

HIGEIA-DA: reconhecimento de neuroImagens para diagnóstico antecipado da Doença de Alzheimer com base em redes neurais

Nathalia C. dos Santos, Matheus Henrique S. Vieira e Dianne S. V. Medeiros

Resumo— O diagnóstico antecipado da Doença de Alzheimer (DA) permite retardar seu efeito degenerativo progressivo. Esse diagnóstico é desafiador devido à grande semelhança entre o cérebro de uma pessoa saudável com envelhecimento natural e o de uma pessoa em estágio muito inicial da doença. Este artigo propõe o HIGEIA-DA, um classificador multiclasse que tem o objetivo de auxiliar médicos especialistas no diagnóstico antecipado da DA. O HIGEIA-DA é composto por três Classificadores Binários (CBs) formados por redes neurais convolucionais e *perceptron* multicamadas para diferenciar três níveis de severidade da DA. A classificação final é obtida pela combinação das classificações dos CBs utilizando a média. Verifica-se a influência de diversos hiperparâmetros e da estrutura da rede neural no desempenho do HIGEIA-DA. Os resultados mostram que os hiperparâmetros taxa de aprendizado, quantidade de blocos e número de camadas densas apresentam influência mais significativa no desempenho do classificador do que os outros hiperparâmetros. A estratégia de combinação é avaliada para o conjunto de teste e em uma simulação de uso real, levando a uma sensibilidade média para cada classe maior do que 70% no conjunto de teste e maior do que 50% na simulação.

Abstract— The early diagnosis of Alzheimer’s Disease (AD) allows delaying the disease’s progressive degenerative effect. This diagnosis is challenging due to the great similarity between the brain of a healthy person with natural aging and the brain of a person at a very early stage of AD. This paper proposes HIGEIA-DA, a multiclass classifier that aims to help specialist physicians in the early diagnosis of AD. HIGEIA-DA is composed of three Binary Classifiers (CBs) formed by convolutional neural networks and multilayer *perceptrons* to differentiate three levels of AD severity. The final classification is obtained by combining the classifications of the CBs using the average. The influence of several hyperparameters and the structure of the neural network on the performance of HIGEIA-DA is verified. Results show that the hyperparameters *batch*, learning rate and number of blocks have a more significant influence on the classifier’s performance. The combination strategy is evaluated for a test dataset and a real use simulation, leading to an average sensibility for each class greater than 70% for the test dataset and greater than 50% for the simulation.

I. INTRODUÇÃO

A crescente expectativa de vida da população tem como consequência o aumento na ocorrência de doenças relacionadas à idade, dentre elas a Doença de Alzheimer (DA). Em seu estágio mais avançado, a DA torna o paciente incapaz de

realizar as tarefas diárias devido à degeneração progressiva dos neurônios. O diagnóstico antecipado da doença permite introduzir tratamentos ao paciente a fim de reduzir os efeitos degenerativos da doença. No entanto, esse diagnóstico antecipado é desafiador porque os exames de Ressonância Magnética por Imagem (RMI) ou de Tomografia Computadorizada (TC) de cérebros saudáveis e no estágio inicial de DA são muito semelhantes [1]. Nesse contexto, é fundamental desenvolver mecanismos que permitam auxiliar no diagnóstico antecipado.

Sistemas para diagnóstico de doenças assistido por computador desenvolvidos entre as décadas de 1970 e 1990 são baseados em regras e dependem fortemente do conhecimento de um especialista humano. Posteriormente, esses sistemas evoluíram tornando-se sistemas supervisionados que utilizam características extraídas de dados médicos para alimentar modelos de aprendizado de máquina, com intervenção de um especialista humano para determinar as características importantes para o diagnóstico [2]. Modelos de aprendizado profundo são amplamente utilizados para análise de imagens visando o reconhecimento de padrões sem a necessidade do envolvimento humano [3] e já foram aplicados com sucesso para análise de imagens de exames médicos, com resultados promissores para detecção e classificação de doenças. No entanto, os conjuntos de dados de neuroimagem disponíveis normalmente são pequenos, tornando-se desafiador desenvolver modelos de aprendizado profundo que possam extrair características úteis para a detecção da DA utilizando poucas imagens.

