no quesito acurácia, todos os outros modelos analisados, que incluíram MLP com duas camadas ocultas e RBFs com neurônios de base radial com raios receptivos $r = 1$ e $r = 2$. A análise detalhada dos melhores modelos mostrou a superioridade da MLP também nos quesitos *recall*, precisão e estabilidade dos resultados nas etapas da validação 10-*fold*.

REFERÊNCIAS

- 1 ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev*, p. 183–193, 1954. Citado na página 4.
- 2 WANG, X.; FENG, B.; BAI, X.; LIU, W.; LATECKI, L. J. Bag of contour fragments for robust shape classification. *Pattern Recognition*, v. 47, n. 6, p. 2116 – 2125, 2014. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0031320313005426>>. Citado na página 4.
- 3 YANG, S. Spectra of shape contexts: An application to symbol recognition. *Pattern Recognition*, v. 47, n. 5, p. 1891 – 1903, 2014. ISSN 0031-3203. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0031320313005062>>. Citado na página 4.
- 4 ANUAR, F. M.; SETCHI, R.; LAI, Y. kun. Trademark image retrieval using an integrated shape descriptor. *Expert Systems with Applications*, v. 40, n. 1, p. 105 – 121, 2013. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S095741741200886X>>. Citado na página 4.
- 5 CAO, J.; PANG, Y.; LI, X. Pedestrian detection inspired by appearance constancy and shape symmetry. *IEEE Transactions on Image Processing*, v. 25, n. 12, p. 5538–5551, Dec 2016. ISSN 1057-7149. Citado na página 4.
- 6 TAREEF, A.; SONG, Y.; CAI, W.; HUANG, H.; CHANG, H.; WANG, Y.; FULHAM, M.; FENG, D.; CHEN, M. Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation. *Neurocomputing*, v. 221, p. 94 – 107, 2017. ISSN 0925-2312. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0925231216311201>>. Citado na página 4.
- 7 ZHU, L.; GAO, Y.; APPIA, V.; YEZZI, A.; AREPALLI, C.; FABER, T.; STILLMAN, A.; TANNENBAUM, A. A complete system for automatic extraction of left ventricular myocardium from ct images using shape segmentation and contour evolution. *IEEE Transactions on Image Processing*, v. 23, n. 3, p. 1340–1351, March 2014. ISSN 1057-7149. Citado na página 4.
- 8 WANG, W.; SHEN, J.; LI, X.; PORIKLI, F. Robust video object cosegmentation. *IEEE Transactions on Image Processing*, v. 24, n. 10, p. 3137–3148, Oct 2015. ISSN 1057-7149. Citado na página 4.
- 9 ZHENJIANG, M. Zernike moment-based image shape analysis and its application. *Pattern Recognition Letters*, v. 21, n. 2, p. 169 – 177, 2000. ISSN 0167-8655. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167865599001440>>. Citado na página 4.
- 10 HU, M.-K. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, v. 8, n. 2, p. 179–187, February 1962. ISSN 0096-1000. Citado na página 4.

- 11 ATAER-CANSIZOGLU, E.; BAS, E.; KALPATHY-CRAMER, J.; SHARP, G. C.; ERDOGMUS, D. Contour-based shape representation using principal curves. *Pattern Recognition*, v. 46, n. 4, p. 1140 – 1150, 2013. ISSN 0031-3203. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0031320312004487>. Citado na página 4.
- 12 WU, W.; WANG, M. Detecting the dominant points by the curvature-based polygonal approximation. *CVGIP: Graphical Models and Image Processing*, v. 55, n. 2, p. 79 – 88, 1993. ISSN 1049-9652. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1049965283710060>. Citado na página 4.
- 13 OSOWSKI, S.; NGHIA, D. D. Fourier and wavelet descriptors for shape recognition using neural networks—a comparative study. *Pattern Recognition*, v. 35, n. 9, p. 1949 – 1957, 2002. ISSN 0031-3203. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0031320301001534>. Citado na página 4.
- 14 LING, H.; JACOBS, D. W. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, n. 2, p. 286–299, Feb 2007. ISSN 0162-8828. Citado na página 4.
- 15 BAI, X.; YANG, X.; YU, D.; LATECKI, L. J. Skeleton-based shape classification using path similarity. *International Journal of Pattern Recognition and Artificial Intelligence*, 2008. Citado na página 4.
- 16 BAI, X.; LATECKI, L. J. Path similarity skeleton graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 30, n. 7, p. 1282–1292, July 2008. ISSN 0162-8828. Citado na página 4.
- 17 JUNIOR, J. J. de M. S.; BACKES, A. R.; BRUNO, O. M. Randomized neural network based descriptors for shape classification. *Neurocomputing*, v. 312, p. 201 – 209, 2018. ISSN 0925-2312. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0925231218306842>. Citado 2 vezes nas páginas 4 e 6.

APÊNDICE A – RESUMO DO ARTIGO FONTE DA TÉCNICA DE DESCRIÇÃO DE CONTORNO

Descritores de Forma com Base em Redes Neurais Randomizadas

Borges, C. D. B.*

* Universidade Federal do Ceará – Campus Sobral
Sobral – Brasil
davidborges@protonmail.com

Resumo—Este trabalho é uma análise acerca do artigo *Randomized Neural Network Based Descriptors for Shape Classification*, escrito por Jarbas Joaci Mesquita de Sá Júnior, André Ricardo Backes e Odemir Martínez Bruno. No artigo, os autores propõem um novo método para gerar descritores de forma invariantes a translação, escalonamento e rotação. O método consiste na modelagem de informações extraídas do contorno do objeto como entradas e saídas para uma rede neural randomizada. Após o treinamento da rede, os pesos da camada de saída são utilizados como vetores descritivos da forma do objeto. Os autores demonstram a eficácia do método em experimentos comparativos com oito técnicas de descrição de formas encontradas na literatura. O uso dos descritores propostos aumenta as taxas de acerto e reduz tempos de classificação na maioria das bases de dados examinadas, fatores que validam a proposta do artigo para uso em tarefas de análise e categorização de formas.

I. INTRODUÇÃO

Os autores [1] descrevem a importância da análise de formas no campo da visão computacional. Os principais pontos motivadores mencionados são a utilização extensiva da forma no reconhecimento de objetos pela visão humana [2] e sua estabilidade em relação à iluminação, cor e textura [3].

A descrição de formas é amplamente utilizada em tarefas de reconhecimento de imagens. Pode-se citar o reconhecimento de símbolos [4], marcas comerciais [5], detecção de pedestres [6] e auxílio na segmentação de imagens de células [7], de estruturas cardíacas em tomografias computadorizadas [8] e de objetos móveis em planos de fundo complexos [9]. Os métodos de descrição de formas podem ser dispostos em três classes:

- 1) baseados em contorno: utilizam apenas informações dos *pixels* de borda do objeto [10], [11], [12], [13];
- 2) baseados em esqueleto: usam informações topológicas extraídas do eixo medial do objeto [14], [15], [16];
- 3) baseados em região: usam todos os *pixels* pertencentes ao objeto na produção dos descritores [17], [18].

A contribuição dos autores trata da utilização dos pesos de redes neurais randomizadas como vetores de atributos para descrição da forma de objetos com base em seu contorno. O restante desta análise organiza-se da seguinte maneira: na seção II são apresentados os fundamentos teóricos e processos utilizados na técnica proposta; a seção III é uma descrição dos experimentos realizados; a seção IV apresenta uma discussão dos resultados obtidos e das comparações feitas. Enfim, na seção V, conclui-se a presente análise.

