

Relatório da Segunda Etapa: Hand Tracking

Processo Seletivo IEEE RAS Voluntários

Objetivos:

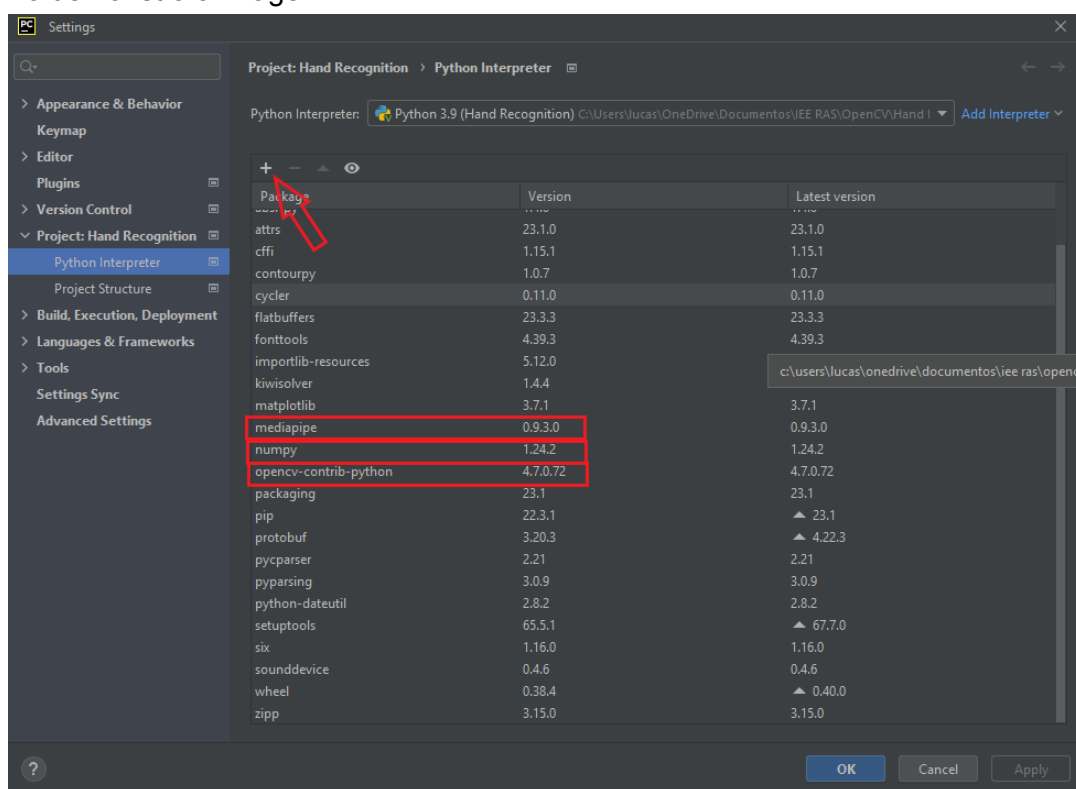
O objetivo dessa etapa é testar o aprendizado da etapa passada, desenvolver novas técnicas e aprender novos tópicos, por meio do desenvolvimento de um projeto dos seis apresentados.

Introdução:

O desafio de Hand Tracking visa criar um algoritmo capaz de contar os dedos em vídeos capturados em tempo real por uma webcam, utilizando as bibliotecas OpenCV e Mediapipe no Python.

Descrição:

Utilizando o IDE PyCharm com o Python 3.9. Utilizando o menu Settings do Pycharm, na seção do interpretador, foi instalado as seguintes bibliotecas utilizando o botão + como demonstra a imagem::



Quando a biblioteca mediapipe for instalada, perceba que ela irá instalar uma série de outras bibliotecas, elas não serão utilizadas mas é pertinente mencionar que elas estão a disposição.

Lembrando que o Mediapipe já vem com diversos algoritmos de visão computacional e inteligência artificial pré-treinadas.

