COMPUTATIONAL INTELLIGENCE AND MACHINE LEARNING
e-ISSN: 2582-7464
Computational Intelligence and Machine Learning
Vol-5, Issue-2, October 2024

Reading Fuel Information Boards at Gas Stations with Artificial Intelligence

Volkan İllik[1], Kadir Yunus Koç[2*], Berk Özsoy[3], Muhammed Gündoğdu[4]

[1, 2, 3, 4] IBSS, Istanbul, Türkiye
*Corresponding Author Email: kadiryunus.koc@ibss.com.tr

Abstract

There have been many studies in the field of image processing from past to present. Studies in the field of image processing are quite useful in cases where a system is automated and the human eye cannot follow. These studies use image processing techniques in order to ensure control during the production phase or in order to ensure correct data flow during the inspection phase. Our work falls within the scope of the inspection phase. Within the scope of our work, it is aimed to transfer the information on the fuel information panels at gas stations to a central system with image processing techniques. In this way, the prices reflected on the panels at gas stations will be controlled. Within the scope of this project, nearby gas station panel images were collected. Labeling was performed for each feature in the collected data set. The Yolov8 model is used for object detection.

Keywords

Automation. central system, computer vision, data collection, feature labeling, fuel price monitoring, gas station, image processing, machine learning, object detection, production control, real-time data processing, technology integration, yolov8.

INTRODUCTION

Today, we see many studies in the field of image processing. We see these studies as license plate reading systems, facial recognition systems, material control in the production line or disease detection studies in the health field. When image processing works with artificial intelligence, it produces very useful results [1].

When we look at these useful results, we use image processing techniques and artificial intelligence together within the scope of our project. In our project, fuel information on the panels at gas stations is read using image processing and artificial intelligence. Fuel price information read from the panel is kept in a central system. In this way, changes in fuel prices at gas stations are controlled from a single system. In the project development step, images of fuel information panels at different gas stations were collected. The parts with fuel information on the 220 images we collected were labeled. The labeled images are trained with the yolov8 model. The Yolo model is a model that provides great benefits in the object detection phase [2]. The areas labeled with the yolo model are detected from the newly received images. After the area detection, keras-ocr is used. In this way, the detected part on the image is converted to text. keras-ocr offers a useful method in the image-to-text conversion phase [3]. As a result of our study on 220 images, it is seen that the text detection rate on fuel information board images is low. For this reason, 220 fuel information board images are multiplied using image processing techniques. Image angles and colors are manipulated on 220 images. In this way, the number of fuel information board images is increased to 800. A new model is trained with 800 fuel information board images. A new model is trained to better read the fuel price information section on each image. With this model, a labeled dataset containing the numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 is prepared. Thanks to this dataset, it is possible to detect numerical data on fuel information board images. It is seen that our models produce more successful results after the data multiplexing and numerical data labeling stages we have performed.

RELATED WORKS

There are many studies on image processing, object detection and Optical Character Recognition (OCR). This study focuses on object detection and OCR. The techniques used enable the automatic reading of price boards in gas stations and check their accuracy.

A research effort focuses on automatically capturing the images of fuel panels while the vehicle is in motion and detecting the prices and fuel types [4]. In this study, appropriate capture settings were made by taking advantage of the features of smart mobile phones. With these settings, the price extraction rate successfully increased by more than 40%. In addition, a preselection threshold value was determined according to the values obtained from the accelerometer of the phone to prevent blurred images due to vibrations in the vehicle. These selections improved the performance of the system by reducing blurred images by 78.57%. This reported that in their study, where they introduced the method of automatic fuel price collection using camera images from roadside price boards at service stations using wireless sensor networks, they detected the price boards with a 92.3% success rate and achieved an accurate price reading of 87.7% [5].

Another study discusses the automatic reading of natural gas meters using object detection, image processing techniques, and OCR [6]. The YOLO model is used to recognize the digits on the meter. For training the model, 10,000 images were used and a success rate of 98% was achieved.