Este artigo propõe o HIGEIA-DA, um classificador baseado em Redes Neurais Artificiais (*Artificial Neural Networks* - ANNs) que reconhece neuroimagens para auxiliar no diagnóstico antecipado da DA. O classificador é composto por três Classificadores Binários (CBs), $C_{i,j}$, formados por uma Rede Neural Convolucional (*Convolutional Neural Network* - CNN) e uma rede *Perceptron* Multicamada (*Multilayer Perceptron* - MLP). Até onde se sabe, não foi encontrado nenhum trabalho que utilize tanto CNN quanto MLP para a construção de um classificador multiclasse, variando diversos hiperparâmetros para avaliar, por meio da simulação de um ambiente de produção, o desempenho do classificador. A ideia é fornecer ao médico especialista uma ferramenta adicional para melhorar a assertividade do diagnóstico antecipado da DA e possibilitar o início do tratamento precoce da doença. Os CBs distinguem três classes de neuroimagens, de acordo com a intensidade da demência: Leve vs Muito Leve ($C_{2,1}$), Leve vs Sem Demência ($C_{2,0}$) e Muito Leve vs Sem De-

Nathalia C. dos Santos, Matheus Henrique S. Vieira, Dianne S. V. Medeiros, Departamento de Engenharia de Telecomunicações, Universidade Federal Fluminense, Niterói-RJ, e-mail: ricardomf@id.uff.br; menezes@midia.com.uff.br; diannescherly@id.uff.br. Este trabalho foi parcialmente financiado por CNPq, CAPES, FAPERJ, FAPESP (2018/23062-5), Prefeitura Municipal de Niterói/FEC/UFF (PDPA 2020) e RNP.

mência ($C_{1,0}$). Cada classificador é treinado individualmente, gerando modelos que são agregados utilizando a integral de *Choquet* [4]. Posteriormente, as classificações são combinadas utilizando a média para gerar a classificação final. Verifica-se a influência de hiperparâmetros e da estrutura da rede no desempenho dos CBs. Os resultados mostram que há uma correlação fraca entre os hiperparâmetros *batch*, taxa de aprendizado e quantidade de blocos e as métricas de análise de desempenho, indicando que esses hiperparâmetros têm influência mais significativa no desempenho dos CBs quando comparados a outros hiperparâmetros testados. O classificador final alcança uma sensibilidade média maior do que 50% para cada classe utilizando a estratégia de combinação em um ambiente que simula o uso real do classificador. Já a sensibilidade média considerando os conjuntos de teste apenas chega a pelo menos 70%.

O restante deste artigo está organizado como segue. A Seção II discute o desafio do diagnóstico da DA e trabalhos relacionados. A Seção III apresenta o classificador HIGEIA-DA proposto. Na Seção IV, a metodologia experimental e o conjunto de dados são apresentados. A Seção V discute os resultados obtidos. Por fim, a Seção VI conclui o artigo e apresenta direções de pesquisa futura.

II. DIAGNÓSTICO ANTECIPADO DA DOENÇA DE ALZHEIMER

A demência é uma manifestação comum da DA, com mais de 80% dos casos provocados por essa doença [3]. A demência provoca perda de memória e outras deficiências nas funções cognitivas [5], gerando transtornos físicos e psicológicos aos indivíduos. No primeiro estágio da DA, há perda de memória, mudança de personalidade e nas habilidades de locomoção. No estágio moderado, há dificuldade na fala e na realização de atividades diárias, além de agitação e insônia. No estágio avançado, o paciente apresenta incontinência urinária e fecal, tendo dificuldades para realizar tarefas diárias devido à deficiência motora progressiva. O último estágio é o terminal, caracterizado pela perda de memória e dependência de terceiros [5]. A DA não tem cura e a confirmação do diagnóstico ocorre no momento da autópsia [6]. No entanto, existem tratamentos que reduzem a sua progressão, melhorando a qualidade de vida do paciente. Para que o tratamento seja mais eficaz, é importante o diagnóstico antecipado da doença [6]. Assim, a ideia da proposta deste artigo é auxiliar os médicos especialistas a identificarem pacientes que estão no início do primeiro estágio da DA para que o tratamento comece precocemente.