II. ALGORITMOS E MÉTODOS

Esta seção descreve redes neurais randomizadas, incluindo o processo de geração de pesos aleatórios e treinamento, a

modelagem das informações de contorno como entradas e saídas da rede e a obtenção dos descritores de forma a partir dos pesos treinados.

A. Redes neurais randomizadas

No contexto deste trabalho, uma Rede Neural Randomizada (RNR) é uma rede de alimentação direta com uma única camada oculta (Figura 1), treinada com o algoritmo *Extreme Learning Machine* (ELM), descrito por Huang et al. [19]. No trabalho aqui analisado, os autores utilizam uma camada oculta composta por Q neurônios com função de ativação sigmoide.

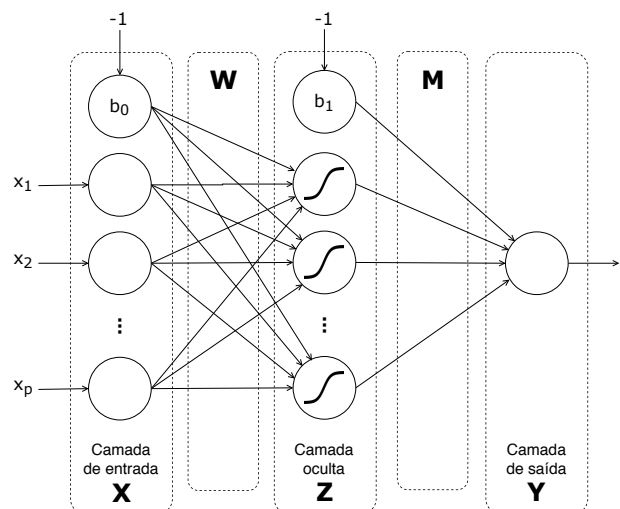


Figura 1. Arquitetura da rede neural randomizada. Fonte: autoria própria

O processo de treinamento da RNR é supervisionado, portanto, exige um conjunto de amostras e seus devidos rótulos. Sejam $X = [\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N]$ e $D = [\vec{d}_1, \vec{d}_2, \dots, \vec{d}_N]$, respectivamente, o conjunto de vetores de treinamento e o conjunto de rótulos. Uma linha composta de valores iguais a -1 , que representa o *bias*, é prefixada à matriz X . A matriz de pesos da camada oculta W é gerada por um processo aleatório e possui dimensões $Q \times p+1$, sendo p o número de atributos das amostras de treinamento e W_{ij} o peso da conexão entre o j -ésimo neurônio da camada de entrada e o i -ésimo neurônio da camada oculta.

O sinal Z , na saída da camada oculta, pode ser obtido através de cálculo matricial, segundo a Equação 1, em que σ representa a função sigmoide aplicada ponto a ponto.

$$Z = \sigma(WX) \quad (1)$$

A matriz Z é também prefixada com uma linha de valores iguais a -1 , representativa do *bias*. Sabe-se, então, que o resultado na camada de saída da rede é $Y = MZ$, em que M é a matriz de pesos da camada de saída. Os dados de treinamento fornecem a matriz D como referência para os valores desejados de Y , portanto, o aprendizado via ELM consiste no ajuste da matriz M em função de D e $Z = \sigma(WX)$. O objetivo do treinamento é minimizar o erro E da equação

$$D = MZ + E \quad (2)$$

através da configuração de M . Uma solução fechada para esse problema, exibida na Equação 3, pode ser obtida com a aplicação do Método dos Mínimos Quadrados. O aprendizado da arquitetura é, assim, concentrado nos pesos da camada de saída da rede neural neural randomizada.

$$M = DZ^T(ZZ^T)^{-1} \quad (3)$$

B. Geração de pesos aleatórios

Os autores do trabalho em análise utilizaram uma variante do Gerador Congruencial Linear (LCG, do inglês, *Linear Congruent Generator*) para produzir a matriz de pesos da camada oculta

$$W = \begin{pmatrix} w_{1,0} & w_{1,1} & \cdots & w_{1,p} \\ w_{2,0} & w_{2,1} & \cdots & w_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ w_{Q,0} & w_{Q,1} & \cdots & w_{Q,p} \end{pmatrix}. \quad (4)$$

O LCG é um gerador de números pseudo-aleatórios iterativo que baseia-se na equação

$$V_{n+1} = (aV_n + b) \bmod c \quad (5)$$

em que V_0 é a semente e os valores a , b e c , denominados multiplicador, incremento e módulo, são parâmetros de ajuste. Ao todo, são produzidos $Q(p+1)$ valores pseudo-aleatórios, quantidade suficiente para preencher W . A modificação inserida pelos autores consiste na substituição de V_{n+1} pela mediana do conjunto $S = \{0, 1, \dots, c-1\} - V$, sempre que for verificado que $V_{n+1} \in V$, em que V é o conjunto de valores já produzidos. Essa medida garante que não haja repetição nos valores gerados, ou seja $V_n \neq V_m \forall n, m \in \mathbb{N}$.

C. Assinaturas de contorno

São apresentados três modos de obtenção de descritores de formas através do treinamento de RNRs via ELM:

1) *Abordagem da vizinhança*: Nessa abordagem os autores utilizam uma assinatura local de cada *pixel* do contorno da forma para produzir os conjuntos de treinamento X e D . O primeiro passo é calcular o centroide da forma \vec{c} . Em sequência, para cada *pixel* $\vec{\pi}$ do contorno, são selecionados $p/2$ vizinhos no sentido horário e outros $p/2$ no sentido anti-horário. As distâncias entre os *pixels* vizinhos e \vec{c} (x_1, x_2, \dots, x_p) são usadas para compor os vetores de entrada, enquanto a distância d entre $\vec{\pi}$ e \vec{c} representa a saída esperada. Adicionalmente, com

o propósito de manter invariância à escala, todas as distâncias calculadas são divididas pelo perímetro do contorno antes da inserção em X e D . Esse processo, ilustrado na Figura 2, gera uma amostra de treinamento para cada *pixel* do contorno.

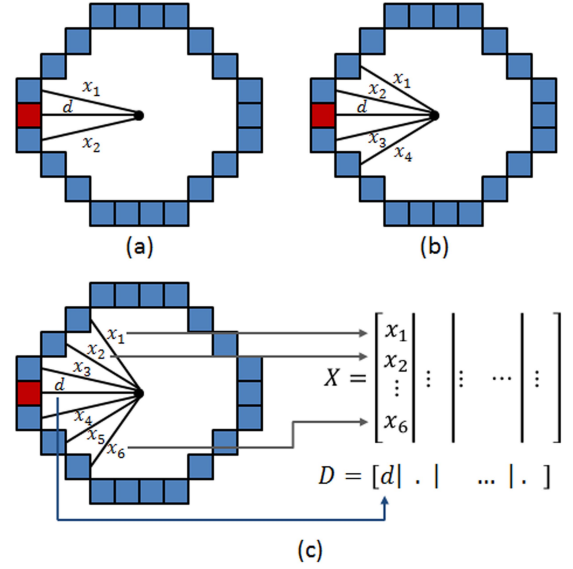


Figura 2. Esquema de obtenção das matrizes X e D na abordagem da vizinhança para (a) 2 vizinhos; (b) 4 vizinhos; (c) 6 vizinhos. Fonte: [1]

