

Rede neural para classificação de imagens

Notebook com todo o processo criação da rede:

<https://colab.research.google.com/drive/1eTLjbAJJvzfVBvjG8a97d8raMy3YcwXT?usp=sharing>

Estrutura da rede:

A Rede Neural criada é do tipo Convolucional e, portanto, possui algumas camadas de convolução para realçar detalhes das imagens e camadas de pooling para reduzir as imagens sem perder os detalhes realçados. Dessa forma, tem-se uma rede com várias camadas de convolução (2D), max pooling (2D), camada flatten e camadas densas de neurônios. Vale ressaltar que foi realizado transfer learning da rede VGG-16 pré-existente. As camadas de convolução e pooling presentes na rede final são de origem da VGG-16.

Ademais, antes das camadas de convoluções de fato começarem, tem-se algumas camadas para preparação dos dados, sendo elas: camada para reescalar os valores das cores de cada pixel para o intervalo 0 a 1 ao invés de 0 a 255; camada de flip que inverte horizontal e verticalmente uma imagem de maneira randômica; camada de rotação para rotacionar uma imagem de forma aleatória; e uma camada que aplica contraste com intensidade aleatória na imagem (dentro do intervalo 0.3 a 0.5).

Imagens utilizadas para treinamento:

O conjunto original dividia as possibilidades nas categorias de condição normal, condição de AVC hemorrágico e condição de AVC isquêmico. Porém, para abordar de forma mais condizente com o trabalho proposto, uma nova distribuição foi feita. As imagens foram movidas para uma nova distribuição de duas classes: AVC (hemorrágico e isquêmico) e não AVC (condição normal).

Sobre o tamanho do conjunto, a classe AVC possui 301 imagens enquanto que a classe não AVC possui 174. A dimensão de cada imagem é de 512 x 512 pixels. A respeito da distribuição dessas imagens para treino e validação, com intuito de evitar sobre ajuste, o conjunto de imagens foi distribuído como 65% dos dados para treinamento e 35% para validação.

Resultados obtidos sobre o desempenho da rede criada:

Desempenho com dados de treinamento

Métrica	Dados de treino	Dados de validação
Loss	0.1936	0.1105
TP (verdadeiro positivo)	90.0000	53.0000
FP (falso positivo)	3.0000	1.0000
TN (verdadeiro negativo)	111.0000	59.0000
FN (falso negativo)	12.0000	2.0000
Acurácia	93.06%	97.39%
Precisão	0.9677	0.9815
Recall	0.8824	0.9636
Área sob a curva ROC	0.9802	0.9944
Taxa de falso negativo	9.76%	3.28%
Taxa de falso positivo	3.23%	1.85%

Considerações importantes:

- Acurácia é a taxa de acertos do modelo com relação ao número total de casos processados
- Taxa de falso negativo = $fn / fn + tn$
- Taxa de falso positivo = $fp / fp + tp$ (1 – precisão)
- Rede criada a partir de uma VGG-16 (*transfer learning*)
- Rede otimizada a partir de Adam com *learning rate* de 0.001
- Treinamento utilizando *batch* de tamanho 32
- As camadas de *data augmentation* (flip, rotação e contraste) apresentaram melhoria significativa no desempenho do modelo

Conclusão

A rede apresentou bom desempenho para o problema o qual aborda. Seus acertos no geral (acurácia) foram de 93%, o equivalente a acertar o diagnóstico de 93 a cada 100 pessoas analisadas. Ademais, com relação a ocorrência de falsos, foi obtido uma taxa baixa para casos negativos e positivos, e a pontuação da área sob a curva ROC também apontou equilíbrio na quantidade desses erros.

O desempenho do classificador com relação a esses falsos foi o equivalente a julgar **incorretamente** três a cada cem indivíduos que possuem quadro compatível com o da doença (pior erro pois nesse caso o modelo indica condição de normalidade quando se tem condição de possível avc) e, **também incorretamente**, dois a cada cem que não possuem quadro da doença.