

EMPREGO DE REDE NEURAL MULTICAMADAS NO RECONHECIMENTO AUTOMÁTICO DE SILHUETAS INFRAVERMELHAS DE NAVIOS

Use of multilayer neural network in automatic recognition of infrared silhouettes of ships

Jorge Amaral Alves¹ , Edgard Braz Alves² 

Resumo: Com o desenvolvimento tecnológico das câmeras de imagem infravermelha presentes em sistemas eletro-ópticos de aeronaves modernas, foram desenvolvidos algoritmos de reconhecimento automático de alvos baseados nas assinaturas infravermelhas de navios. Tais algoritmos auxiliam os pilotos dessas aeronaves nas operações de identificação de contatos em operações navais. Este trabalho apresenta uma nova abordagem para um classificador automático de imagens infravermelhas obtidas a partir de modelos tridimensionais (3-D) de 5 (cinco) classes de navios criados em MATLAB. A abordagem proposta emprega um classificador baseado em uma Rede Neural Multicamadas (*Multilayer Perceptron* — MLP) cujos parâmetros foram otimizados empregando o aplicativo *Classification Learner* (CL) do MATLAB. Os resultados demonstraram que tal arquitetura apresenta melhor desempenho considerando trabalhos similares anteriores.

Palavras-chave: Rede neural artificial. Silhuetas de navios. Reconhecimento automático de alvos. Processamento de imagens infravermelhas.

Abstract: With the technological development of infrared imaging cameras present in electro-optical systems on modern aircraft, automatic target recognition algorithms were developed based on the infrared signatures of ships. Such algorithms assist the pilots of these aircrafts in contact identification operations in naval operations. This work presents a new approach for an automatic classifier of infrared images obtained from three-dimensional (3-D) models of 5 (five) classes of ships created in MATLAB. This proposed approach employs a classifier based on a Multilayer Perceptron (MLP), whose parameters were optimized using the MATLAB Classification Learner (CL) application. The results demonstrate that this architecture presents better performance considering previous similar works.

Keywords: Artificial neural network. Ships Silhouette. Automatic target recognition. Infrared image processing.

1. Coordenador de Eletrônica do Centro Técnico-Científico, Escola Naval, Rio de Janeiro, RJ - Brasil. E-mail: jorge.amaral@marinha.mil.br

2. Mestrando em Ciência da Computação, Instituto Militar de Engenharia, Rio de Janeiro, RJ - Brasil. E-mail: braz.edgard@ime.br

1. INTRODUÇÃO

O cenário atual de guerra busca detectar o inimigo sem ser detectado, bem como atuar sobre o inimigo sem que este possa reagir. Nesse contexto, constata-se que uma maneira de ampliar a capacidade de combate e de vigilância de meios navais e aeronavais consiste em dotar as câmeras infravermelhas aero embarcadas de classificadores automáticos eficazes. Conclui-se que para a manutenção desses preceitos é fundamental dotar tais classificadores de arquiteturas que forneçam altas taxas de eficácia, como, por exemplo, a do classificador apresentado neste trabalho.

A evolução tecnológica das câmeras de imagem infravermelha de aeronaves empregadas em operações navais torna essencial dotar tais sistemas de algoritmos de reconhecimento automático de alvos eficazes. Especificamente, o emprego desses algoritmos de classificação automática de imagens infravermelhas de navios permite maximizar a distância de identificação de um alvo, em operações navais, sem comprometer a segurança da aeronave e de sua tripulação.

Diversas abordagens envolvendo o processo de classificação de navios usando imagens infravermelhas e/ou coloridas podem ser identificadas na literatura. Os classificadores desenvolvidos nessas abordagens possuem aplicações diversas, tais como: vigilância portuária; segurança antipirataria; Sistemas de Guiagem (*seekers*) infravermelho de mísseis antinavio; e identificação automática de navios por aeronave.

Em Alves, Herman e Rowe (2004), foi desenvolvido um classificador baseado em Rede Neural Multicamadas (*Multilayer Perceptron* — MLP), de uma única camada escondida, para reconhecimento de imagens infravermelhas de navios para emprego na identificação de contatos obtidos por câmera infravermelha instalada em helicóptero. Nesse classificador foram consideradas 5 (cinco) classes de navios. Tais classes de navios foram representadas por modelos tridimensionais (3-D) criados em MATLAB a fim de gerar o conjunto de imagens utilizado. Esse conjunto de imagens foi composto por 46 ângulos de elevação e por 180 ângulos de azimute.

