

Revista Brasileira de Computação Aplicada, November, 2020

DOI: 10.5335/rbca.v12i3.10991 Vol. 12, Nº 3, pp. 61-69

Homepage: seer.upf.br/index.php/rbca/index

### ARTIGO ORIGINAL

# PINFI - Ferramenta para classificação de imagens com redes neurais artificiais e lógica fuzzy

# PINFI - Tool for image classification with artificial neural networks and fuzzy logic

Ivan Luís Süptitz<sup>[0],1</sup>, Rejane Frozza<sup>[0],1</sup>, Rolf Fredi Molz<sup>[0],1</sup>

<sup>1</sup>Universidade de Santa Cruz do Sul (UNISC) - Programa de Pós-Graduação em Sistemas e Processos Industriais \*ivansuptitz@unisc.br; frozza@unisc.br; rolf@unisc.br

Recebido: 2020-05-09. Revisado: 2020-07-13. Aceito: 2020-07-20.

### Abstract

This article presents a modeling tool for image classification using artificial neural networks (RNA) and fuzzy logic, developed with focus on the industrial scenario. The design of this tool started with a previous research of RNA and neuro-fuzzy modeling tools and a survey-type research in which it was identified that the industries of a certain region do not know these technologies or do not know how to apply it in their processes. From this, the tool called Neuro-Fuzzy Image Processor for Industry (PINFI) was developed based on the demands of the survey and integrating open source software libraries researched in the bibliographic survey. This tool allows the modeling of projects in five stages: Acquisition of images; pre-processing; fuzzification block; RNA modeling; and the presentation of the results (output). The PINFI was validated in terms of ergonomics and software usability, meeting 91% of the requirements evaluated. As for the accuracy of the classifier, it reached recognition rates of 93% to 100% in the bases evaluated. These bases included images of tobacco leaves, vehicle wheels and human faces, which demonstrates that PINFI has the potential to be used in industrial demands and also in diverse applications.

Keywords: Image processing; modeling tool; neuro-fuzzy

### Resumo

Este artigo apresenta uma ferramenta de modelagem para classificação de imagens utilizando redes neurais artificiais (RNA) e lógica difusa (fuzzy), desenvolvida com foco no cenário industrial. O projeto desta ferramenta iniciou com um levantamento prévio de ferramentas de modelagem de RNA e neuro-fuzzy e de uma pesquisa do tipo survey na qual se identificou que as indústrias de uma determinada região desconhecem ou não sabem como aplicar estas tecnologias em seus processos. A partir disso, foi desenvolvida a ferramenta denominada Processador de Imagens Neuro-Fuzzy para a Indústria (PINFI) com base nas demandas do survey e integrando bibliotecas de software de código aberto pesquisadas no levantamento bibliográfico. Esta ferramenta permite a modelagem de projetos em cinco etapas: Aquisição das imagens; pré-processamento; bloco de fuzzificação; modelagem da RNA; e a apresentação dos resultados (saída). O PINFI foi validado em termos de ergonomia e usabilidade de software atendendo a 91% dos requisitos avaliados. Quanto à acurácia do classificador, alcançou taxas de reconhecimento de 93% a 100% nas bases avaliadas. Estas bases compreenderam, imagens de folhas de tabaco, rodas de veículos e faces humanas, o que demonstra que o PINFI tem potencial para ser utilizado em demandas industriais e, também, em aplicações diversificadas.

Palavras-Chave: Ferramenta de modelagem; neuro-fuzzy; processamento de imagens

# 1 Introdução

A inteligência artificial (IA) é uma área de pesquisa da ciência da computação que tem como objetivo desenvolver métodos e sistemas computacionais com capacidade de resolver tarefas tipicamente humanas, e, por este motivo, consideradas inteligentes (Haykin, 1999). Assim sendo, para essa pesquisa foram utilizadas duas técnicas de IA: redes neurais artificiais (RNAs) e lógica fuzzy em conjunto com técnicas de processamento de imagens visando aplicação em automação e instrumentação em ambiente industrial ou agrícola. A combinação destas técnicas tem apresentado resultados promissores como visto em: Beluco et al. (2015); Ros and Pasero (2013); Taghadomi-Saberi et al. (2014); Geetha (2015); Bouzalmat et al. (2011); Silva et al. (2014); Soares and da Silva (2013); Yen et al. (2019); Sabanci et al. (2017); Gunes and Badem (2016); Tyryshkin and Belyaev (2016); Rao et al. (2018).

As RNAs têm sua arquitetura inspirada no modo como trabalha o cérebro humano, com o objetivo de tentar reproduzir suas habilidades como o reconhecimento de padrões. Consegue-se este efeito por meio de neurônios artificiais interconectados por uma rede de pesos variáveis que são ajustados por meio um algoritmo de aprendizado capaz de aproximar os padrões de entrada aos padrões de saída (Haykin, 1999).

A lógica fuzzy pode ser compreendida como uma extensão da lógica booleana, possibilitando operações com variáveis cujos valores não são absolutos, mas apresentam certo grau de pertinência aos conjuntos definidos. Um típico sistema baseado em lógica difusa é composto de uma etapa de fuzzificação, uma máquina de inferência e uma etapa de defuzzificação (Chi et al., 1996).

