Medicine Classification Using YOLOv4 and Tesseract OCR

Charmaine C. Paglinawan
School of Electrical, Electronics, and Computer Engineering
Mapua University
Intramuros Manila, Philippines
ccpaglinawan@mapua.edu.ph

Marielle Hannah M. Caliolio
School of Electrical, Electronics, and Computer Engineering
Mapua University
Intramuros Manila, Philippines
mhmcaliolio@mymail.mapua.edu.ph

Joshua B. Frias
School of Electrical, Electronics, and Computer Engineering
Mapua University
Intramuros Manila, Philippines
jbfrias@mymail.mapua.edu.ph

***Abstract*—For seniors, consuming the incorrect pharmaceutical drug is a common but serious medical issue. Poor vision, difficulty reading, impaired memory, and problems with fine-motor skills are all factors that might lead to using the incorrect pharmaceutical drug. Previous studies have created systems that can automatically identify medications to address this issue. Although, this research lacked the integration of the latest official version of YOLO and the Tesseract OCR. As such, the researchers developed a system that can classify medicines using YOLOv4 and Tesseract OCR. Through tests, YOLOv4 was established to be the best in accuracy (~95% accuracy), followed by YOLOv4 and Tesseract OCR together (~19% accuracy). The least accurate one was using the Tesseract OCR only (~14% accuracy).**

Keywords—YOLOv4, tesseract OCR, accuracy, medicine

I. Introduction

For seniors, consuming the incorrect pharmaceutical drug is a common but serious medical issue. Poor vision, reading difficulty, impaired memory, and fine motor skills are all factors that might lead to using the incorrect pharmaceutical drug [1], [2]. Around 39% of senior participants admitted to taking their medications inappropriately. Additionally, many seniors with chronic illnesses must take many medications at once [3].

To address this issue, previous studies have created systems that can automatically identify medications [4], [5]. The first research utilized the Inception V3 and, Inception V4 to identify the medicines with accuracies of 92.75% and 94.85% respectively. In contrast, the second research used a previous version of the You Only Look Once algorithm (version 2) for the same purpose with an accuracy of 90%. Thus, in addition to utilizing an older version of the YOLO algorithm, previous research also lacked the integration of an OCR feature to take into account the medication texts present within the medication packaging for medication classification.

For this study, the researchers' objectives are to develop a system that can classify medicines using YOLOv4 and Tesseract OCR. To accomplish the said general objective, the researchers aim to test the accuracy of the different algorithms in identifying and tracking particular pharmaceutical drug text using (1) the YOLOv4 algorithm, (2) Tesseract OCR, and (3) both algorithms together.

This study will only cover the YOLOv4 and Tesseract OCR algorithms. The researchers would be testing both programs in combination with one another and individually. The researchers only utilized pharmaceutical drugs that were stored within the "blister" classification of packaging. In addition, those medicines were only examined one by one since they were already separated individually before being placed within the system.

II. Review of Related Literature

A. Tesseract OCR

Various operating systems can employ optical character recognition (OCR). The primary function of this algorithm is to recognize numbers and letters. Several earlier studies have used the Tesseract OCR tool for that reason [6]. In one study, physical tag information was recognized using Tesseract OCR and the YOLOv3 [7]. Within their second iteration, they achieved an AP (Average-Precision) score of around 93%.

Furthermore, within [8], To circumvent the challenges that manual methods face and to prevent any use of human input, they used OCR in their system for managing pension funds. The findings of their study indicate that several flaws contributed to errors, which they sought to lessen by using strategies like scanning the material at a quality of about 200 dpi. Additionally, the document must follow a specific structure.

B. YOLO Algorithm

Modern object-detection methods like the algorithm YOLO also perform localization and identification. It also performs admirably in circumstances that call for fast outputs. YOLOv4 was used within the research of [9] to overcome the social-distancing problem posed by the pandemic. The system proved very efficient because its accuracy score is about 96%. Additionally, YOLO's latest official version (v4) outperformed faster RCNN and its older iterations in speed and accuracy. In other research, YOLOv4 was utilized to swiftly classify different types of vehicles, which also proved the superior speed and accuracy of the algorithm as mentioned above compared to others within the same category [10]. There have been debates about using the supposedly newest version of YOLO called YOLOv5. However, this study utilized YOLOv4 because there is no official research or published papers for