Para gerar a matriz de pesos W na abordagem da vizinhança, utiliza-se o LCG com parâmetros $V_0 = Q$, $a = p$, $b = L$ e $c = L^2$, em que p é o número de *pixels* da vizinhança, Q o número de neurônios da camada oculta e $L = Q(p+1)$ a quantidade de elementos de W . As equações 1, 2 e 3, para treinamento da RNR, são, então, usadas para produzir um vetor de características de forma $\vec{f}(Q, p) = DZ^T(ZZ^T)^{-1}$. Adicionalmente, os autores definem o descritor da vizinhança $\vec{\Theta}(Q)_{K_1, K_2, \dots, K_n}$ na forma exibida na Equação 6, através da concatenação (operador \parallel) de múltiplos vetores de características com diferentes tamanhos de vizinhança.

$$\vec{\Theta}(Q)_{K_1, K_2, \dots, K_n} = \vec{f}(Q, K_1) \parallel \vec{f}(Q, K_2) \parallel \dots \parallel \vec{f}(Q, K_n) \quad (6)$$

2) *Abordagem da porção do contorno*: Nessa abordagem os autores definem uma porção fixa $P\%$ do contorno e produzem X e D através de uma regra para extração de medidas de distância entre o centroide e *pixels* ao longo do perímetro. Primeiramente, calcula-se o centroide da forma \vec{c} . A seguir, para cada *pixel* $\vec{\pi}$, o algoritmo percorre $P\%$ do contorno no sentido horário e determina o *pixel* \vec{e} na extremidade do segmento perpassado. A distância entre \vec{c} e \vec{e} e entre \vec{c} e $\vec{\pi}$ (x_1, x_2 , respectivamente) são usadas para compor os vetores de entrada, enquanto a distância d entre $\vec{\pi}$ e \vec{e} representa a saída esperada. Adicionalmente, com o propósito de manter invariância à escala, todas as distâncias calculadas são divididas pelo perímetro do contorno antes da inserção em X e D . O processo, exemplificado na Figura 3, gera uma amostra de treinamento para cada *pixel* do contorno.

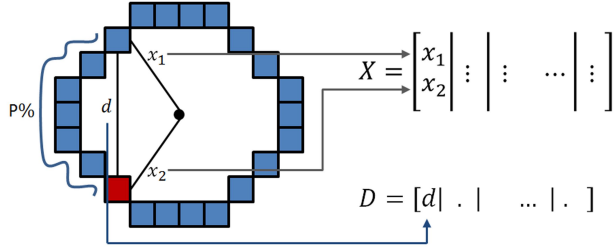


Figura 3. Esquema de obtenção das matrizes X e D na abordagem da porção do contorno para uma porcentagem fixa $P\%$. Fonte: [1]

Para gerar a matriz de pesos W na abordagem da porção do contorno, os autores utilizaram o LCG com parâmetros variantes em função da escolha de $P\%$. Os parâmetros utilizados são exibidos na tabela I, em que Q o número de neurônios da camada oculta, $p = 2$ e $L = Q(p + 1)$.

Tabela I
PARÂMETROS LCG USADOS NA ABORDAGEM DA PORÇÃO DO CONTO

$P\%$	V_0	a	b	c
5	Q	p	L	L^2
10	$Q + 1$	p	$L + 1$	L^2
15	$Q + 2$	p	$L + 2$	L^2
20	$Q + 3$	p	$L + 3$	L^2
25	$Q + 4$	p	$L + 4$	L^2

Novamente, as equações 1, 2 e 3, para treinamento da RNR, são usadas para produzir vetores de características $\vec{f}(Q, P) = DZ^T(ZZ^T)^{-1}$. Os autores definem o descritor da porção do contorno $\vec{\Psi}(Q)_{P_1, P_2, \dots, P_n}$ como na Equação 7, através da concatenação de múltiplos vetores de características com diferentes porcentagens $P\%$.

$$\vec{\Psi}(Q)_{P_1, P_2, \dots, P_n} = \vec{f}(Q, P_1) \parallel \vec{f}(Q, P_2) \parallel \dots \parallel \vec{f}(Q, P_n) \quad (7)$$

3) *Abordagem da junção*: A terceira abordagem mencionada pelos autores é a combinação dos descritores de vizinhança e da porção do contorno através de concatenação, de acordo com a Equação 8.

$$\vec{\Omega}(Q)_{P_1, P_2, \dots, P_n}^{K_1, K_2, \dots, K_n} = \vec{\Theta}(Q)_{K_1, K_2, \dots, K_n} \parallel \vec{\Psi}(Q)_{P_1, P_2, \dots, P_n} \quad (8)$$

III. EXPERIMENTOS

O método proposto foi comparado com 8 técnicas de descrição de formas da literatura: *Fourier Descriptors* [12] *Zernike Moments* [17], *Curvature Descriptors* [11], *Multi-scale Fractal Dimensions* [16], *Complex Networks* [20], *Inner-Distance Shape Context* [13], *Height Functions* [21] e *Radon Histograms* [22]. Os autores testaram os descritores sobre 6 bases de dados de formas e utilizaram um classificador LDA (do inglês, *Linear Discriminant Analysis*) aliado à técnica de validação cruzada *leave-one-out*. O restante desta seção dedica-se a uma breve descrição das bases de dados avaliadas.

A. Kimia

A base *Kimia* é composta por 9 classes de objetos genéricos com 11 imagens cada (Figura 4).

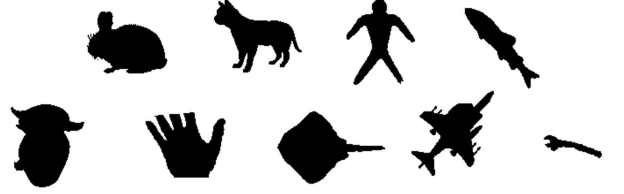


Figura 4. Classes da base de formas *Kimia*. Fonte: [1]

B. Fish

Esta base representa formas associadas a diferentes espécies de peixe e possui 1100 classes com 10 imagens cada, incluindo variações de rotação e escala (Figura 5).

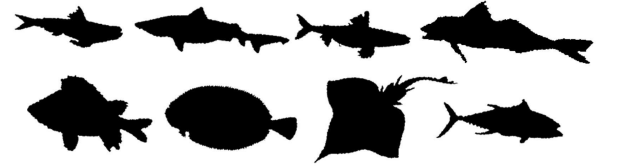


Figura 5. Exemplos da base de formas *Fish*. Fonte: [1]

C. Leaf

Uma base de formas composta por 30 classes de folhas com 20 representantes cada (Figura 6).

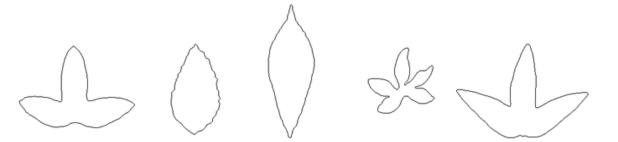


Figura 6. Exemplos de classes da base de formas *Leaf*. Fonte: [1]

D. Rotated Leaf

Esta base foi construída por variações rotacionadas da base *Leaf* sob ângulos de 7° , 35° , 104° , 132° , 201° , 298° . Possui 30 classes com 120 imagens cada (Figura 7).



Figura 7. Exemplos de classes da base de formas *Rotated Leaf*. Fonte: [1]

E. Scaled Leaf

Outra variação da base *Leaf*, dessa vez composto por escalonamentos pelos fatores 1.25, 1.50, 1.75 e 2.00. Possui 30 classes com 80 imagens cada (Figura 7).



