High Speed Barcodes Decoding from Natural Image using SSD and Two-Dimensional FFT

Yuyao Li[1)]　Jin Mitsugi[1)]

1 Introduction

Many parcels and commercial products are labeled with multiple 1D and/or 2D barcodes as shown in Fig. 1. Each barcode is given a dedicated role in its business context such as purchase order identification for business transaction and shipping container code for transportation. Therefore, human operator or interrogating machine needs to capture an appropriate barcode depending on the business context. Such selective capture is often realized by confining the capture area or adjusting the orientation of the object to which a barcode is attached. But such workaround demands human intervention, which slows down the operation.

Figure1: Example of a parcel labelled with multiple barcodes

A barcode comprises a data carrier and an identification key[1]. A data carrier is a graphical representation of an identification key using elemental bars and squares. Popular data carriers are, for example, EAN-13, CODE-128, Data matrix and QR-code. Data carrier is encoded with an identification key such as Global Trade Item Number (GTIN), Serialized Shipment Container Code (SSCC) and Global Returnable Asset Identifier (GRAI). As the combinations of data carrier and identification key are regulated in [1], a desired identification key can be extracted after simultaneously capture multiple barcodes in an image. Such industrial demands trigger the research on multiple barcodes capture[2]-[5].

Multiple barcodes capture comprises two phases, barcode detection and barcode decoding[2]. Traditional methods for the barcode detection is to use an edge detection and morphological transformation[2, 3]. Zhou[6] proposes to detect the rotation of barcode candidate area by accumulating the contour lines of barcode stripes. The accumulator results are also used to detect multiple barcodes in an image. As the image feature of barcode is naturally a cluster of parallel lines and squares, a traditional barcode detection usually handles also the determination of barcode rotation.

The use of object detection using deep learning becomes popular in barcode detection field[3, 4, 7]. Although deep learning can handle even rotated bounding boxes[5], the detection of rotated bounding boxes requires a regression analysis, necessitating a large amount of training data of various aspect ratio and rotation angles. As a deterministic barcode decoding is needed even after a deep-learning-based barcode detection, a reasonable compromise is to apply a traditional image processing using morphological transformation to determine its rotation with respect to a horizontal bounding box after a barcode detection using a deep learning as in [3].

In this study, we reveal that the determination of a barcode rotation in a horizontal bounding box dominates the computation time and the decoding accuracy when we apply a traditional morphological transformation method. To counter this problem, we propose a rotation angle determination technique using two-dimensional Fast Fourier Transform (2D FFT) following the object detection with Single Shot Multibox Detector (SSD). While the application of 2D FFT for rotation angle determination has been used in essentially square character recognition[8], its naive application to rectangular images results in an erroneous rotation angle. We solve the problem by introducing a geometric compensation technique to 2D FFT.

To evaluate the effectiveness of the 2D FFT, we compare the speed of rotation angle determination using 2D FFT with that of a morphological transform-based approach. Further, the 2D FFT method is integrated into a barcode capture system with a SSD based barcode detection with horizontal bounding-box and a custom barcode decoding module written in MATLAB. It was revealed that 2D FFT significantly outperforms the morphological transform in terms of speed and accuracy, making it a more efficient choice for barcode rotation angle determination and decoding.

The contributions of this paper are summarized as follows.

- The invention of 2D FFT based barcode rotation angle determination, achieving 15 times faster than the existing method.
- It is revealed that preparing multiple scanline contributes to the decoding accuracy.
- The whole barcode capture system using the combination of deep learning based object detection and 2D FFT outperforms a traditional method by seven and two times in terms of speed and decoding accuracy, respectively.

The rest of this paper is organized as follows. In Section 2, the whole procedure to detect and decode barcodes from a natural image is explained. In Section 3, the proposal is evaluated using 215 ArTe-Lab[9] barcode dataset. Section 4 concludes

1） Faculty of Environment and Information Studies, Keio University

this paper.