Posteriormente, em Li e Wang (2008), foi apresentada uma nova abordagem considerando um subconjunto de imagens de Alves, Herman e Rowe (2004), e substituindo o classificador original por uma Máquina de Vetores de Suporte (*Support Vector Machine* — SVM), porém empregou apenas duas classes de navios.

Em Kumlu (2012), foi desenvolvido um classificador baseado em SVM empregando imagens coloridas e obtidas a partir de modelos 3-D criados com o aplicativo *Google Sketch Up* para aplicação em câmeras de segurança de portos, sendo testado em apenas um valor de elevação, ou seja, na linha do horizonte ou zero grau.

Em Kechagias-Stamatis, Aouf e Nam (2017), foi desenvolvido um classificador baseado em uma Rede Neural Multimodal empregando imagens infravermelhas obtidas a partir de modelos 3-D criados com um programa de *Computer Aided Design* (CAD) para aplicação em *Seeker* infravermelho de míssil antinavio.

Recentemente, em Westlake (2021), foi apresentado um classificador robusto baseado em uma Rede Neural Convolutiva (*Convolutional Neural Net* — CNN) empregando o mesmo conjunto de imagens utilizado em Kechagias-Stamatis, Aouf e Nam (2017). Entretanto, seu classificador somente foi testado em 03 (três) ângulos de elevação (0°, -10° e -20°), tendo em vista que seu classificador foi desenvolvido para emprego em *Seeker* infravermelho de míssil antinavio.

Este trabalho propõe uma nova arquitetura de classificador baseada em uma MLP com três camadas ocultas e cujos hiperparâmetros foram otimizados de forma automática, a fim de se obter a melhor acurácia.

Os resultados alcançados neste trabalho evidenciam uma melhor capacidade de generalização dessa nova arquitetura de MLP quando comparados com os resultados obtidos em Alves, Herman e Rowe (2004), os quais foram utilizados como *baseline* deste trabalho.

Neste estudo, foi considerada, como uma possível aplicação desse novo classificador automático, a sua instalação em câmeras infravermelhas (*Forward Looking Infrared Cameras* — FLIR) de helicópteros como ferramenta de auxílio a decisão durante operações noturnas para identificação de navios, conforme ilustrado na Figura 1.

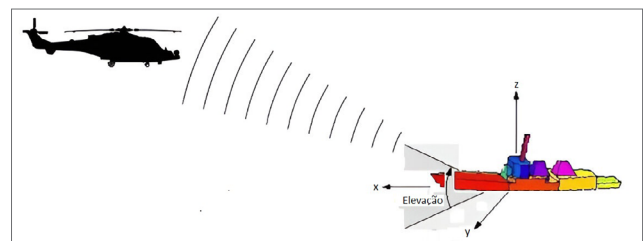


Figura 1. Emprego do classificador instalado em FLIR para identificação de navios.

Nesse contexto, foi empregado o mesmo conjunto de imagens de entrada (*dataset*) utilizado por Alves, Herman e Rowe (2004), o qual foi disponibilizado pelos mesmos para este trabalho. Assim sendo, foram empregadas todas as 05 (cinco) classes de navios, bem como os mesmos ângulos de azimute (total de 180 graus) e de elevação (total de 46 graus). Optou-se por empregar, também, a mesma ferramenta matemática de pré-processamento empregada por Alves, Herman e Rowe (2004), os momentos invariantes de Hu (1962).

Adicionalmente, neste trabalho foi empregado o aplicativo (*application — app*) *Classification Learner* (CL) do MATLAB (2023). Essa ferramenta permitiu identificar o melhor tipo de classificador a ser utilizado, bem como permitiu encontrar, de forma automática, a respectiva configuração otimizada desse tipo de classificador, ou seja, aquela configuração cujos hiperparâmetros foram otimizados de forma automaticamente, a fim de se obter a melhor acurácia total.