Figura 8. Exemplos de classes da base de formas *Scaled Leaf*. Fonte: [1]

F. Noised Leaf

Uma variação da base *Leaf* na qual 4 níveis diferentes de ruído foram aplicados a cada folha. Possui 30 classes com 80 imagens cada (Figura 9).

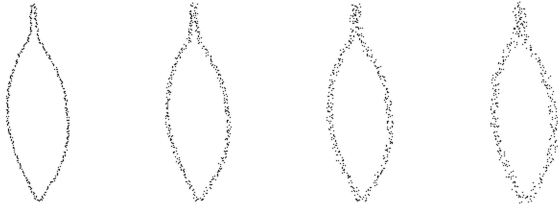


Figura 9. Exemplos de classes da base de formas *Noised Leaf*. Fonte: [1]

IV. RESULTADOS E DISCUSSÕES

O primeiro resultado obtido trata da escolha do descritor mais adequado dentre os três propostos. No artigo, os autores analisaram separadamente os resultados da validação cruzada para cada base de dados com a utilização de $Q = \{2, 3, 4, 5, 6, 7\}$ e três variantes de $\tilde{\Theta}$, $\tilde{\Psi}$ e $\tilde{\Omega}$. Neste trabalho, com o propósito de apresentar um resumo dos resultados alcançados, as tabelas 1, 2, 3, 4 e 5 do artigo em questão foram comprimidas na Tabela II do presente relatório, que apresenta as taxas médias de sucesso obtidas por validação *leave-one-out* nas bases *Kimia*, *Fish*, *Leaf*, *Rotated Leaf* e *Scaled Leaf*.

Tabela II
MÉDIAS DOS RESULTADOS OBTIDOS POR [1] NAS TABELAS 1, 2, 3, 4 E 5

Camada oculta (Q)	2	3	4	5	6	7
$\tilde{\Theta}(Q)_{2,4,6}$	85.61	83.32	85.45	86.27	86.25	87.86
$\tilde{\Theta}(Q)_{2,4,6,8}$	86.82	85.18	85.95	86.56	86.98	88.60
$\tilde{\Theta}(Q)_{2,4,6,8,10}$	88.08	85.98	86.83	87.55	88.09	89.15
$\tilde{\Psi}(Q)_{5,10,15}$	78.48	81.90	84.38	86.07	84.29	86.02
$\tilde{\Psi}(Q)_{5,10,15,20}$	81.96	85.33	86.71	88.43	88.58	89.71
$\tilde{\Psi}(Q)_{5,10,15,20,25}$	82.79	86.43	87.83	88.55	89.44	90.22
$\tilde{\Omega}(Q)_{5,10,15}^{2,4,6}$	88.98	88.71	91.11	90.92	90.84	92.18
$\tilde{\Omega}(Q)_{5,10,15,20}^{2,4,6,8}$	89.93	90.50	91.66	92.84	92.95	93.01
$\tilde{\Omega}(Q)_{5,10,15,20,25}^{2,4,6,8,10}$	90.30	91.01	92.35	93.38	93.57	91.21

A taxa de sucesso do descritor escolhido pelos autores é apresentada em negrito. Apesar de não ser a maior taxa média alcançada, os autores argumentam que a escolha levou em consideração tanto a taxa de acerto quanto o número de dimensões do descritor. Por exemplo, o descritor $\tilde{\Omega}(6)_{5,10,15,20,25}^{2,4,6,8,10}$ obteve a maior taxa média de sucesso, no entanto, a dimensionalidade dos vetores de características produzidos seria $N(Q + 1) = 70$, sendo N o número de parâmetros utilizados e Q o número de neurônios. Enquanto

isso, o descritor genérico escolhido, $\tilde{\Omega}(4)_{5,10,15,20}^{2,4,6,8}$, alcança taxas de sucesso comparáveis utilizando 40 dimensões.

A etapa seguinte foi uma comparação do descritor selecionado com outras técnicas de descrição de formas mencionadas na literatura. Os autores reiteram a importância da escolha de apenas um descritor proposto para essa comparação, pois apenas assim é possível realizar uma análise justa e eliminar a possibilidade de que a proposta tenha obtido melhores taxas por consequência de ajustes finos. Deve-se mencionar que o método proposto foi projetado para contornos contínuos. No entanto, a base *Noised Leaf*, devido ao ruído, apresenta problemas de descontinuidade de contorno. Este problema foi resolvido pelos autores através da aplicação de uma técnica de estimativa de contorno chamada Fluxo do Vetor Gradiente (do inglês, *Gradient Vector Flow*). A comparação exibida na Tabela III mostra que a proposta do artigo supera os métodos da literatura em quase todas as bases de dados analisadas, com exceção de *Fish* e *Noised Leaf*. Mesmo nesses casos, a proposta alcança a segunda melhor taxa de acerto.

Tabela III
COMPARAÇÃO COM DESCRITORES DA LITERATURA. FONTE: [1]

Métodos/Bases	<i>Kimia</i>	<i>Fish</i>	<i>Leaf</i>	<i>R. Leaf</i>	<i>S. Leaf</i>	<i>N. Leaf</i>
Método proposto	97.98	99.07	84.67	87.67	88.92	80.58
C.N. Degree	95.96	99.38	83.67	83.89	84.12	83.50
C.N. J. Degree	86.87	94.80	76.83	78.22	79.45	77.80
Fourier	83.84	99.07	75.00	76.53	81.58	65.04
Zernike	91.92	12.23	68.00	69.92	54.54	70.74
Curvature	76.77	97.55	75.00	78.64	80.00	76.12
M. Fractal Dim.	87.88	37.32	73.00	68.19	74.33	57.46

Tabela IV
COMPARAÇÃO COM MÉTODOS CLASSIFICADORES DA LITERATURA. FONTE: [1]

Métodos/Bases	<i>Kimia</i>	<i>Fish</i>	<i>Leaf</i>	<i>R. Leaf</i>	<i>S. Leaf</i>
Método proposto	92.93	95.46	80.50	85.06	97.63
IDSC	100.0	99.29	87.83	98.44	99.92
Height Functions	100.0	*	83.55	28.07	41.10
Radon Histogram	89.90	98.74	42.17	57.31	72.00

Tabela V
TEMPOS DE CLASSIFICAÇÃO PARA BASE *KIMIA*. FONTE: [1]

Métodos	Tempo
Método proposto	0.7075
C.N. Degree	0.7562
C.N. J. Degree	0.7756
Fourier	0.2867
Zernike	2.3488
Curvature	0.5444
M. Fractal Dim.	2.0546
IDSC	22.2672
Height Functionn	53.4914
Radon Histograms	32.3391

Uma outra avaliação foi realizada, desta vez para comparar a qualidade do método proposto com técnicas da literatura que não geram vetores de características, mas realizam

classificação através de algoritmos próprios de casamento de padrões. Para isso, os autores aplicaram aos descritores do método proposto um simples esquema de casamento 1-NN (do inglês, *1-Nearest Neighbor*) baseado em distância euclidiana. A comparação é mostrada na tabela IV. Verifica-se que o método proposto obtém sempre resultados inferiores ao IDSC (do inglês, *Inner-Distance Shape Context*). Os autores argumentam que a alta eficácia do IDSC é resultante do seu processo de casamento de padrões baseado em programação dinâmica, que otimiza o alinhamento entre as formas para a fase de classificação. Isso, no entanto, gera uma sobrecarga de tempo de processamento. A tabela V mostra uma breve avaliação do tempo de execução para classificar todas as amostras da base de dados *Kimia*. Nota-se neste experimento que o método proposto é cerca de 30 vezes mais rápido que o IDSC. A escolha entre o método proposto e o IDSC é, portanto, uma questão de troca entre velocidade e acurácia.

