Counterfeit Medicine Detection using Deep Learning

PROF. KIRTI MOTWANI[1], RACHAEL DSOUZA[2], RHEA DSOUZA[3], JANICE JOSE[4]
[1] *Professor, Xavier Institute of Engineering*
[2, 3, 4] *Student, Xavier Institute of Engineering*

Abstract*— Counterfeit medicines are becoming a major concern all over the world. Counterfeit medicines or fake medicines are those that are sold with the intent of deceiving the consumer by altering the packaging or tweaking the prices or even changing the ingredients. According to WHO, about 10 % of the pharmaceutical market in India is counterfeit and this poses a threat to the health of the public [1] . In this paper, we propose a method of detecting these counterfeit medicines using deep learning technologies such as object detection and text recognition in order to give the users the ability to detect if the medicine is counterfeit or not by using the image of the medicine package. Using the image of the medicine package we scan the manufacturer's logo and extract useful information from the image and verify it and give users the result of Real or Counterfeit.

Indexed Terms*— Counterfeit Medicines, Deep Learning, Object Detection, text Recognition.

I. INTRODUCTION

The pharmaceutical industry is one of the largest revenues generating and rapidly growing industry. There is an abundance of drugs, medicines being manufactured by various manufacturing companies. However, one of the major problems is identifying if the medicine is original or not. A counterfeit medicine is an item that is sold with the intent of deceiving the consumer which include those with less or none of the stated active ingredients, with added, sometimes hazardous, adulterants, substituted ingredients, completely misrepresented, or sold with a false brand name (wiki).

A Word Health Organization (WHO) 2017 Report claims that 10.5% of the medicines sold in low- and middle-income countries, including India are falsified which is a serious threat to patient safety [2].

False medications may worsen the spread of diseases or have unfavorable side effects, which can have a serious influence on overall healthcare. There is an abundance of counterfeit medicines being sold and our method proposes a way for consumers to determine if their medicine is counterfeit or not. Our solution focuses on the medicine package, using deep learning algorithms to analyze the logo and text on the package for its authenticity.

II. LITERATURE SURVEY

Many anti-counterfeiting technologies with distinct advantages and disadvantages exist today. From a product standpoint the techniques include overt, covert and track and trace technology [3].

- Overt Technology

This involves using holograms or color shift inks for anti-counterfeiting. Several Indian pharmaceuticals have changed their packaging and implemented the addition of holograms on the package in hopes of reducing the impact of counterfeit activity [4]. Even though holograms are difficult to counterfeit and are easy to track or trace through the distribution chain, it has its drawbacks too, one of which is that it's expensive and adds to the MRP of low-end medicines.

- Covert technology

Covert technology involves methods of identification that is not discernible by the naked eye. Like using Ultraviolet (electromagnetic radiation) and this paper [5] that proposes using X-Ray fluorescence analyser to detect counterfeit medicine along with supervised machine learning algorithms. However, they are not consumer friendly in terms of detection and needs the right equipment for identification.

- Track and trace technology

The track and trace system refers to the procedure of allocating each stock unit with a unique identifier during manufacturing that remains with it throughout the supply chain until it's consumption. This information is attached in the form of unique pack

coding, and this enables accessing the same information on a secure database [6]. Radio Frequency Identification (RFID) tags and barcodes are some of the technologies in this category. RFID uses radio frequency to identify products. The product has a tag attached to it, which contains information or data of the product and is transmitted via the radio frequency channel.

Another method to ensure the safety of pharmaceuticals is by tracing the counterfeit supply chain using Blockchain as proposed in this paper [7]. With the advanced features of Blockchain, detecting counterfeit drugs is made simple and serves as a basis for complete traceability of the drug from manufacturer to end consumer.

III. METHODOLOGY

A. Dataset

Our dataset consists of images of medicine packages that we obtained by Web Scraping using tools such as Scrapy and Selenium. We scraped medicines of 10 manufacturers consisting of an equal number of real and fake images. Due to the lack of availability of fake images, we ourselves created by altering the logo and text on the package. The images are self-annotated using the labelling tool LabelImg.

B. System Design

From the input image we let our trained model detect whether the logo is counterfeit or not along with that it detects medicine name and composition. We extract the text from the input medicine image the user uploads. Those texts include medicine name and composition and we query our database and look for them. Database Cloud Storage for Firebase is a powerful, simple, and cost-effective object storage service built for Google scale. Information regarding medicine, its image, composition name etc are stored in this database. This verifies if the contents on the medicine package is authentic.

We combine our results from the logo detection and text verification to in turn determine if the image of the medicine package the user entered is authentic or counterfeit.