O CL permitiu avaliar, dentre outros, os seguintes tipos de classificadores: SVM; Árvores de Decisão; *Ensemble*; *K-Nearest Neighbors* (KNN); e MLP. Após o teste desses tipos de classificadores, identificou-se a MLP como aquela que apresentou o melhor resultado e, portanto, foi utilizada neste trabalho. A arquitetura final da MLP empregada neste trabalho atingiu 93,6% de acurácia, superando os resultados obtidos em trabalhos similares anteriores, como os de Alves, Herman e Rowe (2004) e de Li e Wang (2008).

Inicialmente, são apresentados os principais conceitos envolvidos na etapa de pré-processamento, bem como as vantagens em se empregar os momentos Hu (HU, 1962) como vetor descritor de forma para as silhuetas dos navios empregadas. Em seguida, apresenta-se o processo de escolha do tipo de classificador e de sua respectiva arquitetura a ser utilizada neste trabalho. Posteriormente, faz-se a abordagem de uma possível arquitetura para o classificador baseado na MLP ótima e apresenta-se a eficácia obtida com esse classificador. Por fim, são apresentados os resultados obtidos neste trabalho, bem como é feita uma comparação com resultados de trabalhos anteriores e apresentam-se as conclusões.

2. PRÉ-PROCESSAMENTO

2.1. CONJUNTO DE IMAGENS DE ENTRADA

A premissa inicial deste trabalho foi adotar o mesmo *dataset* de imagens de entrada empregado em Alves, Herman e

Rowe (2004), a fim de permitir uma comparação direta dos resultados obtidos, principalmente em relação à acurácia final obtida pela nova arquitetura do classificador proposto neste trabalho. Adicionalmente, foi empregado o mesmo método descritor de forma adotado em Alves, Herman e Rowe (2004) na etapa de pré-processamento, ou seja, foram empregados os conhecidos momentos Hu (HU, 1962) como descritores de forma das imagens infravermelhas simuladas. Essas estratégias foram adotadas de modo a oferecer melhores condições de comparação com o novo classificador proposto neste trabalho.

O conjunto de imagens de entrada disponibilizado por Alves, Herman e Rowe (2004) é composto por silhuetas de navios obtidas a partir das projeções ortogonais de modelos 3-D criados em MATLAB (2023). Neste trabalho, foram empregados todos os cinco modelos 3-D criados, representando as seguintes classes de navios: Navio Aeródromo; Contratorpedeiro; Fragata; Navio de Pesquisa; e Navio Mercante.

A modelagem 3-D em MATLAB busca representar um objeto pela conexão de faces, ou polígonos, criando, assim, modelos do tipo *wireframe*.

As classes de navios foram modeladas com base nas especificações dos seguintes navios: imagem do Navio Aeródromo *Carl Vinson* (Classe *Nimitz*), escala 1:1,800, obtida no *Jane's Fighting Ships* (SHARPE, 2000); imagem do Contratorpedeiro *Oscar Austin* (Classe *Arleigh Burke*), escala 1:1,500, obtida no *Jane's Fighting Ships* (SHARPE, 2000); imagem da Fragata *Rentz Hgate* (Classe *Oliver Hazard Perry*), escala 1:1,200, obtida no *Jane's Fighting Ships* (SHARPE, 2000); especificações do Manual do Navio de Pesquisa R/V *Pelican* (ALVES; HERMAN; ROWE, 2004); e Navio Mercante baseado nas especificações gerais do petroleiro norte americano *Sea Isle City* (ALVES; HERMAN; ROWE, 2004).

Para simular a visualização de cada classe de navio por uma aeronave, são extraídas as silhuetas dos navios nos diferentes ângulos de azimute e elevação para cada modelo.

A origem de coordenadas de azimute e de elevação, para cada modelo, encontra-se, aproximadamente, nos centros de gravidade (CG) desses modelos.

A Figura 2 mostra um exemplo de visualização modelada para cada classe de navio. Os modelos são visualizados pelo ângulo de $-37,5^\circ$ em azimute e pelo ângulo de 30° em elevação.

