

RESSALVA

Atendendo a solicitação do autor, o texto completo desta dissertação será disponibilizado somente a partir de 27/03/2024.

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE ENGENHARIA
CÂMPUS DE ILHA SOLTEIRA**

GUILHERME PAGANINI CONSTANTINO DE OLIVEIRA

**RECONHECIMENTO DO ESTADO DE SAÚDE DE GLOBOS OCULARES
UTILIZANDO A REDE NEURAL ARTMAP *FUZZY***

Ilha Solteira
2023

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA

GUILHERME PAGANINI CONSTANTINO DE OLIVEIRA

**RECONHECIMENTO DO ESTADO DE SAÚDE DE GLOBOS
OCULARES UTILIZANDO A REDE NEURAL ARTMAP *FUZZY***

Dissertação apresentada à Faculdade de Engenharia de Ilha Solteira – Unesp como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica.

Prof^a. Dr^a. Anna Diva Plasencia Lotufo
Orientadora

FICHA CATALOGRÁFICA
Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

O48r Oliveira, Guilherme Paganini Constantino de.
Reconhecimento do estado de saúde de globos oculares utilizando a rede neural ARTMAP Fuzzy / Guilherme Paganini Constantino de Oliveira. -- Ilha Solteira:[s.n.], 2023
90 f. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2023

Orientador: Anna Diva Plasencia Lotufo

Inclui bibliografia

1. Anomalias oculares. 2. ARTMAP Fuzzy. 3. Codificação de atributos. 4. Extração de atributos de imagens. 5. Reconhecimento de imagens. 6. Redes neurais artificiais.


Amanda Sertori dos Santos

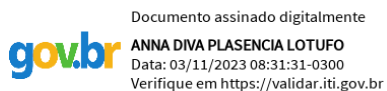
CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: RECONHECIMENTO DO ESTADO DE SAÚDE DE GLOBOS OCULARES UTILIZANDO A REDE NEURAL ARTMAP *FUZZY*

AUTOR: GUILHERME PAGANINI CONSTANTINO DE OLIVEIRA

ORIENTADORA: ANNA DIVA PLASENCIA LOTUFO

Aprovado como parte das exigências para obtenção do Título de Mestre em , área: Automação pela Comissão Examinadora:



Profa. Dra. ANNA DIVA PLASENCIA LOTUFO (Participação Presencial)
Departamento de Engenharia Eletrica / Faculdade de Engenharia de Ilha Solteira - UNESP

Profa. Dra. MARA LUCIA MARTINS LOPES (Participação Presencial)
Departamento de Matematica / Faculdade de Engenharia de Ilha Solteira - UNESP

Prof. Dr. KENJI NOSE FILHO (Participação Virtual)
Centro de Engenharia, Modelagem e Ciências Sociais Aplicadas / Universidade Federal do ABC - UFABC

Ilha Solteira, 27 de setembro de 2023

IMPACTO POTENCIAL DESTA PESQUISA

A pesquisa fomenta possível auxílio diagnóstico na área da oftalmologia em pacientes com eventuais patologias oculares, contribuindo para início precoce de admissível tratamento necessário e podendo favorecer as chances de recuperação. Os resultados podem ser potencializados, futuramente, para implementação em dispositivos inteligentes e estruturados para distinção de anomalias individuais.

POTENTIAL IMPACT OF THIS RESEARCH

The research provides possible diagnostic support in the field of ophthalmology for patients with possible eye pathologies, contributing to the early start of the necessary treatment and potentially improving the chances of recovery. In the future, the results could be used to implement intelligent, structured devices to distinguish individual anomalies.

À Kethelyn Pinheiro, dedico.

AGRADECIMENTOS

Agradeço a Deus por nunca me desamparar e me proporcionar inúmeras bênçãos, mesmo após minhas sucessivas falhas como ser humano. Agradeço por me instigar com ideias que nem mesmo eu achei que poderia arquitetar.

À minha futura esposa, Kethelyn, por conseguir me tornar uma pessoa melhor do que eu jamais merecia ser e me mostrar o lado bom das coisas.

À minha mãe, por ter me criado e proporcionado educação em meio a diversas dificuldades.

À professora Anna Diva, pela atenção, paciência e sugestões. Agradeço pela confiança depositada em mim para realização do trabalho.

Ao professor Minussi, pelas recomendações proporcionadas e pela prontidão em auxiliar.

À Danieli e Giovanni, por contribuírem com dicas em momentos absolutamente cruciais da execução do trabalho, meus mais sinceros agradecimentos. Ao Reginaldo, pelas sugestões de plataformas na seleção do *dataset*.

Agradeço ao amigo Bruno, de longa data, pelos incontáveis conselhos durante os cafezinhos no câmpus. Obrigado pela amizade. Agradeço também ao amigo Victor, pelo companheirismo durante o mestrado.

Ao Sintel e à FEIS, por proporcionarem a estrutura necessária para viabilização desta pesquisa.

À CNPq e à CAPES, pelo apoio financeiro concedido para realização do trabalho.

“ I did not tell half of what I saw, for I knew I would not be believed” - Marco Polo.