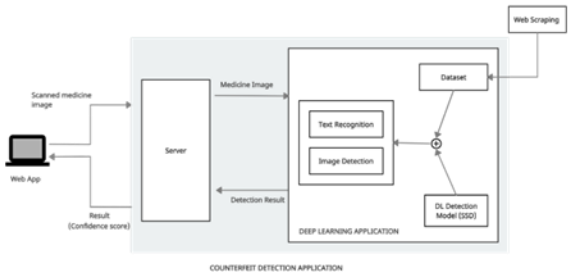

Fig1. Block Diagram

C. Algorithms Used

- Single shot multi-box detector (SSD)

SSD is a method for detecting objects in images using a single deep neural network discretizes the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature map location. At prediction time, the network generates scores for the presence of each object category in each default box and produces adjustments to the box to better match the object shape [8].

Here we use object detection using SSD to detect objects which include different categories like medicine name, medicine logo and medicine composition. It also classifies whether or not the logo on the medicine is fake or real.

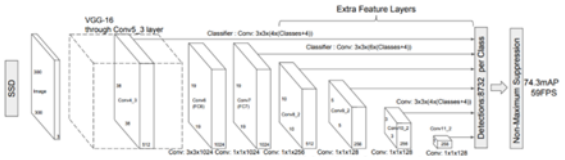

Fig2. SSD architecture

- Azure OCR

Optical Character Recognition has two primary components: text detection and text recognition. Generally, these two components are separate and employ different models for each task. OCR identifies text either written or printed inside an image file or a document, and converts that text into machine-readable form which can further be used for data processing [9].

Azure's Computer Vision service gives us the access to advanced algorithms to process images and returns information based on the visual features. One can use the new Read API to extract printed and handwritten

text from photos and documents. It uses deep-learning-based models and works with text on a variety of surfaces and backgrounds. The APIs support extracting printed text in several languages. Using Azure's OCR API to extract the text which includes medicine name and medicine composition.

D. Flowchart

Fig3. Flowchart

IV. EXPERIMENTAL RESULTS

We can give any image of a medicine package; object detection helps in detecting if the logo is fake or not. If not, we further verify the text to check if it is in the database. Once both return True then we show the result as Real. If any one of them is False it will show Fake.

Fig4. Result

Fig5. Result

Fig6. Result

Fig7. Result

V. CONCLUSION

Medicine counterfeiting is a problem that is addressed by several countries. There is a requirement of multiple measures in order to protect the consumers and the supply chain. Our proposed application makes use of text recognition, image recognition that uses deep learning-based approaches to detect fake medicines. It gives the consumer the power to validate their medicines.

Future works include working on developing a larger dataset containing more images per medicine which would help in the training and increase the accuracy of the model. Secondly, we could employ an alert system that would alert authorities on every counterfeit medicine detected.

Pharma companies tend to change the packaging of the medicine, logos also have different colors for the same

medicine thus we need to work on updating our database real time for verification.

REFERENCES

[1] "DNA," How to differentiate counterfeit medicines from genuine ones, Jun. 24, 2021. https://www.dnaindia.com/analysis/report-dna-special-how-to-differentiate-counterfeit-medicines-from-genuine-ones-2897092

[2] "ETHealthWorld," Counterfeit Drugs: Threat to India's reputation of Indian Self-reliant Model, May 03, 2021. https://health.economictimes.indiatimes.com/news/pharma/counterfeit-drugs-threat-to-indias-reputation-of-indian-self-reliant-model/81342614

[3] E. Daoud, D. Vu, H. Nguyen, and M. Gaedke, "ENHANCING FAKE PRODUCT DETECTION USING DEEP LEARNING OBJECT DETECTION MODELS," IJCISIS, vol. 15, no. 1, pp. 13–24, Oct. 2020, doi: 10.33965/ijcsis_2020150102.

[4] P. Gupta, K. Singhal, and A. Pandey, "Counterfeit (Fake) Drugs & New Technologies to Identify it in India," vol. 3, no. 11, pp. 4057–4064, Oct. 2012.

[5] M. Alsallal, M. S. Sharif, B. Al-Ghzawi, and S. M. Mlkat al Mutoki, "A Machine Learning Technique to Detect Counterfeit Medicine Based on X-Ray Fluorescence Analyser," in 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE), Southend, United Kingdom, Aug. 2018, pp. 118–122. doi: 10.1109/iCCECOME.2018.8659110.

[6] D. Bansal, "Anti-Counterfeit Technologies: A Pharmaceutical Industry Perspective," Sci. Pharm., vol. 81, no. 1, pp. 1–13, 2013, doi: 10.3797/scipharm.1202-03.

[7] R. Kumar and R. Tripathi, "Traceability of counterfeit medicine supply chain through Blockchain," 2019, pp. 568–570. doi: 10.1109/COMSNETS.2019.8711418.

[8] W. Liu et al., "SSD: Single Shot MultiBox Detector," in Computer Vision – ECCV 2016, vol. 9905, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.

[9] G. Boesch, "Optical Character Recognition (OCR) – Overview and Use Cases," viso.ai. https://viso.ai/computer-vision/optical-character-recognition-ocr/