The 2020 CORSMAL Challenge

Multi-modal fusion and learning for robotics

Performance scores

Performance scores

For fullness classification (Task 1) and filling type classification (Task 2), we compute precision, recall and F1-score for each class n across all the configurations belonging to class n, S_n. *Precision* is the number of true positives over the total number of true positives and false positives. *Recall* is the number of true positives over the total number of true positives and false negatives. *F1-score* is the harmonic mean between precision and recall, defined as

$$F_n = \frac{P_n R_n}{P_n + R_n}, \tag{1}$$

with P_n the precision and R_n the recall for each class n. We then compute the Weighted Average F1-score (WAFS) across the N classes, each weighted by the number of configurations in class n, S_n,

$$WAFS = \frac{1}{S} \sum_{n=1}^{N} S_n F_n, \tag{2}$$

with $S = \sum_{n=1}^{N} S_n$ being the total number of recordings. $N = 3$ for fullness classification (T1). $N = 4$ for filling type classification (T2).

For container capacity estimation (Task 3), we compute the relative absolute error between the estimated capacity, $\hat{x}_j^c$, and the annotated capacity, x_j^c, for each configuration, j, of container c,

$$\varepsilon_j^c = \frac{|\hat{x}_j^c - x_j^c|}{x_j^c}. \tag{3}$$

We then compute the Average Capacity Score (ACS), that is the average score across all the configurations S and all the containers,

$$ACS = \frac{1}{S} \sum_{c=1}^{C} \sum_{j=1}^{S_c} \exp(-\varepsilon_j^c), \tag{4}$$

where S_c is the number of configurations with container c. Note that estimated and annotated capacities are strictly positive, $\hat{x}_j^c > 0$ and $x_j^c > 0$. If the capacity of the container c in recording j is not estimated, i.e. $\hat{x}_j^c = -1$, then $\exp(-\varepsilon_j^c) = 0$.

For each configuration j of container c, we then compute the filling mass estimation, $\hat{m}_j^c$, using the estimations of fullness from T1, $\hat{f}_j^c$, filling type from T2, $\hat{y}_j^c$, and container capacity from T3, $\hat{x}_j^c$, and using the prior density of each filling type per container, $D(\cdot)$,

$$\hat{m}_j^c = \hat{f}_j^c \hat{x}_j^c D(\hat{y}_j^c). \tag{5}$$

The density of pasta and rice is computed from the annotation of the filling mass, capacity of the container, and fullness for each container. Density of the water is 1 g/mL. The formula selects the annotated density for container c based on the estimated filling type.

For evaluating the filling mass estimation (overall task), we compute the relative absolute error between the estimated filling mass, $\hat{m}_j^c$, and the annotated filling mass, m_j^c, for each configuration, s_c, of container c, unless the annotated mass is zero (empty),

$$\varepsilon_j^c = \begin{cases} 0, & \text{if } m_j^c = 0 \wedge \hat{m}_j^c = 0, \\ \hat{m}_j^c, & \text{if } m_j^c = 0 \wedge \hat{m}_j^c \neq 0, \\ \frac{|\hat{m}_j^c - m_j^c|}{m_j^c}, & \text{otherwise.} \end{cases} \tag{6}$$

We then compute the Average filling Mass Score (AMS), that is the average score across all the configurations S for all the containers, using Eq. 4. Note that estimated and annotated filling masses are strictly positive, $\hat{m}_j^c > 0$ and $m_j^c > 0$. If the filling mass of the container c in configuration j is not estimated, i.e. $\hat{m}_j^c = -1$, then $\exp(-\varepsilon_j^c) = 0$.