A partir de cada ângulo de visada dos navios modelados, variando-se de 1 em 1 grau, em azimute e em elevação, são

obtidas as respectivas silhuetas. O objetivo é gerar silhuetas similares às que seriam obtidas por uma câmera infravermelha aero embarcada. Esse procedimento gerou 41.400 imagens binárias que foram usadas como conjunto de dados de teste (*Testing Set*) e 240 imagens como o conjunto de dados de treinamento (*Training Set*), mantendo assim a mesma metodologia de treinamento adotada por Alves, Herman e Rowe (2004).

O *Training Set* foi composto por 48 imagens de cada classe de navio. Cada conjunto de 48 imagens foi composto por diferentes ângulos de visada para cada modelo. Os ângulos de visada foram escolhidos variando-se de -90° a 89° , com incrementos de 15° , obtendo-se um total de 12 diferentes ângulos de azimute. Da mesma forma, os ângulos de elevação foram escolhidos variando-se de 0° a 45° , com incrementos de 15° , obtendo-se um total de 4 diferentes ângulos de elevação.

As 41.400 imagens empregadas no *Testing Set* foram obtidas a partir de 5 conjuntos de 8.280 imagens de cada classe de navio. Cada conjunto de 8.280 imagens foi obtido variando-se em azimute de -90° a 89° , com incrementos de 1° , e variando-se em elevação de 0° a 45° , com incrementos de 1° .

2.2. MOMENTOS HU

Cada imagem do *Testing Set* e do *Training Set* foi convertida em um vetor de 12 valores correspondentes a 06 momentos para cada silhueta e 06 momentos para a respectiva borda de cada silhueta, conforme detalhado em Alves, Herman e Rowe (2004).

Foram utilizados os momentos invariantes de Hu (1962) para o pré-processamento das imagens. Esse tipo de pré-processamento tem sido amplamente utilizado no reconhecimento de padrões de conjuntos de imagens (FLUSSER; ZITOVA; SUK, 2009).

O Momento Hu foi escolhido como função de pré-processamento tendo em vista suas características de invariância em rotação, reflexão, escala e translação. Essas características foram consideradas essenciais para este trabalho. Conforme detalhado por Alves, Herman e Rowe (2004), foram calculados os 6 Momentos Centrais Normalizados tanto para a silhueta quanto para a respectiva borda, para cada imagem de cada modelo. Em seguida, os 6 momentos calculados para a silhueta são concatenados com os 06 momentos calculados para a respectiva borda. Desta forma, obtém-se o vetor de 12 valores correspondentes aos dados de entrada do classificador.