RESUMO

Com o desenvolvimento digital tão em pauta atualmente, diversas ferramentas de inteligência artificial puderam ser apropriadamente estabelecidas e passaram a fazer parte importante do cotidiano. Dentro deste tópico, o reconhecimento de imagens tornou-se bastante significativo em inúmeras aplicações, incluindo a área da saúde. Neste escopo de análise, ressalta-se a abordagem destinada à detecção do estado de saúde de globos oculares, constituindo premissa de extrema relevância e com grande possibilidade de exploração na pesquisa científica. O presente trabalho propõe a utilização da rede neural ARTMAP *Fuzzy* para reconhecimento do estado de saúde de globos oculares, tratando-se de uma empregabilidade pioneira da rede. Utilizou-se um banco de dados de imagens de fundos oculares com boa quantidade de amostras de qualidade para possibilitar o reconhecimento de imagem. Manipulações e tratamento nas imagens foram realizados a fim de permitir uma melhor padronização dos dados a serem apresentados à rede. Testou-se três diferentes processos de extração e codificação dos atributos dos dados, que incluíram: transcrição dos pixels das imagens em valores numéricos representativos; contagem de aparecimento de pixels específicos via histograma; e codificação por meio da frequência de palavras visuais existentes (método BoVW - *Bag of Visual Words*). Testes para verificação da capacidade da rede em discernir olhos saudáveis de não saudáveis foram efetuados. A métrica de taxa de classificação correta (CCR - *Correct Classification Rate*) foi empregada para comparação de eficiência dos resultados para cada uma das abordagens. Apesar dos três métodos apresentarem bom percentual de acerto geral, a utilização do BoVW demonstrou melhor distribuição de eficiência em relação aos outros pela congruência de precisão quantos aos olhos saudáveis individualmente.

Palavras-chave: anomalias oculares; ARTMAP *Fuzzy*; codificação de atributos; extração de atributos de imagens; reconhecimento de imagens; redes neurais artificiais.

ABSTRACT

With digital development so much on the agenda these days, several artificial intelligence tools have been properly established and have become an important part of everyday life. Within this topic, image recognition has become very significant in numerous applications, including the health field. Within this scope of analysis, the approach aimed at detecting the state of health of eyeballs stands out, constituting a premise of extreme relevance and with great potential for exploration in scientific research. This work proposes the use of the fuzzy ARTMAP neural network to recognize the state of health of eyeballs, which is a pioneering use of the network. A database of eye fundus images with a good number of quality samples was used to enable image recognition. The images were manipulated and processed in order to better standardize the data to be presented to the network. Three different data attribute extraction and coding processes were tested, which included: transcription of image pixels into representative numerical values; appearance count of specific pixels via histogram; and coding through the frequency of existing visual words (BoVW - Bag of Visual Words). Tests were carried out to verify the network's ability to discern healthy from unhealthy eyes. The Correct Classification Rate (CCR) metric was used to compare the efficiency of the results for each of the approaches. Although all three methods showed a good percentage of accuracy in the general case, the use of BoVW showed a better distribution of efficiency in comparison to the others due to the congruence of accuracy with regard to healthy eyes individually.

Keywords: ocular anomalies; Fuzzy ARTMAP; attribute coding; extraction of attributes from images; image recognition; artificial neural networks.

LISTA DE FIGURAS

Figura 1 - Esquemática de algumas redes vinculadas à família ART.....	26
Figura 2 - Esquemática da arquitetura de uma RNA ART <i>Fuzzy</i> típica.....	29
Figura 3 - Fluxograma do algoritmo de funcionamento da ART <i>Fuzzy</i>	36
Figura 4 - Representação das categorias geradas para o caso unidimensional.	38
Figura 5 - Representação geométrica de uma categoria no caso bidimensional.	38
Figura 6 - Esquemática da arquitetura de uma RNA ARTMAP <i>Fuzzy</i> típica.	41
Figura 7 - Síntese das manipulações iniciais efetuadas no dataset.	51
Figura 8 - Dois exemplares retirados durante o processo de exclusão manual.	52
Figura 9 - Exemplos dos dois tipos de centralização adotadas nas capturas.....	53
Figura 10 - Indicação da retirada do disco óptico.	55
Figura 11 - Esquemática da conversão para escala de preto e branco.....	57
Figura 12 - Estratégia adotada na primeira codificação empregada.	58
Figura 13 - Comparativo de histogramas de dois globos oculares do dataset.	60
Figura 14 - Sumarização do método BoVW.	61
Figura 15 - Pontos de interesse de amostra do dataset via características KAZE. ..	62
Figura 16 - Comparação da quantidade de clusters e erro total no k-means.	66
Figura 17 - Recortes empregados para testes no método matriz em vetor.....	69
Figura 18 - Caráter complexo de distinção em algumas amostras.....	73
Figura 19 - Comparação da melhor taxa de classificação obtida em cada método..	78
Figura 20 - Comparativo das médias dos melhores resultados das codificações. ...	78
Figura 21 – Evolução percentual média com o treinamento em várias épocas.....	79
Figura 22 - Número de categorias criadas nos métodos variando-se o <i>baseline</i>	80

LISTA DE TABELAS

Tabela 1 - Síntese dos parâmetros da rede ART Fuzzy.	31
Tabela 2 - Síntese e contextualização do dataset utilizado no trabalho.	49
Tabela 3 - Síntese dos exemplares remanescentes.	54
Tabela 4 - Informações e parâmetros gerais empregados.	68
Tabela 5 - Resultados principais do método de codificação Matriz em Vetor.	69
Tabela 6 - Síntese dos melhores resultados para o histograma de imagens.	72
Tabela 7 - Síntese dos melhores resultados para codificação BoVW.....	75
Tabela 8 - Comparativo descritivo dos métodos de codificação.	81
Tabela 9 - Distribuição das doenças para o treinamento no <i>dataset</i> inicial.	90

