**Catalogue of Actions to be Detected**

**Action Performance Variables:**

* *Sitting* – This is mostly for when it makes sense. For example, usually cutlery and writing actions are performed while sitting. However, object actions can be performed sitting and standing so for some actions we need to record both but only where it seems necessary.
* *Speed* – The actions need to be performed in the same way, but at a variety of speed levels in order to have enough data to train the machine to recognise the same action being performed by a real test subject regardless of how fast they are performing the action.
* *Lighting* – The Machine needs to be able to detect the action being performed regardless of the lighting condition inside the room (unless it is either very bright or dark which is outside the scope of what we are trying to detect). Actions need to be re-recorded under different lighting conditions that includes different lighting colours along with brightness.
* *Background* – We need to record the actions against a variety of simple backgrounds to help the machine detect the same actions in different environments. This will make it easier for the initial training.
* *Skin shown* – The machine needs to recognise the actions being performed regardless of the clothing worn by the test subject (within reason). All actions should be re-recorded with sleeved actors where the majority of the arm is covered.
* *Left/Right handed* – All the actions need to be repeated and re-recorded for both hands in order to accommodate for the hand being used by real test subjects. The machine needs to be trained to recognise the action regardless of the hand being use to perform the action.

**Definite:**

*These are a select few actions that we will be focusing on and using as a basis to create our data set around. Our project’s current goal is to detect these actions following actions from newly uploaded footage.*

1. (R/L) - Swinging arms – [such as while walking]
2. (R/L) – Door – (Open/Close)
3. (R/L) – Object – (PickUp/PutDown/Pull/Push)
4. (R/L) – Cutlery – (Knife[cut]/Fork[stab]/Spoon[scoop])
5. (R/L) – Handshake
6. (R/L) – Writing
7. (R/L) – Wiping – [like the motion of scrubbing a surface]
8. (R/L) – Turning – (Horizontal/Vertical) – [Like a tap or valve]

**Probable:**

*These are other actions which we currently believe have unique enough features that should allow us to set up rules to detect them.*

(R/L) – Mouse – (Moving/Clicking)

(R/L) - Waving

(R/L) - Switch flick - (on/off)

(R/L) – Pointing – [potentially too similar to ‘Switch flick’ category]

Washing Hands – [as in rubbing together motion]

Typing

(R/L) - Furniture - (PullOut/PushIn) - [Potentially absorbed into the *‘Object’* category]

(R/L) - Packing a Bag (PuttIn/TakeOut) – [Potentially absorbed into the ‘*Object*’ category]

**Questionable:**

*These are actions that we currently believe may be too difficult to set up rules to detect.*

Interactions with specific everyday items (phone, wallet, money, headphones, deodorant, other stationary, etc) – *potentially too hard to differentiate from one another or from some of the other pre-established categories. However, some of them may be possible, we can potentially experiment if the earlier categories prove successful*.

Eating/Drinking – *We potentially lose too much data to accurately detect these types of ‘above camera’ actions at all due to restrictions with this method of recording discussed in more depth in the camera evaluation section below.*

Eyewear (putting on/taking off) – *Another ‘above camera’ that is now in doubt unless we perform object detection.*

Squeezing – *When holding an object it might be too hard to detect that it is being specifically squeezed and not just held normally.*

Pressing Buttons – *might not be intricate enough to differentiate from some of the other actions like flicking a switch or pointing.*

Appling makeup (or other facial actions) – *Another ‘above camera’ that is now in doubt unless we perform object detection.*

Combing (or other hair actions) – *Another ‘above camera’ that is now in doubt unless we perform object detection.*

Stretching/Exercise *– although some specific/unique ones may be distinguishable enough to detect.*

Cooking actions (chopping, slicing, grating, peeling, stirring, etc) – *may be too difficult to differentiate from other actions (the cutlery related ones for example, though there may be some crossover there), we may be able to have a look into it later if time allows.*

Blowing nose – *tissue or handkerchief may obstruct camera, also it is another ‘above camera’ action that we are unsure if will be caught on footage.*

Dressing *- perhaps putting shoes on is one aspect that might be possible, but since the chest strap needs to be on for the camera, we’re not going to get footage of people putting their clothes on.*

*The possibility of using mirrors to help us capture movements that were too difficult to catch from the first person perspective was discussed as a possible option. Although we did have to consider that in a real life scenario, patients would not be performing all of their daily activities in front of mirrors so it may not be practical to devote a large portion of time to exploring this option. At the very least, we should wait until we have succeeded with the “Definite Category” before exploring this method in depth.*

**Ignore:**

*Actions that we’d prefer not to detect on purpose.*

* Random Scratching/fidgeting
* Other people’s movements
* Movements caused by stability of camera/speed or sudden rotation of the user