O diagnóstico é comumente feito por um especialista por meio da análise de exames de RMI ou de TC por Emissão de Póstron, juntamente com o histórico e sintomas do paciente. Modelos de aprendizado de máquina vêm sendo usados nas últimas décadas para desenvolver classificadores de DA. O processo geral do diagnóstico pode ser agrupado em três fases. Na primeira, há uma pré-determinação das regiões de interesse. Na segunda, há a extração de características a partir das regiões de interesse. Na terceira, ocorre a construção do modelo de classificação com base nas características extraídas. As características são geradas e extraídas manualmente dos

dados de imagem de exames RMI [7]. Dessa forma, há grande dependência do fator humano e o conhecimento do especialista pode afetar o desempenho dos classificadores desenvolvidos. Além disso, há uma grande quantidade de características e uma pequena quantidade de amostras de imagens médicas, dificultando o treinamento dos classificadores. Trabalhos propõem soluções para facilitar a seleção de características [8], [9], ao custo de aumentarem a complexidade dos mecanismos de classificação propostos.

A DA apresenta um padrão de progressão de dano ao tecido do cérebro. A redução média anual esperada para o hipocampo em indivíduos saudáveis está entre 0,24 e 1,73%. Em pacientes com DA, essa redução fica entre 2,2 e 5,9%. A degeneração das células neurais apresenta baixa intensidade em imagens de exames de RMI [2] e como é um efeito esperado com o avanço da idade, em alguns casos o especialista não consegue distinguir, por meio do exame, entre pacientes saudáveis mas com envelhecimento e aqueles em estágio inicial de DA. Isso dificulta a extração de características porque as que estão presentes nas imagens dos exames podem não apresentar claramente as evidências da doença [5]. Nesse contexto, modelos de aprendizado profundo auxiliam no diagnóstico da doença, ao detectar padrões dificilmente percebidos por humanos e sem a necessidade de extração manual de características para alimentar um classificador. Dentre as diversas técnicas utilizadas para classificação e diagnóstico da DA, destacam-se redes MLP [10] e CNNs [11]. Um desafio para o desenvolvimento de modelos baseados em aprendizado profundo é a escassez dos conjuntos de dados de neuroimagens disponíveis e a pouca quantidade de amostras contidas neles [2], o que afeta o desempenho dos classificadores.

Abordagens baseadas em CNNs têm alcançado elevada precisão na classificação, tornando esse tipo de rede promissora para aplicações de diagnóstico assistido por computador [7]. Silva *et al.* utilizam CNNs para classificação e diagnóstico de DA. Os autores variam a quantidade de épocas para avaliar o desempenho do classificador [5]. Diferentemente, o HIGEIA-DA utiliza uma CNN associada a uma rede MLP para classificar três níveis de demência. Adicionalmente, este artigo avalia a influência de outros hiperparâmetros no desempenho do classificador, como *batch* e taxa de aprendizado, buscando relações entre o desempenho do classificador e os hiperparâmetros investigados. O HIGEIA-DA conta com dois níveis de agregação a fim de simplificar o desenvolvimento de um classificador multiclasse. Ji *et al.* [7] propõem um conjunto (*ensemble*) de classificadores baseado em CNN, que combina redes conhecidas já pré-treinadas. O problema tratado é construído como uma classificação binária e a proposta alcança sensibilidade maior que 80% e acurácia maior que 88%. Diferentemente, este artigo modela o problema como uma classificação multiclasse e investiga o desempenho da proposta simulando um ambiente de produção. De forma semelhante, Islam *et al.* também utilizam CNN, tratando o problema como uma classificação multiclasse. Os autores utilizam apenas uma CNN, cuja camada de saída utiliza a função de ativação *Softmax*, sendo remodelada para apresentar quatro saídas. A proposta alcança uma acurácia de 73,75%. No entanto, os autores não se preocupam com o desbalanceamento