LISTA DE SIGLAS E ABREVIATURAS

AF	(Rede Neural Artificial) ART <i>Fuzzy</i>
AMF	(Rede Neural Artificial) ARTMAP <i>Fuzzy</i>
ART	<i>Adaptative Resonance Theory</i>
BP	<i>Backpropagation</i>
Cat C	Categorias Criadas (módulo ART_a)
CCA	<i>Canonical Correlation Analysis</i>
CCR	<i>Correct Classification Rate</i>
CNN	<i>Convolutional Neural Network</i>
D.O	Disco Óptico
E/D	Olho: Esquerdo/Direito
HOG	<i>Histogram of Oriented Gradients</i>
LTM	<i>Long Term Memory</i>
NC	Normalização Convencional
NVB	Normalização pelo Valor Base
P/B	Preto e Branco
PCA	<i>Principal Component Analysis</i>
QPV	Quantidade de Palavras Visuais
RFMiD	<i>Retinal Fundus Multi-Disease Image Dataset</i>
RGB	<i>Red, Green, Blue (Additive Color Model)</i>
RNAs	Redes Neurais Artificiais
ROC	<i>Receiver Operating Characteristic</i>
S/NS/G	Caso Ocular: Saudável/Não Saudável/Geral
SGGS	<i>Shri Guru Gobind Singhji Institute of Engineering and Technology</i>
STM	<i>Short Term Memory</i>
TSLN	<i>Ocular Tessellation</i>

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVO E PROPOSTA.....	18
1.2	ORGANIZAÇÃO TEXTUAL.....	18
2	REVISÃO BIBLIOGRÁFICA	20
3	TEORIA DA RESSONÂNCIA ADAPTATIVA	24
3.1	REDE NEURAL ARTIFICIAL ART <i>FUZZY</i>	27
3.1.1	Algoritmo ART <i>Fuzzy</i>	30
3.1.2	Breve Análise da Representação Geométrica das Categorias	35
3.2	REDE NEURAL ARTIFICIAL ARTMAP <i>FUZZY</i>	39
3.2.1	Algoritmo ARTMAP <i>Fuzzy</i>	42
3.2.2	Síntese – Algoritmo de Treinamento AMF	46
4	METODOLOGIA DE IMPLEMENTAÇÃO	48
4.1	<i>DATASET</i> E MANIPULAÇÕES INICIAIS.....	48
4.2	APURAÇÃO E PRÉ-PROCESSAMENTO DAS IMAGENS.....	51
4.2.1	Seleção e Amostras Resultantes	51
4.2.2	Procedimentos Finais Pré-Codificação	54
4.3	EXTRAÇÃO DE ATRIBUTOS E CODIFICAÇÃO.....	56
4.3.1	Matriz da Imagem em Vetor Linha	57
4.3.2	Histograma de Imagens	58
4.3.3	BoVW – <i>Bag of Visual Words</i>	60
5	RESULTADOS EXPERIMENTAIS E DISCUSSÃO	67
5.1	ESTADO DE SAÚDE DE GLOBO OCULAR – MATRIZ EM VETOR....	67
5.2	ESTADO DE SAÚDE DE GLOBO OCULAR – HISTOGRAMA DE IMAGENS.....	71
5.3	ESTADO DE SAÚDE DE GLOBO OCULAR - BOVW.....	75
5.4	ANÁLISE COMPARATIVA.....	77
6	CONCLUSÕES	82

6.1	PROPOSTAS PARA TRABALHOS FUTUROS.....	84
	REFERÊNCIAS.....	85
	APÊNDICE A – INFORMAÇÕES ADICIONAIS DO DATASET.....	89

1 INTRODUÇÃO

Desde a proposição da unidade básica do neurônio artificial (MCCULLOCH; PITTS, 1943), o desenvolvimento dos preceitos vinculados à inteligência artificial ocorreram rapidamente e fundamentaram-se cada vez mais nas atividades que permeiam nosso cotidiano. Sistemas estruturados em aprendizados de máquina estão inerentes cultural e cientificamente, e as RNAs contribuíram significativamente para que este patamar fosse devidamente alcançado.

As RNAs correspondem a idealizações de estruturas de inteligência artificial embasadas no funcionamento cerebral humano. Estas arquiteturas ancoram-se na disposição específica das unidades individuais de aprendizado (neurônios artificiais) e, mediante o processo experimental vinculado a cada uma, permitem que tais unidades, via ajuste progressivo de pesos ponderadores, generalizem uma função descritora do sistema que possibilita sua utilização para análises inéditas (HAYKIN, 1999). A evolução do mundo digital promoveu o aparecimento acentuado das redes neurais artificiais nos diversos âmbitos do conhecimento, sendo comumente empregadas em inúmeras tarefas de classificação e predição de eventos.

Uma vez que as RNAs se fundamentam no funcionamento intrínseco do cérebro humano, é natural buscar que elas idealizem, cada vez mais, as características singulares do aprendizado. É necessário, portanto, que os sistemas possuam aptidão em armazenar conhecimento via treinamento e tenham capacidade de generalizar respostas razoavelmente precisas para dados novos e constituam flexibilidade de aplicação em variadas vertentes (WASSERMAN, 1989), características que se fazem presentes na estrutura cerebral humana.

Mediante a capacidade do cérebro de adequação e organização complexa dos neurônios, tarefas relacionadas a processamentos de percepção, controle motor e reconhecimento de padrões (HAYKIN, 1999) fazem parte essencial das rotinas humanas e permitem constantes interações com o mundo. Dentre estas, vale o destaque para a acentuada distinção que o cérebro pode efetuar das mais inúmeras ocorrências que se apresentam cotidianamente. Independente da entrada visual que uma determinada pessoa esteja recebendo, rapidamente ocorre o reconhecimento e a conseqüente associação em alguma determinada “classe neural” de conhecimento preexistente. Caso esta classe não exista, o aprendizado humano permite que um novo grupo seja criado para que este novo estímulo seja vinculado. A utilização de

sistemas inteligentes e, mais precisamente, de RNAs em aplicações cotidianas, tem como um dos principais objetivos efetivar o relatado anteriormente: estruturar um modelo que possa discernir e reconhecer padrões de imagens e estímulos com acurácia significativa, tal como nosso cérebro é capaz de fazer sistematicamente.