2 Methodology

2.1 Rotation Angle Determination with 2D FFT

2.1.1 Principle

After the object detection using SSD, the major portion in a horizontal bounding box is occupied by an (M × N) image of a θ rotated barcode, which comprises parallel lines either in 1D or 2D barcode as shown in Fig. 2. When we focus on horizontal

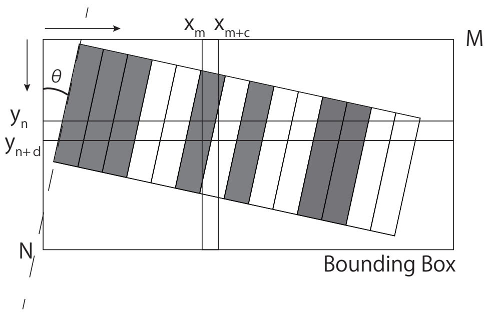


Figure2: A simplified image of barcode in a bounding box

two scan lines y_n, y_{n+d} at n-th and (n+d)-th vertical pixels, they are shifted by $d \sin\theta$ in the x direction such that

$$y_{n+d}(x) = y_n(x - d\sin(\theta)) \tag{1}$$

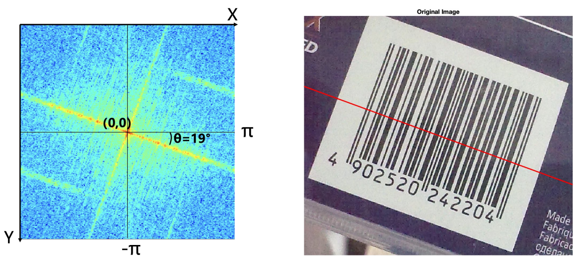
Fourier transform of (1) is given as

$$Y_{n+d}(\omega_x) = Y_n(\omega_x)e^{-j\omega_x d\sin(\theta)}, \tag{2}$$

where $Y_n(\omega_x)$ is the Fourier transform of $y_n(x)$ in the horizontal axis. This linear phase delay in the frequency domain is essentially the group delay. Similarly, vertical two scan lines x_m, x_{m+c} produce the following Fourier transform in the vertical axis.

$$X_{m+c}(\omega_y) = X_m(\omega_y)e^{-j\omega_y c\cos(\theta)} \tag{3}$$

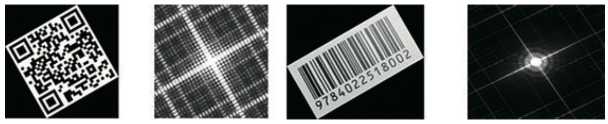
Therefore, a rotated barcode image produces a strong group delay component in its frequency domain signal depending on the rotation angle θ. Fig. 3 depicts a normalized spectrum (left) and the corresponding rectangular image (right). The 2D FFT produces strong components both horizontal and vertical axes in the frequency domain.



(a) After 2D FFT (b) Original barcode image

Figure3: 2D FFT of barcode

This group delay characteristics work also for 2D barcode image as shown in Fig. 4.



Original 2D barcode After 2D FFT Original 1D barcode After 2D FFT

Figure4: 2D FFT of 2D and 1D barcodes.

2.1.2 Geometric compensation

As the unit normalized frequencies in horizontal and vertical axes are $\frac{2\pi}{M}$ and $\frac{2\pi}{N}$, respectively, the physical rotation angle of the barcode, denoted as θ, can be inferred from the group delay measurement in the frequency domain as illustrated in (4). Here, M, N, and Θ represent the horizontal and vertical pixel counts and the rotation angle in the frequency domain, respectively.

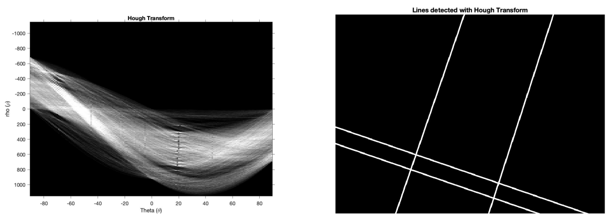
$$\theta = \tan^{-1}\left(\frac{\sin\Theta}{\cos\Theta} \times \frac{M}{N}\right) \tag{4}$$

This is referred to as geometric compensation in this paper.

2.1.3 Group delay determination from 2D FFT result

The dominant group delay angle Θ is calculated from a 2D FFT result using Hough transform and k means cluster analysis.

Hough Transform is a feature extraction technique commonly used in image processing and computer vision for the detection of regular shapes such as lines, circles or ellipses[10]. After performing a 2D FFT on a barcode image, Hough transform is applied. This converts the image from Cartesian space to Hough space as in Fig. 5a, thereby locating the two bright lines in the magnitude spectrum as in Fig. 5b.