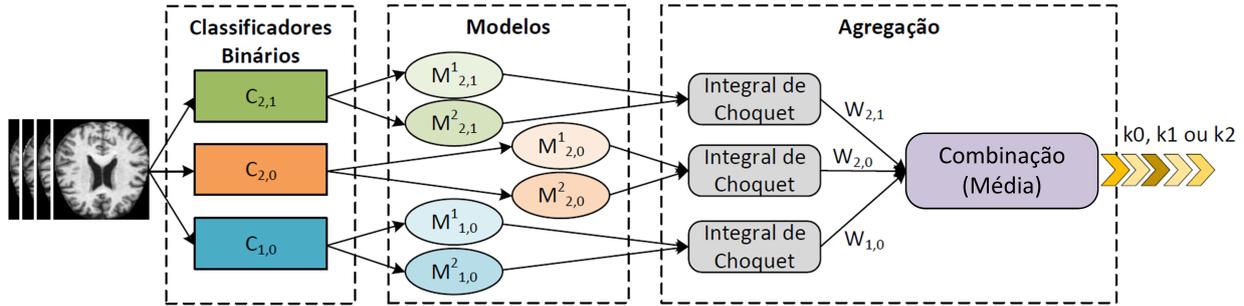


Fig. 1. Esquema do classificador HIGEIA-DA, composto por 3 classificadores binários.

entre as classes.

III. CLASSIFICADOR HIGEIA-DA

Este artigo propõe um classificador baseado em redes neurais artificiais, HIGEIA-DA, que tem o objetivo de classificar neuroimagens de acordo com três níveis de severidade da DA em seu estágio inicial. No primeiro nível, k_0 , o paciente é considerado como Sem Demência. No segundo nível, k_1 , o paciente possui Demência Muito Leve e, no terceiro nível, k_2 , o paciente possui Demência Leve. A fim de simplificar o desenvolvimento do classificador, adota-se a estratégia apresentada por Basaia *et al.* [12], na qual o problema de classificação multiclasse para n classes é dividido em $n(n-1)/2$ problemas de classificação binária. Dessa forma, o classificador HIGEIA-DA é formado por 3 classificadores binários, $C_{2,1}$, $C_{2,0}$ e $C_{1,0}$, cujas redes neurais diferenciam, respectivamente, as classes k_2 e k_1 , k_2 e k_0 , e k_1 e k_0 . A Figura 1 mostra o esquema do classificador proposto. Para cada classificador $C_{i,j}$, obtêm-se dois modelos $M_{i,j}^1$ e $M_{i,j}^2$ a partir do treinamento das redes neurais. Os modelos $M_{i,j}^1$ e $M_{i,j}^2$ são agregados utilizando a integral de Choquet [4], resultando em 3 vetores coluna de classificação $W_{2,1}$, $W_{2,0}$ e $W_{1,0}$, cujos elementos representam a probabilidade de ocorrência de cada classe. Por fim, os 3 vetores são combinados utilizando a média.

A. Arquitetura dos classificadores binários

A CNN que compõe os CBs é construída com um *kernel* de dimensão 3×3 . Utilizam-se 32 mapas de características. Após a convolução, a função de ativação ReLu é aplicada ao mapa de características, extraindo as informações mais relevantes da imagem. A CNN do HIGEIA-DA possui até três blocos dependendo do CB, sendo que cada bloco contém duas camadas de convolução bidimensional. A fim de reduzir a dimensão dos mapas de características, utiliza-se um *pool size* de dimensão 3×3 . Escolhe-se um passo de dimensão 2×2 , reduzindo à metade a dimensão dos mapas de características. As transformações sobre as regiões de dimensão definida pelo *pool size* é feita utilizando o *max pooling*, visto que as imagens utilizadas possuem fundo escuro e os dados de interesse têm *pixels* mais claros. O *dropout* utilizado é de 0,2 [13].

A rede MLP recebe os dados de saída da CNN, gerados pela etapa de *flattening*. Define-se a ReLu como função de ativação para cada camada oculta e a *Softmax* para a camada de saída. Utiliza-se um *dropout* de 0,2 nas camadas densas [13]. O objetivo é reduzir as chances de ocorrer sobreajuste durante

o treinamento [13]. O último hiperparâmetro relacionado à arquitetura do HIGEIA-DA é o *batch*. Deve-se buscar um valor ótimo para o *batch* de maneira que seja maior que 1 e menor que o número total de imagens utilizadas no treinamento.

