

Hand Gesture Tracking in Videos using YOLO_V8

Student Names: Mian Wang, MSDS @ Portland campus

Professor Names: Jeongkyu Lee, Supervisor name: Mohammad Toutiaee



Introduction

In the "Guess Which Hand" game, a player hides an object in one hand, and another guesses its location. This project trains an AI to recognize hand positions and postures using 20 videos with annotated landmarks. The goal is to advance posture recognition for applications in gaming and human-computer interaction.

PROBLEM Definition

The task is to predict which hand holds the object in the "Guess Which Hand" game using video data. The AI must track hand gestures, detect landmarks (e.g., wrist, fingers), and handle challenges like occlusions and lighting variations. The output is a prediction of the object’s likely hand location.

APPROCHES (or Methods)

Dataset Preparation: Preprocessed a dataset of 20 videos, including annotation and augmentation to ensure robust model performance across different scenarios.

Model Selection: Employed the YOLO (You Only Look Once) model for pose estimation and hand gesture tracking, leveraging its real-time object detection capabilities. Tried different versions like V10, V11, and V8 is best on metric.

Landmark Generation: Generated hand landmarks (e.g., wrist, thumb, index finger) and predicted results, which were saved as CSV files for analysis.

OUTCOME (or Results)

Frame	Timestamp	hand_index	hand_label	landmark_index	landmark_name	x	y	z	left_hand_present	right_hand_present	noitiseb9	rtuutT bnuotD	bi_oebiv
0	0.0	0	Right	0	WRIST	0.758034466306710	0.754531350135000	1.00862310148000E-07	FALSE	TRUE	†	†righ	†
0	0.0	0	Right	1	THUMB_CMC	0.768715000878910	0.728202741355896	-0.03405811102880910	FALSE	TRUE	†	†righ	‡
0	0.0	0	Right	2	THUMB_MCP	0.77506380324027	0.750730683403020	-0.045160002185620	FALSE	TRUE	†	†righ	§
0	0.0	0	Right	3	THUMB_IP	0.7684125198242840	0.771088548504070	-0.0509738719819100	FALSE	TRUE	†	†righ	¶
0	0.0	0	Right	4	THUMB_TIP	0.75807325574402800	0.7689101186725190	-0.0531452716190200	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	5	INDEX_FINGER_MCP	0.7981778978301450	0.777878284545360	-0.01844208145141600	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	6	INDEX_FINGER_PIP	0.7752847241401670	0.804400742053860	-0.0370184886425500	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	7	INDEX_FINGER_DIP	0.7458855011940000	0.804189780475190	-0.05178854870457600	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	8	INDEX_FINGER_TIP	0.721728632087140	0.791030155880500	-0.0580708417018800	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	9	MIDDLE_FINGER_MCP	0.781608683330540	0.777342796235860	-0.05540486016112950	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	10	MIDDLE_FINGER_PIP	0.754616787387220	0.7984251842787100	-0.02348030480783500	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	11	MIDDLE_FINGER_DIP	0.7251232865472030	0.796207010740020	-0.03819483321198100	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	12	MIDDLE_FINGER_TIP	0.701842008848150	0.787774716488280	-0.04854814302817500	FALSE	TRUE	0	†righ	⌘
0	0.0	0	Right	13	RING_FINGER_MCP	0.7628886870101200	0.774109442830080	-0.004188780806732000	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	14	RING_FINGER_PIP	0.738020417889580	0.762633788108820	-0.01204476564675700	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	15	RING_FINGER_DIP	0.7158321142196880	0.790889560434340	-0.02031284885802200	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	16	RING_FINGER_TIP	0.6980320700367480	0.782857315444840	-0.02017725785007430	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	17	PINKY_MCP	0.7457074522972110	0.768486078162820	-0.01288082890708000	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	18	PINKY_PIP	0.7288070203108830	0.782025428668820	-0.018646457711800200	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	19	PINKY_DIP	0.71208844270995	0.7818864583989120	-0.00891483125835050	FALSE	TRUE	†	†righ	⌘
0	0.0	0	Right	20	PINKY_TIP	0.6986371278762820	0.773725364868580	-0.01205782185782000	FALSE	TRUE	†	†righ	⌘
0	0.0	1	Left	0	WRIST	0.12247042028391880	0.716843681240080	-1.20834857843400E-07	FALSE	TRUE	†	†righ	⌘

Accuracy: 85%

Conclusions

The YOLO model effectively tracked hand gestures and predicted the presence of the right hand in the "Guess Which Hand" game, demonstrating its potential for posture recognition tasks.

The generated CSV files provided valuable insights into hand landmark positions, enabling precise analysis of hand movements.

References

YOLO: Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). (CVPR).