V. CONCLUSÕES E TRABALHOS FUTUROS

O artigo analisado defende a hipótese de que uma rede neural randomizada é capaz de modelar a forma de objetos bidimensionais e produzir vetores descritivos com boas qualidades discriminativas. Além disso, três abordagens para produção dos descritores são propostas: vizinhança, porção do contorno e junção.

Através dos experimentos realizados sobre múltiplas bases de dados e de comparações com métodos estabelecidos na literatura, verificou-se que a hipótese defendida é válida. Ademais, dentre as três abordagens descritas, a técnica da junção obteve melhor desempenho. Na tarefa de classificação, foram alcançadas taxas de acerto superiores à maioria dos métodos concorrentes. Nos casos em que isso não ocorreu, como na comparação com IDSC, foi mostrado que o método proposto é uma alternativa mais viável quando se consideram os tempos de execução: a proposta é cerca de 30 vezes mais rápida que o IDSC.

Enfim, as ideias de desenvolvimento futuro citadas pelos autores são duas. Em primeiro lugar, o método proposto considera apenas o contorno externo fechado da forma, mas é reconhecível que contornos internos também são importantes na descrição de objetos. Uma possível melhoria pode ser obtida com a extensão do método para abarcar contornos internos ou lidar com curvas abertas. Testes com diferentes distribuições probabilísticas para geração dos pesos da camada oculta da RNR também constituem um caminho interessante para pesquisa futura.

REFERÊNCIAS

- de Mesquita Sá Junior, J. J., Backes, A. R., and Bruno, O. M., "Randomized neural network based descriptors for shape classification," *Neurocomputing*, vol. 312, pp. 201 – 209, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231218306842>
- Attneave, F., "Some informational aspects of visual perception," *Psychol. Rev.*, pp. 183–193, 1954.
- Wang, X., Feng, B., Bai, X., Liu, W., and Latecki, L. J., "Bag of contour fragments for robust shape classification," *Pattern Recognition*, vol. 47, no. 6, pp. 2116 – 2125, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0031320313005426>
- Yang, S., "Spectra of shape contexts: An application to symbol recognition," *Pattern Recognition*, vol. 47, no. 5, pp. 1891 – 1903, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0031320313005062>
- Anuar, F. M., Setchi, R., and kun Lai, Y., "Trademark image retrieval using an integrated shape descriptor," *Expert Systems with Applications*, vol. 40, no. 1, pp. 105 – 121, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S095741741200886X>
- Cao, J., Pang, Y., and Li, X., "Pedestrian detection inspired by appearance constancy and shape symmetry," *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5538–5551, Dec 2016.
- Tareef, A., Song, Y., Cai, W., Huang, H., Chang, H., Wang, Y., Fulham, M., Feng, D., and Chen, M., "Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation," *Neurocomputing*, vol. 221, pp. 94 – 107, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231216311201>
- Zhu, L., Gao, Y., Appia, V., Yezzi, A., Arepalli, C., Faber, T., Stillman, A., and Tannenbaum, A., "A complete system for automatic extraction of left ventricular myocardium from ct images using shape segmentation and contour evolution," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1340–1351, March 2014.
- Wang, W., Shen, J., Li, X., and Porikli, F., "Robust video object cosegmentation," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3137–3148, Oct 2015.
- Ataer-Cansizoglu, E., Bas, E., Kalpathy-Cramer, J., Sharp, G. C., and Erdogmus, D., "Contour-based shape representation using principal curves," *Pattern Recognition*, vol. 46, no. 4, pp. 1140 – 1150, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0031320312004487>
- Wu, W. and Wang, M., "Detecting the dominant points by the curvature-based polygonal approximation," *CVGIP: Graphical Models and Image Processing*, vol. 55, no. 2, pp. 79 – 88, 1993. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1049965283710060>
- Osowski, S. and Nghia, D. D., "Fourier and wavelet descriptors for shape recognition using neural networks—a comparative study," *Pattern Recognition*, vol. 35, no. 9, pp. 1949 – 1957, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0031320301001534>
- Ling, H. and Jacobs, D. W., "Shape classification using the inner-distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 286–299, Feb 2007.
- Bai, X., Yang, X., Yu, D., and Latecki, L. J., "Skeleton-based shape classification using path similarity," *International Journal of Pattern Recognition and Artificial Intelligence*, 2008.
- Bai, X. and Latecki, L. J., "Path similarity skeleton graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 7, pp. 1282–1292, July 2008.
- Backes, A. R. and Bruno, O. M., "Shape skeleton classification using graph and multi-scale fractal dimension," in *Image and Signal Processing*, Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D., and Meunier, J., Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 448–455.
- Zhenjiang, M., "Zernike moment-based image shape analysis and its application," *Pattern Recognition Letters*, vol. 21, no. 2, pp. 169 – 177, 2000. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167865599001440>
- Hu, M.-K., "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, February 1962.
- Huang, G.-B., Zhu, Q.-Y., and Siew, C.-K., "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489 – 501, 2006, neural Networks. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231206000385>
- Backes, A. R., Casanova, D., and Bruno, O. M., "A complex network-based approach for boundary shape analysis," *Pattern Recognition*, vol. 42, no. 1, pp. 54 – 67, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0031320308002793>
- Wang, J., Bai, X., You, X., Liu, W., and Latecki, L. J., "Shape matching and classification using height functions," *Pattern Recognition Letters*, vol. 33, no. 2, pp. 134 – 143, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167865511003412>
- Hasegawa, M. and Tabbone, S., "Histogram of radon transform with angle correlation matrix for distortion invariant shape descriptor," *Neurocomputing*, vol. 173, pp. 24 – 35, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231215010243>

Rodada 2

[illegible]

Rodada 3

[illegible]

Rodada 4

[illegible]

Rodada 5

		Classes preditas																														
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
Classes reais	01	19	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	02	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	03	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	04	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	05	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	06	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
	07	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	08	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	09	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	10	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	11	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	12	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	13	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	17	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	
	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	19	0	0	0	0	0	0	0	0	0	0	0	0	
	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	
	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	
	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	
	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	18	0	0	0	0	0	0	0	1	
	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	
	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	18	0	0	0	0	0	
	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	14	3	2	0	0	
	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	18	0	0	0	
	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	
	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	
	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0
	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20

Rodada 6

[illegible]

Rodada 7

[illegible]

Rodada 8

[illegible]

Rodada 9

[illegible]

Rodada 10

[illegible]

Rodada 2

[illegible]

Rodada 3

[illegible]

Rodada 4

[illegible]

Rodada 6

[illegible]

Rodada 7

[illegible]

Rodada 8

[illegible]

Rodada 9

[illegible]

Rodada 10

[illegible]

APÊNDICE D – CÓDIGOS

mlp10fold.m

```

1 function [success, total] = mlp10fold(Xraw, Yraw, layers)
2     % Função mlp10fold: avalia um modelo MLP com camadas ocultas
3     % definidas pela variável layers, dados de entrada Xraw e
4     % rótulos Yraw com validação 10-fold. Retorna o vetor success,
5     % contendo as taxas de acerto para cada fold.
6
7     % Variáveis calculadas a partir dos dados de entrada e rótulos
8     % Respectivamente, tamanho da camada de entrada, tamanho da camada
9     % de saída, número total de amostras, valor K usado no K-fold e
10    % tamanho de cada fold.
11    inputsize = size(Xraw,1);
12    outputsize = size(unique(Yraw),1);
13    samples = size(Xraw,2);
14    K = 10;
15    foldsize = samples/K;
16
17    % Os valores brutos dos dados de entrada já estão ajustados.
18    X = Xraw;
19
20    % Os valores brutos dos rótulos estão numerados de 0 a 29.
21    % Este trecho de código primeiramente numera de 1 a 30.
22    % Depois altera representação da classe para codificação one-hot.
23    Yraw = Yraw + 1;
24    Yraw = Yraw';
25    Y = zeros(outputsize, samples);
26    for i = 1:samples
27        Y(Yraw(1,i),i) = 1;
28    end
29
30    % Variáveis para armazenamento dos K folds.
31    % KX armazena os dados de entrada separados em K folds.
32    % KY armazena os dados de saída separados em K folds.

