
Radiômica e aprendizado de máquina - Uma assinatura para prever o status da codelecao 1p19q em gliomas de baixo grau

Radiomics and machine learning: A pipeline to predict 1p19q codeletion status in low-grade gliomas

Tony Alexandre Medeiros^{1,2*}, Raimundo Guimarães Saraiva Junior^{1,3}, Guilherme de Souza e Cassia⁴, Francisco Assis de Oliveira Nascimento¹, João Luiz Azevedo de Carvalho¹

RESUMO

Os gliomas compreendem a grande maioria de todos os tumores cerebrais malignos. A identificação do estado de codeleção cromossômica 1p/19q em gliomas de baixo grau (LGG) é de fundamental importância no manejo clínico. Este trabalho propõe uma assinatura de baixa complexidade para avaliação não invasiva do status 1p/19q a partir de imagens de ressonância magnética, usando radiômica e aprendizado de máquina. Os experimentos foram realizados usando imagens ponderadas em T2 de um conjunto de dados público, com 159 pacientes e status 1p/19q comprovado por biópsia. Um conjunto de 68 características quantitativas foi extraído de cada tumor usando a plataforma PyRadiomics. A normalização foi utilizada para que todas as características apresentem uma ordem de grandeza similar. A análise de componentes de vizinhança foi utilizada para redução de dimensionalidade. A previsão do status 1p/19q foi realizada usando um classificador de k-vizinhos mais próximos, com ajuste de hiperparâmetros usando busca em grade com validação cruzada. O desempenho da classificação foi avaliado usando várias métricas. Os resultados mostram que o desempenho da classificação melhora à medida que incorporamos cada etapa adicional a assinatura, e os melhores resultados foram obtidos ao executar a assinatura completa. Quando comparado com vários métodos de ponta da literatura e com resultados obtidos por neurocirurgiões e neurorradiologistas analisando as mesmas imagens, a assinatura proposta apresentou acurácia (81,88%), especificidade (69,25%), precisão (86,04%), área sob a curva (82,70%) e pontuação F1 (86,28%); o método proposto também apresentou alta sensibilidade (88,48%). Este trabalho demonstra técnicas baseadas em radiômica e aprendizado de máquina que fornecem avaliação não invasiva e promissora do status 1p/19q em LGGs. A assinatura proposta pode ser útil para facilitar a tomada de decisões clínicas, cirúrgicas e terapêuticas.

Palavras-chave: Radiômica; Aprendizado de máquina; Codeleção 1p/19q; Gliomas de baixo grau; Análise de componentes vizinhos.

ABSTRACT

Gliomas comprise the vast majority of all malignant tumors. Identification of the 1p/19q chromosomal co-deletion status in low-grade gliomas (LGG) is of fundamental importance in clinical management. This

¹ Departamento de Engenharia Elétrica, Universidade de Brasília

² Departamento de Informática, Instituto Federal Goiano – tony.medeiros@ifgoiano.edu.br

³ Departamento de Eletrônica, Instituto Federal do Ceará

⁴ Departamento de Diagnóstico por Imagem, Rede D'Or – Hospital Santa Luzia

work proposes a low-complexity signature for non-invasive assessment of 1p/19q status from magnetic resonance images, using radiomics and machine learning. The experiments were performed using T2-weighted images from a public dataset, with 159 patients and biopsy-detected 1p/19q status. A set of 68 quantitative features were extracted from each tumor using the PyRadiomics platform. Normalization was used so that all characteristics presented a similar order of magnitude. A neighborhood component analysis was used for dimensionality reduction. Prediction of 1p/19q status was performed using a k-nearest neighbor classifier, with hyperparameter adjustment using degree search with cross validation. Classification performance was evaluated using various metrics. The results show that classification performance improves as we incorporate each additional step into the pipeline, and the best results were obtained when running a full pipeline. When compared with several cutting-edge methods in the literature and with results obtained by neurosurgeons and neuroradiologists analyzing the same images, the proposed pipeline presented accuracy (81.88%), specificity (69.25%), precision (86.04%), area under the curve (82.70%) and F1 score (86.28%); the pipeline proposed also showed high sensitivity (88.48%). This work demonstrates safe techniques in radiomics and machine learning that provide reliable non-invasive assessment of 1p/19q status in LGGs. The proposed pipeline can be useful to facilitate clinical, physiological and therapeutic decision-making.

Keywords: Radiomics; Machine learning; Codeletion 1p/19q; Low-grade gliomas; Neighborhood component analysis.

INTRODUÇÃO

Os gliomas compreendem a grande maioria de todos os tumores cerebrais malignos. A Organização Mundial de Saúde (OMS) recomenda a utilização de uma classificação integrativa dos tumores, considerando as características histopatológicas e moleculares, permitindo assim um melhor manejo clínico (Dubbink et al., 2016). Por exemplo, os oligodendrogliomas são diagnosticados pela identificação das perdas combinadas de todo o braço dos cromossomos 1p e 19q (codeleção 1p/19q). Oligodendrogliomas têm melhor prognóstico do que astrocitomas e glioblastomas; assim, a identificação do status 1p/19q (codeletado ou não deletado) é uma etapa muito importante para avaliar os resultados em pacientes com gliomas (Louis et al., 2016; Weller et al., 2017). Pacientes com glioma de baixo grau (LGG) com status codeletados 1p/19q demonstraram ter taxas de sobrevida global significativamente melhores em comparação com pacientes sem deleção (Eckel-Passow et al., 2015). Assim, o conhecimento do status 1p/19q é de fundamental importância para melhores decisões clínicas, cirúrgicas e terapêuticas.