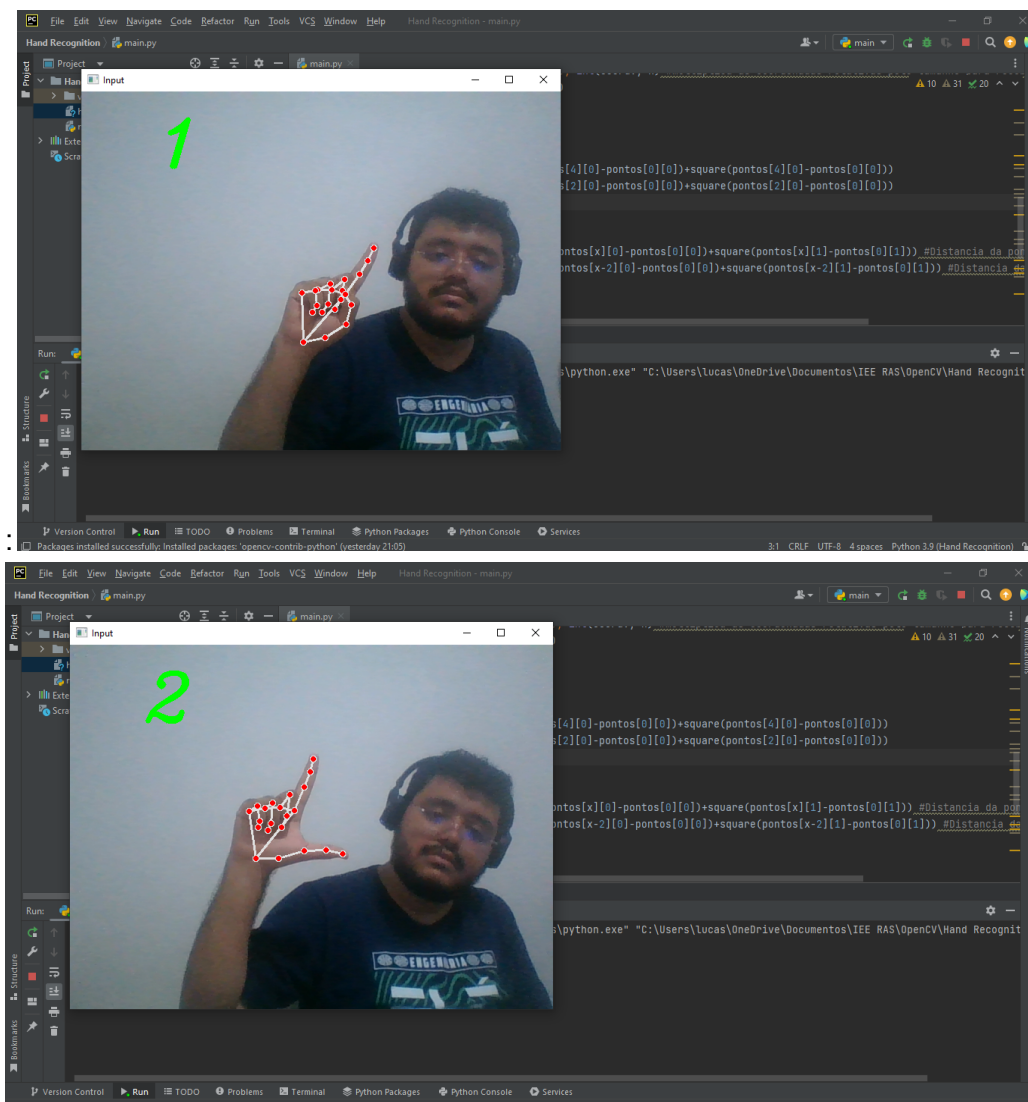
Importando cv2(OpenCV), obtemos a capacidade de utilizar diversas funções necessárias para completar o desafio, uma dessas funções é a cv2.VideoCapture(int. numero da webcam(se tiver mais de uma, caso contrário essa variável é 0)) que deve ser armazenada em uma variável, e utilizando a função var.read() obtemos duas variáveis, a imagem e um check. Outra função importante para esse projeto é a mediapipe.solutions.hands que é o algoritmo de reconhecimento de mãos.

Para não estender o relatório, recomenda-se que o código seja consultado, visto que ele está devidamente comentado explicando a lógica por trás de como o código funciona.

A lógica da contagem dos dedos é bem simples, a distância da ponta de um dedo à base será sempre maior que a da base do dedo à base, isso se aplica a quase todos os dedos, que iremos discutir na conclusão.

Resultados:

Ao fim do projeto, obtemos um programa capaz de contar quantos dedos estão abertos em tempo real, e deve ter o seguinte resultado:



Discussão:

Ao fim do projeto, percebemos que o algoritmo, apesar de ser treinado pela google, ainda falha em alguns momentos mesmo que brevemente.

Outro problema encontrado foi a lógica para o dedão, no início do projeto foi utilizado uma lógica de posição absoluta dos nós do esqueleto na tela, mas isso vinha com uma desvantagem, apesar de ser rápido e “preciso”, essa lógica não levava em conta a inclinação da mão, portanto foi decidido utilizar uma lógica de diferenças de distâncias, mas essa lógica também se prova falha com o dedão em algumas situações onde, mesmo a mão estando fechada, essa distância ainda é maior.

O código ainda tem muito o que melhorar, mas para o objetivo desse projeto, foi decidido que essa implementação será o suficiente.

Conclusão:

A biblioteca openCV é muito interessante, as possibilidades de tratamento e trabalho com imagens é imensa, para esse projeto conclui que apesar de ser bastante técnico, a biblioteca é fácil e intuitiva de se usar.