(a) Hough space (b) Detected lines

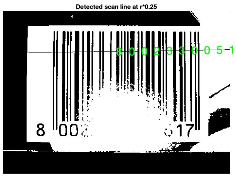
Figure5: Hough transform result

Hough transform of 2D FFT barcode image often produces similar but distinct lines. To merge similar Hough lines into the principal two lines, three class weighted k means cluster analysis was applied. The principal two lines represent the $\sin\Theta$ and $\cos\Theta$ directions. The third class was needed to classify lines produced by noise.

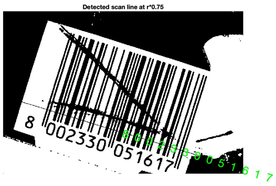
2.1.4 Scanline shift and bi-direction scanning

The dataset [9] contains numerous barcode images affected by various factors, including reflections, pen markings, and obstructions from other objects. These issues partially obscure the barcode lines, thereby posing challenges to the accuracy of barcode decoding. Therefore, to enhance the decoding success rate, we produce multiple candidate scanlines by vertically shifting the nominal scanline. This adjustment aims to choose the best scanline which preserves the integrity of the stripes.

For instance, as illustrated in Fig. 6, in scenarios involving interfering elements such as reflections and obstructions, a scanline exhibiting a complete transition between black and white can be chosen and used in the subsequent decoding.



(a) Reflection



(b) Obstruction

Figure6: Barcodes with interfere elements

Additionally, in instances where the barcode is in an upside-down appearance, conventional left-to-right decoding along the scanline produces a false check digits. Consequently, as demonstrated in Fig. 7, it is necessary to decode by examine the scanline from both directions.

Figure7: Decode along scanline from both sides

2.2 Barcode Detection with SSD

Although this is not our novelty to use a deep learning based barcode detection from a natural image, the barcode detection with SSD is explained for the self-completeness of this paper.

Using a transfer learning, a pre-trained SSD detector[11], originally trained to detect vehicles, was used along with a 1D barcode dataset from ArTe-Lab[9] to train the network to detect horizontally located bounding boxes. In this training, ResNet-50 was utilized for feature extraction, with the classification layers removed due to the homogeneity of the dataset, which exclusively comprises EAN-13 barcodes without other types such as Code 128 or QR codes. To enhance the robustness of the backbone network, 7 additional convolution layers were incorporated. The Image Labeler[12] from MATLAB Apps was used to label the bounding boxes of barcodes from natural images, providing the ground truth data. The entire dataset of 215 images was used, with 60% of the images serving as the training data to train the network and 40% as the testing data to evaluate the precision of trained detector. Data augmentation was utilized to enhance network accuracy by introducing variability into the training data through random transformations, such as horizontal flipping, scaling, and color jitters of images. This approach increased the diversity of the training data without requiring additional labeled samples. The average precision metric was used to evaluate the performance of the trained detector, achieving an average precision of 0.82, as shown in Fig. 8.The precision-recall (PR) curve highlights how precise a detector is at varying levels of recall[11]. For the SSD detector, the total detection time of horizontal bounding box for 215 images was 17.02 seconds, with an average detection time of 0.079 seconds per image. These results demonstrate the efficiency of the SSD model in real-time barcode detection applications, maintaining high performance even under rapid processing demands. Furthermore, this underscores the advantages of utilizing deep learning techniques in barcode detection, such as enhanced detection accuracy and adaptability to varying environmental conditions. These characteristics are crucial for the development of robust automated systems that require minimal human intervention.

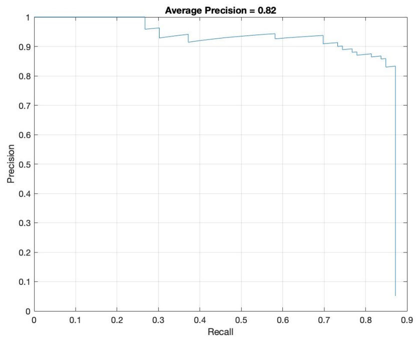
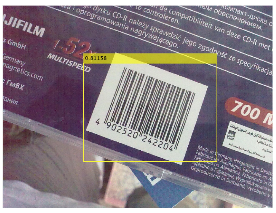


Figure8: Precision-recall (PR) curve

Applying the trained detector to the images in the dataset allows for the cropping of detected bounding boxes, resulting in a new dataset where barcodes occupy the major central area of each image Fig. 9a. This new dataset is then used for the subsequent steps of rotation angle determination and decoding. With this detector, it is possible to individually detect multiple barcodes within one image, as illustrated in Fig. 9b.