A avaliação molecular dos tumores cerebrais é feita por meio de biópsias, que são procedimentos invasivos, sujeitos a erros de amostragem e que envolvem riscos (Wijnenga et al., 2017). Uma alternativa não invasiva para avaliação do status 1p/19q envolve o uso de ressonância magnética (MR), radiômica e aprendizado de máquina. A ressonância magnética é um método de imagem médica não invasivo, pois usa radiação não ionizante. A ressonância magnética fornece contraste superior de tecidos moles e, portanto, é o método padrão-ouro para o diagnóstico e

tratamento de tumores cerebrais (Cha, 2006; DeAngelis, 2001). A radiômica é uma ferramenta de extração de atributos quantitativos de imagens médicas, permitindo a mineração de informações não visuais, que podem se correlacionar com características histopatológicas e genéticas do tumor, possibilitando diagnóstico molecular, avaliação do grau e prognóstico do tumor, além de auxiliar na avaliação clínica, tomada de decisão, orientação de tratamentos e mapeamento pré-cirúrgico (Lambin et al., 2017; Gillies et al., 2016). O aprendizado de máquina fornece técnicas para que um modelo computacional possa ser criado a partir de dados (entrada) com uma resposta conhecida (saída) (Baştanlar e Özuysal, 2014).

Muitos grupos de pesquisa têm trabalhado em métodos de aprendizado de máquina para avaliação do status 1p/19q (Alcaide-Leon, 2017; Akkus et al., 2017; Zhang et al., 2019). O trabalho desses grupos mostrou que o aprendizado de máquina e o aprendizado profundo têm forte potencial para o diagnóstico do status da codeleção 1p/19q. Por exemplo, Akkus et al. (2017) demonstraram um método baseado em redes neurais convolucionais, que teve um desempenho melhor do que os radiologistas que diagnosticam tumores cerebrais em imagens de ressonância magnética. No entanto, existem várias ferramentas de aprendizado de máquina ainda inexploradas para avaliação do status 1p/19q, algumas das quais podem fornecer desempenho ainda melhor para esta aplicação.

Este trabalho propõe uma assinatura promissora de baixa complexidade para avaliação não invasiva do status de codeleção cromossômica 1p/19q a partir de imagens de ressonância magnética ponderadas em T2, usando recursos radiômicos e aprendizado de máquina. O método proposto usa extração de recursos radiômicos, redução de dimensionalidade e um classificador de k-vizinhos mais próximos (KNN), com ajuste de hiperparâmetros usando pesquisa de grade com validação cruzada. Demonstramos que a assinatura proposta fornece resultados promissores em relação aos métodos de ponta da literatura e resultados consideravelmente melhores do que os obtidos por neurocirurgiões e neurorradiologistas analisando as mesmas imagens.

METODOLOGIA

Conjunto de dados

Todos os experimentos foram realizados usando o conjunto de dados LGG-1p19qDeletion do The Cancer Imaging Archive (Erickson et al., 2017; Akkus et al., 2017). Este conjunto de dados consiste em imagens de ressonância magnética adquiridas em 1,5 T ou 3 T, incluindo imagens ponderadas em T1 e T2 pós-contraste (espessura de corte de 1 mm e 3 mm, respectivamente). No entanto, utilizamos apenas as imagens ponderadas em T2, porque elas evidenciam melhor o líquido e o edema em comparação com outras sequências e, portanto, são muito úteis para detalhar vários processos patológicos, inclusive para determinar a extensão da

infiltração de um tumor cerebral. O conjunto de dados contém imagens pré-operatórias de 159 pacientes, todos com status 1p/19q comprovado por biópsia (102 pacientes com LGG codeletado e 57 pacientes com LGG não deletado). Imagens representativas são mostradas na Figura 1. O conjunto de dados também inclui regiões de interesse (ROI) segmentadas para todas as imagens. Cada ROI envolve completamente um tumor e seus limites (Akkus et al., 2017), conforme ilustrado na Figura 2.

Figura 1 – Imagens representativas do banco de dados LGG-1p19qDeletion: (a) grande lesão infiltrativa heterogênea na região frontoparietal direita com áreas de baixo sinal, sugerindo hemorragia ou calcificação, diagnosticada como oligodendroglioma (tumor codeletado 1p/19q); e (b) lesão infiltrativa homogênea no lobo frontal direito com alto sinal, diagnosticada como astrocitoma (tumor não codeletado 1p/19q).

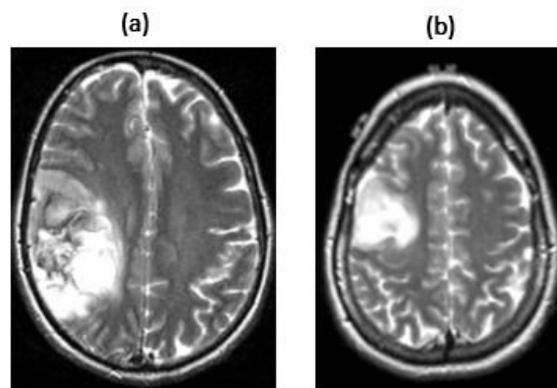
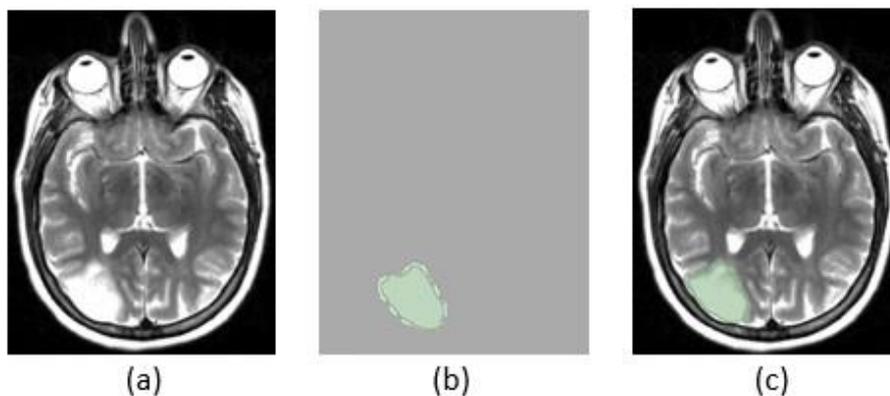


Figura 2 – Exemplo de tumor segmentado do banco de dados LGG-1p19qDeletion: (a) imagem contendo tumor occipital direito com alto sinal homogêneo, diagnosticado como oligodendroglioma (1p/19q codeletado); (b) região de interesse segmentada semi-automaticamente; e (c) região de interesse sobreposta à imagem.



