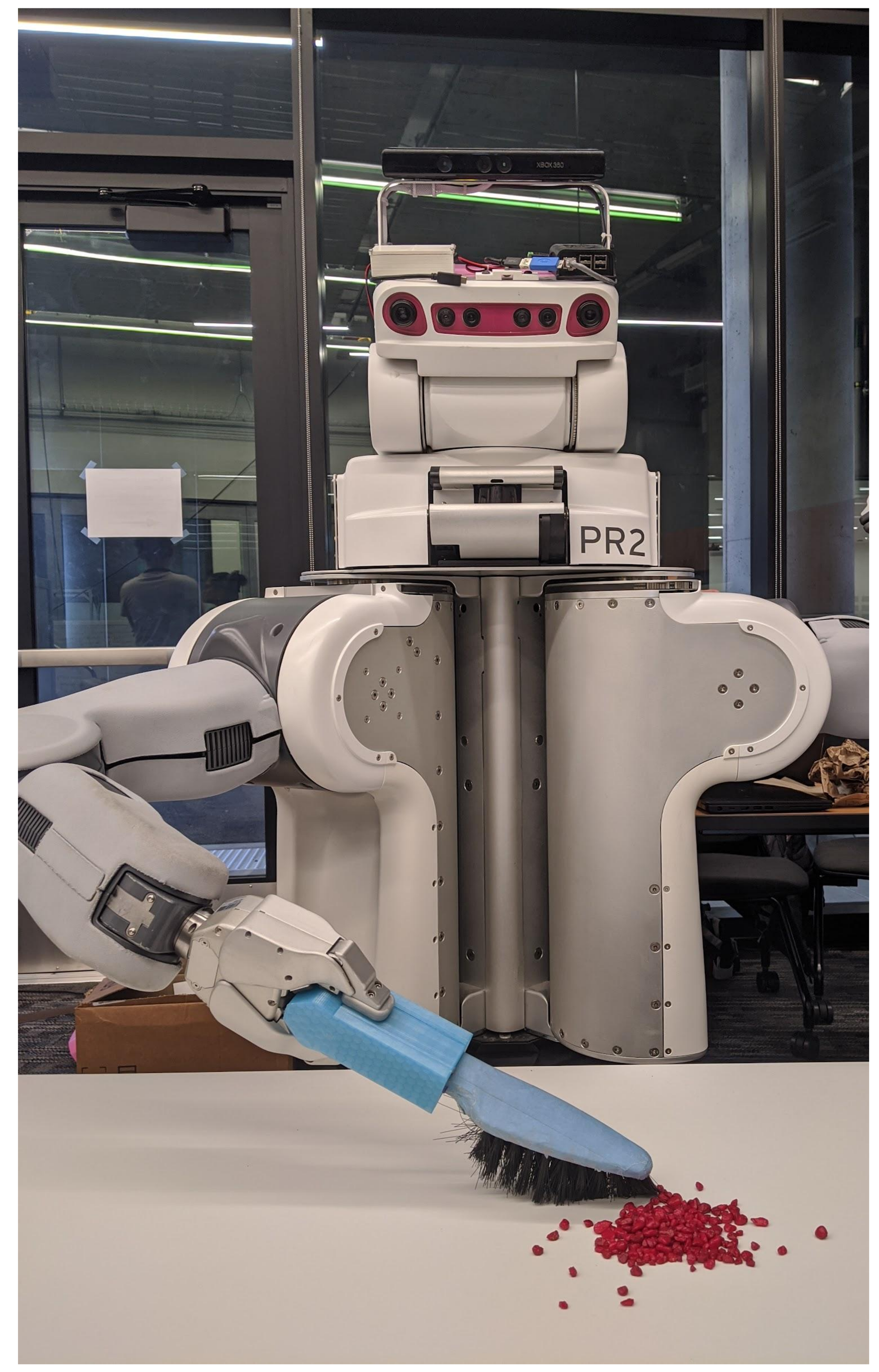
Semantic Segmentation of granular media

(Brian) JoonHo Lee, Rahul Ramanarayanan
joonl4, rr379

Motivation

Our aim is to build a real-time general-purpose dirt detector for robotic cleaning. The task is non-trivial, as it is challenged by lack of largely available data, making generalization difficult.