(a) Single Barcode in one image

(b) Multiple Barcodes in one image

Figure9: Barcodes and bounding boxes detected by SSD

The yellow rectangle represents the detected bounding box, indicating the location of the barcode. The number associated with this bounding box denotes the confidence score, a probability value ranging from 0 to 1 that indicates the likelihood of the detected object being correctly classified.

This means for a single image contains 4 barcodes, the SSD detector is capable of cropping out all 4 barcodes simultaneously and proceeding in parallel with subsequent rotation angle detection and decoding. Consequently, the average time for barcode bounding box detection using SSD would reduce to approximately 0.02 seconds.

2.3 Rotation Angle Determination with morphological transform

In this study, morphological transformations serve as the baseline to evaluate the efficacy of 2D Fast Fourier Transformations (2D FFT). There are two basic morphological operators, erosion and dilation[13]. Erosion is performed using a structural element— typically a small rectangle or circle— applied pixel-by-pixel across the image. A pixel in the original image is retained only if all pixels under the structural element are foreground pixels.This process gradually diminishes areas not fully covered by the structural element, thereby eliminating finer image features and isolated pixels. Therefore, by repeatedly eroding barcode images, noise around the barcode and small fractures can be effectively removed, filling the white gaps between the black bars to fill the barcode region into an unified entity as in Fig. 10.

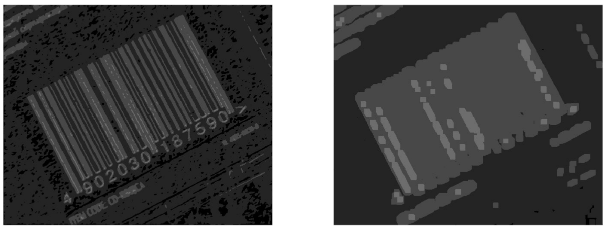


Figure10: Barcodes image after morphological transform

The accuracy of morphological transformations heavily relies on the times of erosion, making it a time-intensive process. Subsequent to the morphological Transformations, edge detection is applied to the modified barcode region to delineate the bounding box as in Fig. 11. A decoding module, analogous to that employed in 2D FFT analyses, is then integrated, enabling effective barcode detection.

Figure11: Bounding box detected by morphological transformation

Furthermore, the orientation of barcode identified by morphological transformation is determined based on the bounding boxes rather than the barcodes themselves. This indirect method prevents scanlines from accurately aligning with the actual rotation angles of the barcodes, consequently degrading the failure rate.

3 Performance evaluation

This study initially examines the detection times using the SSD technique, detailed in Subsection 2.2. Subsequent analysis conducts performance evaluation of 2D FFT explained in Subsection 2.1 compared with a baseline performance produced with morphological transform in Subsection2.3. The performance is quantified in terms of processing speed and decoding accuracy.

The performance comparison was done with MacBook Air with 4.05 GHz clock speed and 16 GB memory. The whole software was built with MATLAB 2024.

3.1 SSD Detection Times for Multiple Barcodes

An analysis of detection time variations reveals that the current SSD detector adeptly identifies up to four barcode bounding boxes concurrently within an image. Notably, the detection time increase minimally as the number of barcodes present escalates. This finding highlights the SSD detector's capability to manage multiple targets simultaneously, significantly reducing the overall decoding time, as illustrated in Fig. 12.

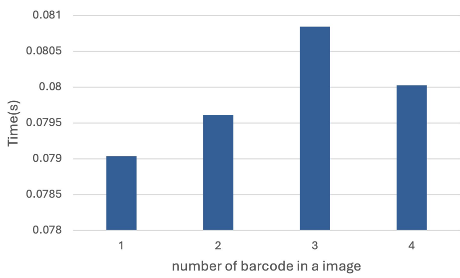


Figure12: Detection time variation for multiple barcodes with SSD

This observation implies that as the number of barcodes within an image increases, the overall processing time per barcode decreases. Assuming the number of barcodes is N, the total processing time can be expressed as $\frac{\mathrm{SSD}}{N}$ + 2D FFT + decoding. Considering that the time for SSD detection remains nearly constant, the overall time per barcode diminishes with increasing N. This efficiency illustrates the effectiveness of the SSD detector in maintaining consistent detection times even as the workload increases.

3.2 2D FFT contribution to determine the rotation angle

Each of the 215 images in the dataset[9] was analyzed using both methods to measure the time required to determine the barcode rotation angles. The findings are presented in Fig. 13. The result reveals that the average time to produce decoded data of 2D FFT is 0.015 seconds, whereas that of the Morphological Transform was 0.235 seconds. This indicates that 2D FFT significantly improves the processing time by approximately 15 times compared with that of the morphological transformation.