Pré-processamento

Extração das características radiômicas

A extração de características quantitativas dos tumores foi realizada usando a plataforma PyRadiomics versão 2.0 (Griethuysen et al., 2017). PyRadiomics é uma plataforma flexível de código aberto, que permite extrair um grande número de características radiômicas de imagens médicas.

As ROIs no banco de dados LGG-1p19qDeletion são fornecidas no formato Neuroimaging Informatics Technology Initiative (NIfTI), que não é compatível com PyRadiomics. Assim, estes foram convertidos para o formato Nearly Raw Raster Data (NRRD) usando a ferramenta de software 3DSlicer (Fedorov et al., 2012).

Extraímos 68 recursos de cada ROI, sem filtros ativados. Os recursos do PyRadiomics são subdivididos em classes, incluindo: estatísticas de primeira ordem (primeira ordem); baseado em forma (forma); matriz de coocorrência de nível de cinza (GLCM); matriz de comprimento de execução de nível de cinza (GLRLM); matriz de zona de tamanho de nível de cinza (GLSZM); e matriz de dependência de nível de cinza (GLDM). Um resumo das características que extraímos é apresentado na Tabela 1.

Tabela 1 - Resumo das características radiômicas extraídas.

Nome das classes	Número de características	Descrição das classes
estatísticas de primeira ordem	0*	Descreve a distribuição de intensidades do voxel dentro da ROI.
forma	0*	Descreve o tamanho tridimensional e a forma da ROI.
GLCM	22	Quantifica a coocorrência de intensidades do voxel dentro da ROI.
GLRLM	16	Quantifica as ocorrências de voxels consecutivos dentro da ROI que possuem o mesmo nível de cinza.
GLSZM	16	Quantifica o número de voxels conectados que compartilham a mesma intensidade de nível de cinza na ROI.
GLDM	14	Quantifica as dependências de nível de cinza na ROI.

*Verificamos que os melhores resultados foram alcançados quando os recursos de primeira ordem e forma não foram usados.

Normalização das características

Diferentes características radiômicas vêm em diferentes ordens de magnitude. Assim, a normalização foi utilizada para que todas as características apresentem intervalos de valores semelhantes. Isso ajuda o algoritmo de classificação, garantindo que todos os recursos sejam ponderados de forma semelhante. O método de padronização foi aplicado a cada conjunto de recursos, de modo que cada um deles tenha média zero e desvio padrão unitário (Kreyszig, 1979; Géron, 2017). Os recursos normalizados são calculados como $z = (x - \mu) / \sigma$, onde x são os recursos não normalizados e μ e σ são a média e o desvio padrão do conjunto de recursos correspondente, respectivamente.

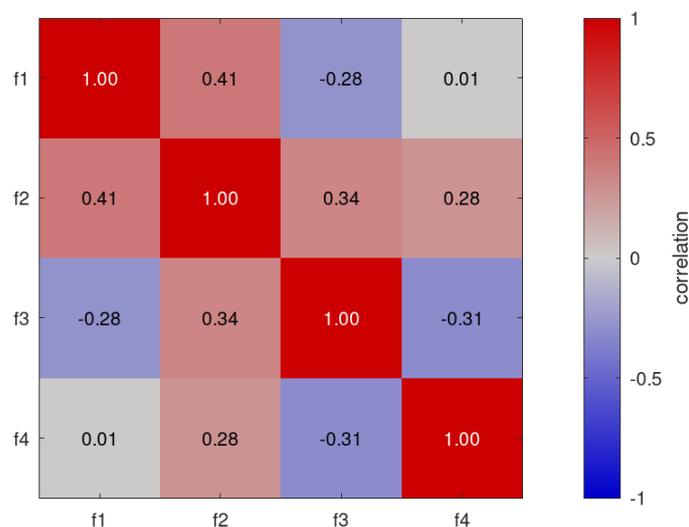
Redução da dimensionalidade dos dados

A análise de componentes de vizinhança (NCA) foi utilizada para a redução de dimensionalidade, que consiste em projetar os dados de características de um espaço de alta dimensionalidade para um subespaço de menor dimensionalidade. NCA (Goldberger et al., 2005)

é um método para aprender uma métrica de distância quadrática que otimiza o erro de classificação do algoritmo KNN, usando uma abordagem de gradiente descendente (Yang et al., 2012; Weinberger e Saul, 2009; Qin e outros, 2015).

Usando o NCA, o conjunto de recursos de 68 dimensões foi projetado em um subespaço de 4 dimensões. A Figura 3 mostra um mapa de calor da matriz de correlação para o conjunto de recursos de 4 dimensões obtido após a redução da dimensionalidade NCA, onde valores próximos a +1 indicam forte correlação positiva, valores próximos a -1 indicam forte correlação negativa e valores próximos a zero indicam correlação fraca. Os valores de correlação entre diferentes componentes do NCA variaram de -0,31 a +0,41. A Figura 4 mostra um gráfico do tipo boxplots dos valores de cada componente NCA para cada grupo de instâncias (azul: *non-deleted*; laranja: *codeleted*). Estes boxplots permitem apreciar o intervalo de valores e a distribuição estatística (mediana, quartis, outliers, etc.) de cada componente, para cada grupo. Eles também nos permitem observar que alguns componentes parecem ser mais propensos a contribuir para discriminar com sucesso entre status codeletado e não deletado do que outros (compare f3 a f1, por exemplo).