```



```

33     % h armazena a quantidade de itens em cada fold.
34     KX = zeros(K,inputsize,foldsize);
35     KY = zeros(K,outputsize,foldsize);
36     h = zeros(1,K);
37
38     % Este trecho de código utiliza a função crossvalind para selecionar
39     % randomicamente um fold para cada amostra. Essa seleção ocorre
40     % separadamente para cada classe, garantindo assim que cada fold
41     % contenha o mesmo número de amostras de cada classe.
42     indices = [];
43     for i = 1:outputsize
44         indices = [indices; crossvalind('Kfold', samples/outputsize, K)];
45     end
46
47     % Assinala cada amostra ao seu fold selecionado previamente.
48     for i = 1:samples
49         h(indices(i)) = h(indices(i)) + 1;
50         KX(indices(i),:,h(indices(i))) = X(:,i);
51         KY(indices(i),:,h(indices(i))) = Y(:,i);
52     end
53
54     % armazena taxas de acerto de cada um dos K folds.
55     success = zeros(K,1);
56     % armazena taxas de acerto sobre todo o conjunto.
57     total = zeros(K,1);
58
59     % Realiza K rodadas de treinamento/ teste, uma para cada fold.
60     for i = 1:K
61         % Separa e ajusta as dimensões do conjunto de teste
62         Xtest = KX(i, :, :);
63         Ytest = KY(i, :, :);
64         Xtest = reshape(Xtest,inputsize,foldsize);
65         Ytest = reshape(Ytest,outputsize,foldsize);
66
67         % Separa e ajusta as dimensões do conjunto de treinamento
68         Xtrain = [];
69         Ytrain = [];
70         for j = 1:K

```

```

71         if j ~= i
72             xx = reshape(KX(j, :, :), inputsize, foldsize);
73             yy = reshape(KY(j, :, :), outputsize, foldsize);
74             Xtrain = [Xtrain xx];
75             Ytrain = [Ytrain yy];
76         end
77     end
78
79     % Configura a MLP
80     net = feedforwardnet(layers);
81     net.divideParam.trainRatio = 1;
82     net.divideParam.valRatio   = 0;
83     net.divideParam.testRatio  = 0;
84     net.trainFcn = 'trainscg';
85     net.trainParam.goal = 0.001;
86     net.trainParam.epochs = 1000;
87     net.trainParam.min_grad = 0;
88     net.trainParam.showWindow = false;
89
90     % Treina a MLP com os folds de treinamento
91     net = train(net, Xtrain, Ytrain);
92
93     % Testa a MLP com o fold de teste. Em seguida, seleciona os neurônios
94     % da camada de saída que tiveram a maior ativação e verifica se estes
95     % correspondem à classe das amostras. Caso positivo, registra acerto.
96     y = net(Xtest);
97     [m1, r1] = max(y);
98     [m2, r2] = max(Ytest);
99     result = (r1==r2);
100     success(i) = sum(result)/foldsize;
101
102     % Grava a matriz de confusao do fold em um arquivo csv caso
103     % o modelo seja MLP com uma camada oculta e 200 neuronios
104
105     if layers == [200]
106         confusionmatrix = zeros(outputsize);
107         for j = 1:length(r1)

```

```

108         confusionmatrix(r2(j), r1(j)) = confusionmatrix(r2(j), r1(j)) ...
            + 1;
109     end
110     csvwrite(sprintf('output/confmatrix-mlp-200-kfold-%d.csv', i), ...
        confusionmatrix);
111 end
112
113 % Testa a MLP com todo o dataset. Em seguida, seleciona os neurônios
114 % da camada de saída que tiveram a maior ativação e verifica se
115 % estes correspondem à classe correta das amostras.
116 % Caso positivo, registra acerto.
117 y = net(X);
118 [m1,r1] = max(y);
119 [m2,r2] = max(Y);
120 result = (r1==r2);
121 total(i) = sum(result)/samples;
122
123 % Grava a matriz de confusao total em um arquivo csv caso
124 % o modelo seja MLP com uma camada oculta e 200 neuronios
125 if layers == [200]
126     confusionmatrix = zeros(outputsize);
127     for j = 1:length(r1)
128         confusionmatrix(r2(j), r1(j)) = confusionmatrix(r2(j), r1(j)) ...
            + 1;
129     end
130     csvwrite(sprintf('output/confmatrix-mlp-200-total-%d.csv', i), ...
        confusionmatrix);
131 end
132 end
133
134 % Grava as taxas de acuracia do fold e totais caso o modelo
135 % seja MLP com uma camada oculta de 200 neuronios
136 if layers == [200]
137     csvwrite('output/success-mlp-200.csv', success);
138     csvwrite('output/total-mlp-200.csv', total);
139 end
140 end

```

rbf10fold.m

```

1 function [success, total] = rbf10fold(Xraw, Yraw, maxneurons)
2     % Função rbf10fold
3     % Avalia um modelo RBF com neurônios definidos pela variável maxneurons,
4     % dados de entrada Xraw e rótulos Yraw com validação 10-fold.
5     % Retorna o vetor success, contendo as taxas de acerto para cada fold.
6
7     % Variáveis calculadas a partir dos dados de entrada e rótulos
8     % Respectivamente, tamanho da camada de entrada, tamanho da camada
9     % de saída, número total de amostras, valor K usado no K-fold e
10    % tamanho de cada fold.
11    inputsize = size(Xraw,1);
12    outputsize = size(unique(Yraw),1);
13    samples = size(Xraw,2);
14    K = 10;
15    foldsize = samples/K;
16
17    % Os valores brutos dos dados de entrada já estão ajustados.
18    X = Xraw;
19
20    % Os valores brutos dos rótulos estão numerados de 0 a 29.
21    % Este trecho de código primeiramente numera de 1 a 30.
22    % Depois altera a representação da classe para a codificação one-hot.
23    Yraw = Yraw + 1;
24    Yraw = Yraw';
25    Y = zeros(outputsize, samples);
26    for i = 1:samples
27        Y(Yraw(1,i),i) = 1;
28    end
29
30    % Variáveis para armazenamento dos K folds.
31    % KX armazena os dados de entrada separados em K folds.
32    % KY armazena os dados de saída separados em K folds.
33    % h armazena a quantidade de itens em cada fold.
34    KX = zeros(K,inputsize,foldsize);
35    KY = zeros(K,outputsize,foldsize);
36    h = zeros(1,K);