B. Agregação dos classificadores binários

Os resultados obtidos para cada CB são combinados para gerar uma classificação final. Para agregar os modelos individuais gerados pelos CBs, este artigo utiliza o método proposto por Chen *et al.* [14], em que a agregação é feita por meio da integral de Choquet, que é capaz de considerar no processo de agregação não somente a importância de cada classificador, mas também a interação entre eles. A integral de Choquet considera a saída individual de cada classificador como base para obtenção de uma medida difusa (*fuzzy*), que considera o grau de incerteza de uma classificação [4]. Assim, para cada classificador binário $C_{i,j}$ são obtidos dois modelos $M_{i,j}^1$ e $M_{i,j}^2$, que são transformados em uma matriz de decisão, que contém o percentual de acerto de cada classe para cada modelo. Para cada linha da matriz de decisão, calcula-se a integral de Choquet, resultando em um vetor coluna $W_{i,j}$. Cada elemento do vetor representa a probabilidade de ocorrência de cada classe após a agregação dos modelos. A classificação dada pelos modelos agregados é definida pelo maior valor dos elementos do vetor. Os vetores coluna são posteriormente combinados utilizando a média para determinar a classificação final.

IV. METODOLOGIA EXPERIMENTAL E CONJUNTO DE DADOS

O ajuste dos hiperparâmetros dos CBs do HIGEIA-DA é feito de forma empírica. Assim, investiga-se a influência no desempenho do classificador da quantidade de camadas convolucionais, do tamanho do *batch*, da taxa de aprendizado, da quantidade de camadas densas e da quantidade de nós em cada camada densa. Essa investigação é feita separadamente para cada variável, variando-se aquela investigada e mantendo-se fixas as demais. Cada experimento é repetido 5 vezes. O ponto de partida para uma nova variação é sempre o melhor resultado do conjunto de variações feitas com o hiperparâmetro investigado, mantendo fixos os demais. São realizados 581 treinamentos da rede neural, totalizando aproximadamente 2105 horas de treinamento. Para cada treinamento, define-se um limite de 200 épocas. Utiliza-se o algoritmo Adam como otimizador, responsável pela atualização dos pesos da rede

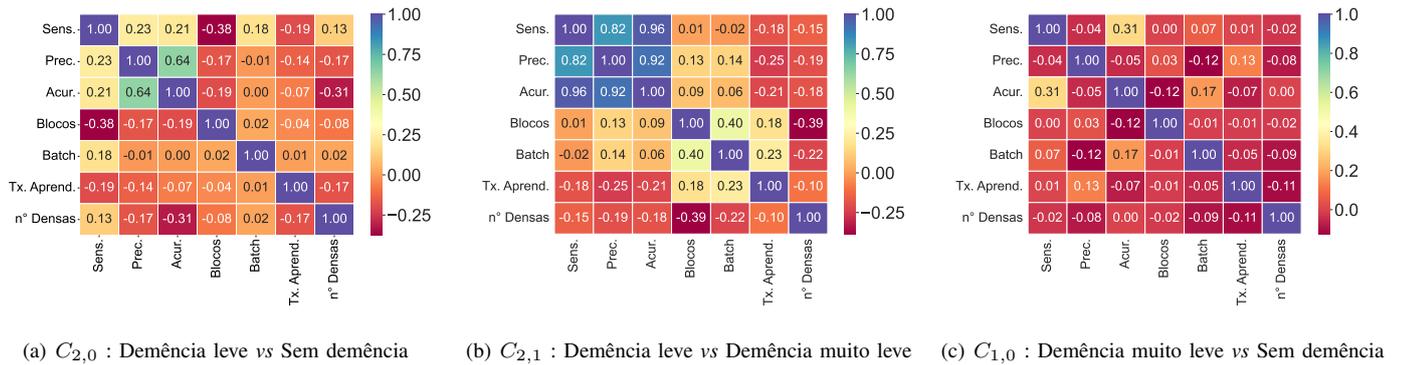


Fig. 2. Mapas de correlação para os CBs $C_{i,j}$. Há correlação fraca entre as métricas de desempenho e a quantidade de blocos, a taxa de aprendizado e o número de camadas densas. A correlação não é uniforme entre os CBs.

neural. A implementação do HIGEIA-DA é feita utilizando a linguagem *Python*. As principais bibliotecas utilizadas são *Pandas*, *NumPy*, *Matplotlib* e *Keras*.