PR-2 on which testing was performed

Key Insight

Given **small data**, more complex **augmentation** methods, **empirical data processing,** and **compact model design** can allow development of an **efficient** segmentation network for the task of **Robotic Cleaning**

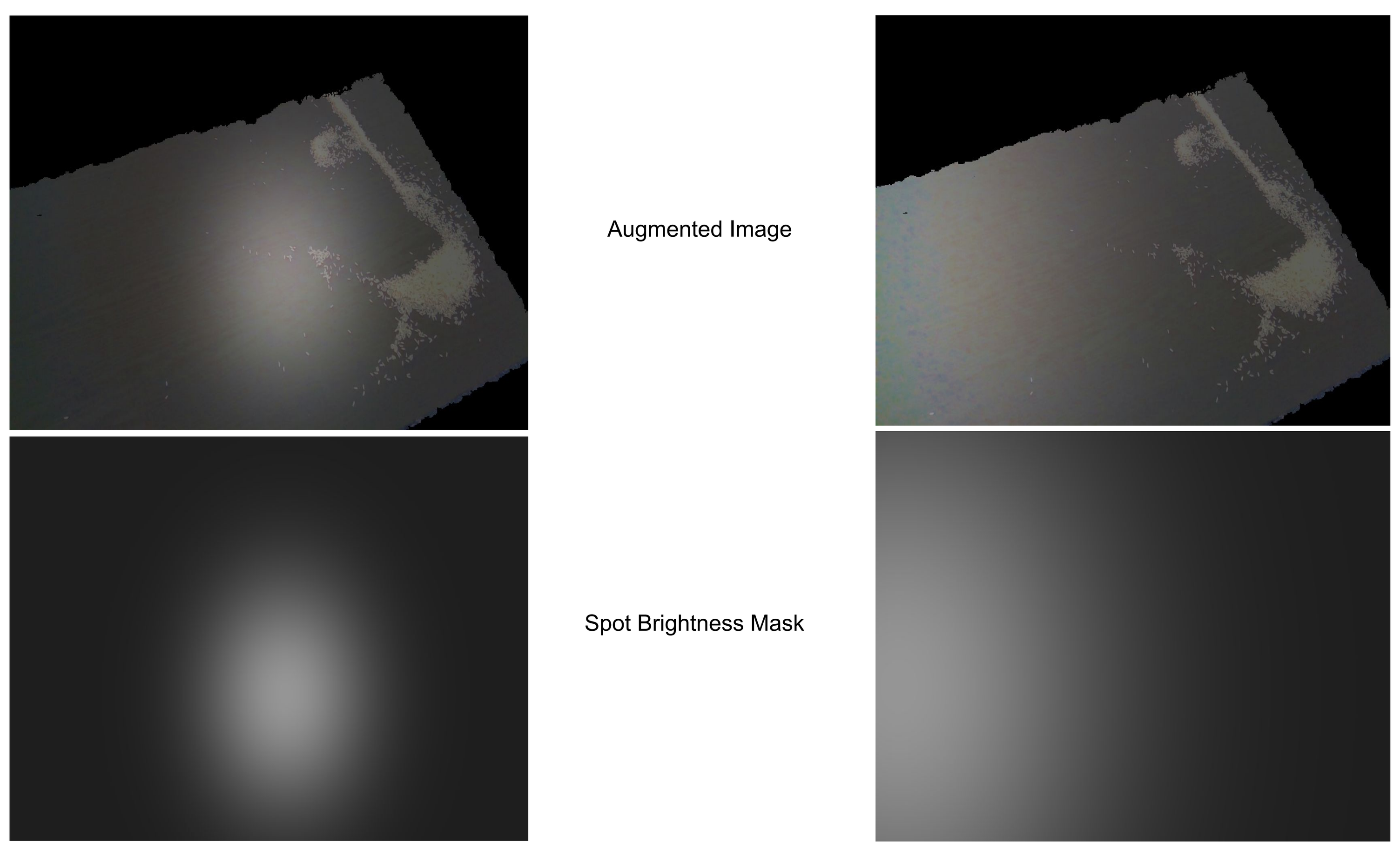
Final Results

Each sample result is in format (Input, model prediction)

All of the dirt types seen here were not seen before.Model is robust to different lighting conditions

