

A novel lightweight skeletal temporal model for real-time, computationally efficient recognition of occupant thermal adaptation behavior

Supporting information to <https://doi.org/10.1007/s12273-025-1335-6>

Table A.1 Steps to Convert the KTH Dataset to the AVA Dataset Format

AVA Folder	Steps to Convert the KTH Dataset to the AVA Dataset Format
Original videos	1. Use the FFmpeg tool to convert all videos in the KTH dataset to 30FPS. 2. Save processed videos as .mp4 files in the ‘Original videos’ folder.
Cropped videos	1. Use the FFmpeg tool to cut the first 11 seconds from each video in the ‘Original videos’ folder. 2. Save processed videos to the ‘Cropped videos’ folder.
Rawframes	1. Use the FFmpeg tool to extract specific frames at a 30 FPS interval for each video in the ‘Cropped videos’ folder. 2. Save processed images as .jpg files to the ‘Rawframes’ folder.
Annotations	1. Divide the images in the ‘Rawframes’ folder into train, validation and test sets according to the individual IDs in the KTH dataset, with a 6:2:2 split ratio. 2. Label the images in the ‘Rawframes’ folder using VIA3 and export the labeled data as a .csv file. 3. Write Python code to read the CSV file, process video frame data, extract position information and action labels, and generate the .pkl and .csv files in the ‘Annotations’ folder for training, validation and testing, along with a label mapping file.

Table A.2 Chosen model’s hyperparameters and ranges

Hyperparameter	Range
hidden_size	[32, 64, 128, 256, 512]
num_layers	[1, 2, 3, 4, 5]
learning_rate	[0.0001, 0.0005, 0.001, 0.005, 0.01]
momentum	[0.8, 0.82, 0.84, 0.86, 0.88, 0.9, 0.92, 0.94, 0.96, 0.98]
dropout	[0.1, 0.3, 0.5, 0.7, 0.9]

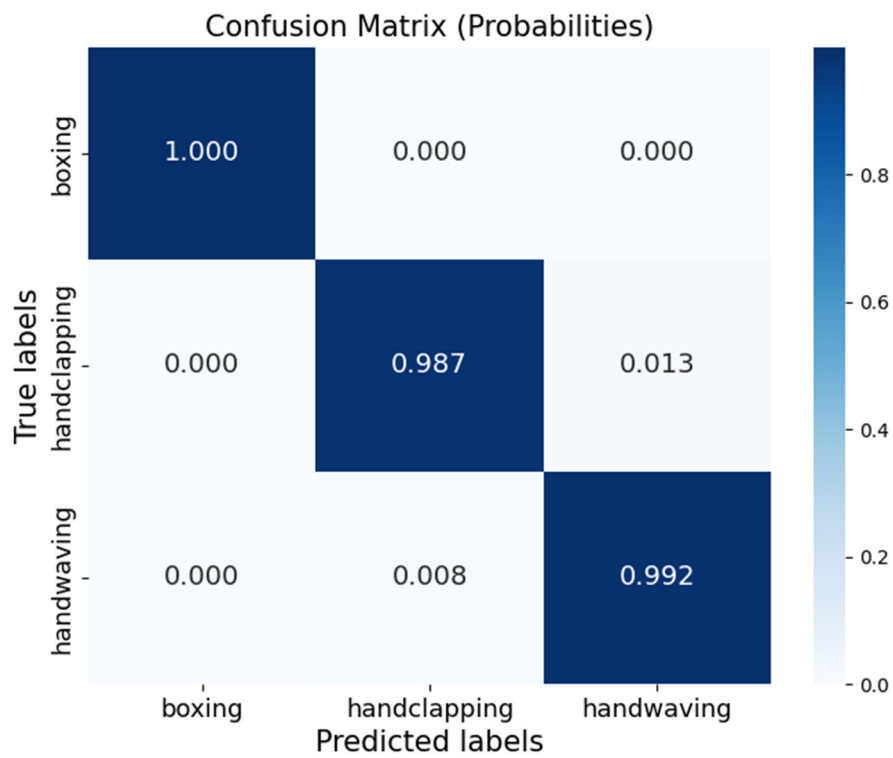


Fig. A.1 Confusion Matrix for all actions (SF)