O conjunto de dados utilizado está disponível publicamente¹ e é composto por neuroimagens rotuladas, separadas em exames de pessoas saudáveis e de pessoas doentes. Existem quatro classes alvo, rotuladas como: Demência Moderada, Demência Leve, Demência Muito Leve e Sem Demência. Para cada classe, estão disponíveis um conjunto de dados para treinamento e um para simulação. O conjunto de treinamento possui 52, 717, 1.792 e 2.560 imagens para as classes Demência Moderada, Leve, Muito Leve e Sem Demência, respectivamente. Já o conjunto de simulação possui 12, 179, 448 e 640 imagens das respectivas classes. A classe Demência Moderada não é considerada neste artigo porque o objetivo é identificar pacientes em estágio inicial de DA e, portanto, com nível de demência leve ou muito leve. A fim de diminuir o risco de sobreajuste na rede devido ao desbalanceamento entre as classes no conjunto de dados utilizado [12], que pode especializar o classificador na detecção da classe com maior número de imagens, apenas um subconjunto das neuroimagens disponíveis é utilizado para treinar cada CB. As imagens no arquivo de treinamento são utilizadas para desenvolver o modelo do classificador. Para tanto, 90% das imagens são separadas para treino, enquanto 10% das imagens são separadas para teste (\mathcal{T}). Dentre as imagens separadas para treino, 90% são utilizadas no treinamento e 10% são utilizadas para validação do modelo a cada época. Ao final do treino, o desempenho dos CBs é avaliado utilizando o conjunto de teste \mathcal{T} . Após obtenção de um modelo que apresente desempenho satisfatório, realiza-se um novo teste utilizando as imagens do arquivo de simulação disponibilizado no conjunto de dados, simulando o uso do classificador binário em um cenário real. Essas imagens são inéditas para os CBs e todas as imagens, independentemente do rótulo, alimentam os CBs.

Cada neuroimagem representa a mesma seção transversal de cérebros de pacientes distintos. As imagens possuem 176×208 pixels e, apesar de terem boa qualidade e estarem pré-tratadas, algumas ações de pré-processamento adicionais são executadas. A moldura preta da imagem é minimizada, reduzindo o tamanho da imagem para 141×174 pixels. As

imagens estão em escala RGB (*Red*, *Green*, *Blue*), porém os canais das cores não carregam informação. Assim, converte-se todas as imagens em escala de cinza, de forma a apresentar apenas um canal, sem trazer prejuízo para o treinamento. Por fim, normaliza-se os valores dos pixels de cada imagem para variar entre 0 e 1.

V. RESULTADOS

O desempenho dos CBs é avaliado a partir da acurácia, da precisão e da sensibilidade. A sensibilidade é a métrica mais relevante neste artigo, visto que os Falsos Negativos (FNs) são considerados mais prejudiciais que os Falsos Positivos (FPs). Verifica-se a correlação entre as métricas de desempenho e os hiperparâmetros tamanho do *batch*, taxa de aprendizagem, quantidade de blocos, número de camadas densas e as quantidade de nós em cada camada densa. Esses hiperparâmetros são variados para cada CB. A ideia é investigar a influência desses hiperparâmetros no desempenho dos CBs antes da composição do classificador final. Os seguintes conjuntos de valores são utilizados para cada hiperparâmetro: tamanho do *batch* $\in \{3; 16; 32; 64; 128; 150; 200; 220; 256\}$, taxa de aprendizagem $\in \{1e^{-5}; 5e^{-5}; 0,0001; 0,004; 0,006; 0,008; 0,001; 0,002; 0,01; 0,1; 0,4\}$, quantidade de blocos $\in \{1; 2; 3; 4\}$, número de camadas densas $\in \{1; 2; 3; 4; 5; 6; 7; 8\}$ e a quantidade de nós em cada camada densa $\in \{0; 20; 50; 100; 200; 400\}$.

A Figura 2 mostra um mapa de calor com a correlação de *Pearson* calculada entre os hiperparâmetros variados e a acurácia, sensibilidade e precisão obtidas para cada CB a partir dos resultados de teste. Observa-se que existe uma correlação positiva fraca entre as métricas de avaliação de desempenho e a quantidade de blocos, taxa de aprendizado e número de camadas densas. Essa correlação depende do CB analisado, não sendo uniforme entre os CBs. A existência dessa correlação, mesmo sendo fraca, é um indicativo de que os hiperparâmetros citados têm maior influência sobre o desempenho do classificador proposto quando comparados aos outros hiperparâmetros testados.

