Reading Barcodes: A Computer-Vision-Based Approach to One-Dimensional Barcode Scanning

A Technical Report submitted to the Department of Computer Science

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Kevin Ming Chung
Spring, 2024

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Rosanne Vrugtman, Department of Computer Science

Reading Barcodes: A Computer-Vision-Based Approach to One-Dimensional Barcode Scanning

CS4991 Capstone Report, 2023

Kevin Chung
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA
kmc5rfu@virginia.edu

ABSTRACT

In the world of distribution and warehousing, where efficiency is the number one priority, scanning 1D barcodes is expensive. By using state-of-the-art computer vision and deep learning techniques, my team at Coros was able to create a solution to 1D barcode scanning that significantly reduces the cost. We combined image segmentation, Hough transforms, and deep learning deblur models, to create a system that was able to locate and decode barcodes in an image around 40% of the time. While this is not acceptable in an industry setting, it is a promising result that is still being experimented with and improved upon. Future improvements should enable us to reach performance levels comparable to the technology being used today.

1. INTRODUCTION

Efficiency in a warehouse is one of the paramount priorities. Many warehouse technologies aim to improve cost, time, and accuracy to reduce the overall resources spent on a task. One task that often is not thought about by people who are not in the industry is barcode scanning. Barcodes are a physically compact and standardized way of encoding lots of information about whatever item carries the barcode.

Anyone who buys products from stores is likely to be familiar with the 1D barcode, a series of thick and thin vertical lines with whitespace in between. Every character to be encoded corresponds to a series of thick and thin lines, similar to Morse code. In order to decode them, people have used laser scanning since the 70s. Today, industry-level handheld scanners can cost more than $2500, and larger rigs can cost much more than that.

With the rapid development of techniques and models in the field of computer vision, one would think we should be able to accurately decode barcodes using cameras instead of lasers, which would significantly reduce the cost. Using cameras would also increase the utility and flexibility of a larger system, since one could read text on a shipping label or on the side of a package in a warehouse. It could also count the number of packages that pass by, or track the package while in the field of view.

My proposed approach to barcode detection and decoding uses state-of-the-art computer vision, such as Hough transforms, deblur models, and detection models. Beyond barcode decoding, this approach can locate and track packages in a video, count the number of packages passing by, and read text printed on the packages.

2. RELATED WORKS

Many previous projects have applied computer vision techniques to decoding 1D barcodes. One example from Zamberletti, et al. [3] applied a Hough transform to an image of a barcode to detect the strong lines in the image. They then used a small neural network to filter for any points in the Hough transform [5] that were likely to indicate barcode lines. By doing this, they then could easily find an angle which had many lines parallel to each other, which almost certainly meant the angle of the barcode. Then they could rotate the image so that the barcode lines were vertical within a few degrees of error. They found that applying this transformation to the barcode image resulted in a higher success rate in decoding barcodes using various 1D barcode decoding libraries publicly available. It is important to note that this project did not actually locate barcodes in an image, but rather focused on improving the decode rate after locating the barcodes.

Another approach done by Wang, et al. uses a rotation-decoupled detector (RDD) [4] to locate barcodes in an image. This provides an advantage over traditional detection models, since the bounding boxes returned by the model will come with an angle of rotation, rather than just having horizontal or vertical bounding boxes. Rotated bounding boxes allow one to crop the image much tighter to the barcode, removing potential noise that could result in an inability to decode. Wang, et al. then fed those barcodes through a generative adversarial network (GAN) based model to improve the sharpness of the barcode image. More specifically, they used a model known as DeblurGAN-v2 [1]. By training this model with blurry barcode images and the corresponding sharp images, they were able to achieve a 14% increase in the decode rate of their barcodes.

3. PROJECT DESIGN

I only worked on the software aspect of the proposed system. Other members of the team were responsible for designing and constructing the hardware, so I will not be able to provide details for those parts. On the software side, the system can be viewed as a pipeline for images. There are three parts of the pipeline: detection, enhancement, and decode.

3.1 Detection

The proposed approach uses YOLOv8, a detection and segmentation model from the YOLO [7] family developed and improved by Ultralytics. YOLOv8 has many capabilities, but this system only uses detection.

Figure 1: Example YOLOv8 detections

YOLOv8 is an extremely large model and requires many thousands of images to converge. However, my model can take advantage of transfer learning, which uses a pretrained version of the model. Theoretically, this significantly reduces the number of training images required, since the model will already be familiar with important features of objects for detection.

For training data, I used a pre-existing dataset at Coros which consisted of hand-labeled images collected from initial trips to warehouses of potential customers. Each image was labeled with bounding boxes for parcels, shipping labels, and 1D barcodes.

After the model was trained, it was placed in the overall pipeline, where parcel and shipping label detections could be used in other projects, and 1D barcode detections were sent further along the pipeline.

3.2 Enhancement

Once the barcodes are detected, the system crops out each of the individual barcodes. It then enhances these images to improve the sharpness and decode rate. Image enhancement can be split into two parts: barcode rotation and deblurring.