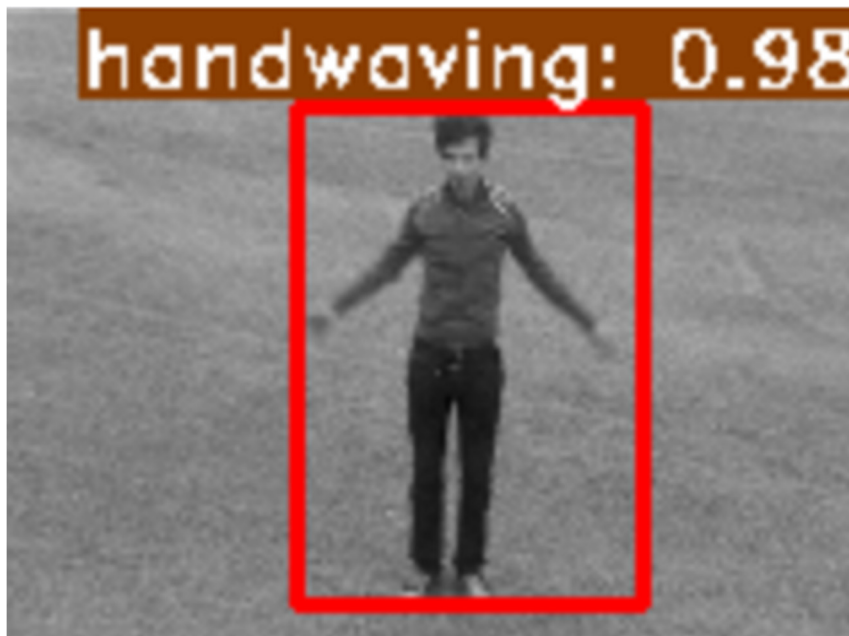


Fig. A.2 SF real-time recognition demo: graphical user interface (GUI)