A slight weakness to extreme cases is seen

Radial brightness Augmentation

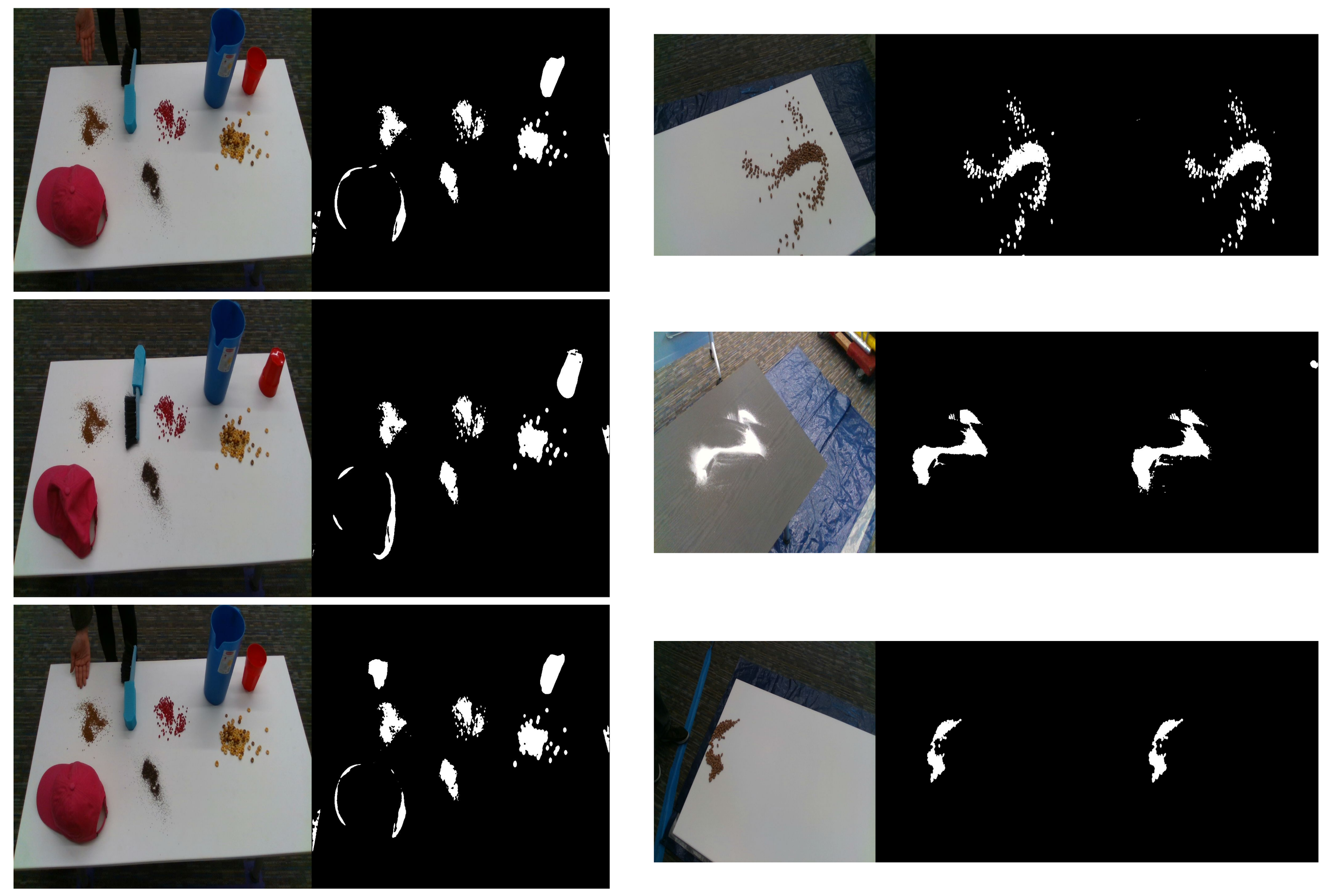


Roll off 10

Roll off 4

The spot brightness augmentation generates a grid of pixel distances from a randomly chosen center position. A mask is generated by using this grid to sample