When the number of data is insufficient, as in our study, various data augmentation techniques are used to generate synthetic data to improve the performance of the model. It has been shown that data augmentation techniques such as random cropping, rotation, color adjustments, and noise addition significantly improve performance in object detection and OCR [7]. This report will review existing image data augmentation techniques used to improve the accuracy of deep learning models in domains where limited datasets are encountered [8].

METHODOLOGY

In this section, we describe the methods we used to identify and read fuel types and their prices from photographs of gas station panels. In this project, we used the YOLOv8 model for object detection and the Keras OCR model for OCR. In addition, model performance was optimized with data (image) augmentation techniques to improve model performance

Dataset

The dataset used in this study consists of images of various gas stations taken from different angles. These images have been carefully selected in order to clearly identify the types of fuel and their prices on the billboards. A sample image is given below.

Figure 1. Sample Image

Image Processing

Image processing is the representation of images in a digital environment using a computer through an algorithm [7]. In addition, it is the process of extracting information from photos, enhancing the quality of the photo, and automating photo-based operations with the help of various algorithms and mathematical models. These processes address each pixel or a group of pixels. The similarities or differences between the addressed pixels or groups provide information about the photo and its content. This information is then used appropriately for various purposes.

Nowadays, since it is difficult to perform these operations by simply looking at pixels or groups of pixels (e.g., reading the text in a photo), deep learning methods are used. The most common uses of these methods include classification, object detection and segmentation. With these methods, we can determine the type of a photo, its content and where that content is located in the photo. This information provides valuable insights in many areas such as autonomous vehicles, MRI results, face recognition and many more. The main reason why these methods are so popular today is that their deep learning models incorporate the CNN architecture [9]. The basic CNN architecture is as follows.

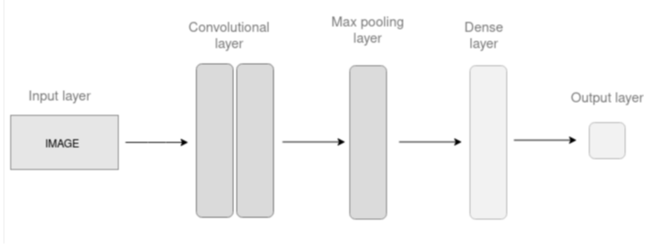

Figure 2. CNN Architecture

Object Detection

Object detection is an important computer vision task that involves identifying and locating objects in an image. This process not only classifies objects but also determines their boundaries by specifying the coordinates of the detected object.

You Only Look Once (YOLO) Algorithm for Object Detection

The YOLO algorithm provides a significant innovation and simplification in object detection. This algorithm was initially developed in 2016 [10]. YOLO both classifies objects and determines the coordinates of an object by processing the objects in the image at once with high speed and accuracy. YOLO's architecture is based on Convolutional Neural Networks (CNNs), which allows it to efficiently extract features and make predictions in a single forward pass. Thanks to its CNN-based structure, YOLO can be used in many sectors such as video analysis, security systems, industrial systems, and automation [11].

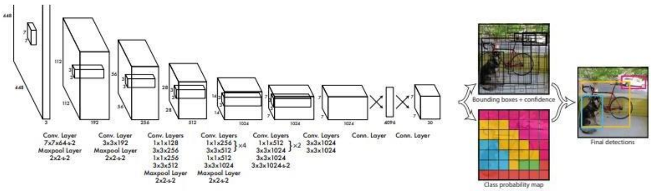

Figure 3. Basic Yolo Architecture

Unlike traditional object detection methods, YOLO's single stage structure reduces computational cost, resulting in faster and more efficient results. Thanks to these advantages, YOLO is widely used in both academic research and industrial applications.

There are several versions of the YOLO algorithm and YOLOv8, one of the most recent and advanced models, was chosen for this study. YOLOv8 takes the single-stage approach even further with advanced architecture and optimization techniques, ensuring high speed and accuracy [12].

