Price detection and recognition from shelves' images

Franko Šikić, Branimir Filipović, Zoran Kalafatić, Marko Subašić, Sven Lončarić
University of Zagreb, Croatia
Faculty of Electrical Engineering and Computing
Image Processing Group
Contact: franko.sikic@fer.hr

1. Introduction

Price detection and recognition is a computer vision problem, where very small instances need to be detected and classified with high accuracy. An efficient price detection and recognition system can be beneficial to both retailers and customers and could be in the future be further integrated into fully automated grocery stores.

2. Problem description

A single shelf image can contain up to 100 price labels. These labels are rather small, with each instance covering just around 0.1% of the image area individually.

Therefore, developing a price detection and recognition module, which can detect each price label instance and read the price off of it, is a challenging task, which is only made harder by the constant rule-exemptions each new example brings.

3. Methodology

1. **A price detection and classification model** was developed. A popular object detection model, YOLOv5, was trained for this purpose.

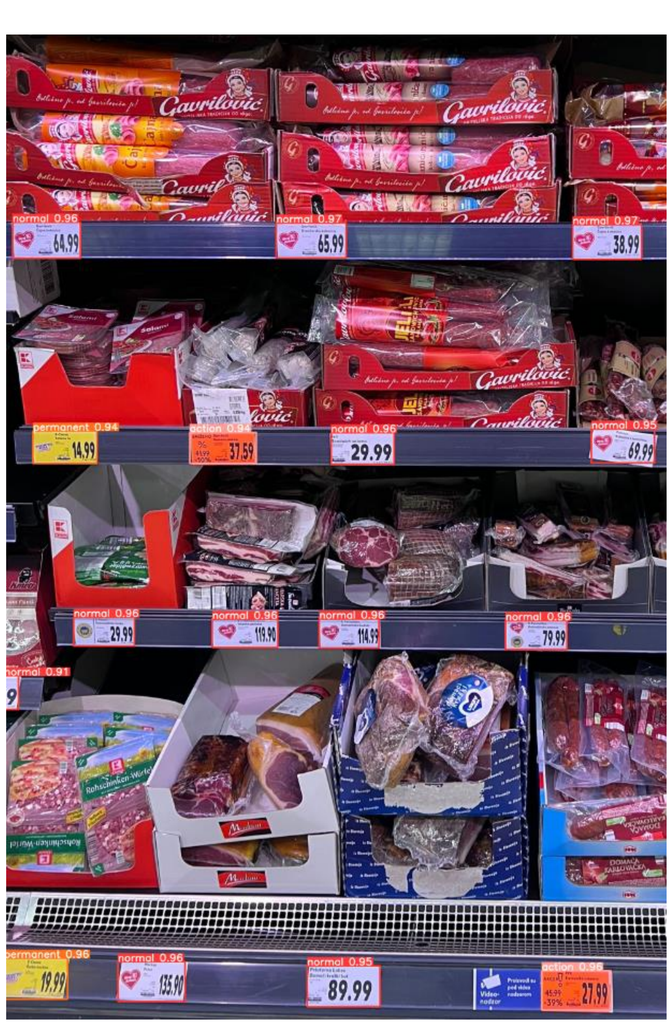

Figure 1: A shelf image with detected and classified prices

2. An artificial prices dataset was generated from existing instances**.** First, all of the digits were cropped from the real prices. Then, the price templates were also cropped and each digit on a template was annotated. Finally, all of the template's digits were replaced with randomly chosen cropped digits which generated completely new price numbers, significantly increasing the size of the available dataset.

Figure 2: A cropped digit (left), an annotated template (center), and a newly generated price (right)

3. A digit detection model was developed by using the artificial prices dataset**.** Once again, we employed the YOLOv5 model for this purpose. However, in order to better simulate the real-world conditions we first downsampled all of the training examples down to 140x70 px resolution and then upsampled it to a higher resolution, 384x384 px, for instance.

Figure 3: The digit detections of the generated image (top left), the real Konzum store image (bottom left), the real Spar store image (top middle), the real Plodine store image (bottom middle), and the real Kaufland store image (right)

4. A prices recogniton module was developed by using the detected digits. First, we clustered the digits by their immediate proximity in order to build numbers from digits. Next, we needed to connect the euros and the cents of the prices, i.e. the kunas and the lipas. We did it by assuring the vertical alignment of the numbers and afterward by clustering by proximity.

4. Results

Price detection class-agnostic YOLOv5 achieved 99% AP, while the performance of the 4-class model was around 80% mAP. Additionally, the following table lists the prices recognition module's results in relation to the resolution of the input price label crop.

As the results show, the module achieves excellent accuracy on generated images, which cannot be said for real-world images. The reason may be that the real-world images are often cropped from low-resolution photographs and therefore the digit detection model cannot function properly on such examples.

Table 1: Price recognition accuracy on generated and real-world images

Image size	Generated images accuracy [%]	Real-world images accuracy [%]
192x192	66.6	65.9
256x256	67.4	67.6
384x384	91.7	75.1
512x512	91.1	71.9

5. Conclusion

Implementing an efficient and accurate price detection and recognition system is a complex problem which consists of several tasks: price detection and classification, artificial prices dataset generation, digits detection and prices recognition. Our deep learning-based solution achieved excellent detection results and very good price recognition results, which shows promise for future research.