from the PDF of normal distribution whose standard deviation is inversely proportional to a "Roll Off" parameter.

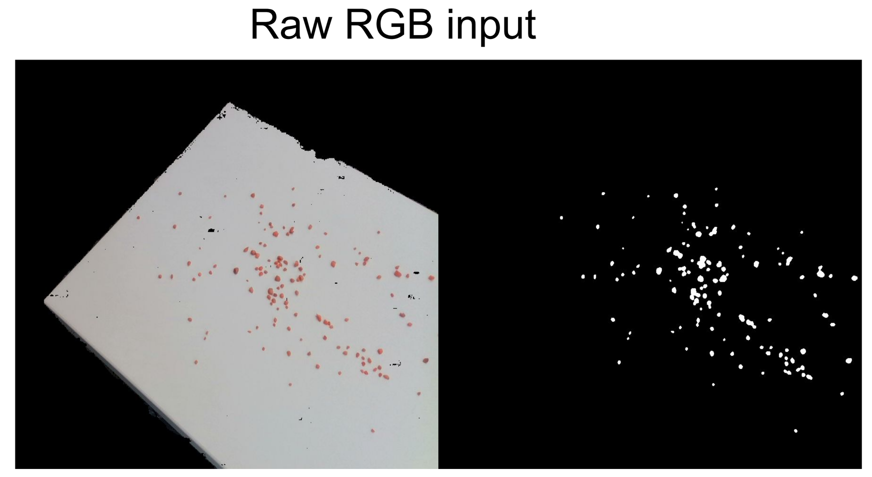
Baseline Results



We see that the baseline has learned to identify the "table region", exemplified by how the hand's position in example.

