Design of an Enhanced Object Detection Algorithm through Image Scaling

Gabriel Avelino Sampedro*, Yohana Jayanti Aruan†, Jae-Min Lee‡, Dong-Seong Kim§
*Computer Engineering Department, Technological Institute of the Philippines (TIP), Quezon City, Philippines
*†‡§Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea
Email: {garsampedro*, yohanajaruan†}@gmail.com; {ljmpaul‡, diskim§}@kumoh.ac.kr

***Abstract*—This paper proposes the implementation of an image scaling algorithm with the use of an object detection algorithm. You Only Look Once (YOLO) algorithm is a quick object detection algorithm that only needs to scan an image once, before detecting objects of interest; however, there are some inaccuracies in the detection of small objects, given the limitation of the clustering boxes generated by the algorithm. Through the implementation of a scaling algorithm, the researchers investigate the increase in detection rates and the design is deemed ready for implementation.**

***Index Terms*—Image scaling algorithm, object detection, You Only Look Once Algorithm (YOLO).**

I. Introduction

Over the years, numerous developments have been made in the field of machine vision. In [1]–[3], machine vision is described to be the ability of a computer to view images in a similar way as humans do. Images are visual and not as easy to analyze as numbers, but with further analysis, images can be represented by wide arrays of numbers describing its various features. Countless researches have been conducted to further enhance the way machines view images and a large number of those researches involve the use of artificial intelligence and machine learning algorithms [4], [5]. In this research, the researchers' goal is to further enhance the accuracy of an object detection algorithm, through its integration with an image scaling algorithm. The system aims to explore the implementation of YOLO with bilinear interpolation (BI).

The YOLO algorithm, as the name suggests, is a quick object detection algorithm that only runs the image through its algorithm once, before generating results from its classification [6]. It is implemented using a convolutional neural network, specifically an improvement to the Region-based Convolutional Neural Network algorithm. The objective of YOLO is to subdivide an image into different boxes, through a grid-based layout system, before analyzing each grid box and comparing it to datasets of each object classification. Given its quick response and output, YOLO is preferred in real-time systems. Despite the accuracy of YOLO, there are still some limitations in the application of the algorithm with images that contain small objects of interest. The bounding boxes are limited to the grids on the image, thus objects which are very small compared to the grid box may be missed out. Fig. 1 illustrates the implementation of the algorithm from the application of a bounding box, to the generation of a class probability map and bonding boxes, to the final detection of objects of interest.

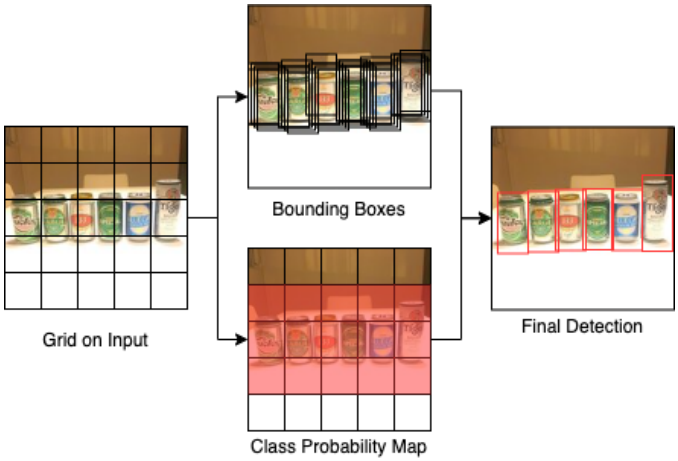

Fig. 1. YOLO implementation shows how the image is subdivided into different segments before object detection and classification.

In order to address the limitations set by small images, the researchers explore the use of BI to scale up the image. BI is an extension of linear interpolation wherein each pixel in an image is expanded and interpolation techniques are used to fill up the gaps caused by expansion [7]. The algorithm makes use of distance-weighted averages of the various pixels surrounding the gap and uses the closest value to determine the pixel color to fill it with. Common applications of BI would include image scaling and grid mapping problems. Fig. 2 illustrates the implementation of the scaling algorithm and the enlargement process.

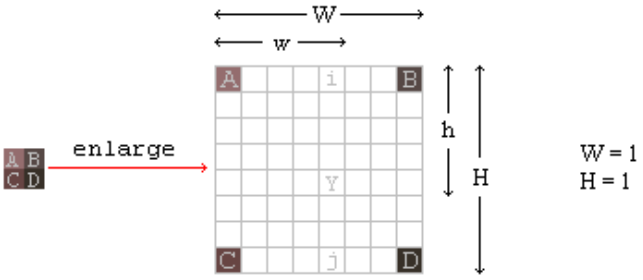

Fig. 2. Image enlargement using BI

The general objective of this paper is to plan out the

implementation of a system to further improve the YOLO algorithm to identify small objects within a given image. Specifically, this design aims to integrate the use of BI to scale up the image and apply filtering methods to clean images. The use of BI shall allow the resizing of the image and thereby allow the detection of small objects within a given image.

