

MAIS 202 - PROJECT DELIVERABLE 1

Home Security App

Description:

1. Dataset: <https://cocodataset.org/#home>
 - a. Why this dataset?
 - i. Captures everyday scenery (useful when detecting intruders in a home setting)
 - ii. 80 categories, large-scale dataset, includes object segmentation
2. Methodology
 - a. Data Preprocessing
 - i. Identify different categories using histograms
 - ii. Instructions for data preprocessing (for COCO dataset) presented here: <https://towardsdatascience.com/master-the-coco-dataset-for-semantic-image-segmentation-part-1-of-2-732712631047#:~:text=%E2%80%9CCOCO%20is%20a%20large%2Dscale,segmentation%2C%20and%20captioning%20dataset.%E2%80%9D&text=COCO%20provides%20multi%2Dobject%20labeling,versatile%20and%20multi%2Dpurpose%20dataset.>
 - b. Machine Learning Model
 - i. Identify an intruder (considering no occupants of the house are present)
 - ii. Machine learning model/algorithm: classify human (intruder) vs not human. Due to this being a classification problem, the following algorithms may be favourable (considering there may be several different characteristics that define a human subject): Decision Tree or Naive Bayesian model (particularly useful for large datasets). Logistic Regression could be an option, but it would be difficult to consider multiple factors.
 - iii. Evaluation Metric: Since there may be an imbalance in the classification (more chances of no intruder than one being present), evaluation should be done based on recall-precision metrics (F1 Score) and a confusion matrix instead of accuracy alone.
 - iv. Final webpage/app: Simple user interface that, when given an image, identifies intruder (yes/no). If yes, notifies user along with a picture of the suspected intruder, and options to confirm (correct identification by algorithm/incorrect identification)