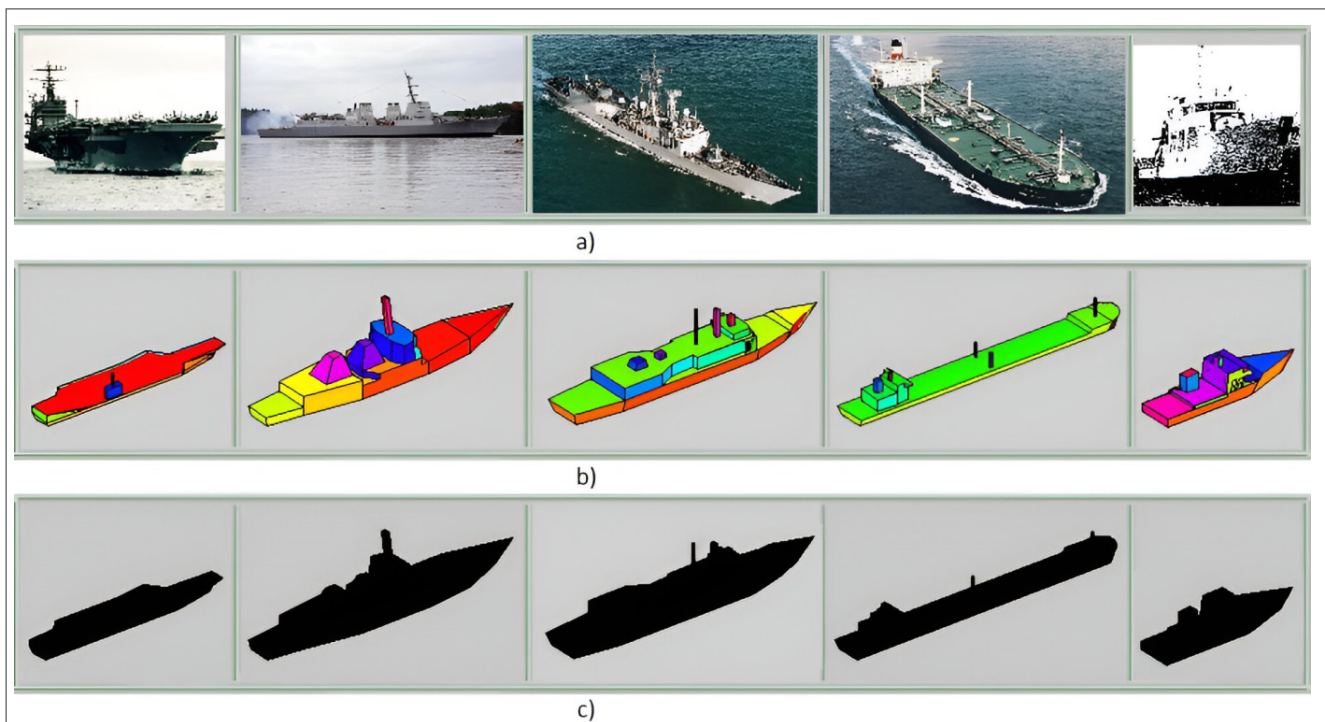


Figura 2. Modelos gerados: (A) Imagens reais; (B) Modelos 3-D; (C) Silhuetas.

Para a extração da borda da silhueta de cada imagem binária foi escolhido o método *Canny* (CANNY, 1986), que resultou em bordas adequadas a este trabalho. A Figura 3 ilustra a sequência necessária para se obter essas bordas.

3. CLASSIFICADOR

Para identificar o tipo de classificador mais adequado e, portanto, aquele que fornecesse a melhor acurácia, empregou-se um aplicativo do MATLAB denominado CL. O emprego do CL permitiu testar diferentes tipos de classificadores de forma automática e, assim, identificar aquele mais adequado ao presente trabalho.

Dentre os diferentes tipos de classificadores testados, destacam-se: Árvores de Decisão; SVM; *Ensemble*; KNN; e MLP.

Dentre os diversos tipos de classificadores testados, aquele que obteve o melhor desempenho foi a MLP. Esse resultado foi baseado nos valores obtidos nas diferentes Matrizes de Confusão apresentadas pelo CL. A MLP que apresentou os melhores resultados possui 3 camadas escondidas: 296 neurônios na primeira camada; 36 neurônios na segunda camada; e 141 neurônios na terceira camada. Essa arquitetura é representada na Figura 4.

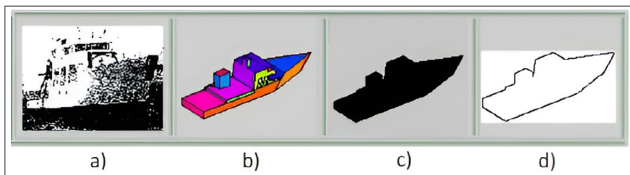


Figura 3. Borda do R/V Pelican: (A) Imagem real; (B) Modelo; (C) Silhueta; (D) Borda.

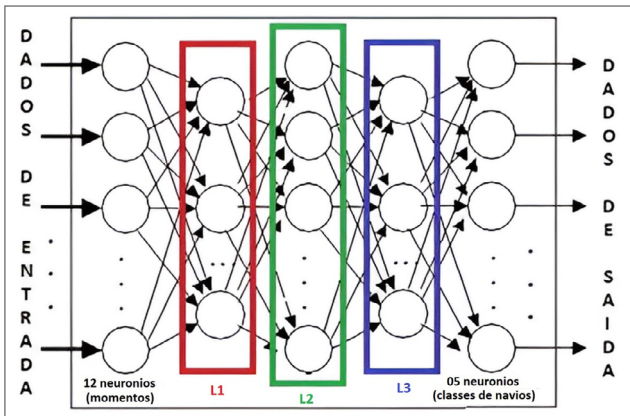


Figura 4. Arquitetura da MLP: L1 = 296 neurônios; L2 = 36 neurônios; e L3 = 141 neurônios.