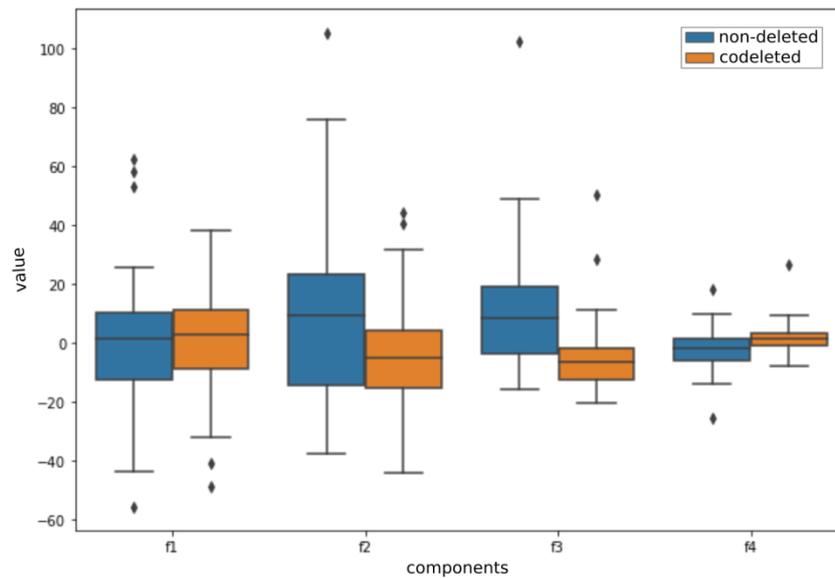
Figura 3 - Matriz de correlação para o conjunto de dados, obtida após a redução da dimensionalidade NCA. Valores próximos a +1 indicam correlação positiva forte, valores próximos a -1 indicam correlação negativa forte e valores próximos a zero indicam correlação fraca.



Implementação e Classificação

Nossa assinatura proposta para classificação do status 1p/19q usa o algoritmo KNN (Bonaccorso, 2017). KNN é um classificador supervisionado, projetado por Fukunaga e Narendra (1975). O KNN determina o rótulo de classificação de uma amostra, com base nas amostras de vizinhos, onde a variável k indica o número de vizinhos mais próximos aos quais uma medida de distância será aplicada, a fim de identificar a classe à qual a nova amostra pertence.

Figura 4 - Boxplots dos valores de cada componente NCA (f1 a f4) para cada grupo de instâncias (azul: não deletado; laranja: codeletado).



Após a etapa de pré-processamento, o conjunto de dados com 159 instâncias (57 não deletados e 102 codeletados) e 4 recursos (obtidos após redução da dimensionalidade NCA) foi dividido aleatoriamente em dois subconjuntos: 90% para treinamento KNN (143 instâncias) e 10% para teste KNN (16 instâncias).

As instâncias de treinamento são usadas para encontrar os parâmetros ótimos para o classificador KNN, ou seja, número de vizinhos mais próximos e abordagem de ponderação. Este processo de ajuste de hiperparâmetros foi realizado usando *grid-search* sobre uma grade de parâmetros (Géron, 2017). Esse processo avalia sistematicamente várias combinações dos parâmetros do modelo e retorna o ajuste que resultou na melhor acurácia (conforme definido mais adiante, na Eq. 1). Para a abordagem de ponderação, avaliamos “uniforme” e “baseado na distância”. Para o número de vizinhos mais próximos, avaliamos k variando de 1 a 30.

Durante o ajuste de hiperparâmetros, a avaliação de cada configuração de parâmetro foi feita usando validação cruzada estratificada *k-folds* repetida (Berrar, 2019; Géron, 2017). Isso envolve particionar aleatoriamente o conjunto de dados em k subconjuntos de tamanho igual (não confundir com o número de vizinhos mais próximos em KNN), em que cada subconjunto contém aproximadamente as mesmas proporções de cada classe de instâncias (não excluídas e codeletadas). Um subconjunto é retido para validação, enquanto os subconjuntos restantes são usados para treinamento. Esse processo é repetido k vezes (para que cada subconjunto seja usado uma vez para validação) e os k resultados são calculados para produzir uma estimativa da precisão da classificação. Todo esse processo (partição aleatória seguida por k repetições das etapas de treinamento e validação) é repetido n vezes e a acurácia da classificação é calculada em média nessas n execuções. Usamos $k = 5$ dobras (*folds*) e $n = 10$ repetições. Observe que esta etapa faz

parte do treinamento KNN e, portanto, apenas o conjunto de dados de treinamento (143 instâncias) é usado. As 16 instâncias de teste não são vistas pelo validador cruzado. Assim, o validador cruzado usou dois subconjuntos de 28 instâncias e três subconjuntos de 29 instâncias.

Finalmente, as 16 instâncias de teste são aplicadas ao classificador KNN, usando o número ótimo de vizinhos mais próximos e a abordagem de ponderação encontrada durante o ajuste de hiperparâmetros. O número de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos é calculado e o desempenho é avaliado de acordo com várias métricas (descritas abaixo).

Avaliação da performance da assinatura

As seguintes métricas amplamente aceitas foram usadas para avaliar o desempenho da assinatura proposta na previsão do status 1p/19q: acurácia (Acc), sensibilidade (Se), especificidade (Spe), precisão (Pre), F1-Score (F1), e área sob a curva (AUC). A AUC mede a probabilidade de uma amostra positiva ser rotulada como positiva pelo classificador; é a área sob a curva ROC (Receive Operating Characteristic) (um gráfico da taxa de verdadeiros positivos em relação à taxa de falsos positivos), que ilustra como os resultados da classificação mudam à medida que o limite de discriminação do classificador varia). A AUC foi calculada usando a biblioteca Scikit-learn 0.24.2 (Géron, 2017). As outras métricas são calculadas como:

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \quad (1)$$

$$Se = \frac{TP}{TP + FN} \quad (2)$$

$$Spe = \frac{TN}{TN + FP} \quad (3)$$

$$Pre = \frac{TP}{TP + FP} \quad (4)$$

$$F1 = 2 \cdot \frac{Pre \cdot Se}{Pre + Se} \quad (5)$$

onde TP, FP, TN e FN representam o número de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos, respectivamente (as instâncias codeletadas foram tratadas como "positivas" e as instâncias não deletadas foram tratadas como "negativas"). Observe que convertemos essas métricas em porcentagens (multiplicando cada uma por 100%).

