

# Portfólio – Processamento de Imagens

Gabriel Gian, 6º período

O artigo “*A multi-scale gated multi-head attention depthwise separable CNN model for recognizing COVID-19*” por Geng Hong et al trata do uso recente de uma rede neural convolucional, classe de redes neurais amplamente utilizada no quinhão do processamento de imagens, com características de escala múltipla e profundidade de atenção – modelo nomeado MGMADSCNN-3 – para o processamento de imagens obtidas em raios-X e tomografias para a identificação de sintomas e sequelas da COVID-19.

O estudo constatou que o modelo convolucional de redes neurais com 3 camadas de atenção tem apresentado resultados convincentes para o processamento de exames, com acurácia de 96,75% em imagens de raio-X e 98,25% em imagens de tomografia. Com apenas 43.6 Mb de tamanho e velocidade de detecção de 6,09 ms e 4,23 ms para imagens de raio-X e tomografia respectivamente, o modelo de camada tripla possibilitou a detecção e a classificação dos efeitos da COVID-19 com precisão e tempo recordes. A detecção funciona de modo a identificar manchas nos pulmões, mudanças nos tecidos intersticiais e turbidez nas imagens.

No início de 2017, um dataset de imagens de raio-X de regiões peitorais foi compilado por Xiaosong Wang et al, assim como a proposição de um algoritmo capaz de identificar 14 tipos diferentes de doenças respiratórias. De 2020 em diante, a pesquisa acerca da identificação de COVID-19 tem ganhado cada vez mais força frente ao número crescente de casos e mortes pelo vírus.

Um dos métodos de classificação é a busca binária em imagens de raio-X entre casos confirmados negativos e positivos para a COVID-19. Uma rede baseada neste princípio, a COVIDX-Net, obteve 91% de acurácia de classificação, todavia o dataset era pequeno e desbalanceado. Outro método é a classificação multiclasse das imagens que atraem mais atenção de médicos e pesquisadores. Foram propostos modelos de 3 classes (COVID-19, não infectados e pneumonia) como o DarkCovidNet, CoroNet e COVID-Net, o último atingindo uma eficácia geral de 92.4% baseada em 13.975 imagens de raio-X. Quando testes de 4 classes (COVID-19, pneumonia viral, pneumonia bacteriana e normal) foram executados, houve a obtenção de uma eficácia geral de 83,5%.

O MGMADS-3 é estruturado da seguinte forma:

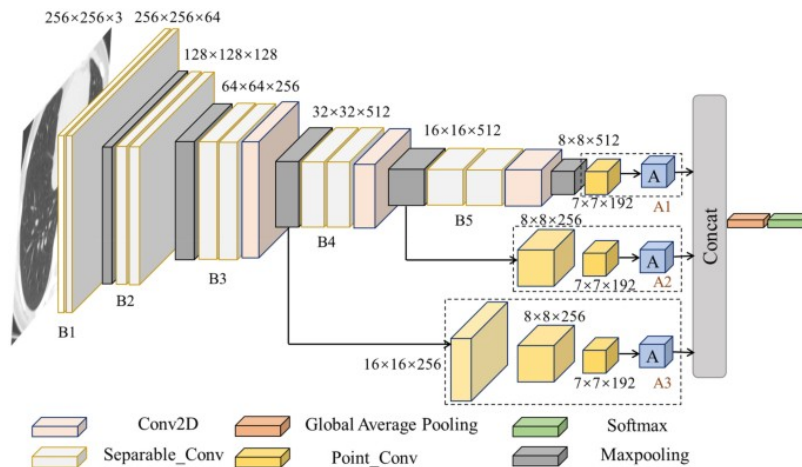


Figura 1. Modelagem do MGMADS-3

Constitui-se da extração de convoluções em blocos e módulos de atenção, sendo os dois primeiros blocos compostos de duas camadas convolucionais com profundidade e uma camada de *pooling* máximo. Os próximos três blocos convolucionais possuem duas camadas convolucionais com profundidade e uma camada convolucional comum, além de uma camada de *pooling* máximo. A multi-escala do MGMADS-CNN significa que há a possibilidade de operação com diferentes blocos de atenção, além de um *pooling* global médio para assegurar a integridade das entradas.

O modelo proposto pelo artigo mostra que há vantagens em termos de tamanho, velocidade rápida de detecção e alta precisão. O modelo consegue extrair pequenas partes de informação alvo de diferentes subespaços, o que aumenta bastante a acurácia das detecções. A convolução com profundidade permite reduzir o número de parâmetros de cálculo, reduzindo também tanto o tamanho do modelo quanto a latência entre detecções. A proposta melhorará o futuro dos diagnósticos assistidos por computador em termos gerais, e não apenas no combate contra a COVID-19.