4. METODOLOGIA PROPOSTA

A metodologia proposta neste trabalho foi desenvolvida em ambiente computacional MATLAB e pode ser visualizada através da Figura 5.

O principal objetivo deste trabalho foi identificar o tipo de classificador mais adequado, bem como sua respectiva arquitetura. A escolha desse classificador foi feita adotando-se como critério principal a melhor taxa de acurácia obtida no reconhecimento de imagens infravermelhas de navios considerando o *dataset* idêntico ao empregado em Alves, Herman e Rowe (2004). Para atingir esse objetivo, diversas arquiteturas de classificadores foram testadas por meio do emprego da ferramenta CL (MATLAB, 2023). Esse aplicativo possibilitou testar uma ampla variedade de algoritmos de forma automática. Esse procedimento forneceu condições adequadas para a escolha do tipo de classificador ideal para este trabalho. O emprego do CL também permitiu acompanhar as etapas intermediárias associadas ao treinamento e teste dos diversos classificadores avaliados. As ferramentas do CL empregadas no acompanhamento do processo de treinamento foram: Matriz de Confusão; Curva de Característica de Operação do Receptor (ROC) e, principalmente, o valor de acurácia de cada arquitetura testada.

5. RESULTADOS

A Tabela 1 apresenta os valores de acurácia das 5 melhores arquiteturas obtidas pelo CL durante o processo de avaliação das diversas arquiteturas de classificadores testados.

Com base nos resultados apresentados na Tabela 1, optou-se por empregar a MLP como classificador ideal para este trabalho tendo em vista seu melhor desempenho, especificamente, por obter 93,6% de acurácia, considerando as 41.400 imagens *do testing set*.

A Figura 6 ilustra o processo iterativo de otimização dos hiperparâmetros em busca da MLP ótima. Nota-se que

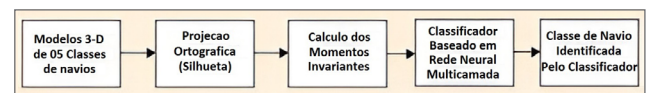


Figura 5. Diagrama de blocos da metodologia utilizada.

com 16 interações foi possível atingir a melhor configuração que corresponde ao modelo de MLP adotado neste trabalho.

A Tabela 2 apresenta as características principais da arquitetura da MLP otimizada.

A Matriz de Confusão do classificador proposto neste trabalho, empregando a MLP otimizada obtida, é apresentada na

Tabela 1. Acurácias obtidas pelo Classification Learner.

Posição	Algoritmo	Acurácia (%)
1º	MLP	93,6
2º	SVM	92,4
3º	Ensemble	90,9
4º	KNN	87,6
5º	Árvores de Decisão	78,7

MLP: rede neural multicamadas; SVM: máquina de vetores de suporte; KNN: *K-Nearest Neighbors*.

Tabela 3. É possível observar nessa tabela o valor de acurácia total do classificador proposto, bem como as taxas de verdadeiros positivos (*true positive rate* — TPR) e os Valores Previstos Positivos (PPV) para cada modelo testado. Pode-se observar

Tabela 2. Parâmetros da rede neural multicamadas otimizada.

Parâmetros	Valor
Neurônios de Entrada	12
Neurônios da 1ª Camada Escondida	296
Neurônios da 2ª Camada Escondida	36
Neurônios da 3ª Camada Escondida	141
Neurônios de Saída	6
Duração do Treinamento	283,3 s
Acurácia da Validação	92,3%
Acurácia do Teste	93,6%