Todo o processo descrito anteriormente foi simulado 30 vezes, com cada nova simulação começando com uma nova divisão aleatória do conjunto de dados (90% para treinamento e 10% para teste), seguida de ajuste de hiperparâmetros (usando repetição por validação cruzada estratificada *k-fold*), seguida pela classificação KNN (usando o número ideal de vizinhos mais

próximos e a abordagem de ponderação encontrada durante o ajuste de hiperparâmetros). O desempenho alcançado pela assinatura em cada simulação foi avaliado (usando as métricas descritas acima) e a média foi calculada. Também usamos os resultados dessas 30 simulações para calcular intervalos de confiança para cada métrica, no nível de confiança de 95%, com base em uma distribuição *t-Student*.

Para avaliar a importância de cada etapa da assinatura, avaliamos também os resultados obtidos com versões “incompletas” da assinatura. No total, foram avaliadas 4 versões de assinatura:

1. classificador KNN com $k = 1$ vizinho mais próximo e ponderação baseada na distância;
2. normalização dos dados (padronização), seguida de classificador KNN com $k = 1$ vizinho mais próximo e ponderação baseada na distância;
3. normalização de dados (padronização), seguida de classificador KNN com ajuste de hiperparâmetros; e
4. assinatura completa, ou seja, normalização de dados (padronização), seguida de redução de dimensionalidade (NCA), seguida de classificador KNN com ajuste de hiperparâmetros.

Observe que cada versão da assinatura foi avaliada usando 30 simulações, cada simulação usando uma nova divisão aleatória do conjunto de dados (90% para treinamento e 10% para teste). No entanto, as mesmas 30 divisões foram usadas para avaliar cada versão da assinatura.

RESULTADOS

A Tabela 2 apresenta o desempenho de classificação alcançado após a incorporação de cada etapa na execução da assinatura (média de 30 simulações). Os valores mostram que os resultados melhoram à medida que incorporamos cada etapa adicional a assinatura. Os melhores resultados são obtidos ao executar a assinatura completa (linha 4).

Tabela 2. Desempenho de classificação alcançado após a incorporação de cada etapa na execução da assinatura. A assinatura completa (versão 4) está destacada em negrito. Os resultados em cada linha correspondem a uma média de 30.

ETAPAS DA ASSINATURA	ACC	SE	SPE	PRE	AUC	F1
1. KNN	56.25	67.92	35.17	67.80	56.70	66.35
2. Normalização + KNN	73.33	85.15	47.36	77.80	73.80	80.79
3. Normalização + KNN + hiperparâmetros	76.04	80.76	66.51	83.23	73.60	81.31
4. Normalização + NCA + KNN + hiperparâmetros	81.88	88.48	69.25	86.04	82.70	86.38

A Tabela 3 apresenta os resultados individuais obtidos em cada uma das 30 simulações da assinatura completa; valores médios e intervalos de confiança (IC) para cada métrica também

são apresentados. Esses resultados demonstram a consistência da assinatura na obtenção de bons resultados de classificação. A Tabela 4 mostra uma matriz de confusão obtida pela média das matrizes de confusão obtidas em cada uma das 30 simulações da assinatura completa, o que mostra que o número médio de erros de classificação (FP e FN) é pequeno, comparado ao número de classificações corretas (TP e TN).

Tabela 3 - Resultados da classificação individual obtidos em cada uma das 30 simulações da assinatura completa, valores médios e intervalos de confiança (ao nível de confiança de 95%).

Métricas	Média	IC.
TP	9.4	9.0-10.9
FP	1.6	0.9-1.9
TN	3.7	3.2-5.1
FN	1.3	0.9-2.1
Acc	81.88	74.4-89.3
Se	88.48	79.7-97.2
Spe	69.25	55.2-83.3
Pre	86.04	77.9-94.1
AUC	82.70	77.5-87.9
F1	86.38	80.4-92.3

Tabela 4 - Matriz de confusão obtida pela média das matrizes de confusão obtidas em cada uma das 30 simulações da assinatura completa.

		PREDICTO	
		NÃO-DELETADO	CODELETADO
REAL	NÃO-DELETADO	3.7 (TN)	1.6 (FP)
	CODELETADO	1.3 (FN)	9.4 (TP)

Métricas como sensibilidade, especificidade e precisão dependem do limiar de discriminação do classificador, portanto, muitas vezes é útil analisar gráficos como a curva ROC (Figura 5) e a curva de precisão-sensibilidade (PR) (Figura 6). Essas ferramentas ilustram como os resultados da classificação mudam à medida que o limite de discriminação do classificador varia. A curva ROC é um gráfico da taxa de verdadeiros positivos (Se) contra a taxa de falsos positivos (1-Spe), enquanto a curva PR é um gráfico de precisão (Pre) contra a sensibilidade (Se). Essas curvas são calculadas com base em estimativas de probabilidade da classe positiva (status codeletados 1p/19q). No entanto, ao usar a assinatura completa, todas as estimativas de probabilidade em todas as 30 simulações foram 0,0 ou 1,0, de modo que a variação do limite de discriminação não alteraria os valores na matriz de confusão de cada simulação. Isso resultou em apenas dois pontos e um ponto, respectivamente, para a construção das curvas ROC e PR para cada simulação (Katostaras e Katostara, 2013). Como resultado ilustrativo, apresentamos as duas curvas construídas para o resultado médio da assinatura completa, ou seja, usando $Se = 0,88$, $Spe = 0,69$ e $Pre = 0,86$. A área sob a curva ROC (Figura 5) foi de 0,8270, que corresponde ao valor AUC de 82,70%. Esta é uma estimativa da probabilidade de rotular com precisão uma amostra positiva (status codeletado 1p/19q). Embora a curva ROC seja mais comumente usada, ela pode não ser o indicador de desempenho mais apropriado ao lidar com conjuntos de dados altamente

desbalanceados; nesses casos, pode ser benéfico usar a curva PR. A área sob a curva PR (Figura 6) foi de 0,8638, o que corresponde ao *score* F1 de 86,38%. Como demonstrado na Eq. 5, esta é a média harmônica entre precisão e sensibilidade.

