Effects of Image Transformation Techniques for Character Detection of EasyOCR

Noriel R. Achero
New Era University
acheronorielr@gmail.com

Abstract

This research investigates how different image transformation techniques impact the performance of EasyOCR in character detection. The objective was to know how transformations such as Translation, Rotation, Scaling, Affine Transformation, and Projective Transformation impact the accuracy and confidence of text recognition. We applied each of these transformations to a set of images and used EasyOCR to detect and read the characters. Comparing the confidence levels of the recognized text across the different transformations helped us evaluate which methods worked best. The results show that certain transformations, particularly scaling and affine transformation, led to improvements in recognition accuracy and confidence, offering valuable insights for optimizing OCR performance, especially in challenging conditions.

1. Introduction

Character detection is one of the most important elements in many real-world applications, including scanning documents, recognizing street signs, and enhancing security systems. Yet, it is not so easy to get accurate character recognition, especially when the images are distorted, blurry, or poorly lit. In this study, we are going to explore how various image transformation techniques can enhance the accuracy of character recognition using EasyOCR. We experiment with techniques like Translation, Rotation, Scaling, Affine Transformation, and Projective Transformation to see how they affect OCR performance. We apply these transformations to a variety of images and compare how each one impacts the confidence levels of the detected characters. Our final goal is to find the most effective method to enhance character detection for more accurate and reliable OCR systems, even in less-than-ideal conditions.

2. Methodology

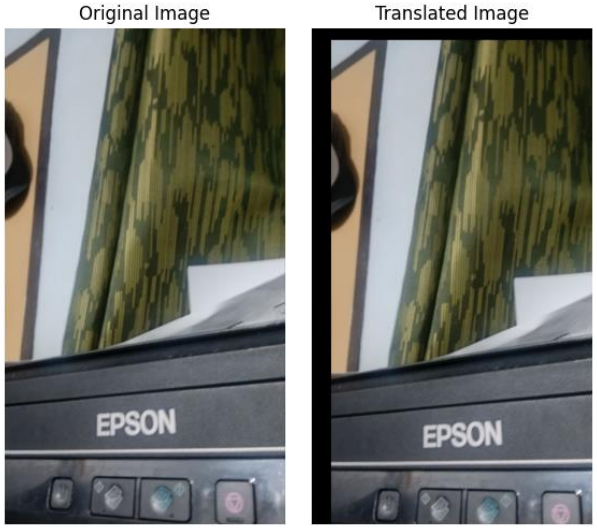
2.1 Dataset

The images or dataset used for testing the character recognition of the easyOCR with different image transformation techniques was collected using a smartphone camera. It captured a printer with a logo text

present. These images were processed with different image transformation techniques with a goal to see the effects and versatility of the OCR model.

2. 2 Translation

The first image processing technique used for the study is translation. Translation is a process of shifting an image in the x and y directions without altering its shape or size. This is achieved by adding a constant value to the pixel coordinates.



2. 2 Rotation

Another image processing techniques used for the study is Rotation. Rotation is a geometric transformation that repositions an image by rotating it around a specified point, typically the center, by a certain angle. This process changes the orientation of the image without altering its shape or size.

2.3 Scaling

Scaling is a transformation that scales the size of an image by changing the number of pixels, either scaling up or down. The scaling factors determine whether the image is magnified or reduced, it simply changes its overall dimensions.

2.4 Affine Transformation

An affine transformation is a method of linear mapping that preserves points, straight lines, and planes. Parallel lines remain parallel after transformation. It has components of translation, rotation, scaling, and shearing.

2.5 Projective Transformation

Projective transformation, sometimes referred to as homography, converts straight lines to straight lines, however parallelism isn't always maintained. It is frequently employed for adjustment of perspective.

2.6 easyOCR

To detect the text in the image, we utilized a publicly available Optical Character Recognition (OCR) the easyOCR. It is a ready-to-use OCT Ready-to-use OCR with 80+ supported languages and all popular writing scripts including Latin, Chinese, Arabic, Devanagari, Cyrillic, etc.

3. Results and Discussion

This section discussed the detailed analyzation of each image processed in different image translation methods. The objective of this discussion is to find the effected of these image translation techniques in the easyOCR. We will compare the different processed images based on the output text and confidence level of the recognition.