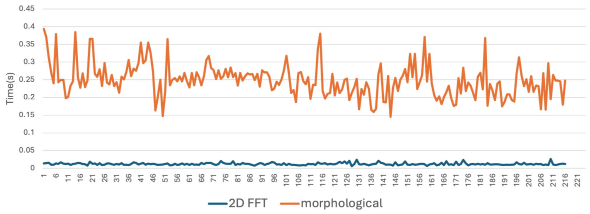
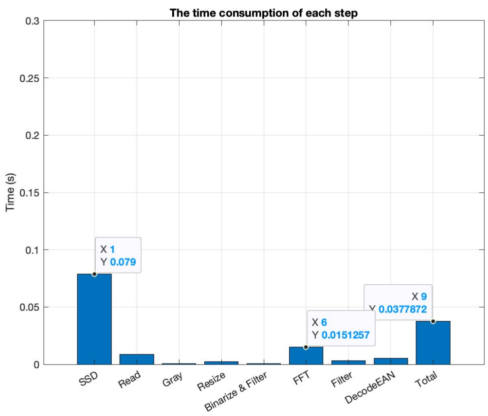


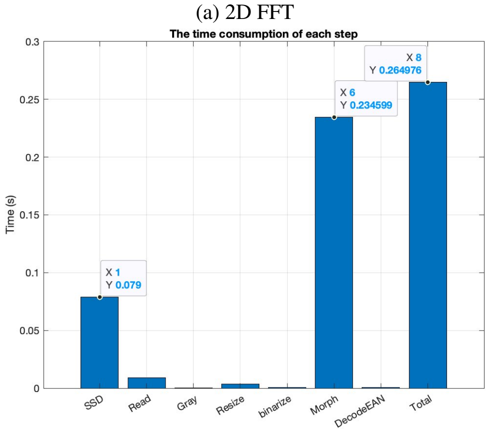
Figure13: Processing time consumed by rotation angle determination of 215 dataset

The total processing time per image of both methods is further analyzed to reveal which segment of the computation dominate the total processing time.

The result is shown in Fig. 14. In the morphological transformation, the dominant segment is "Morph". The result shows that the total time for 2D FFT is 0.038 seconds, while morphological transform requires 0.265 seconds, approximately 7 times longer.



(a) 2D FFT



(b) Morphological

Figure14: Segment duration

3.3 Accuracy comparison

In case of the erroneous detection of bounding box or the failure of rotation angle recovery, the decoding produces erroneous result which can be evaluated by an error checking facility. 1D barcodes, such as EAN-13 and CODE 128 are usually protected with a check digit. 2D barcodes, such as QR code and Data matrix are protected with a CRC.

Through the analysis and implementation of the methods described in Subsection 2.2 and Subsubsection 2.1.4, which involve detecting the position of multiple bounding boxes and implementing scanline-shifting and bi-directional decoding, the decoding accuracy has improved from 54% to 82% incrementally, as illustrated in Fig. 15.

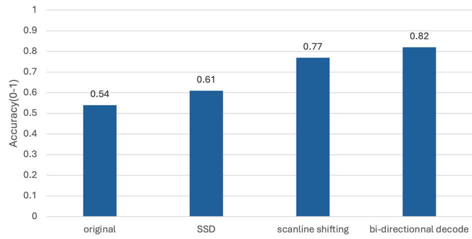


Figure15: Stepwise improvement in accuracy

Fig. 16 provides the accuracy comparison of the two methods, showing that the accuracy of the 2D FFT method is approximately 82%, while that of the morphological transformation method is 34%.

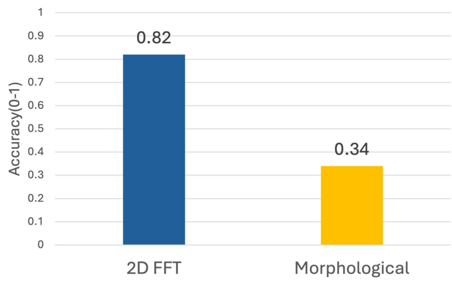


Figure16: Accuracy comparison

The results are also presented in the confusion matrix in Table **??**, which contrasts the classification accuracies of the two methods. This matrix clearly illustrates the number of images correctly classified by both methods, as well as discrepancies between them, allowing for a straightforward assessment of their performance.