```

```

38 % Este trecho de código utiliza a função crossvalind para selecionar
39 % randomicamente um fold para cada amostra. Essa seleção ocorre
40 % separadamente para cada classe, garantindo assim que cada fold
41 % contenha o mesmo número de amostras de cada classe.
42 indices = [];
43 for i = 1:outputsize
44     indices = [indices; crossvalind('Kfold', samples/outputsize, K)];
45 end
46
47 % Assinala cada amostra ao seu fold selecionado previamente.
48 for i = 1:samples
49     h(indices(i)) = h(indices(i)) + 1;
50     KX(indices(i), :, h(indices(i))) = X(:, i);
51     KY(indices(i), :, h(indices(i))) = Y(:, i);
52 end
53
54 % Vetor para armazenar as taxas de acerto de cada um dos K folds.
55 success = zeros(K, 1);
56 % Vetor para armazenar as taxas de acerto sobre todo o conjunto.
57 total = zeros(K, 1);
58
59 % Realiza K rodadas de treinamento e teste, uma rodada para cada fold;
60 for i = 1:K
61     % Separa e ajusta as dimensões do conjunto de teste
62     Xtest = KX(i, :, :);
63     Ytest = KY(i, :, :);
64     Xtest = reshape(Xtest, inputsize, foldsize);
65     Ytest = reshape(Ytest, outputsize, foldsize);
66
67     % Separa e ajusta as dimensões do conjunto de treinamento
68     Xtrain = [];
69     Ytrain = [];
70     for j = 1:K
71         if j ~= i
72             xx = reshape(KX(j, :, :), inputsize, foldsize);
73             yy = reshape(KY(j, :, :), outputsize, foldsize);
74             Xtrain = [Xtrain xx];

```

```

75         Ytrain = [Ytrain yy];
76     end
77 end
78
79 % Configura e treina a RBF
80 % Utiliza-se a função evalc para capturar o output de texto da
81 % função newrb na variável T e assim desativar a impressão de
82 % texto no terminal MATLAB.
83 [T, net] = evalc('newrb(Xtrain, Ytrain, 0, 2.0, maxneurons, ...
84                 maxneurons+10)');
85
86 % Testa a RBF com o fold de teste. Em seguida, seleciona os neurônios
87 % da camada de saída que tiveram a maior ativação e verifica se estes
88 % correspondem à classe das amostras. Caso positivo, registra acerto.
89 y = net(Xtest);
90 [m1,r1] = max(y);
91 [m2,r2] = max(Ytest);
92 result = (r1==r2);
93 success(i) = sum(result)/foldsize;
94
95 % Grava a matriz de confusao do fold em um arquivo csv caso
96 % o modelo seja RBF com uma camada oculta e 160 neuronios
97
98 if maxneurons == 160
99     confusionmatrix = zeros(outputsize);
100     for j = 1:length(r1)
101         confusionmatrix(r2(j), r1(j)) = confusionmatrix(r2(j), r1(j)) ...
102             + 1;
103     end
104     csvwrite(sprintf('output/confmatrix-rbf-160-kfold-%d.csv', i), ...
105         confusionmatrix);
106 end
107
108 % Testa a RBF com todo o dataset. Em seguida, seleciona os neurônios
109 % da camada de saída que tiveram a maior ativação e verifica se
110 % estes correspondem à classe correta das amostras.
111 % Caso positivo, registra acerto.

```

```

109     y = net(X);
110     [m1,r1] = max(y);
111     [m2,r2] = max(Y);
112     result = (r1==r2);
113     total(i) = sum(result)/samples;
114
115     % Grava a matriz de confusao total em um arquivo csv caso
116     % o modelo seja RBF com uma camada oculta e 160 neuronios
117
118     if maxneurons == 160
119         confusionmatrix = zeros(outputsize);
120         for j = 1:length(r1)
121             confusionmatrix(r2(j), r1(j)) = confusionmatrix(r2(j), r1(j)) ...
122                 + 1;
123         end
124         csvwrite(sprintf('output/confmatrix-rbf-160-total-%d.csv', i), ...
125             confusionmatrix);
126     end
127
128     % Grava as taxas de acuracia total e do fold em um arquivo csv caso
129     % o modelo seja RBF com uma camada oculta e 160 neuronios
130     if maxneurons == 160
131         csvwrite('output/success-rbf-160.csv', success);
132         csvwrite('output/total-rbf-160.csv', total);
133     end
134
135 end

```

mlp.m

```

1 % Carrega os datasets.
2 Xraw = csvread('X.csv');
3 Yraw = csvread('Y.csv');
4
5 % Configura o gerador aleatório para uma semente fixa.
6 rng('default');
7 rng(1);
8
9 % Define o intervalo de número de neurônios a ser investigado.
10 % Cria variáveis para armazenar os mínimos,
11 % máximos e médias das taxas de acerto.
12 range_ = 0:10:400;
13 min_ = zeros(1, length(range_));
14 max_ = zeros(1, length(range_));
15 mean_ = zeros(1, length(range_));
16 min_total = zeros(1, length(range_));
17 max_total = zeros(1, length(range_));
18 mean_total = zeros(1, length(range_));
19 i = 1;
20
21 % Se o intervalo definido possuir mais de um valor,
22 % salvar os dados em um arquivo, caso contrário,
23 % apenas exibir no terminal.
24 savedata = false;
25 if length(range_) > 1
26     % Define o número mínimo de neurônios para 5.
27     range_(1) = 5;
28     savedata = true;
29 end
30
31 % Para todos os valores do intervalo.
32 for n = range_
33
34     n
35
36     % Calcula as taxas de acerto com K-fold da MLP

```



```

37     % com n neurônios na camada oculta.
38     [success, total] = mlp10fold(Xraw, Yraw, [n]);
39
40     % Calcula mínimos, máximos e médias
41     min_(i) = min(success);
42     max_(i) = max(success);
43     mean_(i) = mean(success);
44     min_total(i) = min(total);
45     max_total(i) = max(total);
46     mean_total(i) = mean(total);
47
48     i = i + 1;
49
50     % Salva ou apenas imprime os dados no terminal,
51     % de acordo com a variável savedata.
52     if savedata
53         data = [range_; min_; max_; mean_; min_total; max_total; mean_total];
54         csvwrite('output/mlpdata.csv', data);
55     else
56         n
57         success
58         min(success)
59         max(success)
60         mean(success)
61         total
62         min(total)
63         max(total)
64         mean(total)
65     end
66
67 end

```

rbf.m

```
1 % Carrega os datasets.
2 Xraw = csvread('X.csv');
3 Yraw = csvread('Y.csv');
4
5 % Configura o gerador aleatório para uma semente fixa.
6 rng('default');
7 rng(1);
8
9 % Define o intervalo de número de neurônios a ser investigado.
10 % Cria variáveis para armazenar os mínimos,
11 % máximos e médias das taxas de acerto.
12 range_ = 0:10:400;
13 min_ = zeros(1, length(range_));
14 max_ = zeros(1, length(range_));
15 mean_ = zeros(1, length(range_));
16 min_total = zeros(1, length(range_));
17 max_total = zeros(1, length(range_));
18 mean_total = zeros(1, length(range_));
19 i = 1;
20
21 % Se o intervalo definido possuir mais de um valor,
22 % salvar os dados em um arquivo, caso contrário,
23 % apenas exibir no terminal.
24 savedata = false;
25 if length(range_) > 1
26     % Define o número mínimo de neurônios para 5.
27     range_(1) = 5;
28     savedata = true;
29 end
30
31 % Para todos os valores do intervalo.
32 for n = range_
33
34     n
35
36     % Calcula as taxas de acerto com K-fold da RBF
```

```

37     % com n neurônios na camada oculta.
38     [success, total] = rbf10fold(Xraw, Yraw, n);
39
40     % Calcula mínimos, máximos e médias
41     min_(i) = min(success);
42     max_(i) = max(success);
43     mean_(i) = mean(success);
44     min_total(i) = min(total);
45     max_total(i) = max(total);
46     mean_total(i) = mean(total);
47
48     i = i + 1;
49
50     % Salva ou apenas imprime os dados no terminal,
51     % de acordo com a variável savedata.
52     if savedata
53         data = [range_; min_; max_; mean_; min_total; max_total; mean_total];
54         csvwrite('output/rbfdata.csv', data);
55     else
56         n
57         success
58         min(success)
59         max(success)
60         mean(success)
61         total
62         min(total)
63         max(total)
64         mean(total)
65     end
66
67 end