Sistemas fuzzy funcionam bem para modelar o conhecimento explícito fornecido por especialistas humanos, ao passo que redes neurais artificiais são apropriadas para desenvolver modelos baseados em conhecimentos implícitos (Figueiredo and Gomide, 1999). Neste trabalho, estas técnicas são combinadas por meio de uma abordagem descrita na literatura como redes neurais fuzzy, na qual o foco está na rede neural artificial e as regras do sistema difuso servem de apoio, a fim de melhorar o desempenho da RNA (Berenji and Khedkar, 1992).

A entrada de dados para a ferramenta desenvolvida, ocorre por meio de imagens. O processamento de imagens pode ser dividido em: processamento de baixo nível, envolvendo a aquisição das imagens e préprocessamentos como redução de ruído e compensações; nível intermediário, no qual são realizadas tarefas de segmentação e descrição; alto nível, envolvendo tarefas de cognição inteligente, como reconhecimento e interpretação de padrões (Gonzalez and Woods, 2000).

O artigo está estruturado nas seguintes seções: A seção 2 apresenta a metodologia aplicada no desenvolvimento do trabalho; a seção 3 aborda o projeto da ferramenta desenvolvida; a seção 4 relata a validação e os testes realizados; na seção 5 é descrita a avaliação de ergonomia e usabilidade da ferramenta e a seção 6 apresenta a conclusão da pesquisa.

O objetivo principal desta pesquisa foi desenvolver uma ferramenta de modelagem de aplicações com o uso combinado de RNA, lógica fuzzy e processamento de imagens para aplicação na indústria.

## 2 Metodologia

Esta pesquisa iniciou com um levantamento de ferramentas de modelagem de RNAs sendo avaliados 13 softwares que possibilitam modelagem de redes neurais artificiais. Destes, apenas 3 apresentam todos os requisitos avaliados, sendo, no entanto, pagas e voltadas para problemas específicos. Também há existência de componentes de software de código aberto em 11 das 13 ferramentas avaliadas. Estas APIs (Interfaces de Programação de Aplicações) podem ser utilizados na confecção de uma ferramenta de modelagem com os requisitos esperados.

A justificativa para este trabalho foi obtida a partir das constatações de um levantamento do tipo pesquisa de avaliação publicado em Suptitz et al. (2017). O trabalho teve por objetivo identificar o estado da arte referente à aplicação de IA (em especial RNA e lógica fuzzy) nas indústrias do Rio Grande do Sul, a fim de contribuir com a possibilidade de expansão do uso destas técnicas em diferentes contextos, bem como a expectativa dos usuários em relação a uma ferramenta com estas tecnologias. Este tipo de pesquisa foi aplicado para avaliar a utilização, por parte da indústria, de técnicas amplamente conhecidas no meio acadêmico.

Para se chegar a estes resultados, três questões foram propostas: i) Qual o nível de contato que as indústrias gaúchas têm com técnicas de IA, em especial com RNA e lógica fuzzy?; ii) Quais são os empecilhos para a utilização de IA nas indústrias gaúchas?; iii) Quais são os requisitos que uma ferramenta de modelagem de RNAs e lógica fuzzy deve ter para ser aplicada na indústria?

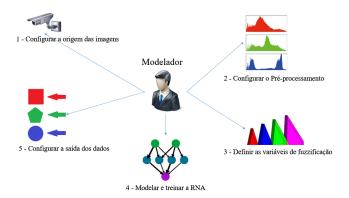
A maior representatividade das indústrias do RS é das Indústrias de Transformação, com 86

A partir desta pesquisa foi possível confirmar a afirmação de Bravo et al. (2014), que relata que o conhecimento público das técnicas de RNAs e lógica fuzzy é baixo. Ao mesmo tempo, foi possível verificar que os requisitos mais importantes em uma ferramenta que se proponha a oferecer estas tecnologias são o baixo custo e a facilidade de uso. Este dado sinaliza que a indústria aceitaria mais facilmente uma ferramenta de uso gratuito desenvolvida com atenção especial para questões de ergonomia e usabilidade de software.

Após a etapa de estudos e definição das necessidades para desenvolvimento de uma nova ferramenta, iniciou-se a etapa de implementação dos módulos de processamento de imagens, redes neurais artificiais e lógica fuzzy para compor a ferramenta PINFI. No desenvolvimento da ferramenta foram abordados aspectos de usabilidade e ergonomia para, então, realizar o processo de uso e validação do PINFI.

## 3 Projeto do software PINFI

Atendendo às necessidades levantadas na survey e levando em consideração o padrão de organização de sistemas semelhantes, foram definidos para o PINFI dois casos de uso: modelagem de projetos e execução de projetos. Na Fig. 1 está representada uma visão geral da modelagem de projetos.



**Figura 1:** Caso de uso: modelagem de um projeto PINFI.

Como se pode observar, o ator humano deste caso de uso é o modelador, que pode ser um especialista humano que pretende utilizar a ferramenta, ou um consultor externo com algum conhecimento do processo no qual a ferramenta será aplicada. Ele será responsável pela criação e manutenção dos projetos devendo configurar as 5 etapas. A seguir estão especificados os requisitos de cada uma destas etapas.

