

Aumento de dados para detecção de câncer de mama em imagens histopatológicas utilizando redes adversárias generativas

VII Encontro de Programas de Educação Tutorial

Andressa Gomes Moreira, Ialis Cavalcante de Paula Junior

O câncer de mama é uma doença maligna, predominante em mulheres, sendo a principal causa de mortalidade em todo mundo. É caracterizado pelo crescimento exorbitante de células que formam um tumor, chamado de neoplasia. Os tumores podem ser cancerosos (malignos) ou não cancerosos (benignos). Dessa forma, existem diferentes maneiras de diagnosticar o câncer de mama, seja por mamografia, ultrassom, biópsia ou técnicas mais recentes baseadas em aprendizado de máquina, subcampo da inteligência artificial (IA). É notável o sucesso da utilização de técnicas de aprendizado profundo, como Redes Neurais Convolucionais (CNN), para classificar anomalias benignas ou malignas em imagens histopatológicas, a fim de realizar a detecção do câncer de mama. Entretanto, embora as CNNs manifestem êxito em categorizar as neoplasias, ainda existem fatores que interferem na classificação final, como a escassez de informações, uma vez que a maioria dos conjuntos de dados são limitados ou não estão disponíveis publicamente. Desse modo, para mitigar a carência de dados é proposto o uso de método de aumento de dados (DA). Assim sendo, foi utilizado o conceito de Redes Adversárias Generativas (GAN), introduzido pela primeira vez em um artigo de Ian Goodfellow (2014), cujo objetivo é gerar imagens artificiais semelhantes às imagens reais. Portanto, fez-se uso de GAN para expandir a quantidade de imagens histopatológicas provenientes do Banco de dados BreakHis (2014), de tal modo que os dados sintéticos gerados pela rede são anexadas ao conjunto de dados original para compor a entrada da CNN e melhorar o desempenho da rede, a fim de classificar o câncer de mama e categorizar as neoplasias.

Palavras-chave: Redes Adversárias Generativas, Aumento de dados, Aprendizagem profunda.