O reconhecimento de imagens e padrões trata-se de uma das pautas mais presentes no contexto de evolução digital, cujo desenvolvimento propiciou grandes melhoras em diversas áreas da sociedade. Efetuar adequadamente o reconhecimento de imagens, pautado em um determinado objetivo, permite otimizar e automatizar tarefas que outrora necessitavam inspeção visual (MATHWORKS, 2023a; GONZALEZ; WOODS, 2009). Além disso, sendo possível identificar padrões em imagens e usar estas informações para tomar decisões assertivas, fica assegurada alta confiabilidade do sistema e o mesmo pode ser redirecionado a ramos essenciais da vida humana.

Por intervenção desta propriedade do reconhecimento de imagens, sistemas podem monitorar ambientes e detectar situações de risco ou perigo; aparelhos móveis pessoais podem ser vinculados ao usuário e desbloqueados, garantindo potencialização da segurança; drones com câmeras podem sobrevoar regiões de matas densas e detectar possíveis focos de incêndio; consumidores podem buscar mais facilmente produtos específicos via verificação de preferências em compras online; profissionais da saúde podem utilizar ferramentas de detecção de anomalias; dentre muitas outras aplicações. Este progresso foi possível, em grande parte, à constante criação de algoritmos poderosos capazes de assimilar padrões de dados e efetuar predições de eventos (PIURI et al., 2020).

Uma das vertentes mais importantes de aplicação do reconhecimento de imagens trata-se da área da saúde. Utilizar inteligência artificial para auxiliar na detecção de anomalias fisiológicas corresponde a um tópico de extrema relevância, visto que um sistema devidamente treinado e com boa acurácia pode colaborar com o diagnóstico de um paciente e, por consequência, permite otimizar o tempo para tomada de ações no eventual tratamento (PACHADE et al., 2021). Neste contexto, a detecção de anomalias oculares e do estado de saúde do olho como um todo destaca-se pela quantidade de avarias inerentes que podem aparecer no decorrer da vida do ser humano. Além disso, muitas destas anomalias podem estar vinculadas a ocorrências raras, sendo necessária uma boa estruturação de dados e amostras a fim

de possibilitar o uso de técnicas computacionais de reconhecimento, uma vez que o leque de possibilidades é considerável (PACHADE et al., 2021).

Para estas aplicações, é comum utilizar-se de técnicas de aprendizado profundo (*deep learning*) (DENG; YU, 2014), como a rede neural convolucional (CNN). A vantagem destes métodos reside no processamento de grande quantidade de dados para melhor análise e algoritmos auto adaptativos para lidar com uma maior quantidade de amostras, possuindo boa aptidão em extrair características complexas dos componentes do banco de dados e favorecendo utilização nestas aplicações (PIURI et al., 2020).

Além das estruturas vinculadas ao aprendizado profundo, uma classe de RNAs que merece destaque pertence às redes da família ART, ancoradas na teoria da ressonância adaptativa. Estas estruturas foram altamente estudadas por Carpenter e Grossberg e, devido ao alto caráter de flexibilidade das redes, foi possível obter diversas variações de arquiteturas, cada qual destinando-se à manipulação de dados diferentes e efetuando assimilação de padrões adaptativos e auto estabilização em resposta a estímulos de entrada complexos (GROSSBERG, 1988). Além disso, por serem altamente interativas, é factível a integração de módulos e a implementação de alterações nas arquiteturas, algo que possibilita cada vez mais atingir novos objetivos.

Muito do que torna esta família de redes interessante corresponde à possibilidade de lidar com o dilema da estabilidade e plasticidade, além de possuir estruturas de memória de curto e longo prazo (GROSSBERG, 1976a). Esta capacidade de lidar com este dilema permite à rede ancorar-se nos dois preceitos simultaneamente: de um lado, garante que ocorrerá a convergência para uma determinada configuração de pesos que generalize o processo para o qual ela foi treinada; de outro, assegura que a estrutura será capaz de aprender continuamente com a inclusão de novos padrões sem se desfazer do aprendizado obtido até o momento. Esta capacidade é vital dado que nosso cérebro, ao qual as redes fundamentam-se, possui tais aptidões e frequentemente lida com estímulos ambientais variantes e não estacionários, aos quais distinguimos e lidamos continuamente (CARPENTER; GROSSBERG, 1988).

Dito isso, é interessante ressaltar que estes princípios permeiam duas grandes classificações básicas de modelos de arquiteturas destas redes: os aprendizados não-supervisionados e os supervisionados. No primeiro, caracterizado por um módulo ART independente, ocorre a categorização dos padrões de entrada sucessivos

apresentados à rede por meio de associações dos que apresentam características similares (CARPENTER; GROSSBERG, 1987a). No segundo caso, os treinamentos supervisionados exigem que se possua um par treinado para a execução do processo. A cada entrada de dados deve-se associar uma saída, justamente para permitir que a estrutura elabore um mapeamento assertivo dos pares. Neste caso, as categorias de entrada e saída serão relacionadas, e mais de um módulo ART é necessário para efetuar o procedimento. Destaca-se, portanto, as variações da rede ARTMAP (CARPENTER; GROSSBERG; REYNOLDS, 1991).