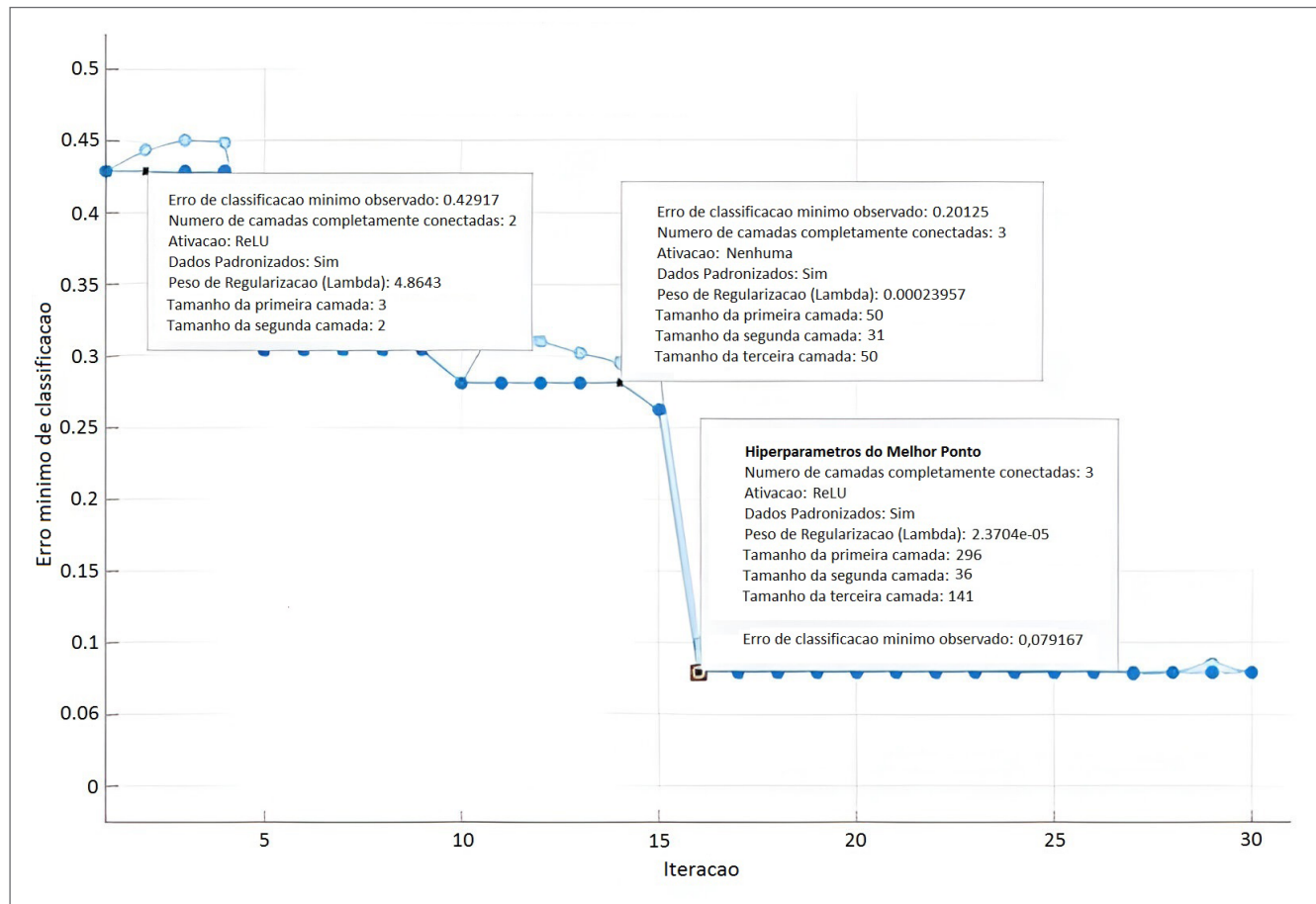


Figura 6. Evolução do processo iterativo em busca da rede neural multicamadas ótima.

também que a Classe Contratorpedeiro foi a que obteve os melhores resultados, ou seja, 97,7% de TPR e 97,0% de PPV. Em contrapartida, nota-se que a Classe Fragata foi a que obteve os piores resultados, ou seja, 89,3% de TPR e 89,9% de PPV.

A Tabela 4 apresenta os valores obtidos por Alves, Herman e Rowe (2004).

Ao compararmos as Tabelas 3 e 4, pode-se concluir que a arquitetura proposta neste trabalho permitiu melhorar a precisão e a abrangência de Alves, Herman e Rowe (2004) em todas as 5 classes de navios propostas, além de obter melhor acurácia total.

6. CONCLUSÕES

Neste trabalho foi possível avaliar diversas arquiteturas e tipos de classificadores com potencial para serem empregados na classificação automática de imagens infravermelhas de navio simuladas.