- · Aquisição de dados: Nesta primeira etapa, o modelador escolhe a origem das imagens que serão utilizadas para o treinamento e para a execução. O sistema disponibiliza as opções: TCP, UDP, diretório de arquivos, câmera ou ainda streaming (fluxo de mídia) JPEG ou MJPEG. TCP e UDP são protocolos de comunicação de dados comumente disponíveis em equipamentos industriais. Nesse caso, é preciso especificar o endereço IP e a porta do dispositivo. Em caso de diretório de arquivos, é preciso especificar a pasta onde estão os arquivos e um filtro de busca como, por exemplo, \*.jpg. Caso a origem escolhida seja uma câmera de vídeo, é preciso definir o dispositivo de captura e o número de frames por segundo que deve ser capturado. Nas opções de streaming de vídeo é preciso definir a URL (endereço web) da origem do vídeo e a frequência de captura.
- Pré-processamento de imagens: A segunda etapa da modelagem prepara as imagens e extrai as características necessárias para as etapas seguintes. Inicialmente, o sistema exibe uma das imagens provenientes da aquisição de dados para que seja possível acompanhar os tratamentos de forma visual. O modelador pode adicionar, remover e editar filtros. Os filtros podem ser adicionados em cascata. Cada filtro tem suas próprias configurações. Um ou mais

- dos filtros escolhidos terá opções de extrair os dados para as etapas seguintes. O modelador terá que selecionar quais dados desta etapa serão utilizados.
- Definição das variáveis de fuzzificação: A terceira etapa do projeto é opcional, se o modelador quiser utilizar somente a RNA, pode desativar o bloco fuzzy e não terá que configurá-lo. Caso contrário, o sistema lista todas as características provenientes da etapa anterior (que serão as variáveis fuzzy) com um nome e uma breve descrição. O modelador precisa especificar o universo e os termos linguísticos para cada variável. Para cada termo, é preciso escolher uma das funções de pertinência disponíveis e configurar os seus valores. À medida que o usuário define as configurações, o sistema exibe um gráfico com as funções de todos os termos.
- Modelar e treinar a RNA: Esta é a etapa mais importante do projeto PINFI. O modelador realiza testes com diversas opções de configuração para verificar qual melhor se aplica ao seu processo. As configurações são as seguintes: Algoritmo de treinamento, função de ativação, número de camadas e neurônios em cada camada, taxa de aprendizado, erro mínimo aceitável e ainda opções específicas de cada algoritmo. Também é necessário definir o conjunto de treinamento. Para isso, o sistema carrega a lista de imagens disponíveis a partir da etapa de aquisição e o modelador cria as classes1 e vincular as imagens nas suas respectivas classes. Em seguida, o modelador deve iniciar o treinamento da RNA, podendo interrompê-lo a qualquer momento. A cada iteração, o sistema informa o progresso do treinamento por meio do erro médio. Por fim, é possível realizar testes definindo um percentual do conjunto de treinamento para validação. O sistema reporta o percentual de acertos da RNA.
- Configurar a saída dos dados: A última etapa do projeto consiste em especificar como ocorre a saída do PINFI. O sistema disponibiliza as opções: TCP, UDP, diretório de arquivos ou saída em tela. As opções TCP e UDP enviam o resultado (nome da classe reconhecida) através do protocolo de rede, sendo necessário especificar a porta que irá atender ao serviço. A opção de diretório de arquivos cria uma pasta para cada classe reconhecida no diretório configurado e salva as imagens classificadas no interior da respectiva pasta. Por fim, a opção de exibir em tela lista as imagens classificadas em uma interface do próprio sistema, permitindo ao usuário acompanhar de forma visual o processo.

O segundo caso de uso da ferramenta PINFI é a execução do projeto configurado no caso de uso da modelagem do projeto. O ator humano neste caso é o operador do sistema. O PINFI recebe como entrada as imagens de uma parte do processo, realiza cada uma das etapas configuradas e apresenta a saída. Esta pode ser apresentada ao operador que vai tomar determinada decisão na linha de produção, ou pode ser conectada a um atuador ou outro software. De qualquer forma, a execução pode ser supervisionada pelo operador por meio da visualização dos logs gerados. As etapas são

executadas da seguinte forma:

- 1. Aquisição: As imagens são lidas pelo sistema a partir da origem configurada e organizadas em uma fila
- 2. Pré-processamento: Cada imagem adquirida recebe a aplicação dos filtros definidos. Os filtros apropriados realizam a extração das características da imagem.
- 3. Fuzzificação (caso definida): Para cada característica da etapa anterior, o sistema calcula o valor da pertinência de todos os seus termos linguísticos conforme as funções definidas.
- 4. Classificação: A RNA classifica cada uma das imagens a partir do treinamento previamente realizado. Se o bloco fuzzy estiver ativo, a origem do vetor de dados de entrada da RNA será a lista dos valores das pertinências de cada termo de todas as caraterísticas escolhidas. Se não estiver ativo, a entrada da RNA será alimentada diretamente com os valores das características extraídas na etapa 2, sendo realizada, inicialmente, a normalização dos valores.
- 5. Saída: As classes reconhecidas pela etapa 4 são organizadas em uma fila aguardando serem encaminhadas ao destino configurado no projeto, podendo ser exibida ao operador, ou ser enviado um comando para outro sistema.