Dada toda esta conjuntura, muitos trabalhos têm sido feitos a fim de estudar as propriedades intrínsecas da AMF e como a variação de seus parâmetros atua efetivamente no desempenho da mesma (ANAGNOSTOPOULOS; GEORGIPOULOS, 2002; GEORGIOPULOS; HUANG; HEILMAN, 1994). Por ser uma rede altamente flexível, com boa característica de convergência e possibilitar excelente tratativa de variados tipos de entradas, a AMF pode ser explorada em inúmeras aplicações, incluindo classificação, predição de eventos e até mesmo aproximação de funções (CANO-IZQUIERDO et al., 2013).

Com todo este panorama da AMF, é curioso que estudos envolvendo reconhecimento de imagens não tenham usualmente empregado suas arquiteturas. Apesar das características interessantes presentes nas técnicas de *deep learning* (DENG; YU, 2014), um questionamento específico faz-se necessário: dada a facilidade de manipulação e flexibilidade da AMF, esta poderia fornecer uma alternativa sólida a despeito da eficiência quando em tratativas de reconhecimento de imagens?

Tratando-se especificamente do reconhecimento do estado de saúde de globos oculares, é instigante a suposição de que, se as amostras forem devidamente codificadas e modeladas, o processo de treinamento e as características próprias da rede AMF devem permitir boa generalização, eficiência e acurácia em teste com amostras inéditas. Se a rede é capaz de trabalhar bem com diversos tipos de dados, basta que sejam apresentados adequadamente para que o desempenho seja satisfatório.

O grande desafio, portanto, não reside inteiramente no processo de treinamento da estrutura, que é capaz de se auto organizar e maximizar generalização com redução do erro preditivo (CARPENTER; GROSSBERG; REYNOLDS, 1991): a forma como as amostras serão codificadas e padronizadas será o principal

determinante no desempenho final da rede. Se por um lado, nas técnicas de *deep learning*, a extração de características das imagens representativas para reconhecimento posterior pode ser efetuadas diretamente nos entrelaçamentos de conexões e camadas das redes (RAN et al., 2021), por outro, no caso da utilização da rede AMF, é necessário que a extração de atributos seja realizada anteriormente e de forma individual, possibilitando o treinamento adequado da estrutura.

O presente trabalho propõe explorar tal pioneirismo de utilização da rede, analisando sua eficácia para esta aplicabilidade não usual de detecção de estado de saúde ocular por meio de reconhecimento de imagens e verificando se pode constituir alternativa sólida para tal.

1.1 OBJETIVO E PROPOSTA

A proposta de trabalho ancora-se na utilização da rede ARTMAP *Fuzzy* para o tema de reconhecimento de imagens, tal como discutido anteriormente. Almeja-se utilizar esta arquitetura para a finalidade específica de efetuar a distinção do estado de saúde de globos oculares entre saudáveis e não saudáveis. A classificação “saudável” refere-se a olhos isentos de qualquer tipo de doença; em contrapartida, olhos classificados como “não saudáveis” correspondem a exemplares acometidos por uma ou mais das 45 condições presentes no *dataset* (PACHADE et al., 2021).

A eficiência da rede será testada para diferentes formas de extração de atributos e codificação das imagens: conversão da matriz de dados em vetor linha, frequência de aparecimento de pixels via histograma de imagens, e aplicação do conceito de sacola de palavras visuais (BoVW).

Testar-se-á a eficácia recorrendo a amostras inéditas alheias às do treinamento. Se disporá de um *dataset* com imagens de fundos de retina de globos oculares (PACHADE et al., 2021) e aplicar-se-á a métrica de taxa de classificação correta (CCR).

1.2 ORGANIZAÇÃO TEXTUAL

O referido trabalho é organizado em seis capítulos, incluindo a presente introdução.

No Capítulo 2 serão expostos e discutidos alguns trabalhos da literatura relacionados ao tema de reconhecimento de imagens, assim como algumas técnicas usuais para efetivação das análises e estudos sobre parâmetros da rede AMF.

O Capítulo 3 detalha os preceitos e fundamentos que permeiam a teoria da ressonância adaptativa. Aqui, é discutido o teorema da estabilidade e plasticidade, apresentando os conceitos principais de aprendizado das redes da família ART. Posteriormente, descreve-se o funcionamento prático da arquitetura da rede ART e, em seguida, da rede ARTMAP *Fuzzy*, evidenciando como tais algoritmos de treinamento possibilitam o ajuste progressivo dos pesos.

O Capítulo 4 apresenta toda a metodologia que foi empregada para viabilizar a implementação do tema na arquitetura da rede AMF. São descritos os passos efetuados no que diz respeito ao tratamento, modelagem e extração de atributos das imagens. Além da descrição geral do *dataset*, são expostas todas as manipulações das imagens a fim de possibilitar obtenção da maior quantidade de amostras adequadas para o treinamento. Além disso o capítulo descreve, passo a passo, todos os métodos de extração de atributos e codificações das imagens empregados, mostrando como cada um foi viabilizado e suas respectivas fundamentações.

No Capítulo 5 são apresentados os resultados computacionais obtidos mediante os testes com amostras inéditas. São descritos os parâmetros utilizados para o treinamento da rede e os atributos específicos obtidos após sua conclusão. Os melhores resultados sugeridos para cada caso são comentados e analisados, verificando a eficiência de cada abordagem.

No Capítulo 6 expõe-se as conclusões a respeito da aplicabilidade da rede ARTMAP *Fuzzy* no tema de reconhecimento de imagens e detecção de estado de saúde de globos oculares, trazendo-se perspectivas e sugestões de eventuais pesquisas futuras a serem exploradas na área a fim de potencializar o estudo.

Por fim, no Apêndice A é evidenciado, de forma breve, a visualização da distribuição de patologias oculares existentes nas amostras pertencentes ao banco de dados inicialmente.

