

Relatório Visão Computacional

Modelos treinados para detecção facial

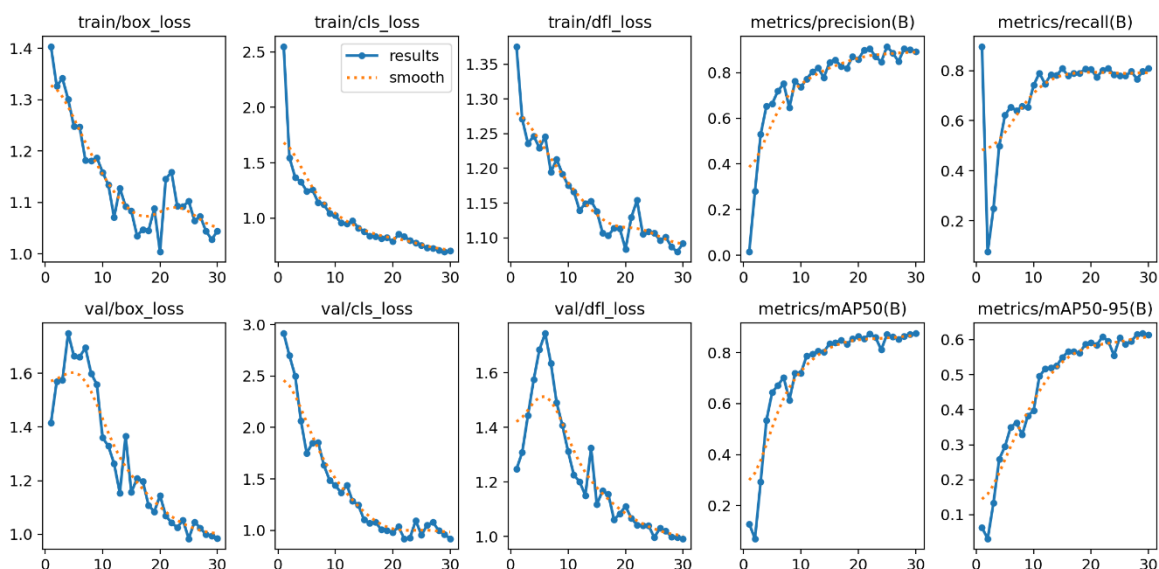
Ciro Guilherme Nass

Introdução

Para o treinamento do modelo pré-treinado YOLO yolo11n.pt, foram usados três datasets criados no site roboflow, um com 500 imagens, outro com 3 mil e o último contendo 5mil imagens. Todos montados a partir do de 500 com junções de datasets já existentes com distribuição de imagens para 80% treino e 20% validação, os testes foram feitos somente com uma imagem que serão mostrados neste relatório. Os modelos foram treinados em 30 épocas e estão contidos no zip enviado, junto com a descrição por gráficos que o modelo autogera. Apesar de nem todos serem em si serem precisos, resultou que seus índices r e mAP50 excedem os 85% já o mAP50-95 está por volta dos 70% de precisão.

Dataset 500 imagens

Este dataset que foi criado para o treinamento do modelo, se encontra no link: <https://universe.roboflow.com/ciogn/face-identifier500>, após treinar, foi obtida esta precisão.



Já treinado o modelo, logo em seguida foi feita uma predição desta imagem com confiança 0.5 desta imagem:



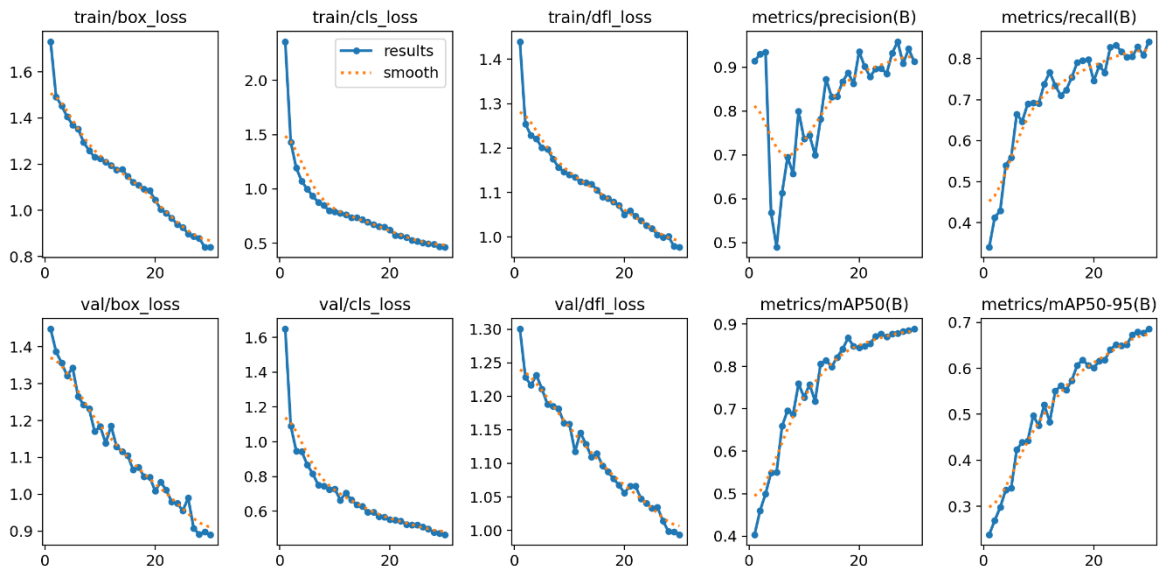
E obtivemos como resultado:



Vemos que nem todas as faces foram reconhecidas, isso pode ocorrer pelas imagens labeladas do dataset, que podem ser muito parecidas, sem variedade como distância, ângulo, identificação com acessórios, oclusões totais ou parciais dos rostos, iluminação ou até mesmo falta de imagens com ruído para serem usadas no treinamento. Outra condição que também implica na predição, é a confiança, que quanto mais perto de 0, mais o modelo vai labelar objetos que ele identifica como face, podendo também acontecer imprecisões com mais facilidade, identificamos isso pelo valor ao lado da face labelada na predição, quanto mais próximo de 0, menos chance de ser um rosto ao modelo.

Dataset 3000 imagens

Dataset disponível em: <https://universe.roboflow.com/cirogn/identiface-3k/dataset/>, com as configurações padrão de 80 e 20, foram obtidos estes resultados:



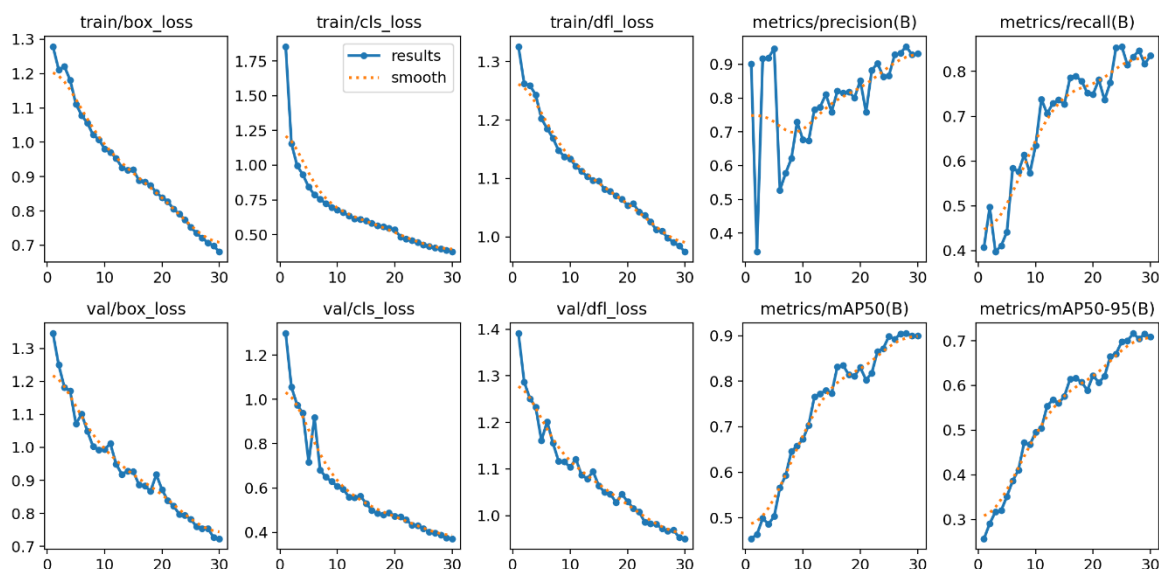
Novamente o r e mAP50 estão localizados acima de 80% de precisão e o índice mAP50-95 perto dos 70%. Com este modelo treinado e setado com confiança de 0.5, obtivemos estes resultados.



Como pode-se ver, alguns rostos antes identificados pelo modelo de 500 imagens, agora, não foram identificados. Isso pode-se dar por causa do dataset conter as imagens do dataset anterior somadas a outro de 2.5k imagens, independente disso, a classe é a mesma, que é 'face', isso para todos os datasets.

Dataset 5000 imagens

Dataset disponível em: <https://universe.roboflow.com/cirogn/faceidentifier5k/dataset/>, com as configurações padrão de 80 e 20, foram obtidos estes resultados:



Novamente os índices seguem o mesmo padrão de precisão. Após treinado, aqui o resultado da predição:



Desta vez, o modelo treinado por 5000 imagens identificou todos os rostos como esperado, com uma confiança de 0.5 foi alcançado o objetivo.

Conclusão

Os modelos aqui treinados, podem não alcançar a melhor precisão desejada em outras imagens, isso se dá por conta das imagens selecionadas nos datasets, quanto maior a quantidade de imagens assim como sua variedade em um dataset, melhor o modelo vai ser treinado para assim identificar os rostos.