A especificação destes casos de uso resultou na elaboração do diagrama de classes2 as quais são utilizadas para persistência dos dados e para execução dos processos.

A classe principal é a classe "Projeto" que reúne todos os dados e algoritmos para configurar e executar um projeto no PINFI. A partir desta classe principal são instanciadas as demais classes, uma para cada etapa da ferramenta: "AquisicaoDados", "PreProcessamento", "Fuzzy", "RedeNeural" e "Saida".

A classe responsável pela aquisição das imagens gerencia a fila de imagens obtidas, por este motivo ela faz referência a uma lista de objetos "ImagemPinfi". Esta classe por sua vez possui a estrutura de dados necessária para armazenar em memória a imagem adquirida e as suas características por meio da classe "Caracteristicalmagem".

A classe de preparação das imagens possui uma lista de filtros, representados pela referência à interface3 "IFilter". Todos os filtros aplicáveis no PINFI implementam esta interface que está definida no componente AForge.NET (Kirillov, 2016). Para a extração das características da imagem, é preciso que o filtro herde ainda a classe abstrata "Estatisticas\_Pinfi\_Base".

O bloco de lógica fuzzy é implementado por meio da classe "Fuzzy" que possui uma lista de objetos da classe "FuzzyVariavel" para representar as variáveis de fuzzificação e cada uma delas possui ainda uma lista de termos linguísticos por meio da classe "FuzzyTermo".

A classe responsável pela RNA faz referência a uma lista de camadas (classe "RNACamada") e ao conjunto de treinamento (classe "RNATreinamento"). Esta é a maior classe do projeto, reunindo referências a diversas classes e rotinas de componentes de terceiros, como: Neurônios, Funções de Ativação e Algoritmos de Treinamento.

Por fim, a classe denominada "Saida" possui o có-

digo necessário para comunicar o resultado da classificação conforme a opção de saída configurada.

Foram utilizados componentes de software de código aberto disponíveis, visando reduzir o tempo de desenvolvimento e partir de algoritmos já validados. Os levantamentos realizados, indicaram que o componente AForge.NET (Kirillov, 2016) é o mais adequado para desenvolvimento de modelos de RNAs. Os mesmos levantamentos também embasaram a decisão de desenvolver a ferramenta PINFI na linguagem de programação C# da plataforma Microsoft .NET Framework.

Verificou-se que o componente AForge.NET disponibiliza também métodos para manipulação de lógica fuzzy e que, apesar de não contemplarem todas as situações de modelagens, atendem aos requisitos do PINFI, dispensando a necessidade de outro componente de software.

Em relação à etapa de pré-processamento das imagens, foram utilizados filtros disponíveis nas bibliotecas AForge.NET (Kirillov, 2016) e OpenCV (Bradski, 2000). OpenCV é uma biblioteca de código aberto com mais de 2500 algoritmos voltada para visão computacional (Bradski, 2000).

A seguir, são apresentadas as principais interfaces do PINFI. A primeira delas permite configurar a etapa de pré-processamento das imagens e está apresentada na Fig. 2.

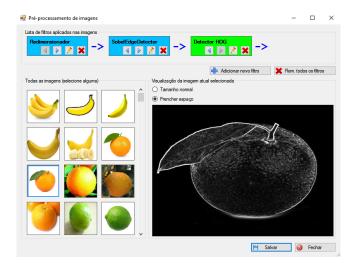


Figura 2: Tela de pré-processamento das imagens.

Pode-se observar que esta tela possui 3 regiões delimitadas por caixas rotuladas. Na parte superior da tela, existe uma região destinada a abrigar a lista de filtros que devem ser aplicados nas imagens. Na esquerda da tela, são exibidas todas as imagens obtidas a partir do módulo de aquisição. O usuário pode clicar sobre uma destas imagens e, no lado direito da tela, na caixa indicada pelo rótulo "Visualização da imagem atual selecionada", a mesma será exibida em tamanho maior com todos os filtros já aplicados.

Para adicionar um novo filtro na lista é preciso acionar o botão "Adicionar novo filtro", o que leva a uma outra tela permitindo escolher qualquer filtro disponibilizado pela ferramenta.

Dentre a lista de filtros disponibilizados na etapa de preparação, cinco deles possibilitam extrair características das imagens, as quais servem de entrada para as etapas seguintes (fuzzificação e RNA). Na Fig. 3 é apresentada a tela de modelagem da RNA.

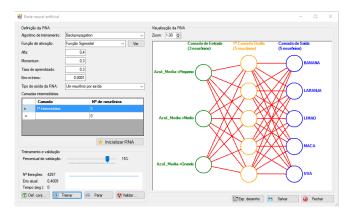


Figura 3: Tela de modelagem da RNA.

Como se pode observar, o lado esquerdo é destinado ao preenchimento das configurações e o lado direito apresenta as informações na forma de um grafo da RNA definida. As entradas, que podem ser provenientes do bloco fuzzy ou diretamente dos filtros de extração de caraterísticas, são representadas pela coluna de nodos de cor verde à esquerda do desenho juntamente com rótulos com seus nomes. As camadas ocultas são representadas por colunas de círculos amarelos e a camada de saída por uma coluna de círculos azuis juntamente com o respectivo nome da classe. As conexões são representadas por linhas vermelhas.