Figura 5 - Curva ROC para o resultado médio da assinatura completa. A área sob a curva ROC foi de 0,827.

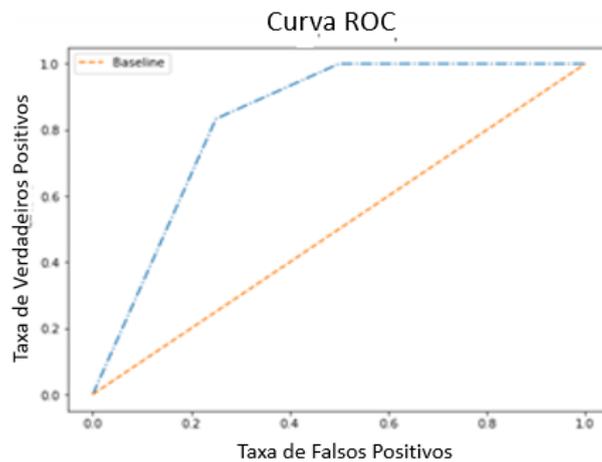
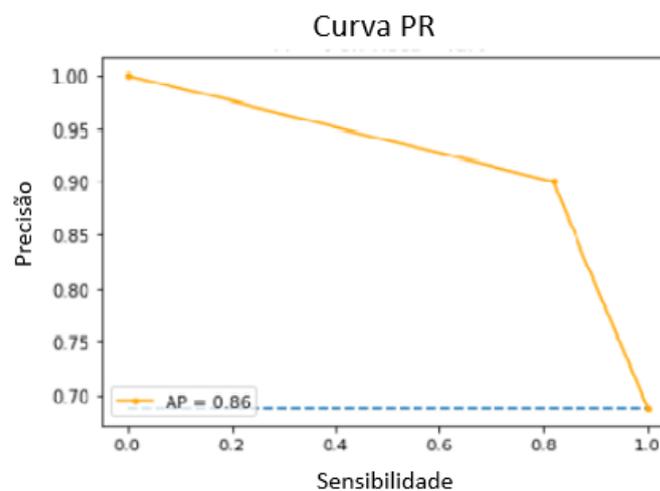


Figura 6 - Curva PR para o resultado médio da assinatura completa. A área sob a curva PR foi de 0,863.



CONCLUSÃO

Este estudo demonstrou que, por meio do uso de técnicas de aprendizado de máquina, a radiômica pode ser uma ferramenta valiosa para a identificação não invasiva do status de codeleção 1p/19q em gliomas de baixo grau. Resultados promissores foram obtidos usando características radiômicas extraídas de imagens de ressonância magnética ponderadas em T2. A biblioteca PyRadiomics se mostrou adequada para extrair recursos radiômicos do conjunto de dados LGG-1p19qDeletion, exigindo apenas uma conversão de formato de arquivo (de NIFTI para NRRD).

No aprendizado de máquina, o uso de conjuntos de dados com um pequeno número de instâncias representa um desafio para a tarefa de classificação. Xing e outros. (2019) sugerem uma relação de 10 a 15 instâncias por característica radiômica utilizada na etapa de classificação. Portanto, a etapa de redução de dimensionalidade foi importante para o bom desempenho da nossa assinatura proposta. De fato, os resultados da Tabela 2 demonstram que cada etapa contribuiu para melhorar o desempenho do classificador.

A Tabela 5 apresenta uma comparação entre a assinatura proposta e outros métodos da literatura. Eles representam o estado da arte na classificação do status da codeleção 1p/19q com base em radiômica e aprendizado de máquina. A Tabela 5 também inclui métricas de um estudo comparativo de Van der Voort et al. (2019), em que dois neurocirurgiões e dois neurorradiologistas avaliaram as imagens ponderadas em T2 do conjunto de dados LGG-1p19qDeletion. Os avaliadores classificaram cada imagem em relação à capacidade de determinar com segurança o status 1p/19q do tumor, usando uma escala de 1 a 5 (onde 1 correspondia a "inseguro" e 5 a "confiante"). Essa pontuação foi convertida em um marcador de previsão de codeleção, a partir do qual foram calculadas as métricas apresentadas.

Ao avaliar o status 1p/19q, é mais importante evitar falsos negativos (diagnosticar incorretamente um tumor codeletado) do que evitar falsos positivos (diagnosticar incorretamente um tumor não deletado); ou seja, alta sensibilidade é mais crítica do que alta especificidade. O diagnóstico de codeleção positiva é importante porque os pacientes com tumores codificados têm melhor prognóstico e maior sobrevida, pois esses tumores respondem bem à quimioterapia dedicada (à base de procarbanize, lomustina e vincristina) e à radioterapia. Nossa assinatura desenvolvida alcançou a segunda maior sensibilidade (88,4%) do que a maioria dos métodos da literatura, ao custo de menor especificidade. Acreditamos que isso pode ser devido ao pequeno tamanho do conjunto de dados LGG-1p19qDeletion, particularmente ao pequeno número de instâncias não excluídas. Usar um conjunto de dados desequilibrado pode ser um problema, pois os padrões de dados da classe majoritária podem se destacar daqueles da classe minoritária. Um modelo treinado com um conjunto de dados desequilibrado pode projetar maior acurácia na classe majoritária e, possivelmente, exibir desempenho de classificação ruim para a classe minoritária.