II. Methodology

The formulation and conceptualization of this research was derived from one of the underlying problems in the implementation of YOLO as an object detection algorithm – limitation in the detection of small objects of interest. This research aims to apply a scaling algorithm before applying object detection to increase the detection rates of small-sized objects.

A. Software Development Design

In the development of the software, Python is used together with an optimized computer vision library (OpenCV) and a CNN algorithm (YOLO). The development of the proposed solution entails a comparative analysis between a manual inspection of images and a manual count of objects of interest versus the object count generated by the implementation of the proposed solution.

In the system, an input image shall be fed together with a dataset to be used for comparing detected objects with samples of objects of interest. Once entered, a series of processes which include image preprocessing, image scaling, YOLO, and image descaling will be applied. At the last stage of the process, the output shall be the objects detected by the system.

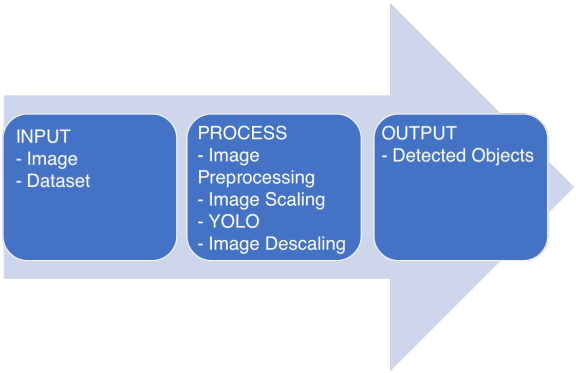

Fig. 3. Software design flow shows the different processes applied to the inputs to produce the output

B. Performance Analysis

To further validate the results of the proposed system, a comparison shall be made between manual counts done on sample images, object detection count made using YOLO only, and an automated object detection count made using the proposed solution. From the results, the percent difference (PD) between the manual count and automated count using both YOLO and YOLO with BI will be cobtained. The testing of the algorithm would be applied to 50 sample images. The desired outcome is that the results from the proposed solution would be closer to the manual count results, as compared to the outcome of using YOLO only. To compare if the results from the proposed method is statistically close to the manual count, a t-test shall be used to validate the test results. Through the use of a t-test, outliers which may distort the results can be avoided.

III. Conclusion

The proposed system aims to detect more objects of interest compared to the use of YOLO alone. The next step in this research would be for the implementation of the proposed system. Furthermore, the researchers aim to search through different scaling algorithms and compare the results with that of BI.

The proposed system is an ideal model that can be used for searching for objects of interest that may be small in size. Applications of such may include monitoring of people captured through a CCTV, automating tedious tasks such as manual counts in medical fields, and even in industry applications. The increase of the accuracy of YOLO, after the application of the proposed system, would greatly advance researches in those fields.

Acknowledgement

This research work was supported by Priority Research Centers Program through NRF funded by MEST (2018R1A6A1A03024003) and the Grand Information Technology Research Center support program (IITP-2021-2020-0-01612) supervised by the IITP by MSIT, Korea.

References

[1] D. P. Penumuru, S. Muthuswamy, and P. Karumbu, "Identification and classification of materials using machine vision and machine learning in the context of industry 4.0," *Journal of Intelligent Manufacturing*, pp. 1–13, 2019.

[2] S. Puttemans, T. Goedemé, A. Mian, T. B. Moeslund, and R. Gade, "Special issue on advanced machine vision," 2020.

[3] C.-B. Moon, J.-Y. Lee, D.-S. Kim, and B.-M. Kim, "Inspection system for vehicle headlight defects based on convolutional neural network," *Applied Sciences*, vol. 11, no. 10, p. 4402, 2021.

[4] S. Raschka and V. Mirjalili, "Python machine learning: Machine learning and deep learning with python," *Scikit-Learn, and TensorFlow. Second edition ed*, 2017.

[5] C. B. Moon, B. M. Kim, and D.-S. Kim, "Real-time parallel image-processing scheme for a fire-control system," *IEIE Transactions on Smart Processing & Computing*, vol. 8, no. 1, pp. 27–35, 2019.

[6] M. B. Ullah, "Cpu based yolo: A real time object detection algorithm," in *2020 IEEE Region 10 Symposium (TENSYMP)*. IEEE, 2020, pp. 552–555.

[7] E. J. Kirkland, "Bilinear interpolation," in *Advanced Computing in Electron Microscopy*. Springer, 2010, pp. 261–263.