## 4 Testes e validação

Concluído o desenvolvimento do software, foram realizados os testes finais e a validação da ferramenta desenvolvida. Para isso, foram utilizadas as seguintes bases de dados:

 Rodas de veículos: Base com 2079 fotografias extraídas de uma gravação de vídeo em uma praça de pedágio. O objetivo é identificar frames do vídeo que correspondem a rodas e os que não são rodas, possibilitando a contagem do número de eixos do veículo. Desta forma, são duas classes distintas: "roda" e "nãoroda". • Folhas de tabaco: Esta base é composta de 104 fotografias de folhas de tabaco do tipo Virgínia, distribuídas nas seguintes classes de cores: cor "L" - cor de limão, admitindo-se manchas acastanhadas ocupando até 50% da superfície; cor "O" - folhas de cor laranja, admitindo-se manchas acastanhadas em até 50% da superfície; cor "R" - folhas nas quais a cor castanho claro e escuro ocupam mais de 50% da superfície. Esta base de dados foi coletada por estudantes da mesma instituição autores para um trabalho ainda não publicado

sobre análise multivariada. · Faces frontais de pessoas: É um conjunto de 4 bases de fotografias frontais de rostos em graus de dificuldade crescente (Spacek, 2007). O download das bases foi realizado a partir do site indicado na referência. Após a descompactação, alguns arquivos ficaram corrompidos, razão pela qual o número total de amostras indicado pelo autor (7900) diverge da quantidade utilizada nos testes (7856). A base "faces94" é composta de 3040 fotografias de rostos de 152 pessoas, com plano de fundo estático, sem variação de distância em relação à câmera e sem variações de luminosidade, sendo, portanto, o conjunto com menor grau de dificuldade. A base "faces95" possui 1440 fotografias de 72 pessoas com variação de distância em relação à câmera, tendo, desta forma, um grau de dificuldade maior do que a anterior. As bases faces96" e "grimace" possuem o maior grau de dificuldade de acordo com o autor. A primeira delas é composta por 3016 imagens de 151 pessoas com variação de luminosidade e distância em relação à câmera. O último conjunto é formado por 360 imagens de 18 pessoas fazendo caretas e movimentos (Spacek, 2007).

A escolha das bases de dados levou em consideração a diversidade dos objetos alvos da classificação, visando demonstrar que o PINFI é uma ferramenta de ampla aplicação. A escolha das bases de rodas e de folhas de fumo foi motivada pela disponibilidade do material pelo desenvolvimento de pesquisas paralelas a esta na mesma instituição. A pesquisa com folhas de tabaco tem ainda um forte apelo econômico para a região. A base de faces de pessoas foi escolhida em razão da dificuldade que reside em classificar rostos humanos, da relevância deste tipo de pesquisa para a área de visão computacional e pela qualidade metodológica presente na produção dos conjuntos de imagens.

Para cada uma das bases de dados, realizou-se uma exaustiva bateria de experimentos, cruzando combinações de filtros de pré-processamento, bloco fuzzy e configurações da RNA.

Os melhores resultados obtidos estão dispostos na Tabela 1. Cada linha resume uma validação. A primeira coluna indica o nome da base. A segunda coluna denota como ocorreu a extração das características. A terceira coluna, indica o algoritmo utilizado na RNA e a quarta indica o número de neurônios por camada (o primeiro número indica a quantidade de neurônios na camada de entrada, o último número indica os neurônios de saída e os intermediários, quando há, se referem às camadas ocultas). A última coluna indica o percentual de reconhecimento obtido em cada teste.

Em todos os testes utilizou-se o mesmo percentual de validação de 15% do conjunto de treinamento, sendo este um valor frequentemente utilizado na literatura. A função de ativação Sigmoidal com alfa = 2 apresentou o melhor resultado em todos os casos, razão pela qual esta informação foi omitida da tabela. Também não aparece na tabela a informação dos filtros de redimensionamento e ajuste de contraste para preparação das imagens. Em todos os casos, a taxa de aprendizado utilizada foi de 0,3 e o tipo de saída foi "Um neurônio por saída".

Tabela 1: Resultados dos testes de validação

Método	Algoritmo	Camadas	% de
de	da RNA	da RNA	Rec.
extração			
HOG	Perceptron	3780; 2	100,00%
Médias	Back-	8; 10; 5;	94,44%
de Histo-	propagation	1 3	
gramas			
Detector	Perceptron	3780;	99,78%
de rosto;		152	
HOG			
Detector	Perceptron	3780; 72	97,22%
de rosto;			
HOG			
Detector	Perceptron	3780; 151	93,75%
de rosto;			
HOG			
Detector	Perceptron	3780; 18	98,15%
de rosto;			
HOG			
	de extração  HOG Médias de Histo- gramas Detector de rosto; HOG Detector de rosto; HOG Detector de rosto; HOG Detector de rosto; HOG Detector de rosto;	de extração  HOG Perceptron  Médias de Histo-gramas  Detector de rosto; HOG  Detector de rosto; HOG	de extração  HOG Perceptron 3780; 2  Médias Back- 8; 10; 5; propagation 3  Detector de rosto; HOG  Detector de rosto; HOG

Verifica-se que, em todos os testes apresentados, o percentual de reconhecimento foi superior a 90%, chegando a 100% no caso base de rodas de veículos. O resultado referente à classificação das folhas de tabaco foi levemente superior ao obtido pelos autores da base, que obtiveram 92% de predição utilizando PCA -Análise de componentes principais.

