



**INSTITUTO METRÓPOLE DIGITAL
BACHARELADO EM TECNOLOGIA DA INFORMAÇÃO**

PEDRO HENRIQUE TEIXEIRA E SILVA / 20180117642

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE - UFRN

IMD1114 - Aprendizado Profundo - 2022.2

Respostas das questões conceituais da avaliação da 2a unidade

**NATAL
2022**

1. Na arquitetura AlexNet, há um total 62,416,616 parâmetros, dos quais 19,136 (aproximadamente 0,31%) são não treináveis. Para a VGG, a quantidade mais do que dobra, passando para 138,423,208 parâmetros, das quais 62,8% são treináveis. Já na no caso da ResNet a quantidade reduz drasticamente para 11,511,784 dos quais cerca 500,000 não são treináveis. Para a NiN, continua a tendência de redução, indo para 6,414,360. Sendo muito próxima a GoogLeNet, que possui pouco menos de 6,800,000 parâmetros. No caso da NiN e GoogLeNet, essa grande redução foi possível graças a remoção de camadas totalmente conectadas e a utilização de camadas FC densas para o armazenamento dos pesos. Tais modificações permitiram uma grande economia de memória.
2. O tamanho mínimo da imagem de entrada da GoogLeNet é 224x224. É necessário por conta da profundidade do modelo, e após um certo número de camadas de convolução e re-escalamento, o resultado gerado será pequeno demais para poder ser processado a depender se seu tamanho for menor do que isso.
3. A principal diferença entre as duas reside na forma como elas realizam o pooling. No caso da Residual, praticamente não existem camadas de pooling para re-escalamento, ao invés disso, esse processo ocorre por meio de camadas convolucionais específicas com stride de 2, resultando numa redução em fator de 2. A Inception faz esse processamento por meio de diversas entradas paralelas de processamento de dados orientadas a um processamento multi-escalar. Este bloco se torna eficiente pela forma como são implementadas as camadas convolucionais 1x1.
4. Transfer learning é um método para reutilização de modelos de aprendizado de máquina, utilizando modelos prontos como ponto de partida para desenvolvimento de novos algoritmos com tarefas diferentes para as quais foram desenvolvidos inicialmente. É especialmente utilizado em problemas de visão computacional e processamento de linguagem natural, por conta de suas amplas necessidades de recursos e tempo para a criação de modelos eficientes.