

## Joint 2D and 3D Pose Estimation and Action Recognition

The aim of this research project is to enhance the accuracy of 2D and 3D pose estimation from still images and human action recognition from video sequences by developing a multitask framework that leverages common underlying information. The proposed methodology consists of data collection from diverse datasets, such as MPII, Human3.6M, Penn Action, and NTU, followed by preprocessing to eliminate noise, outliers, and artifacts. Subsequently, features will be extracted using CNNs and RNNs, and a novel multitask framework will be developed for joint 2D and 3D pose estimation and action recognition. The framework will use shared feature representations, attention mechanisms, and end-to-end optimization. The effectiveness of the developed framework will be evaluated on benchmark datasets using metrics like accuracy, precision, and recall. The project will also involve applying the framework to real-world problems, such as human behavior analysis in surveillance and activity recognition in sports.

### References

- [1] Luvizon, Diogo C., David Picard, and Hedi Tabia. "2d/3d pose estimation and action recognition using multitask deep learning." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
- [2] Luvizon, Diogo C., David Picard, and Hedi Tabia. "Multi-task deep learning for real-time 3D human pose estimation and action recognition." *IEEE transactions on pattern analysis and machine intelligence* 43.8 (2020): 2752-2764.
- [3] Diogo Carbonera Luvizon, Hedi Tabia, and David Picard. "Learning features combination for human action recognition from skeleton sequences." In: *Pattern Recognition Letters* 99 (2017), pp. 13–20.
- [4] Diogo C. Luvizon, David Picard, and Hedi Tabia. "Consensus-Based Optimization for 3D Human Pose Estimation in Camera Coordinates." In: *Int. J. Comput. Vis.* 130.3 (2022), pp. 869–882.
- [5] Diogo C Luvizon, Hedi Tabia, and David Picard. "Human pose regression by combining indirect part detection and contextual information." In: *Computers & Graphics* 85 (2019), pp. 15–22.
- [6] Luvizon, Diogo Carbonera, Hedi Tabia, and David Picard. "Ssp-net: Scalable sequential pyramid networks for real-time 3d human pose regression." *Pattern Recognition* (2023): 109714.
- [7] Akremi, Mohamed, Rim Slama, and Hedi Tabia. "SPD Siamese Neural Network for Skeleton-based Hand Gesture Recognition." *17th International Conference on Computer Vision Theory and Applications VISAPP 2022*. SCITEPRESS-Science and Technology Publications, 2022.
- [8] Ben Charrada, Tarek, et al. "TopoNet: Topology Learning for 3D Reconstruction of Objects of Arbitrary Genus." *Computer Graphics Forum*. Vol. 41. No. 6. 2022.