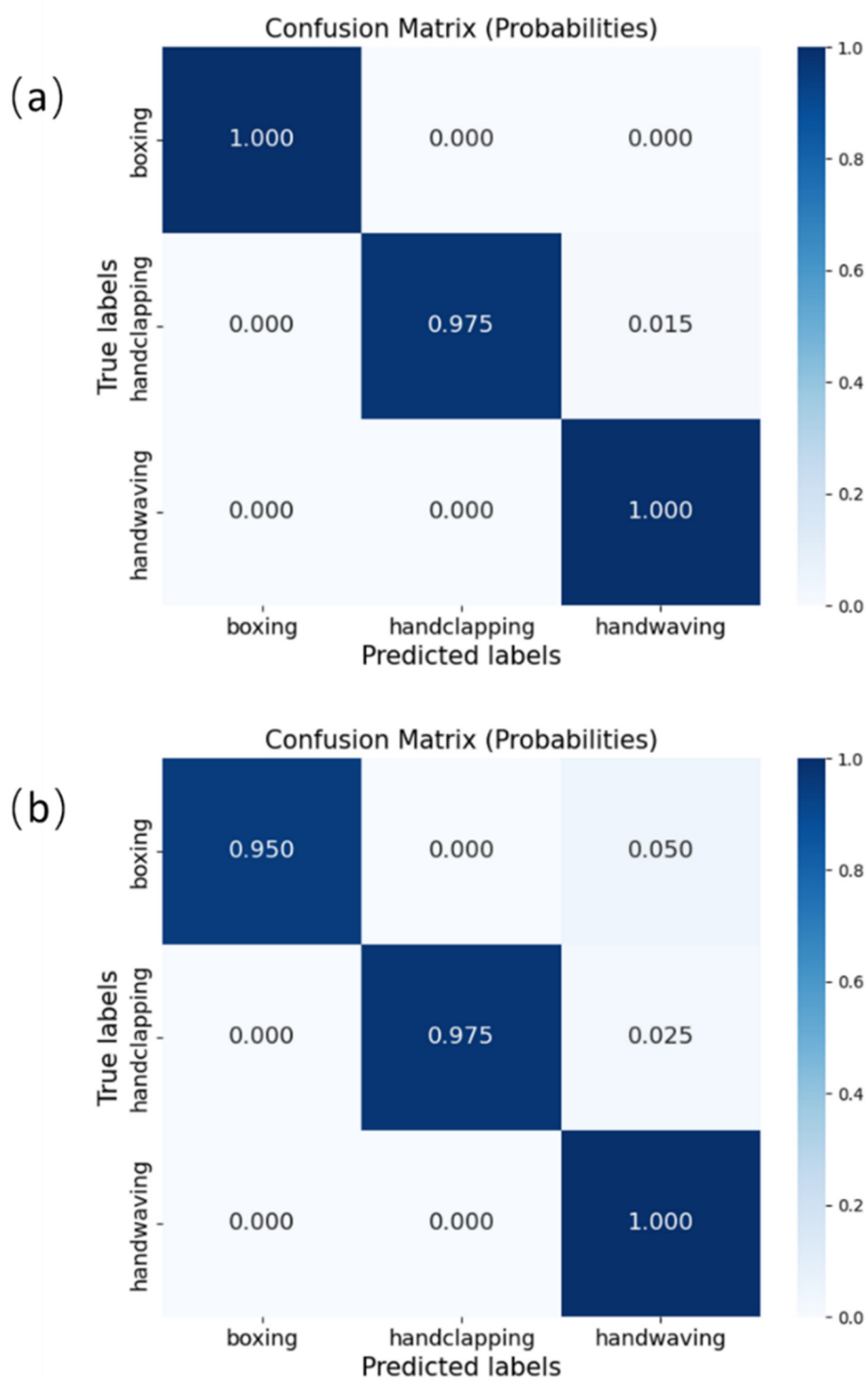


Fig. A.3 Confusion Matrix for all actions: (a) OpenPose; (b) LightOpenpose

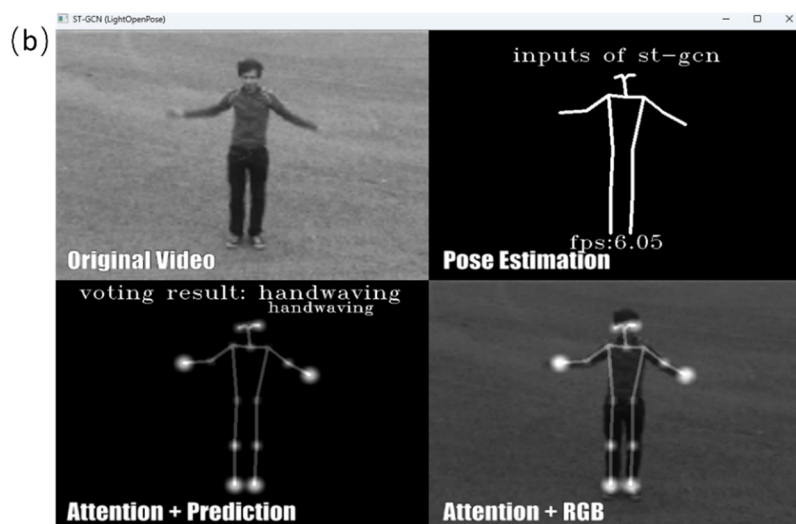
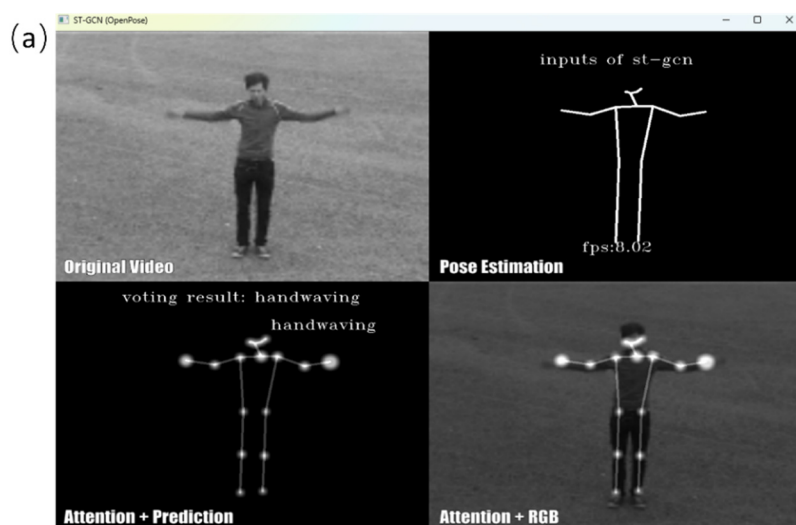


