**Step 1**: `Identify the customer may processed by many techniques as`:

- capture the face of user and identifie him to recognize him at his next visit using: feature extraction then some classification operation such as dense-SIFT, Histogram of Oriented Gradients (HoG) features and classifiers can be k-nearest neighbor, support vector machine (svm) classifiers
- Via QR generator and reader on smartphones

**Step 2**: `know user's Action well`:

- The camera will taking several pictures of customer in different poses such as: walking with free hands, holding a product using Action recognition by CNN as stablished here :https://papers.nips.cc/paper/5353-two-stream-convolutional-networks-for-action-recognition-in-videos.pdf
- Estimate all users poses by this Real-Time beautiful technique https://github.com/michalfaber/keras_Realtime_Multi-Person_Pose_Estimation/blob/master/README.md?fbclid =IwAR3af5C5dYOvN9XkQlJkJAa4X8XDrnWQu1hL7mLTH Atnvm9YuYdV8CY7nbc

**Step 3**: `detect products movement -customer usually buy/need  at specific times-` using:

- histogram equalization or segmentation and classifier as SVM, Logistic regression or Naïve bayes to recognize it
- Sensors on shelfs to detect the product movement if it's token or returned back

**The final step** : `Match the person again at the walkout door` to their previously computed feature map using sift to end the tracking.

Now we have the needed data from each user so can create a feature map and feed our CNN.