Os resultados com os 4 conjuntos de faces confirmam a expectativa do autor das bases em relação à ordem de dificuldade para classificação devido à qualidade da aquisição, chegando a 99,78% na base com menor grau de variação e 93,75% no caso da base com mais dificuldade (Spacek, 2007). Para efeito de comparação, Belghini et al. (2012) alcançam 98% de reconhecimento com imagens das bases faces 94, faces 95 e grimace. Em outro trabalho, Toygar and Acan (2003) utilizaram o conjunto completo (faces94, faces95, faces96 e grimace) e obtiveram 93% de reconhecimento com a técnica de PCA.

Pode-se observar que o método de extração de histograma orientado por gradientes (HOG) aparece em 5 dos 6 testes mais bem-sucedidos, o que determina este como sendo o algoritmo de extração mais eficiente do PINFI. De fato, ele foi utilizado com sucesso para alimentar a RNA na base de rodas, e nas 4 bases de faces, sendo que nestas foi precedido pela aplicação do filtro detector de rosto com a função de recortar somente a região do rosto das imagens.

O filtro de extração de histograma convencional foi bem-sucedido no caso da base de folhas de tabaco, na qual a análise da cor é essencial. Foram extraídos os valores da média de azul, vermelho e verde do histograma RGB, além das médias dos componentes do histograma YCbCr e as médias de Saturação e Luminância do histograma HSL.

Pode-se observar ainda que o algoritmo Perceptron aparece na maioria dos testes da tabela. Isso ocorre devido ao fato de que o elevado número de entradas fornecido pelo filtro de extração HOG inviabiliza a utilização dos demais algoritmos. Durante os testes verificouse que os algoritmos Backpropagation, Deep Neural Network e Resilient Backpropagation atingem percentuais de reconhecimento maiores do que o Perceptron, mas isso somente é verificado após um número muito grande de iterações (em torno de 200.000). Em caso de um conjunto de poucas entradas (menos de 20) e poucos neurônios nas camadas intermediárias, este número de iterações, e consequentemente o erro aceitável, é atingido rapidamente. Mas no caso das 3780 entradas fornecidas pelo algoritmo HOG, o tempo (estimado em semanas para cada teste) é impraticável com o hardware de que se dispõe neste trabalho (computador com processador "Intel(R) Core(TM) i7-5500U de 2.40GHz e Memória RAM de 8GB). Como o algoritmo Perceptron é mais simples, não permitindo utilização de camadas intermediárias e como a atualização dos pesos ocorre em maior velocidade, o tempo para realizar cada iteração é menor e o número de iterações necessárias para atingir o erro aceitável é menor.

Nos casos em que foi possível aplicar os algoritmos Deep Neural Network e Resilient Backpropagation, estes apresentaram resultados de percentual de acertos similares ao Backpropagation, apresentando um desempenho superior apenas no tempo de convergência da RNA, ou seja, o tempo em que a mesma leva para aproximar o erro médio de zero.

Outro resultado a relatar, diz respeito à utilização da etapa de fuzzificação, que está resumida na Tabela 2. Foram utilizadas as mesmas bases e as mesmas configurações de extração de dados e da RNA da Tabela 1. Na coluna central estão os resultados obtidos com a RNA recebendo diretamente os dados do filtro de extração. Na última coluna estão os resultados obtidos quando foi definida a etapa de fuzzificação como entrada para a RNA.

**Tabela 2:** Validação com e sem fuzzificação

Base	Rec. sem fuzzy	Rec. com fuzzy
rodas	100,00%	100,00%
fumo	94,44%	72,22%
faces94	99,78%	98,03%
faces95	97,22%	89,81%
faces96	93,75%	93,56%
grimace	98,15%	96,30%

Como se pode observar em nenhuma situação a utilização do bloco fuzzy trouxe resultados superiores em relação ao modelo que não o utiliza. Os resultados são sempre similares ou levemente inferiores.

Uma razão para isso pode ser o fato de que os modelos que fuzzificação foram construídos pelos autores deste trabalho, sem a presença de um especialista na área, o que costuma ser um grande diferencial da lógica fuzzy. Ainda assim, deve-se afirmar que a presença da etapa de fuzzificação não apresenta ganhos para a classificação de imagens por RNA para os casos estudados.

De qualquer forma, o bloco de fuzzificação permanece como uma ferramenta opcional no PINFI, possibilitando que sejam realizados testes mais acurados com presença de especialistas.

## 5 Avaliação de ergonomia e usabilidade

Além da validação da acurácia do classificador e testes no sistema como um todo, também foi realizada uma avaliação das características de ergonomia e usabilidade do PINFI por meio da técnica de checklist. A lista de questões aplicadas foi extraída do software ErgoList (Cybis, 2003b).

A Tabela 3 demonstra o resultado da aplicação da checklist totalizado por grupo de questões. O significado das colunas é o seguinte:

- T: Número total de questões no grupo.
- S: Número de questões que atendem às exigências.
- N: Número de questões que não atendem às exigências.
- · NA: Número de questões não aplicáveis ao PINFI.

