Project No.	Project Title	
2022-01-173	Pose Estimation tools comparison on swimmers' data	
Academic Advisor	**Co-Advisor**	
Prof. Raziel Riemer	Prof. Gera Weiss	
Team Members		
Orit Rabani	-	-
oritrab@post.bgu.ac.il	-	-

Abstract

Pose Estimation is a method to find human joints location in a picture. This study examines several Pose Estimation tools and their performance on under water pictures of swimmers. This research learning the fundamentals of how each algorithm and architecture works and evaluating the performance of Pose Estimation platforms with and without training on our swimmers' data. This aim will be achieved by adapting and improving some of the models so that the models can deal with a new type of data - swimming videos. Where the long goal is to be able to analyze swimming technique and provide feedback on the performance.

Pose estimation refers to computer vision techniques that detect human figures in images and videos, so that one can determine the location of the human joints. A human pose skeleton represents the orientation of a person in a graphical format. Essentially, it is a set of coordinates that can be connected to describe the pose of the person. Because Pose Estimation architectures were trained on human data outside of the pool which resulted in different view of the body (i.e., observing a human from above) they have difficulty to find the swimmer's joints in a swimming picture due to missing joints. This let to the need to study and modify Pose Estimation architectures and to find the best one to use for our data.

Our data contains six swimming videos of two males and one female, a total of 1,187 frames. Five videos were used for training (1,043 frames), and one video was used for validation (144 frames). Several Pose Estimation implementations were examined including "MediaPipe", "MoveNet", and "Detectron2" with and without training. We analyzed the performance on our data by using common metrics – Percentage of Detected Joints (PDJ), Object Keypoint Similarity (OKS) as well as visually. Both metrics represent how close the predicted joint is from the ground truth based on different elements of the picture. The model which provided the best performance was Detectron2 after re-training on swimmers' data. All the implementations of the different models were performed with relevant Python packages and were run in "Google Colab", on a Tesla P100-PCIE-16GB GPU.

Our results show a positive indication that existing pose estimation implementations can detect swimmers' joints accurately. Future work will focus on continuing to modify these architectures to enhance their performance on swimmers' data and analyzing swimmers' movement to be able to provide an automatic and precise feedback.

Keywords: Pose Estimation, Computer Vision, Swimming ,Swimmers' data