A escolha pelos melhores modelos dos CBs depende da combinação feita entre eles para compor o classificador final. São utilizados dois modelos de cada CB. Os 2 modelos, não necessariamente ótimos, de cada CB são unificados pela

¹<https://www.kaggle.com/tourist55/alzheimers-dataset-4-class-of-images>

TABELA I
HIPERPARÂMETROS UTILIZADOS NA OBTENÇÃO DOS MODELOS
UTILIZADOS NO CLASSIFICADOR FINAL.

$C_{i,j}$	Blocs.	Batch	Tx. Aprend.	# nós cam. densas	Sensib.
$C_{2,0}$	1	128	0,001	[100, 100, 150]	0,82
$C_{2,0}$	1	128	0,0001	[50, 100, 150]	0,69
$C_{2,1}$	2	32	0,006	[400, 200, 100]	0,60
$C_{2,1}$	2	32	0,006	[400, 200, 100]	0,60
$C_{1,0}$	3	32	0,001	[400, 200, 100]	0,72
$C_{1,0}$	2	32	0,004	[400,200, 100]	0,63

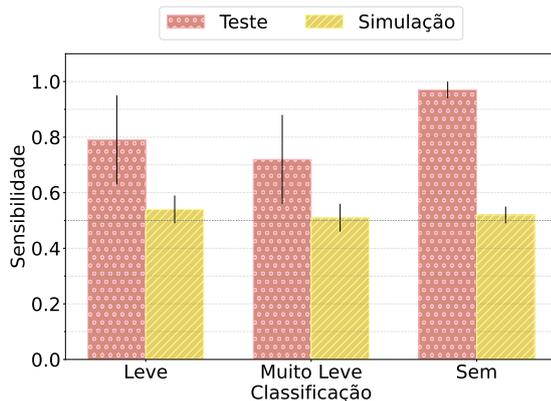


Fig. 3. Sensibilidade do HIGEIA-DA para o conjunto de teste e para a simulação.

integral de Choquet. A utilização apenas de modelos ótimos não resulta em bom desempenho para o classificador final. Dessa forma, são combinados seis modelos para gerar o classificador final. Os hiperparâmetros que geram os modelos individuais unificados são destacados na Tabela I. A escolha é feita em um total de 195 treinamentos de $C_{2,0}$, 90 de $C_{2,1}$ e 124 de $C_{1,0}$. Os resultados obtidos para a sensibilidade a partir da média dos modelos individuais estão apresentados na Figura 3 tanto para o conjunto de teste quanto para a simulação. Destaca-se que na simulação os CBs individuais são alimentados com imagens que eles não necessariamente sabem classificar, mas espera-se que o classificador final seja capaz classificá-las por meio da combinação dos resultados dos CBs individuais utilizando a média.

O intervalo de confiança do teste representado na Figura 3 é obtido por meio das 5 repetições realizadas em cada modelo. Já o da simulação é obtido dividindo-se o conjunto de imagens em subconjuntos de 10 imagens e verificando a sensibilidade para cada subconjunto, considerando a quantidade de imagens para simulação. Logo, utilizam-se 170 imagens k_2 , 440 k_1 e 640 k_0 . A sensibilidade é calculada usando a heurística “Um-vs-Todos” (One-vs-All - Ova).

VI. CONCLUSÃO

Este artigo propôs o HIGEIA-DA, um classificador multiclasse que visa diferenciar níveis de severidade da Doença de Alzheimer (DA) para auxiliar no diagnóstico antecipado da doença. O classificador é construído com base em classificadores binários formados por uma CNN e uma MLP, cujos hiperparâmetros são ajustados de forma empírica por meio de diversos experimentos. O processo de ajuste foi feito

empiricamente e verificou-se a correlação entre os hiperparâmetros e as métricas de desempenho avaliadas. Essa análise mostrou que os hiperparâmetros quantidade de blocos, taxa de aprendizado e número de camadas densas têm maior influência no desempenho do que outros hiperparâmetros. O classificador final formado a partir de CBs construídos com modelos não necessariamente ótimos permitiu alcançar uma sensibilidade média acima de 50% para cada classe utilizando a abordagem da média, enquanto o teste alcançou sensibilidade média acima de 70%. Vislumbra-se como trabalhos futuros, o refinamento do classificador utilizando pesos para unificar os modelos obtidos. Ademais, o uso de neuroimagens 3-D para treinar o classificador pode ajudar a alcançar melhor desempenho.