2023 15th International Conference on Computer and Automation Engineering (ICCAE) | 979-8-3503-9622-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCAE56788.2023.10111387

YOLOv5 yet as of the time of this researche's making, unlike YOLOv4 which has been well-researched and tested. Using YOLOv5 without proper research and peer review may not be reliable [11], [12]. With that being said, the motivation for this research to utilize YOLOv4 is that as previously stated, it is the latest official version of the YOLO algorithm with the best performance. Innovations that were made for the said algorithm was that the research utilized its own custom dataset, as such parameters such as the width and height and fine tuning for the number of classes to be accommodated were made for the system's configuration.

III. Methodology

A. Methodological Approach

The system will require data broken into three distinct classifications: the training dataset (300 images for each medicine and negative training data), metadata using labeling software specifically for YOLOv4, and the testing dataset from the system itself.

B. Conceptual Framework

The diagram's input section shows the many types of information that will be provided to the system, expressly the ones mentioned in the methodological approach previously. By comparison, there will be five different procedures in the process segment of the diagram, which include: (1) localizing the placement of the medicine within the image provided through YOLOv4, (2) performing non-max suppression to remove all unnecessary bounding boxes, (3) performing text detection and recognition through the usage of Tesseract OCR. Lastly, YOLOv4 will utilize its built-in improvements called (4) BoF and (5) BoS to increase efficiency. These processes will lead to the output, which is proper medicine classification.

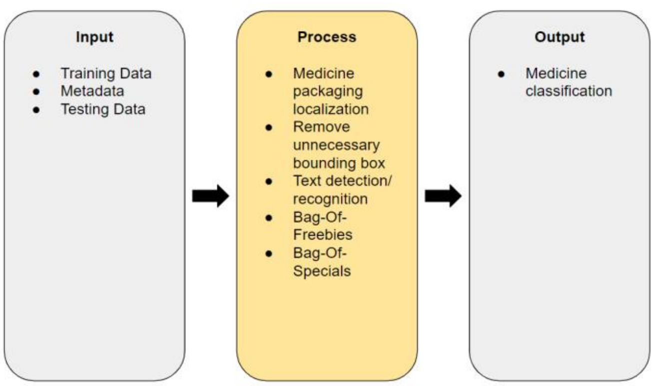


Fig. 1. Conceptual Framework.

C. Flow Chart

For the flowchart it can be seen that the medication must first be placed on the viewing platform for the system to be able to capture proper pictures of the medication. Afterwards, the system can then perform localization of the placement of the medication within the photo taken and perform text detection and prediction to finally identify the medication itself. This process will repeat 8 times for the 8 medications

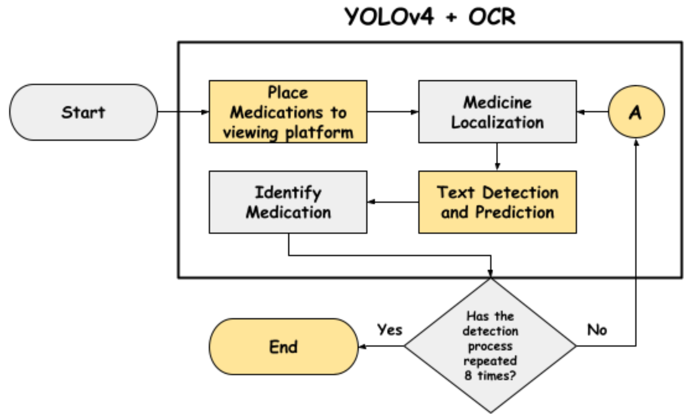
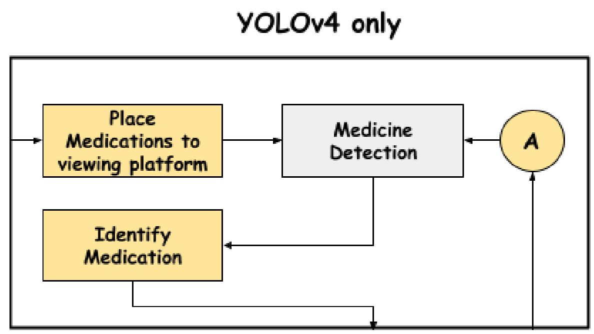


Fig. 2. System flowchart.

