RELATED WORK  
Data association of objects and SLAM are typically solved as decoupled problems in the literature. some paper showed that when the SLAM solution is known, and thus there is no uncertainty in robot poses, robot poses provide good prior information about object locations and can achieve better recalls than frame by frame detections.  
other work used a SLAM solver to build a 3D map of a room, and then fixed the map and manually labeled objects in the room. On the other hand, object detection can improve localization as well. we can pre-mapped doors and chairs as landmarks. During the navigation stage, these pre-mapped objects are detected online and their location information is used to localize the robot. However, in the scenario considered here, neither data association of objects nor robot poses are perfectly known. Algorithms that solve object detection and SLAM jointly can be categorized into front-end approaches and back-end approaches.

Front-end Data Association  
In front-end approaches, objects detected in new images are compared with previous images. If matches between new and old images are found, then corresponding objects are associated to the same unique identifier. These data associations by front-end procedures are taken as reliable and true, and then passed to a SLAM solver. In this work, instead of creating exact templates for  
objects, deep learning is used to detect objects in the environment. Deep learning generalizes much better than templatebased approaches. However, the detections have significant ratio of false positives and partial occlusions, thus are very challenging for front-end algorithms to produce reliable data associations.

Back-end Robust SLAM  
Robust SLAM is a line of research that explicitly use backend approaches to deal with outliers in the data. In robust SLAM, when some measurement is incorrectly associated, it will be inconsistent with other object measurements of the same identifier. Robust SLAM instead maximizes a set of measurements that are consistent with each other in both identifiers and predicted locations. Only the consistent measurements are plugged into a SLAM solver to recover the robot poses and landmark locations.  
By nature robust SLAM relies on the assumption that inlier measurements with unique identifier associations are the majority compared to outlier measurements. Under this assumption, eliminating outliers can still give good SLAM results. However, in object SLAM, it is often the case that there are multiple instances of the same object class. If all object measurements with same class are associated to the same identifier, different object instances will always give inconsistent measurements. If only one set of consistent  
measurements for each object class is kept, the algorithm will eliminate the majority of the data and fail to identify any repetitive instances of the same class.

The algorithm presented in this paper is a back-end approach where there are multiple instances of the same object class. The data association of object measurements to unique identifiers are considered unknown and must be established while doing SLAM. We exploit the coupling between data association and SLAM, jointly optimize both, and achieve better performance on both.