**Tabela 3:** Resumo da checklist de usabilidade por

grupo de questões						
Grupo de questões	Т	S	N	NA		
Presteza	17	10	3	4		
Agrupamento por loca- lização	11	10	0	1		
Agrupamento por formato	17	15	0	2		
Feedback	12	10	0	2		
Legibilidade	27	20	1	6		
Concisão	14	11	1	2		
Ações Mínimas	5	5	0	0		
Densidade Informa- cional	9	8	0	1		
Ações Explíci- tas	4	4	0	0		
Controle do Usuário	4	3	1	0		
Flexibilidade	3	0	3	0		
Experiência do Usuário	6	3	0	3		
Proteção contra erros	7	5	1	1		
Mensagens de erro	9	6	2	1		
Correção de erros	S5	1	1	3		
Consistência	11	11	0	0		
Significados	12	10	0	2		
Compatibilida	d <b>e</b> 21	18	1	2		
Total 194	150	14	30			

São 18 grupos de questões, totalizando 194 perguntas. Destas 30 não foram aplicáveis ao PINFI, 150 foram atendidas e 14 não foram atendidas. A especificação de cada grupo de questão pode ser consultada em (Cybis, 2003b).

Os resultados da checklist apontam que em 77% das questões de avaliação de ergonomia o PINFI atendeu a expetativa, 16% foram consideradas não aplicáveis e somente 7% não atenderam a exigência. Considerando somente as 164 questões aplicáveis ao PINFI, observase que 91% dos itens avaliados foram atendidos com sucesso pela ferramenta desenvolvida.

Segundo Cybis (2003a), um software será considerado em não conformidade em termos de usabilidade quando os critérios avaliados se encontrarem abaixo de 50%. Desta forma, o resultado obtido é considerado satisfatório e válido para o PINFI em termos de usabilidade de software. A título de comparação, outro software avaliado com a mesma checklist em outro trabalho apresentou percentual de 85% (Posselt et al., 2015).

### 6 Conclusão

Os resultados da avaliação do percentual de reconhecimento do PINFI demonstraram que a ferramenta se apresenta como um bom classificador de imagens em termos de capacidade de reconhecimento. Ao mesmo tempo, mostra-se como uma ferramenta aplicável a diferentes situações, inclusive em ambiente industrial ou agrícola, com numerosas opções de aquisição, preparação, processamento e saída das imagens.

Quanto a avaliação de ergonomia e usabilidade, foi possível demonstrar que o projeto e desenvolvimento do software foram bem-sucedidos e que o PINFI está pronto para ser utilizado em ambiente real. Por um lado o operador do chão de fábrica terá uma interface simples para interagir ao passo que o especialista (responsável pela montagem da sequência de técnicas) terá a disposição todos os recursos desenvolvidos. Ainda que, cada novo contexto exigirá nova bateria de testes por parte do especialista, estes são de fácil realização nas interfaces da ferramenta.

Desta forma, o PINFI se apresenta como uma solução para melhorar processos industriais modelando sistemas que precisam do processamento de imagens com redes neurais artificiais e lógica fuzzy. Por ser uma ferramenta com diversas opções de configuração e integração destas três tecnologias, o PINFI pode ser utilizado para testar diferentes possibilidades de aplicação destas técnicas, contribuindo assim para a disseminação da aplicação prática e o conhecimento público da inteligência artificial. Podem ser utilizadas bases de imagens de diferentes áreas do conhecimento, como plantações afetadas por pragas e doenças, alimentos em condições não adequadas para consumo humano, entre outros.

Destaca-se também o Registro da Ferramenta PINFI já concedido pelo INPI (Instituto Nacional da Propriedade Industrial) em 2017.

## Referências

- Belghini, N., ZARGHILI, A., Kharroubi, J., Majda, A., Sidi, D. and Ben, M. (2012). Learning a backpropagation neural network with error function based on bhattacharyya distance for face recognition, *International Journal of Image*, *Graphics and Signal Processing* **4**: 8-14. https://doi.org/10.5815/ijigsp.2012.08.02.
- Beluco, A., Engel, P. M. and Beluco, A. (2015). Classification of textures in satellite image with gabor filters and a multi layer perceptron with back propagation algorithm obtaining high accuracy, *International Journal of Image*, *Graphics and Signal Processing* **6**: 437–460. https://doi.org/10.5815/ijigsp.2012.08.02.
- Berenji, H. R. and Khedkar, P. (1992). Learning and tuning fuzzy logic controllers through reinforcements, *IEEE Transactions on Neural Networks* **3**(5): 724-740. https://doi.org/10.1109/72.159061.
- Bouzalmat, A., Belghini, N., ZARGHILI, A. and Kharroubi, J. (2011). Face detection and recognition using back propagation neural network and fourier gabor filters, Signal & Image Processing: An International Journal 2. https://doi.org/10.5121/sipij.2011.2302.
- Bradski, G. (2000). The opency library, Dr Dobb's J. Software Tools 25: 120-125. Disponível em https://ci.nii.ac.jp/naid/10028167478/en/.
- Bravo, C., Saputelli, L., Rivas, F., Perez de Rivas, A., Nickolaou, M., Zangl, G., Guzmán, N., Mohaghegh, S. and Nunez, G. (2014). State of the art of artificial intelligence and predictive analytics in the e&p industry: A technology survey, SPE Journal 19: 547–563. https://doi.org/10.2118/150314-PA.
- Chi, Z., Yan, H. and Pham, T. (1996). Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition, Advances in fuzzy systems applications and theory, World Scientific. https://doi.org/10.1142/3132.
- Cybis, W. d. A. (2003a). Engenharia de usabilidade: uma abordagem ergonômica, Laboratório de Utilizabilidade de Informática. Disponível em http://www.labiutil.inf.ufsc.br/hiperdocumento/conteudo.html.
- Cybis, W. d. A. (2003b). ERGOLIST: Software para avaliação de interfaces, Laboratório de Utilizabilidade de Informática, Florianópolis. Disponível em http://www.labiutil.inf.ufsc.br/ergolist.
- Figueiredo, M. and Gomide, F. (1999). Design of fuzzy systems using neurofuzzy networks, *IEEE Transactions on Neural Networks* **10**(4): 815–827. https://doi.org/10.1109/72.774229.
- Geetha, S. (2015). Imaging based tablet inspection in drug industry, International Journal of Pharmacy and Technology 7: 8610-8617. Disponível em https://www.ijptonline.com/wp-content/uploads/2015/10/8610-8617.pdf.