O classificador proposto neste trabalho permitiu alcançar uma acurácia total de 93,6% empregando uma arquitetura baseada em uma MLP otimizada e, dessa forma, melhorando o resultado de 87,3% apresentado em Alves, Herman e Rowe (2004). Conclui-se, portanto, que o presente trabalho oferece como contribuição para a comunidade científica um classificador automático de imagens infravermelhas de navios simuladas de melhor desempenho quando comparado a trabalhos similares anteriores.

Também foi demonstrada a importância do emprego de uma ferramenta capaz de otimizar os hiperparâmetros de diversos classificadores de forma automática. Tal ferramenta permitiu o teste de diversos tipos e diversas arquiteturas de classificadores possibilitando a escolha daquele que apresentou a melhor acurácia.

Como sugestão de trabalhos futuros, pode-se citar o teste do classificador proposto em imagens infravermelhas reais e o aumento do número de classes de navios a serem modeladas e testadas.

Tabela 3. Matriz de Confusão do classificador proposto neste trabalho.

Classes de Navios de Entrada	Classes de Navios Estimadas na Saída					TPR (%)
	NAe	CT	F	M	R/V	
Navio Aeródromo (NAe)	7.515	28	341	310	86	90,8
Contratorpedeiro (CT)	34	8.098	87	0	70	97,7
Fragata (F)	315	222	7.398	1	344	89,3
Mercante (M)	318	0	206	7.748	8	93,6
Navio de Pesquisa (R/V)	41	1	198	56	7.984	96,6
PPV (%)	91,4	97,0	89,9	95,5	94,0	93,6

TPR: taxa de verdadeiros positivos (*true positive rate*); PPV: valor preditivo positivo (*positive predictive value*).

Tabela 4. Matriz de Confusão obtida por Alves, Herman e Rowe (2004).

Classes de Navios de Entrada	Classes de Navios Estimadas na Saída					TPR (%)
	NAe	CT	F	M	R/V	
Navio Aeródromo (NAe)	6.711	201	425	550	397	81,1
Contratorpedeiro (CT)	318	7.301	397	7	257	88,2
Fragata (F)	345	788	6.809	121	217	82,2
Mercante (M)	297	188	291	7.455	49	90,0
Navio de Pesquisa (R/V)	67	146	177	17	7.873	95,1
PPV (%)	86,7	84,7	84,1	91,5	89,6	87,3

TPR: taxa de verdadeiros positivos (*true positive rate*); PPV: valor preditivo positivo (*positive predictive value*).

REFERÊNCIAS

ALVES, J. A.; HERMAN, J.; ROWE, N. C. *Robust recognition of ship types from an infrared silhouette*. Technical report. Monterey: Naval Postgraduate School, 2004.

CANNY, J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, n. 6, p. 679-698, 1986. <https://doi.org/10.1109/TPAMI.1986.4767851>

FLUSSER, J.; ZITOVA, B.; SUK, T. *Moments and moment invariants in pattern recognition*. Chichester: John Wiley & Sons, 2009.

HU, M. K. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, v. 8, n. 2, p. 179-187, 1962. <https://doi.org/10.1109/TIT.1962.1057692>

KECHAGIAS-STAMATIS, O.; AOUF, N.; NAM, D. Multi-modal automatic target recognition for anti-ship missiles with imaging infrared capabilities. *In: SENSOR SIGNAL PROCESSING FOR DEFENCE CONFERENCE*, 2017. *Anais [...]*. p. 1-5, 2017. <https://doi.org/10.1109/SSPD.2017.8233244>

KUMLU, D. *Autonomous ship recognition from color images*. Tese (Doutorado) - University of Southern California, California, 2012.

LI, H.; WANG, X. Automatic recognition of ship types from infrared images using support vector machines. *In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER SCIENCE AND SOFTWARE ENGINEERING*, 6., 2008. *Anais [...]*. 2008. p. 483-486. <https://doi.org/10.1109/CSSE.2008.1647>

MATLAB. Massachusetts: The MathWorks Inc., 2023.

SHARPE, R. *Jane's Fighting Ships*. Alexandria: Jane's Information Group, 2000.

WESTLAKE, S. T. *The Application of Deep Learning Algorithms to Longwave Infrared Missile Seekers*. Tese (Doutorado) - Cranfield University, Cranfield, 2021.