3.2.1 Barcode Rotation

To rotate the barcode straight up and down, we use a slightly modified method from Zamberletti, et al. [3] In their method, they first applied a Canny edge filter to collect edges, or points of high contrast in the image. Afterwards, they applied a Hough transform which is essentially a collection of all possible lines through every edge point. By determining which lines have many points that intersect with it, one can deduce which lines are prevalent in the image. Zamberletti, et al. use a small feed-forward network to filter out these lines. My method, however, simply uses a threshold. If a line is above the 50th percentile in strength relative to the other lines, then keep it. After filtering out weak lines, the barcode angle can be found by searching for many parallel lines. Once the angle is found, the system rotates the barcode to be completely vertical.

3.2.2 Deblurring

Deblurring is simply a matter of feeding the barcode images through a pretrained DeblurGAN-v2 [1]. Our current method takes the publicly available version of DeblurGAN-v2 and runs the rotated barcodes to improve the contrast between the black and white sections of the barcode.

3.3 Decode

To decode the barcodes, our system uses Pyzbar, a Python wrapper for the zbar library, which is built for decoding barcodes. For our purposes, we only use the 1D barcode functionality of zbar. The zbar library claims to have performance independent of barcode rotation, but we found that rotating the barcode improves the decode rate. For the decode step, we feed both the image without deblur and with deblur to Pyzbar, since the deblur model sometimes created artifacts in the barcode that did not exist.

4. RESULTS

Any quantitative results collected are proprietary to Coros, so I will discuss them qualitatively. For every portion of the system, we performed ablations to determine how effective each individual part is.

YOLOv8 was very successful in detecting 1D barcodes in images similar to the ones in the training dataset, but struggled slightly with generalization. The previous iteration of the system used a Mask-RCNN model [2], which performed worse than YOLOv8 in experiments.

Barcode rotation was extremely accurate, with the removal of the feed-forward network from [3] resulting in an immense speedup without loss of accuracy. Previously, the system tried to use the tightest rectangle that could fit around a detection mask, which was faster but significantly less accurate.

DeblurGAN-v2 had very promising initial results, but did not reach an acceptable level without training it specifically for barcode deblurring. Comparing this to a smaller U-Net model trained specifically for barcode deblurring, it performed slightly worse, but this should be remedied by training DeblurGAN-v2 on 1D barcodes.

Overall, the system did not perform at an acceptable level for industry, but had very promising results for the future.

5. CONCLUSION

The proposed system will have a significant impact on how retail and distribution companies manage warehouses. By replacing laser-based barcode scanners with camera-based ones, it allows companies to have the same level of surveillance of parcels in the warehouse, while decreasing cost and increasing the flexibility of their systems. Another benefit of the proposed design is that the parts of the pipeline are modular. As more powerful techniques and models are developed, one can easily substitute part of the pipeline for something else. If the system reaches a point where the performance is industry standard, it will almost certainly result in a transition towards vision-based systems in warehouses.

6. FUTURE WORK

Development of this system is not nearly completed. Future work on it should focus on the enhancement and decode parts of the pipeline. Image enhancement is still a very new innovation and is being constantly improved and researched. I would do more research and experiments on various methods and models to determine what works best for 1D barcodes. Since the current system uses a premade barcode decoding library, more work could be done to make custom decode software that is more robust than the publicly available libraries.

7. ACKNOWLEDGMENTS

I would like to acknowledge Coros for allowing me the opportunity to work on this project with them. I would also like to specifically acknowledge Bill Moore (President) and Stephen Hahn (Vice President) for working with me on this project and providing guidance with the entire process.

REFERENCES

[1] O. Kupyn, T. Martyniuk, J. Wu, and Z. Wang, “DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better.” arXiv, Aug. 10, 2019. doi: 10.48550/arXiv.1908.03826.

[2] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN.” arXiv, Jan. 24, 2018. doi: 10.48550/arXiv.1703.06870.

[3] A. Zamberletti, I. Gallo, and S. Albertini, “Robust Angle Invariant 1D Barcode Detection,” in *2013 2nd IAPR Asian Conference on Pattern Recognition*, Nov. 2013, pp. 160–164. doi: 10.1109/ACPR.2013.17.

[4] B. Zhong and K. Ao, “Single-Stage Rotation-Decoupled Detector for Oriented Object,” *Remote Sensing*, vol. 12, no. 19, Art. no. 19, Jan. 2020, doi: 10.3390/rs12193262.

[5] R. O. Duda and P. E. Hart, “Use of the Hough transformation to detect lines and curves in pictures,” *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972, doi: 10.1145/361237.361242.

[6] C. Wang, N. Guevara, and D. Caragea, “Using Deep Learning to Improve Detection and Decoding Of Barcodes,” in *2022 IEEE International Conference on Image Processing (ICIP)*, Oct. 2022, pp. 1576–1580. doi: 10.1109/ICIP46576.2022.9897371.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection.” arXiv, May 09, 2016. doi: 10.48550/arXiv.1506.02640.