

Accurate On-Device Recognition of Handwritten Numbers

The high-level goal of this project is to develop algorithms that accurately recognize and classify structured (and semi-structured) handwritten numbers. The algorithms developed will be incorporated into our ODK Scan software and deployed with remote users in a variety of low-resource environments. The project is in conjunction with VillageReach, a Seattle-based NGO working in the vaccine delivery space, with support from the Bill & Melinda Gates Foundation.

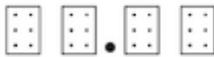
Starting Point

Currently, we have developed a prototype, web-based tool that allows people to create their own paper forms. Numbers on the forms are represented using boxes that have six dots (see figures below) intended to guide users to write structured numbers (similar to seven-segment displays) that can be automatically recognized by the software. Our current number recognition algorithm achieves approximately 88% accuracy (some digits better than others).

CHILD'S PATIENT ID:



Weight at Birth (kg):



Child's birthdate:



Project Goals

1. Explore alternate machine-learning algorithms and techniques to improve the accuracy of the number recognition. Realistically, we need to reach 98-99%.
2. Investigate the potential for more 'relaxed', freeform number recognition that does not force users to structure the numbers too rigidly.
3. Investigate methods for automatically constraining the values that make sense for specific number types. For example, if the field represents a date, it does not make sense to have a day outside the 1-31 range, or a month outside the 1-12 range etc.

Other Information

- We would be happy to provide guidance and advice on the different things we've tried and on approaches that we think may have potential and can provide starting data sets.
- It is essential that all of the computation be done locally on an off-the-shelf Android device.

- The ODK Scan software currently uses OpenCV and C++, so for ease of integration with the rest of the platform it would be beneficial to develop techniques with these development tools.
- We have partner organizations that are anxious to deploy this in the field as soon as possible. This project will be used by and have impact on a variety of real people in diverse and remote environments.