Fig. A.4 ST-GCN real-time recognition demo: GUI (a) OpenPose; (b) LightOpenpose



Fig. A.5 NLST model real-time recognition demo: GUI

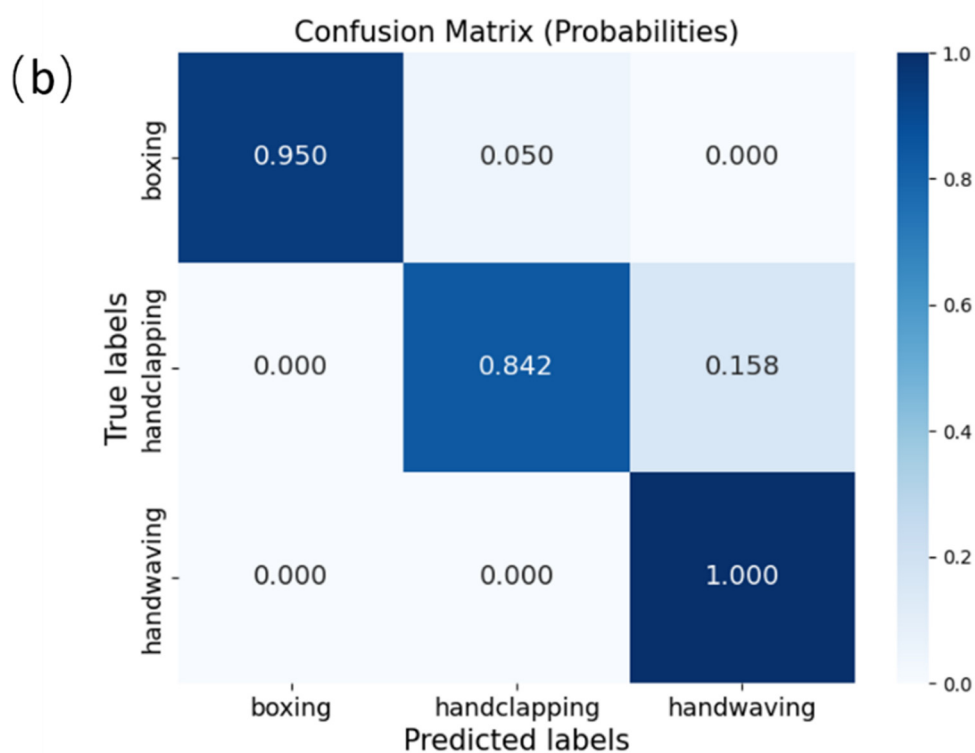
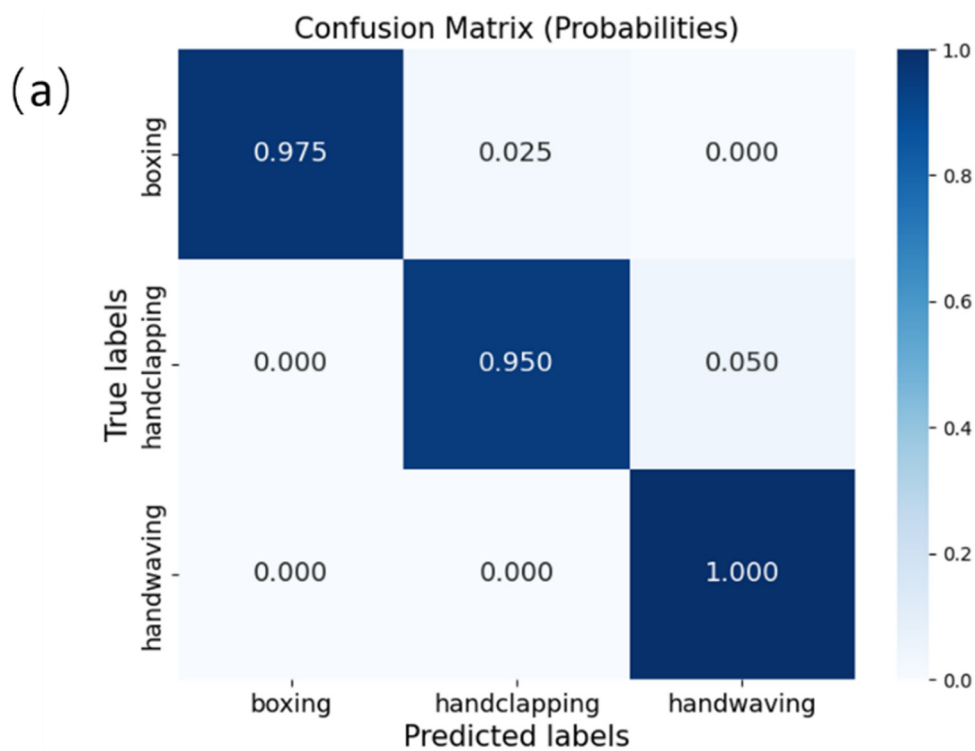