Only minor changes were made to the diagram above to create two brand-new flowcharts because the researchers also used YOLOv4 and Tesseract OCR to classify the medicines independently.



Modification for YOLOv4 only

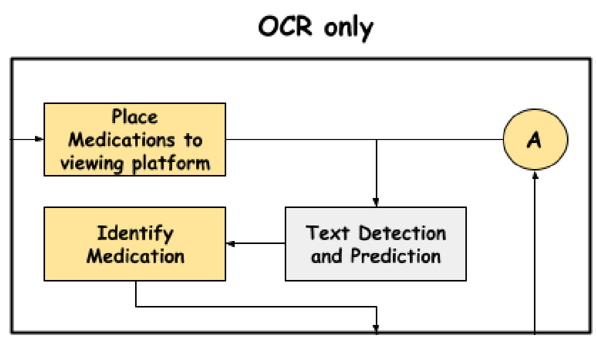


Fig. 3. Modification for Tesseract OCR only.

D. Engineering Standards

A specific standard utilized within the study is the 1118.1 standard from IEEE. This standard was all about the definition of microcontrollers and their primary usage. A microcontroller is a technology that, by nature, can send commands through a serial connection across other sensing devices and electronic parts that might or might not be associated with the microcontroller. As for the system within the study, an Arduino microcontroller was used to control/instruct several of the system's other parts via serial connection.

E. Statistical Analysis

In Table I, is the confusion matrix for guiding the researchers regarding what events are considered as True-Positive, False-Negative, etc., which would be vital for knowing if the system is actually able to tell whether the medication should be given to the patient or not within a pre-established list of correct and incorrect medications.

TABLE I. CONFUSION MATRIX

Value	Description
True Positive	Medication is correct and is detected as correct
True Negative	Medication is incorrect and is detected as incorrect
False Positive	Medication is incorrect but is detected as correct
False Negative	Medication is correct but is detected as incorrect; Medication is incorrect and is detected as incorrect, but medication texts are not detected properly (When using OCR)

F1-score, Recall (R), Accuracy (A), and Precision (P) were utilized in the study to evaluate the performance of the algorithms. The formulae for each one are displayed below:

$$R = \frac{TP}{TP+FN};\ P = \frac{TP}{TP+FP};\ F1\,score = 2x\frac{PxR}{P+R}$$

$$A = \frac{TP+TN}{\qquad\qquad} x100\%$$

where TP is the True Positive, TN is the True Negative, FP is the false positive and FN is the False Negative. F1 score is the harmonic mean machine learning metric utilizing precision and recall, and is good, if close to a value of 1

IV. RESULTS AND DISCUSSION

The combined algorithms are YOLOv4 and Tesseract OCR, were evaluated for the first test. Furthermore, the next tests would also examine how effectively each algorithm performed alone in selecting whether or not to approve or deny medications. For the combined algorithm, the Tesseract OCR was used to recognize the alphanumeric character on the pharmaceutical medicine's packaging, whereas YOLOv4 was employed to localize it. In comparison, if only YOLOv4 had been used, the pharmaceutical drug's acceptance depended solely on YOLOv4's capacity to recognize medicines through their packaging. Lastly, using Tesseract OCR alone, the researchers would test if it could read the writing on the medicine's packaging. The medicines to be accepted and rejected are displayed in Fig. 4.

Fig. 4. Medications within the list/Correct Medication.

Fig. 5. Medications outside the list/Incorrect Medication.

A. Algorithm Accuracy Test

The results presented above proved the accurate result of YOLOv4 in recognizing and identifying medications based on their packaging's numerous visual elements and the outlines of any writings or letterings included. Furthermore, it showed that the information provided to YOLOv4 during training on primary medications was to be accepted. The negative photos and information adequately identified all the tested medications correctly.