A assinatura proposta neste estudo foi projetada após a experimentação de outros algoritmos e técnicas de aprendizado de máquina para previsão de status 1p/19q. Em nossos trabalhos anteriores (Silva et al., 2020; Cassia et al., 2021; Silva et al., 2021), experimentamos redução de dimensionalidade com base na análise de componentes principais, aumento de dados baseado em ruído gaussiano e SMOTE e quatro classificadores diferentes : perceptron multicamadas, regressão logística, máquina de vetores de suporte e floresta aleatória. Embora razoavelmente bem-sucedidas, essas técnicas não forneceram resultados tão bons quanto os

obtidos com a assinatura proposta neste trabalho, que usa redução de dimensionalidade NCA e classificação KNN, com ajuste de hiperparâmetros usando pesquisa em grade baseada em validação cruzada estratificada repetida de k dobras.

Tabela 5 - Comparação com métodos de ponta da literatura para classificação de status 1p/19q com base em radiômica e aprendizado de máquina.

REFERÊNCIAS	DATASET	ALGORITMOS	ACC	SE	SPE	PRE	AUC	F1
Van der Voort et al. (2019)	LGG-1p19qDeletion	neurosurgeons*	45.7	45.9	45.5		44.9	
		neuroradiologists*	57.4	45.9	79.5		79.2	
Akkus et al. (2017)	LGG-1p19qDeletion	SVM	81.0	80.0	82.0			
Lu et al. (2018)	LGG-1p19qDeletion	Quadratic SVM	80.0	66.7	100.0			
Han et al. (2018)	Own	RF	70.0	68.3	71.2		76.0	
Van der Voort et al. (2019)	LGG-1p19qDeletion (test) Own - EMC/HMC (train)	SVM	69.3	73.2	61.7	78.7	72.3	69.7
Kocak et al. (2020)	LGG-1p19qDeletion	MLP	83.8	87.5	75.8	90.5	86.9	88.3
Kong et al. (2020)	Own – PUMCH	RF	83.3	100.0	75.0			
Shboul et al. (2020)	TCGA-LGG	XGBoost		75.0	85.0		80.0	
Decuyper et al. (2021)	LGG-1p19qDeletion TCGA-LGG BraTS 2019	CNN – 3D U-Net	83.3	86.7	79.2		82.1	
	Own - GUH		75.0	58.3	82.1		86.6	
Assinatura proposta	LGG-1p19qDeletion	KNN + NCA	81.8	88.4	69.2	86.0	82.7	86.3

* Van der Voort et al. (2019) apresentam um estudo comparativo no qual dois neurocirurgiões e dois neurorradiologistas classificaram as imagens do conjunto de dados com relação à capacidade de determinar com segurança o status 1p/19q do tumor, usando uma escala de 1 a 5 (onde 1 correspondia a "incerto", e 5 correspondeu a "confiante"). Essa pontuação foi convertida em um marcador de previsão de codeleção, a partir do qual essas métricas foram calculadas.

Acreditamos fortemente que nossa assinatura poderia ter obtido melhor desempenho se um conjunto de dados maior estivesse disponível.

Em conclusão, apresentamos uma assinatura para classificação não invasiva do status de codeleção 1p/19q em gliomas de baixo grau com base em radiômica e aprendizado de máquina. O método proposto obteve resultados promissores aos métodos de ponta da literatura e resultados consideravelmente melhores do que os obtidos por neurocirurgiões e neurorradiologistas analisando as mesmas imagens. Pacientes com codeleção 1p/19q têm melhores taxas de sobrevida e prognóstico do que pacientes sem deleção. Portanto, este trabalho demonstra que as técnicas baseadas em radiômica e aprendizado de máquina representam uma oportunidade positiva para auxiliar na melhor tomada de decisão clínica, cirúrgica e terapêutica.

REFERÊNCIAS

Alcaide-Leon P, Dufort P, Geraldo AF, Alshafai L, Maralani PJ, Spears J, et al. Differentiation of enhancing glioma and primary central nervous system lymphoma by texture-based machine learning. *AJNR Am J Neuroradiol.* 2017; 38:1145–50. <https://doi.org/10.3174/ajnr.A5173>

Akkus Z, Ali I, Sedlář J, Agrawal JP, Parney IF, Giannini C, Erickson BJ. Predicting Deletion of Chromosomal Arms 1p/19q in Low-Grade Gliomas from MR Images Using Machine Intelligence. *J Digit Imaging.* 2017; 30(4): 469–476. Published online 2017 Jun 9. PMID: PMC5537096. <https://doi.org/10.1007/s10278-017-9984-3>

Baştanlar Y., Özuysal M. (2014) Introduction to Machine Learning. In: Yousef M., Allmer J. (eds) *miRNomics: MicroRNA Biology and Computational Analysis*. Methods in Molecular

Biology (Methods and Protocols), vol 1107. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-62703-748-8_7

Berrar D. Cross-validation. Encyclopedia of Bioinformatics and Computational Biology, pp. 542–545. Academic Press, Oxford 2019. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>

Bonaccorso G. Machine Learning Algorithms: A Reference Guide to Popular Algorithms for Data Science and Machine Learning, Packt Publishing, Birmingham, UK, 2017.

Cassia GS, Silva TAM, Carvalho JLA. Radiomics to predict 1p/19q chromosomal codeletion status of low-grade gliomas. In: GliMR - Virtual Annual Meeting 2021, 2021, Rotterdam. GliMR - Virtual Annual Meeting, 2021.

Cha S. Update on brain tumor imaging: from anatomy to physiology. AJNR Am J Neuroradiol. 2006; 27(3): 475–87.

DeAngelis LM. Brain tumors. N Engl J Med. 2001; 344(2):114–23. <https://doi.org/10.1056/nejm200101113440207>

Decuyper M, Bonte S, Deblaere K, and Van Holen R. Automated MRI based pipeline for segmentation and prediction of grade, IDH mutation and 1p19q co-deletion in glioma. Computerized Medical Imaging and Graphics, Vol. 88, 2021, 101831, ISSN 0895-6111, <https://doi.org/10.1016/j.compmedimag.2020.101831>.

Dubbink HJ, Atmodimedjo PN, Kros JM, French PJ, Sanson M, Idbaih A, et al. Molecular classification of anaplastic oligodendroglioma using nextgeneration sequencing: a report of the prospective randomized EORTC Brain Tumor Group 26951 phase III trial. Neuro Oncol. 2016; 18:388–400.