REFERÊNCIAS

- [1] M. K. Andrade, M. Santana, and W. Santos, “Avaliação do desempenho de classificadores inteligentes na detecção da doença de Alzheimer em imagens de ressonância magnética utilizando extratores de forma e textura,” in *Anais do II Simp. de Inovação em Eng. Biomédica, Recife, Brazil*, 2018.
- [2] J. Islam and Y. Zhang, “A novel deep learning based multi-class classification method for Alzheimer’s disease detection using brain MRI data,” in *Proc. of the Int. Conf. on Brain Informatics*. Springer, 2017, pp. 213–222.
- [3] S. Bringas, S. Salomón, R. Duque, C. Lage, and J. L. Montaña, “Alzheimer’s disease stage identification using deep learning models,” *Journal of Biomedical Informatics*, vol. 109, p. 103514, 2020.
- [4] A. G. Pacheco and R. A. Krohling, “Agregação de elenco de classificadores utilizando integral de Choquet com respeito a medida λ -fuzzy,” 2016.
- [5] I. R. Silva, R. G. de Souza, G. Silva, C. S. de Oliveira, L. H. Cavalcanti, R. S. Bezerra, A. d. A. Roberta, and W. P. dos Santos, “Utilização de redes convolucionais para classificação e diagnóstico da doença de Alzheimer,” 2018, pp. 73–76.
- [6] M. C. Maia, W. P. dos Santos, P. H. Pequeno, W. W. Azevedo, G. S. Silva, and M. A. de Santana, “Inteligência artificial para o apoio ao diagnóstico da doença de Alzheimer utilizando imagens de ressonância magnética,” *Anais do III Simp. de Inovação em Eng. Biomédica*, p. 47, 2019.
- [7] H. Ji, Z. Liu, W. Q. Yan, and R. Klette, “Early diagnosis of Alzheimer’s disease using deep learning,” in *Proc. of the 2nd Int. Conf. on Control and Computer Vision*, 2019, pp. 87–91.
- [8] X. Zhu, H.-I. Suk, L. Wang, S.-W. Lee, D. Shen, A. D. N. Initiative et al., “A novel relational regularization feature selection method for joint regression and classification in AD diagnosis,” *Medical image analysis*, vol. 38, pp. 205–214, 2017.
- [9] F. Liu, C.-Y. Wee, H. Chen, and D. Shen, “Inter-modality relationship constrained multi-modality multi-task feature selection for Alzheimer’s disease and mild cognitive impairment identification,” *NeuroImage*, vol. 84, pp. 466–475, 2014.
- [10] C. R. Munteanu, C. Fernandez-Lozano, V. M. Abad, S. P. Fernández, J. Álvarez-Linera, J. A. Hernández-Tamames, and A. Pazos, “Classification of mild cognitive impairment and Alzheimer’s disease with machine-learning techniques using 1H magnetic resonance spectroscopy data,” *Expert Systems with Applications*, vol. 42, no. 15-16, pp. 6205–6214, 2015.
- [11] M. Raju, V. P. Gopi, V. Anitha, and K. A. Wahid, “Multi-class diagnosis of Alzheimer’s disease using cascaded three dimensional-convolutional neural network,” *Physical and Engineering Sciences in Medicine*, vol. 43, no. 4, pp. 1219–1228, 2020.
- [12] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, A. D. N. Initiative et al., “Automated classification of Alzheimer’s disease and mild cognitive impairment using a single MRI and deep neural networks,” *NeuroImage: Clinical*, vol. 21, p. 101645, 2019.
- [13] C. C. Aggarwal et al., “Neural networks and deep learning,” *Springer*, vol. 10, pp. 978–3, 2018.
- [14] J.-F. Chen, Q. He, and Y. Li, “Different types of classifiers combination based on Choquet integral,” in *Proc. of the Int. Conf. on Systems, Man and Cybernetics*. IEEE, 2010, pp. 4056–4061.