TABLE II. TESTING A.1: ALGORITHM ACCURACY (YOLOV4 ONLY)

Subjects in all trials: Ibuprofen Paracetamol 200mg/325mg, (~9.5%), Phenylephrine HCL Chlorphenamine Maleate Paracetamol 10mg/2mg/500mg, (~9.5%), Paracetamol 500mg, (~9.5%), Ascorbic Acid & Zinc 500mg/10mg, (~9.5%), Simvastatin 40mg (~9.5%), Silymarin 125mg, and (~4.76%), Phenylpropanolamine HCL Paracetamol 25mg/500mg, (~4.76%), Simvastatin 20mg (~14.29%), Ibuprofen 400mg (~14.29%), and Tramadol Hydrochloride Paracetamol 37.5mg/325mg (~14.29%)

Trial#	True Positive	True Negative	False Positive	False Negative	Sorting Time
1	4	2	1	0	4m & 12s
2	4	3	0	0	4m & 15s
3	4	3	0	0	4m & 11s
Total:	**12**	**8**	**1**	**0**	-

Accuracy Metrics

Accuracy	Precision	Recall	F1 score
95.2381%	0.9231	1	0.9600

TABLE III. COMPARING YOLOV4 ACCURACY TO PREVIOUS STUDIES

Comparing Accuracy of YOLOv4 to other algorithms	
	Accuracy
YOLOv4	95.2381%
YOLOv2	90%
Inception v3	92.75%
Inception v4	94.85%

It can be seen that the results of YOLOv4 performed slightly better than the object detection frameworks in idenitifying medications.

TABLE IV. TESTING A.2: ALGORITHM ACCURACY (TESSERACT OCR ONLY)

Subjects in all trials: Ibuprofen Paracetamol 200mg/325mg, (~9.5%), Phenylephrine HCL Chlorphenamine Maleate Paracetamol 10mg/2mg/500mg, (~9.5%), Paracetamol 500mg, (~9.5%), Ascorbic Acid & Zinc 500mg/10mg, (~9.5%), Simvastatin 40mg (~9.5%), Silymarin 125mg, and (~4.76%), Phenylpropanolamine HCL Paracetamol 25mg/500mg, (~4.76%), Simvastatin 20mg (~14.29%), Ibuprofen 400mg (~14.29%), and Tramadol Hydrochloride Paracetamol 37.5mg/325mg (~14.29%)					
Trial#	**True Positive**	**True Negative**	**False Positive**	**False Negative**	**Detection Time**
1	0	0	0	7	3m & 46s
2	0	1	0	6	3m & 48s
3	2	0	0	5	3m & 49s
Total:	**2**	**1**	**0**	**18**	-
Accuracy Metrics					
Accuracy	**Precision**	**Recall**	**F1 score**		
14.2857%	1	0.1	0.1818		

This effectiveness can be associated with several factors, including (a) Tesseract's hypersensitivity in terms of what it might detect (since occasionally the printed information is already substantial and comprehensible; however, the algorithm was still not seeing it), (b) the various typefaces used in the majority of prescription drugs are not properly detached from one another or not printed adequately, and (c) the differing specular lighting of the blister wrappers that impact the detection performance.

TABLE V. TESTING A.3: ALGORITHM ACCURACY (YOLOV4 AND TESSERACT OCR)

Subjects in all trials: Ibuprofen Paracetamol 200mg/325mg, (~9.5%), Phenylephrine HCL Chlorphenamine Maleate Paracetamol 10mg/2mg/500mg, (~9.5%), Paracetamol 500mg, (~9.5%), Ascorbic Acid & Zinc 500mg/10mg, (~9.5%), Simvastatin 40mg (~9.5%), Silymarin 125mg, and (~4.76%), Phenylpropanolamine HCL Paracetamol 25mg/500mg, (~4.76%), Simvastatin 20mg (~14.29%), Ibuprofen 400mg (~14.29%), and Tramadol Hydrochloride Paracetamol 37.5mg/325mg (~14.29%)					
Trial#	**True Positive**	**True Negative**	**False Positive**	**False Negative**	**Sorting Time**
1	2	0	0	5	4m & 21s
2	0	1	0	6	4m & 23s
3	1	0	0	6	4m & 23s
Total:	**3**	**1**	**0**	**17**	-
Accuracy Metrics					
Accuracy	**Precision**	**Recall**	**F1 score**		
19.0476%	1	0.15	0.2609		

The prior technique was only marginally enhanced by combining both algorithms. Within trials, it was observed that YOLOv4 was virtually always effectively doing its task of localizing the medications. Although, even if Tesseract OCR's extremely high sensitivity in recognizing the pharmaceutical drug's inscriptions were still wildly wrong.

