

TP555 - Inteligência Artificial e Aprendizado de Máquina

Projeto Prático

Proposta

Nome: Ranyeri do Lago Rocha

Curso: Doutorado em telecomunicações

Matrícula: 40

Github: https://github.com/rlrocha90/TP555_projeto_final.git

Descrição: neste projeto prático será realizada a reprodução do trabalho chamado “MixFormer: End-to-End Tracking with Iterative Mixed Attention”, reproduzindo com apenas as sequências envolvendo barcos que existem no dataset LaSOT. O algoritmo é de 2022 e tem atualização em fevereiro de 2023, que será utilizado neste projeto prático. O MixFormer usa as principais técnicas atuais utilizadas em visão computacional, como Visual Transformer e Mixed Attention Module (MAM, visando extração de features de forma progressiva a partir do target objetivo e a área de busca. Junto com MAM são utilizadas camadas de MLP. Além, na localização dos cantos da “bounding box” do objeto trackeado, são utilizadas várias camadas Conv-BN-ReLU para os cantos esquerdo-superior e direito-inferior. O trabalho original utilizou quatro datasets e neste projeto prático será utilizado apenas parte de um deles. Há a possibilidade de inclusão de outras sequências de vídeos, se disponível com as informações de bounding box verdadeiro.

Artigo: MixFormer: End-to-End Tracking with Iterative Mixed Attention

Autores: Yutao Cui, Cheng Jiang, Limin Wang e Gangshan Wu.