Optical Character Recognition – OCR

Optical character recognition (OCR) is the process of converting characters in photographs into text. The goal is to detect the characters within the content of a photo, regardless of whether it is a document, passport, invoice, landscape, city, house, or similar, and convert them into text. In doing so, it must take into account different fonts (including handwriting), lighting conditions, colors, and the angle at which the photo was taken. One of the most well-known engines is Tesseract [13].

In this project, we utilized the Keras OCR model in addition to the Tesseract engine.

Keras OCR is a text recognition model developed on TensorFlow and Keras platforms. Using deep learning methods, it can successfully process various fonts and complex text types. This model offers the ability to recognize text with high accuracy, especially in complex and variable environments.

Evaluation of Object Detection and Optical Character Recognition (OCR)

Intersection over Union (IoU) is used to measure the accuracy of object detection. IoU calculates the correspondence ratio between the area estimated by the object detection model and the actual area [14].

$$\text{IoU} = \frac{\text{Intersection Area}}{\text{Union Area}} \quad (1)$$

The IoU value is between 0 and 1, with 1 representing perfect overlap. This metric is used to assess the accuracy of predicted objects.

Character Error Rate (CER) and Word Error Rate (WER) are two key metrics used to evaluate the accuracy of OCR systems. Character Error Rate (CER) measures the proportion of erroneous characters in text recognized by OCR. The Word Error Rate (WER) detects the error rate at the word level.

$$\text{CER} = \frac{\text{Number of Incorrect Characters}}{\text{Total Number of Characters}} \quad (2)$$

$$\text{WER} = \frac{\text{Insertions} + \text{Deletions} + \text{Substitutions}}{\text{Total Words}} \quad (3)$$

First of all, in our study, we trained the YOLOv8 model for object detection using 80% of the 220 photographs we took of price information boards at gas stations. However, training with this limited number of images did not yield satisfactory results in model testing. To overcome this and improve the model's performance, we utilized image augmentation techniques. We increased our dataset from 220 images to 800 images by changing the angles and colors of the available images we had. We retrained our object detection model with the large dataset we obtained and achieved significant success in object detection. An example image where we performed object detection using our trained model is shown below.

Figure 4. Sample Object Detection

Using our model, we successfully identified the location of panels, fuel types and prices in photos taken at gas stations. Before applying OCR, we applied preprocessing steps such as bitwise notation and gray scaling to detect the characters of each fuel type and price more successfully. In addition to these techniques, we aimed to enhance the success of the OCR application by using morphological transformations such as erosion and dilation to reduce noise in areas related to fuel types and prices. Following these preprocessing steps, we used the Tesseract model; however, it did not achieve the expected success in recognizing prices and fuel types. One of the main reasons for this was the variation in panel types across different gas stations. While some panels were LED displays, others were made up of seven-segment digital panels. This variation showed us that the preprocessing settings applied to one photograph were not suitable for others. Additionally, the angles and colors of the photographs further complicated the reading of the text and digits. Next, we deployed the KerasOCR model and achieved successful results in recognizing fuel types. However, we still did not achieve the expected success in detecting prices. Despite these steps, since we did not attain the desired accuracy in digit recognition, we decided to adopt a different approach.

To accurately detect prices, we decided to use object detection instead of applying OCR to the prices. First, we cropped the sections of the photos corresponding to the prices by determining their coordinates. Then, we labeled each digit from 0 to 9 as a separate object. Using the 770 images consisting solely of prices, we created a separate object detection model with YOLOv8. This allowed us to develop a new object detection model capable of detecting each digit accurately, even in different types of signs, angles, and colors. With this approach, we were able to achieve high accuracy in detecting prices by directly identifying digits as objects. In the simulation we prepared, an example image along with the detected fuel types and prices is presented below.