Data

Raw RGB input

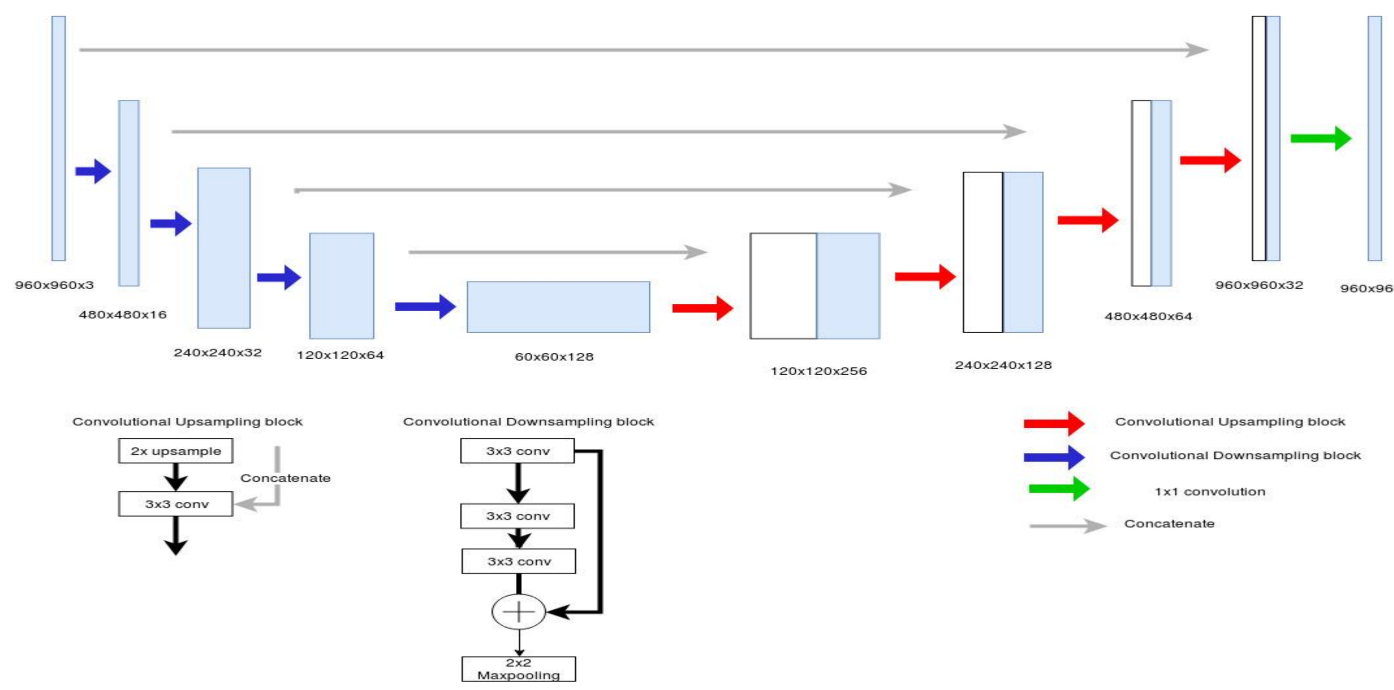


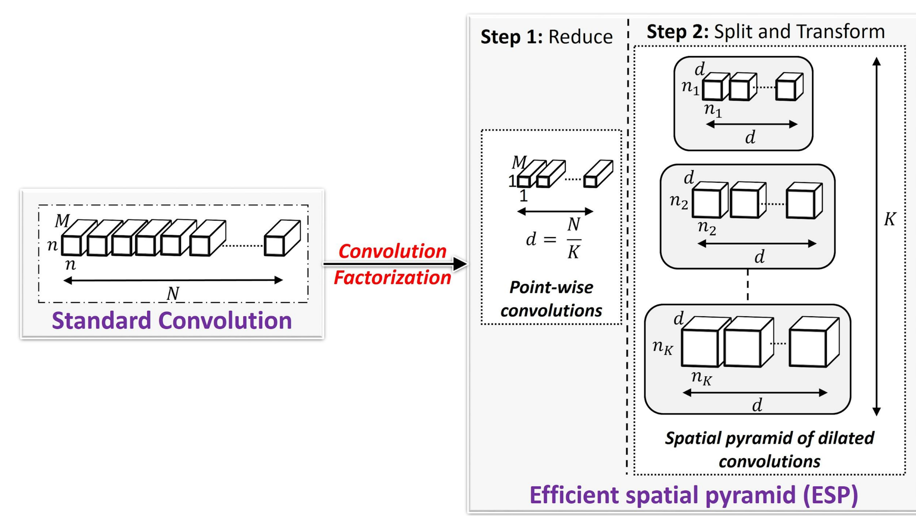
Masked RGB input(left) label(right)

The dataset comprised 2000 labelled images of different dirt spread across a standard long table. The images captured from multiple camera perspectives and table surface textures.

Model Architecture

Our model is a compact version of Unet, with 16 starting neurons, increasing by factor of 2 at each downsampling





For ~5 FPS gain in inference speed with minimal loss in accuracy, we implemented ESP modules in our encoder, modifying the Unet architecture.

Source: https://sacmehta.github.io/ESPNet/

Acknowledgments: Vinitha Ranganeni, Aaron Walsman