V. CONCLUSION

To sum up, the study assessed how well Tesseract OCR and YOLOv4 identified and listed all necessary inscriptions on pharmaceutical packages, precisely the dosage as well as the generic name of the medicine. It is found that the best method is to use YOLOv4 by itself. YOLOv4 also considers other visual aspects of pharmaceutical packaging to improve its ability to recognize prescription drugs. The research's main limitations were the use of a limited number of medical classifications and the program's capacity only to handle those sorts of data.

VI. RECOMMENDATIONS

For recommendations, it is suggested that future researches expand on the number of prescription drugs to be handled by the YOLOv4 algorithms and expand the ability of the system to also take into account other inscriptions within the packaging of the prescription drugs; for example, the expiration date.

REFERENCES

[1] D. SchwArtz, M. Wang, L. Zeitz, and M. E. Goss, "Medication errors made by elderly, chronically ill patients," *Am. J. Public Health Nations. Health*, vol. 52, no. 12, pp. 2018–2029, Dec. 1962, doi: 10.2105/AJPH.52.12.2018.

[2] A. C. Paglinawan *et al.*, "Detection of three visual impairments: Strabismus, blind spots, and blurry vision in rural areas using Raspberry PI by implementing hirschberg, visual field, and visual acuity tests," *HNICEM 2017 - 9th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag.*, vol. 2018-January, pp. 1–9, 2018, doi: 10.1109/HNICEM.2017.8269550.

[3] P. Midlöv *et al.*, "Medication report reduces number of medication errors when elderly patients are discharged from hospital," *Pharm. World Sci.*, vol. 30, no. 1, pp. 92–98, Jan. 2008, doi: 10.1007/S11096-007-9149-4.

[4] N. Hnoohom, S. Yuenyong, and P. Chotivatunyu, "MEDiDEN: Automatic Medicine Identification Using a Deep Convolutional Neural Network," *2018 Int. Jt. Symp. Artif. Intell. Nat. Lang. Process. iSAI-NLP 2018 - Proc.*, Jul. 2018, doi: 10.1109/ISAI-NLP.2018.8692824.

[5] H. W. Ting, S. L. Chung, C. F. Chen, H. Y. Chiu, and Y. W. Hsieh, "A drug identification model developed using deep learning technologies: Experience of a medical center in Taiwan," *BMC Health Serv. Res.*, vol. 20, no. 1, pp. 1–9, Apr. 2020, doi: 10.1186/S12913-020-05166-W/FIGURES/4.

[6] R. Smith, "An Overview of the Tesseract OCR Engine."

[7] A. Bazerque, D. Moraes, and M. Souza, "Using Object Detection Algorithm and Optical Character Recognition to Read Data from alphanumeric tags in text," 2020, Accessed: Oct. 12, 2021. [Online]. Available: https://arc.cct.ie/ict.

[8] M. J. C. Samonte, A. M. L. Bejar, H. C. L. Bien, and A. M. D. Cruz, "Senior Citizen Social Pension Management System Using Optical Character Recognition," *ICTC 2019 - 10th Int. Conf. ICT Converg. ICT Converg. Lead. Auton. Futur.*, pp. 456–460, Oct. 2019, doi: 10.1109/ICTC46691.2019.8940013.

[9] J.K.D Lagman, A.B. Evangelista A.and C.C. Paglinawan, "Unmanned Aerial Vehicle with Human Detection and People Counter Using YOLO v5 and Thermal Camera for Search Operations," *2022 IEEE International Conference on Automatic Control and Intelligent Systems, I2CACIS 2022* 10.1109/I2CACIS54679.2022.9815490.

[10] J. A. Kim, J. Y. Sung, and S. H. Park, "Comparison of Faster-RCNN, YOLO, and SSD for Real-Time Vehicle Type Recognition," *2020 IEEE Int. Conf. Consum. Electron. - Asia, ICCE-Asia 2020*, Nov. 2020, doi: 10.1109/ICCE-ASIA49877.2020.9277040.

[11] "YOLOv5 Is Here! Is It Real or a Fake? - viso.ai." https://viso.ai/deep-learning/yolov5-controversy/ (accessed Sep. 19, 2021).

[12] "GitHub - ultralytics/yolov5: YOLOv5." https://github.com/ultralytics/yolov5 (accessed Sep. 19, 2021).