		morphological	
		TRUE	FALSE
2D FFT	TRUE	63	114
	FALSE	8	30

Table1: Confusion matrix of 2D FFT and morphological transform

Upon analyzing the 215 images from dataset [9], that failed to decode, it was found that, among the 39 images where the 2D FFT method failed, 7 images were incorrectly cropped by the SSD detector, indicating inaccuracies in the bounding box positioning and resulting in the barcode region being partially excluded. For 1D barcodes, missing lines inevitably lead to read failures. The precision of deep learning can be improved by using larger dataset as training data. The remaining 32 images failed to decode due to blurriness or distortion, causing the color transitions on the scanlines to be unclear and leading to incorrect ratios.

Analyzing the failure cases of the morphological transformation method revealed that most failures were due to insufficient iterations of dilation or erosion resulting in incomplete filling of gaps between barcode lines. This led to deep recursion causing stack overflow. However, increasing the times of morphological transformations would further extend the processing time, creating a trade-off between execution time and accuracy.

4 Conclusion

In order to capture and decode multiple barcodes in a natural image, deep learning based methods, such as SSD, can produce horizontal bounding boxes with about 120 training data. Furthermore, utilizing SSD for simultaneous detection of 4 barcode bounding boxes in one image can reduce the average detection time from 0.079s to 0.02s. The utilization of methods enhancing scanline detection has increased the accuracy in 2D FFT from 54% to 82%. Determination of the rotation angle of a barcode within a horizontal bounding box can be effectively handled by the proposed 2D FFT with the geometric compensation. The proposal outperforms the traditional morphological transform based rotation determination methods. Preparing multiple candidate sanlines contributes the accuracy revealing the robustness to image noise and obscurement. The evaluation demonstrates that 2D FFT method significantly accelerates the processing time for rotation angle determination by 15 times, and reduces the total processing time by 7 times. Moreover, the accuracy was improved from 34% to 82%.

References

[1] GS1, "GS1 General Specifications Standard", Release 24.0, Jan. 2024, https://ref.gs1.org/standards/genspecs/, visited on June 2, 2024.

[2] P.Bodnar and L.G.Nyul, "Barcode detection with Morphological operations and clustering", Signal processing, pattern recognition and applications 2012.

[3] Q.Yang, G.Golwata, S.Sundarm, P.Lee and J. Allebach, "Barcode detection and decoding in on-line fashion images", doi.org/10.2352/ISSN.2470-1173.2019.8.IMAWM-413

[4] A. Zamberletti, I. Gallo, M. Carullo, and E. Binaghi, "Neural image restoration for decoding 1-D barcodes using common camera phones," in Proc. VISAPP' 10, (2010), pp. 5–11.

[5] Z. Liu, H. Wang, L. Weng and Y. Yang, "Ship Rotated Bounding Box Space for Ship Extraction From High-Resolution Optical Satellite Images With Complex Backgrounds," IEEE Geoscience and Remote Sensing Letters, vol. 13, no. 8, (2016), pp. 1074-1078.

[6] R. Zhou and X. Guo, "A new method of angle-robust multiple 1D-barcodes detection," 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 2016, pp. 433-438.

[7] Y. Ren and Z. Liu, "Barcode detection and decoding method based on deep learning," 2019 2nd ICISCAE, (2019), pp. 393-396.

[8] P.W.Dennis, S.A.Mills, R.B,Dydyk, "Rotational Correlation and Duplicate Image Identification by Fourier Transform Correlation", European Patent EP1 185 951 B1, (2000).

[9] http://artelab.dista.uninsubria.it/downloads/datasets/barcode/medium_barcode_1d/medium_barcode_1d.html, accessed on May 29, 2024.

[10] A. Zamberletti, I. Gallo and S. Albertini, "Robust Angle Invariant 1D Barcode Detection," 2013 2nd IAPR Asian Conference on Pattern Recognition, Naha, Japan, 2013, pp. 160-164.

[11] "Object Detection Using SSD Deep Learning," https://ww2.mathworks.cn/help/vision/ug/object-detection-using-single-shot-detector.html, accessed on May 29, 2024.

[12] https://ww2.mathworks.cn/help/vision/ref/imagelabeler-app.html?searchHighlight=image%20labeler&s_tid=srchtitle_support_results_1_image%20labeler, accessed on May 29, 2024.

[13] https://docs.opencv.org/4.x/d9/d61/tutorial_py_morphological_ops.html,accessed on May 30, 2024.