Images below are the processing image, and the results of the OCR detected:

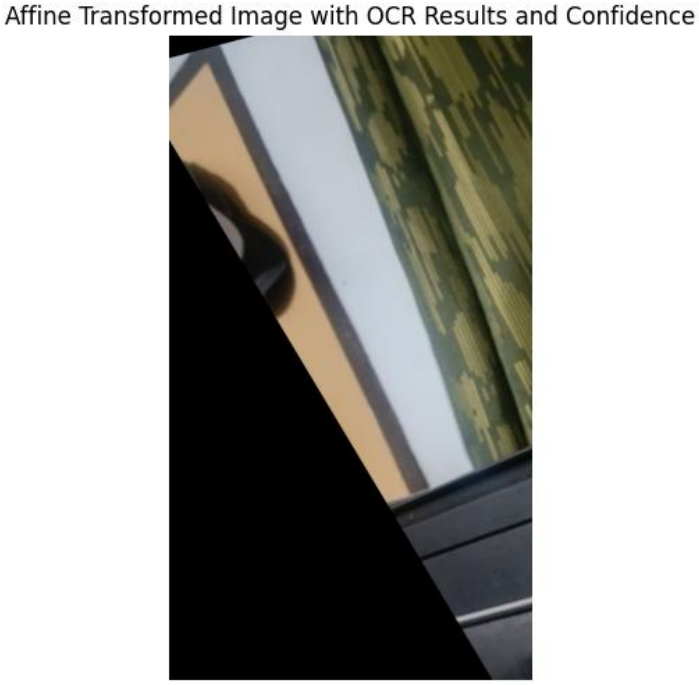
Original Image

Translated Image

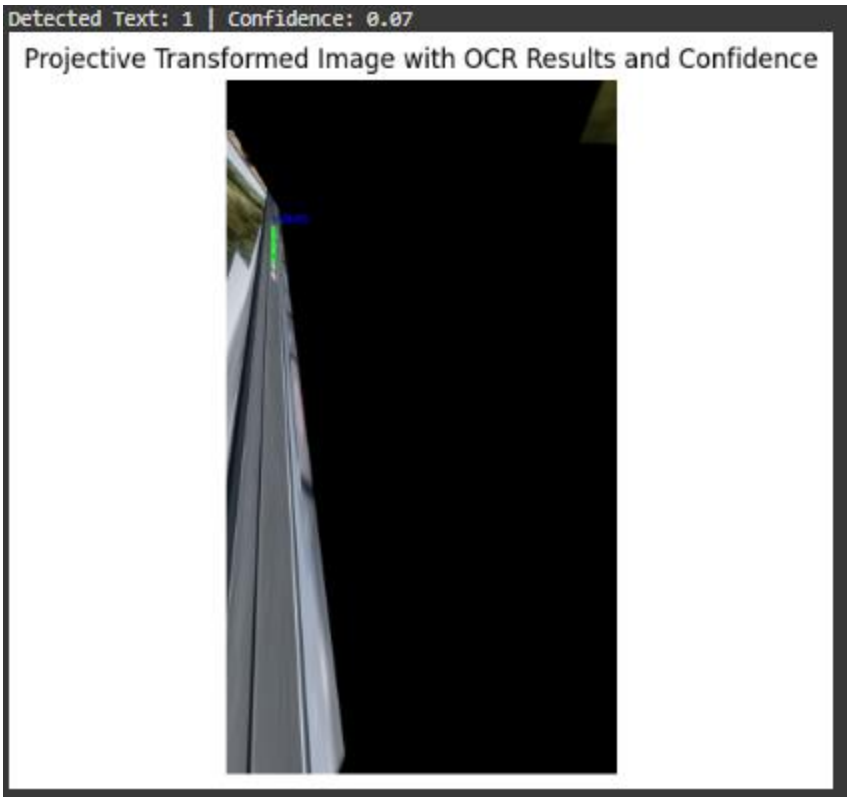
Rotated Image

Scaled Image

Affine Transformed Image



Projective Transformed Image



Results

Image	Detected Text	Confidence Level
Original	EPSON	1.00
Translated	EPSON	1.00
Rotated	8	0.13
Scaled	EPSON	1.00
Affine Transformed	null	null
Projective Tranformed	1	0.07

The table presents the results of various image transformation techniques applied to detect text using easyOCR with their corresponding confidence levels. The original image had a confidence level of 1.00 with the detected text as " EPSON," same with the translated and scaled images. The other techniques detect other text and has a low confidence level.

4. Conclusion

This research investigates the impacts of various transformation techniques of an image: Translation, Rotation, Scaling, Affine Transformation, and Projective Transformation-on the character detection of EasyOCR. The findings of the study indicate that with the translations and scalings of the images, the EasyOCR retained its accuracy and confidence level because the extracted text remained identical to the source text with a confidence score of 1.00. However, techniques such as rotation and projective transformation resulted in considerable drops in confidence and accuracy, with incorrect or partially detected text. The affine transformation did not detect any text at all, underlining the adverse impact on OCR performance.

These results show that EasyOCR is robust to translations and scaling but vulnerable under more complex transformations. This calls for further research into new preprocessing methods or hybrid approaches that could improve OCR resistance against challenging transformations, with a final goal of recognition accuracy in real-world environments.