

## Instruções

---

1. Esta avaliação deve ser feita de duas à 4 pessoas.
2. Data de entrega: **06/12/2023 até 18:50**. Trabalhos não podem ser entregues em atraso.
3. Esta avaliação tem por objetivo consolidar o aprendizado sobre reconhecimento de imagem e processamento digital de imagens.
4. A implementação deverá ser desenvolvida em Python.
5. O sistema deve ser entregue funcionando corretamente.
6. Deve ser apresentado uma a solução, no dia 06/12:
  - Identificação do autor e do trabalho.
  - Enunciado do projeto
  - Explicação e contexto da aplicação para compreensão do problema tratado pela solução
  - Desenvolvimento
  - Códigos importantes da implementação.
  - Resultados obtidos com a implementação
  - Análise e discussão sobre os resultados finais (inclusive de problemas)
7. Deve ser disponibilizado os códigos da implementação juntamente com a relatório (salvo o caso da disponibilidade em repositório aberto do aluno, que deve ser fornecido o link). O repositório deve estar aberto do momento da entrega em diante, sendo que o professor não responsabiliza caso o projeto não esteja disponível para consulta no momento da correção, sendo do(s) aluno(s) essa responsabilidade de manter disponível.
8. A não entrega e apresentação do trabalho acarretará nota zero para todos os integrantes. Todos os integrantes deverão ter conhecimento de todas as partes da implementação. O integrante que não apresentar, terá sua nota zerada no quesito apresentação. Todos os alunos deverão responder os questionamentos na apresentação.
9. Postar Slide + Repositório/Colab/Código no material da disciplina.
10. Cada dupla/grupo terá de 10 à 20 minutos para apresentar seu desenvolvimento. A inferência estará passível de ser testada com qualquer imagem no momento da apresentação.

## Descrição do projeto a ser desenvolvido

---

Um dos desafios do século XXI, no campo do *Machine Learning*, foi desenvolver modelos de reconhecimento facial que funcionassem para uma grande variedade de pessoas. Originalmente, podemos pensar que cada pessoa fosse um indivíduo único, sendo assim uma classe. Logo, os primeiros modelos foram baseados em modelos de classificação.

Contudo, havia um grande inconveniente nesse processo: toda vez que uma pessoa nova surgia para ser cadastrada no banco de dados, era necessário treinar o modelo novamente, uma vez que você incluiu uma nova classe no seu problema. Antes havia  $N$  classes, com a introdução de uma nova pessoa  $N + 1$ . Conforme o sistema fosse aumentando, acabava complicando o procedimento para colocar o sistema em produção novamente. Uma saída para esse problema, que acabou virando um campo específico de *Deep Learning*, é o Aprendizado de Métrica, no inglês *Metric Learning*. O método consiste em invés de trabalharmos com um problema de classificação, treinamos o modelo para aprender uma métrica de distância entre duas entidades. A saída de um modelo que foi treinado mediante o método de *Metric Learning* é um vetor descritor. Esse vetor é responsável por computar as características, ou features, que descrevem a face de uma pessoa. O método de *Metric Learning* tornou-se promissor,

possibilitando a inclusão de novas pessoas no banco de dados, de forma que não fosse necessário re-treinar o modelo.

Contudo, apesar de todos nós possuirmos características únicas em nossas faces, a pandemia do Covid-19 trouxe um grande problema para os modelos de reconhecimento facial, as máscaras faciais. Cobrindo parte importante do nosso rosto, do qual os modelos de *deep learning* extraíam características para computar o vetor descritor. Com isso, os modelos de reconhecimento facial apresentavam um alto índice de falhas. Algumas soluções foram encontradas para contornar esse problema.

Nessa tarefa, você precisará fazer o seguinte:

1. Treinar um modelo de rede neural nesse conjunto de dados disponível no material da disciplina.
  - a. O conjunto de dados possui uma grande variedade de celebridades.
2. Uma vez treinado o modelo, vocês devem criar um banco de dados com as imagens das celebridades com o output do vetor descritor do seu modelo.
3. Após o treinamento da rede neural, vocês devem incluir a foto de 1 integrante do grupo no banco de dados.
4. Uma vez concluídas as etapas (1), (2) e (3), vocês devem tirar a pessoa do item (3) usando máscara facial e fazer o reconhecimento facial dela.

## OBSERVAÇÕES

- Caso você tenha problemas em usar a GPU do Colab, pode utilizar uma rede pré-treinada para criar o vetor descritor de características dos rostos.
- É obrigatório exibir a imagem que será usada para inferência, isto é, a imagem do item (4), e exibir qual é o score da métrica que você usará e, também, qual foi a pessoa com a qual ela foi identificada no banco de dados.
- Caso você nunca tenha trabalhado com *metric learning*, aqui está um post detalhado sobre o assunto: [Link](#).
- Para possibilitar a pontuação completa do trabalho, é importante que seu algoritmo consiga fazer o reconhecimento da pessoa do item 4.
- Para a extração de descritor vector (ou feature vector), pesquise sobre MetricEmbedding. Você também pode entender sobre o conceito nesse material: [Link](#).