Eckel-Passow JE, Lachance DH, Molinaro AM, Walsh KM, Decker PA, Sicotte H, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. N. Engl. 2015, J. Med. 372, 2499–2508. <https://doi.org/10.1056/NEJMoa1407279>

Erickson B, Akkus Z, Sedlar J, Korfiatis P. Data From LGG-1p19q Deletion. The Cancer Imaging Archive. 2017. <https://doi.org/10.7937/K9/TCIA.2017.dwehtz9v>

Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging. 2012;30: 1323–41.

Fukunaga K, Narendra PM. A branch and bound algorithm for computing k-nearest neighbors. IEEE Transactions on Computers, 1975, v. 100, n. 7, p. 750–753

Géron A. Hands-on Machine Learning With Scikit-Learn and TensorFlow: Concepts Tools and Techniques to Build Intelligent Systems, Sebastopol, CA, USA:O'Reilly, 2017.

Gillies RJ, Kinahan PE, Hricak H. Radiomics: imagens are more than pictures, they are data. Radiology. 2016; 278:563–577 CrossRefMedline.

Goldberger J, Roweis S, Hinton G, and Salakhutdinov R. Neighbourhood components analysis. In Advances in Neural Information Processing Systems 17. MIT Press, 2005, pp. 513–520.

Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts HJWL. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research, 2017. 77(21), e104–e107. <https://doi.org/10.1158/0008-5472.CAN-17-0339>

Han Y, Xie Z, Zang Y, Zhang S, Gu D, Zhou M, et al. Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas. *J Neurooncol.* (2018) 140:297–306. <https://doi.org/10.1007/s11060-018-2953-y>

Katostaras T, Katostara N. Area of the ROC curve when one point is available. *Stud Health Technol Inform.* 2013;190:219-21. PMID: 23823428.

Kocak B, Durmaz ES, Ates E, et al. Radiogenomics of lower-grade gliomas: machine learning-based MRI texture analysis for predicting 1p/19q codeletion status. *Eur Radiol*, 2020;30:877–86 CrossRef Medlin.

Kong Z, Jiang C, Zhang Y, Liu S, Liu D, Liu Z, Chen W, Liu P, Yang T, Lyu Y, Zhao D, You H, Wang Y, Ma W, Feng F. Thin-slice magnetic resonance imaging-based radiomics signature predicts chromosomal 1p/19q Co-deletion status in grade II and III gliomas. *Front. Neurol.*, 11 (2020), p. 551771.

Kreyszig E. *Advanced Engineering Mathematics*, 1979 (Fourth ed.). Wiley. p. 880, eq. 5. ISBN 0-471-02140-7.

Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017; 14:749–62. <https://doi.org/10.1038/nrclinonc.2017.141>

Louis DN, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2016; 131:803–820. <https://doi.org/10.1007/s00401-016-1545-1>

Lu C-F, Hsu F-T, Hsieh KL-C, Kao Y-CJ, Cheng S-J, Hsu JB-K, et al. Machine learning-based radiomics for molecular subtyping of gliomas. *Clin Cancer Res.* 2018; 24:4429–36. <https://doi.org/10.1158/1078-0432.CCR-17-3445>

Shboul ZA , Chen J, Iftekharuddin MK. Prediction of molecular mutations in diffuse low-grade gliomas using MR imaging features. *Sci Rep.*(2020);10(1):3711.

Qin C, Song S, Huang G, Zhu L. Unsupervised neighborhood component analysis for clustering. 2015, *Neurocomputing* 168:609–617.

Silva TAM, Cassia GS, Carvalho JLA. Utilização de Características Radiômicas para Classificar Gliomas de Baixo Grau com Relação ao Status da Codeleção Cromossômica 1p/19q. In: XXVII Brazilian Congress in Biomedical Engineering, 2020, Vitória - ES. *Anais do XXVII Congresso Brasileiro de Engenharia Biomédica*, 2020. p. 23-28.

Silva TAM, Cassia GS, Nascimento FAO, Carvalho JLA, Predição da codeleção 1p/19q utilizando atributos radiômicos e aprendizado de máquina, In: XXXIX Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, 2021.

Van der Voort SR. et al. Predicting the 1p/19q co-deletion status of presumed low grade glioma with an externally validated machine learning algorithm. *Clinical Cancer Research*, 2019, 1127.

Weinberger K and Saul L. Distance metric learning for large margin nearest neighbor classification. *The Journal of Machine Learning Research*, vol. 10, pp. 207–244, 2009.

Weller M, van den Bent M, Tonn JC, et al. European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. *The Lancet Oncology.* 2017; 18(6):e315-e329. ISSN 1470-2045. [https://doi.org/10.1016/S1470-2045\(17\)30194-8](https://doi.org/10.1016/S1470-2045(17)30194-8)

Wijnenga MMJ, Mattni T, French PJ, et al. Does early resection of presumed low-grade glioma improve survival? A clinical perspective. *Journal of neuro-oncology*. 2017; 133(1):137-146. ISSN 1573-7373. <https://doi.org/10.1007/s11060-017-2418-8>

Xing L, Li R, Napel S, Rubin D. *Radiomics and Radiogenomics: Technical Basis and Clinical Applications*. Boca Raton, FL: CRC Press, 2019

Yang W, Wang K, Zuo W. Neighborhood component feature selection for high-dimensional data. 2012, *JCP* 7(1):161–168.

Zhang Y, Chen C, Cheng Y, Teng Y, Guo W, Xu H, Ou X, Wang J, Li H, Ma X and Xu J. Ability of Radiomics in Differentiation of Anaplastic Oligodendroglioma From Atypical Low-Grade Oligodendroglioma Using Machine-Learning Approach. *Front. Oncol.* 2019; 9:1371. <https://doi.org/10.3389/fonc.2019.01371>

Recebido em: 02/12/2022

Aprovado em: 03/01/2023

Publicado em: 07/02/2023