Universidade Federal do Amapá - UNIFAP Inteligência Artificial - 2020.3

Aluno: Patrício Víctor Marinho Dias

Nesse documento, eu explicarei cada passo que fiz para a matriz confusão inicial ter se tornado como ficou no estágio final.

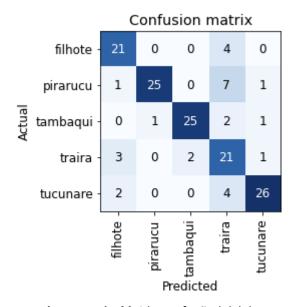


Imagem 1 - Matriz confusão inicial

No primeiro treinamento, vemos uma média de 30 imagens testadas para cada espécie de peixe, isso ocorre por conta da função **RandomSplitter** que, ao criar o data block, utiliza 20% das 150 imagens disponíveis de cada espécie apenas para testes, ou seja, 30 imagens.

A matriz confusão mostrava vários erros de predição, visto que haviam muitas imagens incorretas no dataset, com poucos detalhes, elementos extras que poderiam levar a máquina ao erro, como pescadores segurando o peixe, formas desenhadas, o peixe já pronto para o consumo ou outras espécies de peixe.

Uma análise dos dados da matriz inicial baseado nos erros e acertos:

Predições corretas: 118
Predições incorretas: 29

Taxa de erro: 0.2457627118644068

Após remover várias imagens incorretas da base de dados usando o **ImageClassifierCleaner**, a matriz mostrou-se mais precisa.

No final, restaram poucas imagens da espécie **filhote**, pois o Bing fez download de muitas imagens incorretas. Talvez isso deve ter ocorrido por considerar

também **filhotes** de qualquer espécie peixes e outros animais na busca, criando essa inconsistência, dessa forma, apenas 3 imagens sobraram para serem testadas.

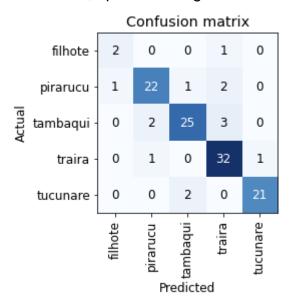


Imagem 2 - Matriz confusão final.

Mesma análise de dados, dessa vez com a matriz final:

Predições corretas: 102 Predições incorretas: 14

Taxa de erro: 0.1372549019607843

Dessa forma, conclui-se que a máquina obteve uma precisão aprimorada após a correção do dataset. Tivemos uma taxa de erro reduzida de 24% para 13%.