Figure 5. Image of a Gas Station

fuel_type	price
oto ipg aygoz otogoz +	18.39
Kursunsuz benzin 95 ultraforce 95	35.22
Motorin ultraforce	37.01
Motorin ecoforce	36.97

Figure 6. OCR Results

CONCLUSION AND FUTURE WORKS

In this study, the object detection performance of the YOLOv8 model was evaluated using photographs of panels from fuel stations. Initially, the amount of data used for training the model was insufficient, resulting in unsatisfactory test results. However, by employing image augmentation techniques, the dataset size was increased, leading to a significant improvement in the YOLOv8 model's object detection performance.

The model successfully identified the locations of panels, fuel types, and prices. We initially attempted to read prices using the Tesseract OCR model but did not achieve success. Using the Keras-OCR model, we achieved good results in reading fuel types; however, detecting prices remained challenging. To address these difficulties caused by different types of signs, we individually labeled each digit and retrained the YOLOv8 model, enabling accurate reading of prices on the panels.

And concise, and they should not be complete sentences. with the tab key. to its appearance in the text. Cite each table in numerical order. be placed in parentheses to the right of the equation. Do not create equations as pictures. Use MathType or insert symbols as normal text.(1) by MathType:(2) should be numbered consecutively with Arabic numerals to avoid ambiguities, if they will be referred to in text. Citation for an equation should be made by using "(1)," not "Eq. (1)" or "equation (1)," except at the beginning of a sentence: "Equation (1) is…" appendix may be included (and is often helpful) in mathematical or computational modeling. We are very grateful to experts for their appropriate and constructive suggestions to improve this template. Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

[1] C. Kaur and U. Garg, "Artificial intelligence techniques for cancer detection in medical image processing: A review," Materials Today: Proceedings, vol. 81, pp. 806-809, 2023.

[2] Y. Swathi and M. Challa, "YOLOv8: Advancements and Innovations in Object Detection," in International Conference on Smart Computing and Communication, Singapore: Springer Nature Singapore, 2024.

[3] H. Moussaoui et al., "Enhancing automated vehicle identification by integrating YOLO v8 and OCR techniques for high-precision license plate detection and recognition," Scientific Reports, vol. 14, no. 1, p. 14389, 2024.

[4] Y. F. Dong, S. Kanhere, C. T. Chou, and R. P. Liu, "Automatic Image Capturing and Processing for PetrolWatch," Technical Report UNSW-CSE-TR-1109, University of New South Wales, Australia, May 2011.

[5] W. Wahlster, "Verbmobil: Foundations of Speech-to-Speech Translation," in Foundations of Speech-to-Speech Translation, Springer, Berlin, Heidelberg, pp. 33-60, 2000, doi:10.1007/978-3-540-69170-9_10.

[6] A. Iqbal, A. Basit, I. Ali, J. Babar, and I. Ullah, "Automated Meter Reading Detection Using Inception with Single Shot Multi-Box Detector," Department of Computer Science, University of Balochistan, Quetta, Pakistan, 2020.

[7] M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A Comprehensive Survey of Image Augmentation Techniques for Deep Learning," Pattern Recognition, 2023, doi: 10.1016/ j.patcog.2023.109347.

[8] N. E. Khalifa, M. Loey, and S. Mirjalili, "A comprehensive survey of recent trends in deep learning for digital images augmentation," Artificial Intelligence Review, vol. 55, pp. 2351–2377, 2022, doi:10.1007/s10462-021-10066-4.

[9] R. C. Gonzalez, "Digital Image Processing," New York, NY, USA: Pearson, 2018, ISBN: 978-0-13-335672-4.

[10] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," 2015. doi:10.48550/arXiv.1511.08458.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

[12] Y. Redmon, A. Farhadi, and S. Divvala, "YOLO9000: Better, Faster, Stronger," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 7263-7271.

[13] Y. Swathi and M. Challa, "YOLOv8: Advancements and Innovations in Object Detection," in International Conference on Smart Computing and Communication, Singapore: Springer Nature Singapore, 2024.

[14] R. Smith, "An Overview of the Tesseract OCR Engine," in Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Curitiba, Brazil, 2007, pp. 629-633, doi: 10.1109/ICDAR.2007.4376991.