Fig. A.6 Confusion Matrix for all actions: (a) Trail 15; (b) Trail 19

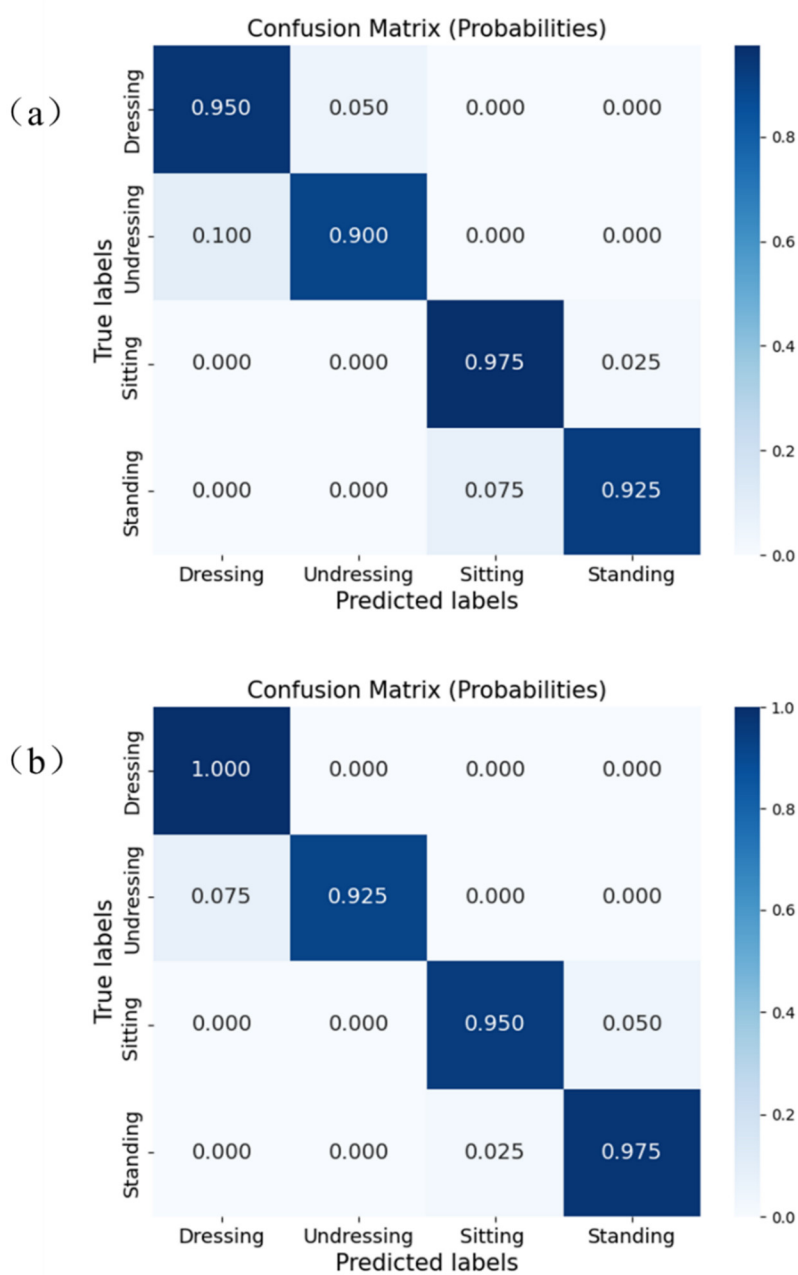


Fig. A.7 Confusion Matrix for all actions: (a) LSTM with attention; (b) Double LSTM with attention

Dataset download link:

<https://drive.google.com/file/d/1neKPpnLcLz7knicW1pgRCe9J5jkGytgJ/view?usp=sharing>

Video 1 link:

<https://drive.google.com/file/d/1NX2FTBRfSXu4nXOMklQ7amRx-RCRgMuH/view?usp=sharing>

Video 2 link:

https://drive.google.com/file/d/1clKfmjpvddjobxpN_rsmSWE74kBRDsGz/view?usp=sharing