6 CONCLUSÕES

O presente trabalho apresentou uma metodologia pioneira para detecção de estado de saúde de globos oculares baseada no reconhecimento de imagens. Almejou-se empregar a rede ARTMAP *Fuzzy* para efetuar as referidas classificações, correspondendo a uma abordagem inédita para tal tema. Utilizou-se um banco de dados estruturado com imagens de fundos de retina de globos oculares. Diversas seleções manuais e manipulações foram efetuadas no *dataset* até obter o banco resultante e padronizado.

Três processos diferentes de extração de atributos das imagens foram aplicados a fim de testar e comparar as respectivas eficiências por meio da métrica de taxa de classificação correta. O primeiro método, proposto no presente trabalho, referiu-se à conversão completa da matriz de dados equivalente a cada captura em um vetor unidimensional. No segundo, as frequências de aparecimento de uma determinada tonalidade representativa de pixel foram contabilizadas e estruturadas na forma de vetor de características para treinamento. Por último, a terceira abordagem abrangeu a utilização do conceito de *Bag of Visual Words*, elaborando-se um dicionário de palavras visuais mediante detecção de pontos de interesse KAZE, extração de características e aplicação do algoritmo de agrupamento *k-means*.

Para o treinamento da rede no caso do método Matriz em Vetor, três tipos de recortes foram efetuados nas imagens do *dataset* mencionado com diferentes graus de incidência: um excetuando-se apenas o disco óptico, outro as porções mais externas e, por último, um com focalização central nas capturas. Pelos resultados obtidos, priorizou-se a análise dos métodos subsequentes atrelada na terceira abordagem de recorte.

Pelos resultados obtidos com as amostras pertencentes ao banco de testes, os três métodos atingiram percentuais satisfatórios para o caso geral (analisando conjuntamente olhos saudáveis e não saudáveis), com desempenho médio ligeiramente superior no caso do BoVW, atingindo aproximadamente 90% e ficando muito próximo do modelo do histograma. No caso do desempenho absoluto máximo obtido para um teste individual, obteve-se 93,51% na análise com *baseline* 0 no caso da codificação matriz em vetor. No entanto, pelo desbalanceamento dos dados, é importante se analisar o comportamento médio de todos os testes efetuados a fim de garantir a confiabilidade do sistema, e não apenas o maior valor obtido isoladamente

em um dos testes. Por este enfoque, o método BoVW apresentou claro melhor desempenho.

Observando individualmente a média de classificação correta dos olhos não saudáveis, fica evidente novamente a assertividade dos métodos, com pequena superioridade da abordagem do histograma (atingindo quase 94%) seguida muito próxima pelo BoVW. No caso da Matriz em Vetor, atingiu-se média de 86% para este caso.

Por fim, analisando o panorama dos olhos saudáveis individualmente é que se percebe uma discrepância maior dos métodos de extração de atributos empregados. O histograma de imagens apresentou uma média de resultados abaixo dos 10% de classificação correta, representando o método com maior dificuldade de discernir corretamente este aspecto. Isso se deve provavelmente à forma de codificação em si, na qual a contagem de aparecimento de tonalidades estruturada no vetor de características não traz informações espaciais diretas de cada captura, apenas informações qualitativas que podem provocar associações incorretas. Olhando pelo enfoque dos saudáveis exclusivamente, percebe-se como o método BoVW se sobressai e mantém a melhor distribuição de acertos em relação aos outros, chegando a um valor máximo de acerto de quase 70%.

Os resultados também mostram como, de uma maneira geral, a média de taxa de classificação correta dos métodos cresce ligeiramente mediante o treinamento da rede em sucessivas épocas até atingir 100% de precisão nas amostras de treinamento. No entanto, a pequena melhoria percentual obtida não justifica o custo computacional superior decorrente do treinamento sucessivo para tal implementação. É possível analisar, por fim, como os métodos detentores da maior quantidade de elementos por vetor de características acarretam uma maior criação de categorias no módulo ART_a com o aumento do *baseline*, chegando a 255 no método da Matriz em Vetor e apenas 12 no BoVW (que possui vetores com apenas 100 ou 300 elementos).

A comparação das formas de extração de atributos e codificações evidencia como a robustez do BoVW permite uma melhor distribuição de acertos em cada uma das análises efetuadas, ainda que o processo para criação dos vetores representativos de cada imagem seja mais árduo e com maior custo computacional, envolvendo uma boa quantidade de etapas para implementação. Método como o histograma de imagens, que envolve codificação mais facilitada e vetor não muito extenso, permite acerto geral satisfatório e custo computacional reduzido, mas acaba

deixando a desejar em assertividades individuais e carece de algum complemento na codificação. O modelo de matriz em vetor acaba figurando entre os dois supracitados, exigindo um certo custo computacional e envolvendo uma codificação simples, mas com muitos elementos.

6.1 PROPOSTAS PARA TRABALHOS FUTUROS

Os métodos demonstraram desempenho satisfatório no que tange a eficiência geral da taxa de classificação do estado de saúde dos globos oculares, com destaque para a codificação BoVW por apresentar também boa assertividade no que diz respeito à classificação dos saudáveis individualmente. Com a implementação de alguns tópicos, é possível potencializar cada uma das abordagens para resultados ainda mais promissores.