```

mlpplot.m

```

1 mlpdata1 = csvread('output/mlpdata-1L.csv');
2 mlpdata2 = csvread('output/mlpdata-2L.csv');
3
4 rangemlp1 = mlpdata1(1,:);
5 minmlp1   = mlpdata1(2,:);
6 maxmlp1   = mlpdata1(3,:);
7 meanmlp1  = mlpdata1(4,:);
8 mintotalmlp1 = mlpdata1(5,:);
9 maxtotalmlp1 = mlpdata1(6,:);
10 meantotalmlp1 = mlpdata1(7,:);
11
12 rangemlp2 = mlpdata2(1,:);
13 minmlp2   = mlpdata2(2,:);
14 maxmlp2   = mlpdata2(3,:);
15 meanmlp2  = mlpdata2(4,:);
16 mintotalmlp2 = mlpdata2(5,:);
17 maxtotalmlp2 = mlpdata2(6,:);
18 meantotalmlp2 = mlpdata2(7,:);
19
20 [m1 i1] = max(meanmlp1);
21 [m2 i2] = max(meanmlp2);
22
23 best1 = rangemlp1(i1);
24 best2 = rangemlp2(i2);
25
26 if m1 > m2
27     neurons = best1;
28     fprintf('Best architecture has 1 layer and %d neurons. K-fold result: ...
           %.4f\n', neurons, m1) ;
29 else
30     neurons = best2;
31     fprintf('Best architecture has 2 layers and %d neurons. K-fold result: ...
           %.4f\n', neurons, m2) ;
32 end
33
34 hFig = figure(1);

```

```
35 set(hFig, 'Position', [450 200 1000 600])
36 axis tight
37
38 plot(rangemlp1, meanmlp1, 'b—s', 'MarkerFaceColor','b');
39 hold on
40 plot(rangemlp2, meanmlp2, 'r—d', 'MarkerFaceColor','r');
41
42 set(gca, 'FontSize', 16)
43 xlabel('Número de neurônios nas camadas ocultas', 'FontSize', 16);
44 ylabel('Taxa de acurácia no K-fold', 'FontSize', 16);
45 legend({'MLP com 1 camada oculta', 'MLP com 2 camadas ocultas'}, ...
        'Location','southeast', 'FontSize', 14);
46 ylim([0 1])
47
48 saveTightFigure(hFig, 'figuras/mlp-graph.pdf')
```

rbfplot.m

```

1 rbfdata1 = csvread('output/rbfdata-1S.csv');
2 rbfdata2 = csvread('output/rbfdata-2S.csv');
3
4 rangerbf1 = rbfdata1(1,:);
5 minrbf1   = rbfdata1(2,:);
6 maxrbf1   = rbfdata1(3,:);
7 meanrbf1  = rbfdata1(4,:);
8 mintotalrbf1 = rbfdata1(5,:);
9 maxtotalrbf1 = rbfdata1(6,:);
10 meantotalrbf1 = rbfdata1(7,:);
11
12 rangerbf2 = rbfdata2(1,:);
13 minrbf2   = rbfdata2(2,:);
14 maxrbf2   = rbfdata2(3,:);
15 meanrbf2  = rbfdata2(4,:);
16 mintotalrbf2 = rbfdata2(5,:);
17 maxtotalrbf2 = rbfdata2(6,:);
18 meantotalrbf2 = rbfdata2(7,:);
19
20 [m1 i1] = max(meanrbf1);
21 [m2 i2] = max(meanrbf2);
22
23 best1 = rangerbf1(i1);
24 best2 = rangerbf2(i2);
25
26 if m1 > m2
27     neurons = best1;
28     fprintf('Best architecture has %d neurons and spread = 1.0. K-fold result: ...
           %.4f\n', neurons, m1);
29 else
30     neurons = best2;
31     fprintf('Best architecture has %d neurons and spread = 2.0. K-fold result: ...
           %.4f\n', neurons, m2);
32 end
33
34 hFig = figure(1);

```

```
35 set(hFig, 'Position', [450 200 1000 600])
36 axis tight
37
38 plot(rangerbf1, meanrbf1, 'b—s', 'MarkerFaceColor','b');
39 hold on
40 plot(rangerbf2, meanrbf2, 'r—d', 'MarkerFaceColor','r');
41
42 set(gca, 'FontSize', 16)
43 xlabel('Número de neurônios nas camadas ocultas', 'FontSize', 16);
44 ylabel('Taxa de acurácia no K-fold', 'FontSize', 16);
45 legend({'RBF com raio receptivo r = 1', 'RBF com raio receptivo r = 2'}, ...
        'Location','southeast', 'FontSize', 16)
46 ylim([0 1])
47
48 saveTightFigure(hFig, 'figuras/rbf-graph.pdf')
```

mlprbfplot.m

```

1 mlpdata1 = csvread('output/mlpdata-1L.csv');
2 rbfdata2 = csvread('output/rbfdata-2S.csv');
3
4 rangemlp1 = mlpdata1(1,:);
5 minmlp1   = mlpdata1(2,:);
6 maxmlp1   = mlpdata1(3,:);
7 meanmlp1  = mlpdata1(4,:);
8 mintotalmlp1 = mlpdata1(5,:);
9 maxtotalmlp1 = mlpdata1(6,:);
10 meantotalmlp1 = mlpdata1(7,:);
11
12 rangerbf2 = rbfdata2(1,:);
13 minrbf2   = rbfdata2(2,:);
14 maxrbf2   = rbfdata2(3,:);
15 meanrbf2  = rbfdata2(4,:);
16 mintotalrbf2 = rbfdata2(5,:);
17 maxtotalrbf2 = rbfdata2(6,:);
18 meantotalrbf2 = rbfdata2(7,:);
19
20 hFig = figure(1);
21 set(hFig, 'Position', [450 200 1000 600])
22 axis tight
23
24 plot(rangemlp1, meanmlp1, 'b—s', 'MarkerFaceColor','b');
25 hold on
26 plot(rangerbf2, meanrbf2, 'r—d', 'MarkerFaceColor','r');
27
28 set(gca, 'FontSize', 16)
29 xlabel('Número de neurônios nas camadas ocultas', 'FontSize', 16);
30 ylabel('Taxa de acurácia no K-fold', 'FontSize', 16);
31 legend({'MLP com 1 camada oculta', 'RBF com raio receptivo r = 2'}, ...
        'Location','southeast', 'FontSize', 16)
32 ylim([0 1])
33
34 saveTightFigure(hFig, 'figuras/mlp-rbf-graph.pdf')

```