- Gonzalez, R. and Woods, R. (2000). Processamento de Imagens Digitais, Editora Edgard Blucher, São Paulo, SP.
- Gunes, M. and Badem, H. (2016). Detecting direction of pepper stem by using cuda-based accelerated hybrid intuitionistic fuzzy edge detection and ann, *Journal of Sensors* **2016**. https://doi.org/10.1155/2016/4052101.
- Haykin, S. (1999). *Neural Networks. A Comprehensive Foundation*, 2 edn, Prentice Hall, Sebastopol, CA.
- Kirillov, A. (2016). Aforge.net framework. Disponível em http://www.aforgenet.com.
- Posselt, E., Frozza, R. and Molz, R. (2015). Infuzzy: Ferramenta para desenvolvimento de aplicações de sistemas difusos, *Revista Brasileira de Computação Aplicada* 7(1): 42-52. https://doi.org/10.5335/rbca.2015.3960.
- Rao, S., Venkatesan, K. and Karthikeyan, R. (2018). Tracing a weld line using artificial neural networks, International Journal of Networked and Distributed Computing 6: 216. https://doi.org/10.2991/ijndc.2018.4.6.4.
- Ros, P. M. and Pasero, E. (2013). Defects Detection in Pistachio Nuts Using Artificial Neural Networks, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 147–156. https://doi.org/10.1007/978-3-642-35467-0\_16.
- Sabanci, K., Toktas, A. and Kayabasi, A. (2017). Grain classifier with computer vision using adaptive neuro-fuzzy inference system, *Journal of the Science of Food and Agriculture* **97**(12): 3994–4000. https://doi.org/10.1002/jsfa.8264.
- Silva, A. A. V. d., Silva, I. A. F., Teixeira Filho, M. C. M., Buzetti, S. A. and Teixeira, M. C. M. (2014). Estimativa da produtividade de trigo em função da adubação nitrogenada utilizando modelagem neuro fuzzy, Revista Brasileira de Engenharia AgrÃcola e Ambiental 18: 180 187. https://doi.org/10.1590/S1415-43662014000200008.
- Soares, P. L. B. and da Silva, J. P. (2013). Redes neurais aplicadas na redução de ruído impulsivo de imagens digitais, *Revista Brasileira de Computação Aplicada* 5(1): 2-13. https://doi.org/10.5335/rbca.2013.2454.
- Spacek, L. (2007). Collection of facial images: Faces94. computer vision science and research projects, university of essex, uk. Disponível em http://cswww.essex.ac.uk/mv/allfaces.
- Suptitz, I. L., Frozza, R. and Molz, R. F. (2017). Estado da arte da indústria gaúcha referente ao uso de técnicas de inteligência artificial por meio de um survey, iSys-Brazilian Journal of Information Systems 10(3): 36–48. https://doi.org/10.5753/isys.2017.345.
- Taghadomi-Saberi, S., Omid, M., Emam-Djomeh, Z. and Ahmadi, H. (2014). Evaluating the potential of artificial neural network and neuro-fuzzy techniques for estimating antioxidant activity and anthocyanin

- content of sweet cherry during ripening by using image processing, *Journal of the Science of Food and Agriculture* **94**(1): 95–101. https://doi.org/10.1002/jsfa.6202.
- Toygar, Ö. and Acan, A. (2003). Face recognition using pca, lda and ica approaches on colored images, Istanbul University-Journal of Electrical & Electronics Engineering 3(1): 735-743. Disponivel em https://electricajournal.org/Content/files/sayilar/49/735-743.pdf.
- Tyryshkin, A. and Belyaev, A. (2016). Dual clustering in vision systems for robots deployed for agricultural purposes, MATEC Web of Conferences 79: 01083. https://doi.org/10.1051/matecconf/20167901083.
- Yen, V. T., Nan, W. Y. and Van Cuong, P. (2019). Robust adaptive sliding mode neural networks control for industrial robot manipulators, *International Journal of Control*, *Automation and Systems* 17(3): 783-792. https://doi.org/10.1007/s12555-018-0210-y.