Algumas sugestões para eventuais melhorias são listadas a seguir:

- 1) **Matriz em Vetor:** Redução da taxa de dimensionalidade aplicada e inserção de mais elementos no vetor representativo de cada fotografia. Empregar técnicas de preservação dos componentes principais, como o PCA, para trabalhar com uma menor quantidade de elementos;
- 2) **Histograma de Imagens:** Complementar a codificação com algum outro método que relacione características espaciais das figuras, podendo até ser considerada detecção de pontos de interesse e extração de características;
- 3) **BoVW:** Teste com outras estruturas de detecção de pontos de interesse além das características KAZE. Integrar as imagens centralizadas no disco óptico no *dataset* para treinamento, testando também abordagem sem recortes;
- 4) **Classificação de Doenças:** Empregar as técnicas do presente trabalho para classificação de vasta gama de patologias oculares em olhos não saudáveis;
- 5) **Dataset:** Integrar a utilização de outro *dataset* a fim de aumentar a quantidade de amostras sem ruído existentes para treinamento e consequente melhoria da eficiência, dado que inúmeras exclusões foram necessárias para eliminar exemplares ruins do banco empregado;
- 6) **Implementação:** Futuramente, pode-se buscar integrar o presente sistema em dispositivos inteligentes que façam escaneamento imediato da retina e permitam a detecção de possíveis anomalias em pacientes.

REFERÊNCIAS

- ALCANTARILLA, P. F.; BARTOLI, A.; DAVISON, A. J. KAZE Features. *In: EUROPEAN CONFERENCE ON COMPUTER VISION – ECCV, 21., 2012, Florença. Anais [...]* Heidelberg: Springer, 2012. p. 214-227.
- AL-OBAYDY, W. N. I.; SUANDI, S. A. Open-set single-sample face recognition in video surveillance using fuzzy ARTMAP. **Neural Computing and Applications**, Londres, v. 32, p. 1405-1412, 2020.
- AMIDI, A.; AMIDI, S. **The evolution of image classification explained**. Stanford Educational Blog. Disponível em: <https://stanford.edu/~shervine/blog/evolution-image-classification-explained>. Acesso em: 02/06/2023.
- AMORIM, D. G. **Redes ART com categorias internas de geometria irregular**. 2006. 352 p. Tese (Doutorado em Física) – Departamento de Eletrônica e Computação, Universidade de Santiago de Compostela, Santiago de Compostela, 2006.
- ANAGNOSTOPOULOS, G.; GEORGIOPOULOS, M. Category regions as new geometrical concepts in Fuzzy-ART and Fuzzy-ARTMAP. **Neural Networks**, Oxford, v. 15, n. 10, p. 1205-1221, 2002.
- ARTHUR, D.; VASSILVITSKII, S. k-means++: The Advantages of Careful Seeding. *In: PROCEEDINGS OF THE EIGHTEENTH ANNUAL ACM-SIAM SYMPOSIUM ON DISCRETE ALGORITHMS (SODA '07), 18., 2007, Nova Orleans. Anais [...]* Philadelphia: Society for Industrial and Applied Mathematics, 2007. p. 1027-1035.
- ATWANY, M. Z.; SAHYOUN, A. H.; YAQUB, M. Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey. **IEEE Access**, Piscataway, v. 10, p. 28642-28655, 2022.
- BAY, H.; ESS, A.; TUYTELAARS, T.; GOOL, L. V. Speeded-Up Robust Features (SURF). **Computer Vision and Image Understanding**, Nova Iorque, v. 110, n. 3, p. 346-359, 2008.
- CANO-IZQUIERDO, J. M.; SÁNCHEZ, E. G.; BRAVO, M. J. A.; PINZOLAS, M.; IBARROLA, J. How well Fuzzy ARTMAP approximates functions? **Journal of Intelligent & Fuzzy Systems**, Amsterdam, v. 25, n. 2, p. 335-350, 2013.
- CARPENTER, G. A.; GROSSBERG, S. A massively parallel architecture for a self-organizing neural pattern recognition machine. **Computer Vision, Graphics, and Image Processing**, San Diego, v. 37, n. 1, p. 54-115, 1987.
- CARPENTER, G. A.; GROSSBERG, S. ART2: Stable self-organization of pattern recognition codes for analog input patterns. **Applied Optics**, Washington, v. 26, n. 23, p. 4919–4930, 1987.
- CARPENTER, G. A.; GROSSBERG, S. Self-organizing neural network architectures for real-time adaptive pattern recognition. *In: INTERNATIONAL SYMPOSIUM ON*

NEURAL AND SYNERGETIC COMPUTERS, 1., 1988, Krün. **Anais [...]** Heidelberg: Springer-Verlag, 1988. p. 42-74.

CARPENTER, G. A.; GROSSBERG, S.; ROSEN, D. B. Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. **Neural Networks**, Oxford, v. 4, n. 6, p. 759–771, 1991.

CARPENTER, G. A.; GROSSBERG, S.; REYNOLDS, J. H. ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. **Neural Networks**, Oxford, v. 4, n. 5, p. 565-588, 1991.

CARPENTER, G. A.; GROSSBERG, S.; REYNOLDS, J. H.; ROSEN, D. B.; MARKUZON, N. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. **IEEE Transactions on Neural Networks**, Piscataway, v. 3, n. 5, p. 698-713, 1992.

CHEN, T.; GUESTRIN, C. XGBoost: A Scalable Tree Boosting System. *In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, 22., 2016, San Francisco. **Anais [...]** Nova Iorque: Association for Computing Machinery, 2016. p. 785-794.

DALAL, N.; TRIGGS, B. Histograms of Oriented Gradients for Human Detection. *In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR'05)*, 1., 2005, San Diego. **Anais [...]** [S.l]: IEEE, 2005. p. 886-893.

DENG, L.; YU, D. **Deep Learning: methods and applications**. Boston: Now Publishers, 2014. v. 7. 195 p.

EUROPEAN CONFERENCE ON MACHINE LEARNING – ECML-95, 8., Iráclio. **Anais [...]** Heidelberg: Springer, 1995. v. 912.

