Supplementary materials

Gesture	Applications	Gesture	Applications
one, peace, three, three2, four, palm	numeric value input, e.g. volume scale, light level, etc., also possible their combinations for range extension	call	– incoming call acceptance – making a call on the intercom in the smart home
like, dislike	– like / dislike or save/remove something (e.g. music track or video) – evaluating content to improve recommendations – expressing approval/disapproval for videoconference participants	mute	– switching applications to silent mode – mute the microphone during a videoconference – switching the smart home to "night mode"
stop, stop inverted	– stop something (e.g. music or video playback) – swipe to scroll content when combined with the gesture "stop inverted"	ok	– command confirmation – confirmation of smart home mode switching
peace, peace inverted	– switching the smart home to relax mode – activation something – rotation of the object when combined with the gesture "peace inverted"	fist	– expressing applause, approval or encouragement for videoconference participants – dragging objects when combined with the gesture "palm"
two up, two up inverted	– swipe to scroll content when combined with the gesture "two up inverted"	rock	– launching entertainment mode for smart home – changing the numerical value of any smart home/application characteristic to a maximum

Table 4. Possible uses of gestures. Gesture "three2" has the same meaning as the gesture "three" as the last is inconvenient to show by some people.

Model	Weight Decay	Learning Rate	Scheduler	Scheduler' Params.
ResNet	1^{-4}	1^{-1}	ReduceLROnPlateau	mode: min, factor: 0.1
MobileNetV3	5^{-4}	5^{-3}	StepLR	step size: 30, gamma: 0.1
VitB16	5^{-4}	5^{-3}	CosineAnnealingLR	T max: 8
RetinaNet	1^{-4}	1^{-2}	StepLR	step size: 30, gamma: 0.1
SSDLite	5^{-4}	1^{-3}	StepLR	step size: 30, gamma: 0.1
YoloV7	5^{-4}	1^{-2}	LambdaLR	sinusoidal function

Table 5. Training hyperparameters.

Figure 5. Examples of labeled samples from HaGRID. Gestures and “no gestures” are highlighted in yellow and green bounding boxes, respectively.