GEORGIOPOULOS, M.; HUANG, J.; HEILEMAN, G. Properties of learning in ARTMAP. **Neural Networks**, Oxford, v. 7, n. 3, p. 495-506, 1994.

GEORGIOPOULOS, M.; FERNLUND, H.; BEBIS, G.; HEILEMAN, G. L. Order of Search in Fuzzy ART and Fuzzy ARTMAP: Effect of the Choice Parameter. **Neural Networks**, Oxford, v. 9, n. 9, p. 1541-1559, 1996.

GONZALEZ, R. C; WOODS, R. E. **Processamento Digital de Imagens**. 3. ed. São Paulo: Pearson Universidades, 2009. 624 p.

GROSSBERG, S. Adaptive Pattern Classification and Universal Recoding: I. Paralel development and coding of neural feature detectors. **Biological Cybernetics**, Heidelberg, v. 23, p. 121-134, 1976.

GROSSBERG, S. Adaptive Pattern Classification and Universal Recoding: II. Feedback, Expectation, Olfaction, Illusions. **Biological Cybernetics**, Heidelberg, v. 23, n. 4, p. 187-202, 1976.

GROSSBERG, S. Nonlinear neural networks principles, mechanisms and architectures. **Neural Networks**, Oxford, v. 1, n. 1, p. 17-61, 1988.

GROSSBERG, S.; STONE, G. Neural Dynamics of Word Recognition and Recall: Attentional Priming, Learning, and Resonance. **Psychological Review**, Washington, v. 93, n. 1, p. 46-74, 1986.

HAYKIN, S. **Neural networks: a comprehensive foundation**. 2. ed. Upper Saddle River: Prentice-Hall, 1999. 842 p.

HIRSCH, R. **Exploring colour photography: a complete guide**. Londres: Laurence King Publishing, 2004. 360 p.

JOLLIFE, I. T. **Principal component analysis**. 2. ed. Nova Iorque: Springer, 2002. 488 p.

KARTALOPOULOS, S. V. **Understanding neural networks and fuzzy logic: basic concepts and applications**. 1. ed. Hoboken: Wiley-IEEE Press, 1995. 232 p.

KIM, J. et al.; Machine learning predicting myopic regression after corneal refractive surgery using preoperative data and fundus photography. **Graefe's Archive for Clinical and Experimental Ophthalmology**, Heidelberg, v. 260, n. 11, p. 3701-3710, 2022.

KRZANOWSKI, W. J. **Principles of multivariate analysis**. Revised edition. Oxford: Oxford University Press, 2000. 608 p.

LIU, C.; WECHSLER, H. Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition. **IEEE Transactions on Image Processing**, Nova Iorque, v. 11, n. 4, p. 467-476, 2002.

LOPES, M. L. M. **Desenvolvimento de redes neurais para previsão de cargas elétricas de sistemas de energia elétrica**. 2005. 169 p. Tese (Doutorado em Engenharia Elétrica) — Faculdade de Engenharia, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Ilha Solteira, 2005.

LOWE, D. G. Distinctive Image Features from Scale-invariant Keypoints. **International Journal of Computer Vision**, Amsterdam, v. 60, n. 2, p. 91-110, 2004.

MATHWORKS Image Recognition. What is Image Recognition? Disponível em: <https://www.mathworks.com/discovery/image-recognition-matlab.html>. Acesso em: mai. 2023.

MATHWORKS Help Center. Imread: Read image from graphics file. Disponível em: <https://www.mathworks.com/help/matlab/ref/imread.html>. Acesso em: mar. 2023.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, Nova Iorque, v. 5, n. 4, p. 115–133, 1943.

OKAWA, M. From BoVW to VLAD with KAZE features: Offline signature verification considering cognitive processes of forensic experts. **Pattern Recognition Letters**, Amsterdam, v. 113, p. 75-82, 2018.

PACHADE, S. et al. Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research. **MDPi Data**, Basel, v. 6, n. 2, 14, 2021.

PIURI, V.; RAJ, S.; GENOVESE, A.; SRIVASTAVA, R. **Trends in deep learning methodologies: algorithms, applications, and systems**. 1. ed. Cambridge: Academic Press, 2020. 306 p.

RAN, A. R et al. Deep learning in glaucoma with optical coherence tomography: a review. **Eye: The Scientific Journal of the Royal College of Ophtalmologists**, Londres, v. 35, n. 1, p. 188-201, 2021.

SENGAR, N.; JOSHI, R. C.; DUTTA, M. K.; BURGET, R. EyeDeep-Net: a multi-class diagnosis of retinal diseases using deep neural network. **Neural Computing and Applications**, Londres, v. 35, p. 10551-10571, 2023.

TUFAIL, A. B. et al. Diagnosis of Diabetic Retinopathy through Retinal Fundus Images and 3D Convolutional Neural Networks with Limited Number of Samples. **Wireless Communications and Mobile Computing**, Londres, v. 2021, 15 p, 2021. Disponível em: <https://doi.org/10.1155/2021/6013448>. Acesso em: 02 fev. 2023.

VIMINA, E. R.; JACOB, K. P. Feature fusion method using BoVW framework for enhancing image retrieval. **IET Image Processing**, Londres, v. 13, p. 1979-1985, 2019.

WASSERMAN, P. D. **Neural computing: theory and practice**. Nova Iorque: Van Nostrand Reinhold, 1989. 230 p.

WEENINK, D. Category ART: a Variation on Adaptive Resonance Theory Neural Networks. *In*: IFA PROCEEDINGS, 21, 1997, Amsterdam. **Anais [...]**. Amsterdam: [s.n], 1997. p. 117-129.

ZADEH, L. A. Fuzzy sets. **Information and control**. Pequim, v. 8, n. 3, p. 338-353, 1965.