IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 11, Iss 10, 2022

FAKE PRODUCT IDENTIFICATION USING MACHINE LEARNING

[1]R Divya Vani,[2]G Anitha,[3]P Avaniketh,[4]K Rishitha
[1,2,3]Assistant Professor,[4]Student

Department of CSE

Christu Jyothi Institute of Technology &Science,Colombo Nagar, Telangana

ABSTRACT
The problem of recognizing counterfeit (fake) products is a tedious task in certain cases and can be dangerous when it comes to medical products. It becomes easier to produce and sell fake products if an individual does not check the details of the product properly. This paper suggests a better solution using Artificial Intelligence for non-tech-savvy customers who can scan the product with the help of a mobile application to identify, if the product received is fake or original. The major focus will be on the detection of logos (both image and textual representation).

Keywords — Counterfeit products, Artificial Intelligence, non tech-savvy, logos.

I. INTRODUCTION

A logo is a visual representation commonly used by companies and organizations to endorse public acknowledgment. Logos may be either purely graphical (only symbols), purely textual (only name of the organization), or textual-graphical (a combination of symbols and text). Logos are allied with the shared values and brands of organizations and hence their designs are protected by several Intellectual Property Rights (IPR) protocols [2]. This makes a logo exclusive to a particular organization and delivers a visual indication for distinguishing and identifying it [1]. A product's logo is often recognized as one of the most vital elements of a product. Possibly, that is one of the key reasons why logos of renowned organizations are counterfeited. The counterfeit trade is becoming a growing hazard in the evolving markets and many successful products are falling prey to it. Counterfeit logos not only betray customers looking to procure the original product, financial loss but can also have an adverse impact on sales of the original [3].

Currently, the main applications of logo detection systems are in various security and recognition agencies where they can be used to track or recognize organizations [4]. The encounters in designing a logo detection system include building a dependable data model to represent random logo shapes, font style, the color of the logo, the spelling of the text included in a logo, and finding ways of linking the models with accuracy in real-time. Other challenges include handling rotation and scaled-variations of the original image [5].

The organization of the paper is as follows: segment 2 offers an overview of related work, segment 3 offers an outline on the proposed approach with discussions on overview, feature extraction, and classification schemes, methodology and the flow of the working with implementation strategy and segment 4 offers an overall conclusion.

II. LITERATURE REVIEW

Fake Product Monitoring system improves the efficiency of the testing phase of original or fake products. Most fault prediction models combine well-known methodologies and algorithms such as statistical techniques and machine learning. They require historical data in terms of faulty information as training data to learn which models seem to be a fake product. Based on the knowledge from training data, such tools can estimate fake product modules. A recent study on fault prediction models shows that the AI-based fake product monitoring system can detect 70% of all faults while manual code reviews can detect between 35 to 60% of faults [6].

Existing methods for fake Logo Detection are as follows:

ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 11, Iss 10, 2022

Naïve Bayes Classifier

Naïve Bayes classifier algorithms are used till date for predicting fake or originality. The Bayes Theorem describes the posterior probability as comparative to the prior probability of the class p(Ci), and the likelihood of traits, p(X/Y=Ci) [7]. In binary classification problems such as fault prediction, Naïve Bayes computes the posterior probability of a module being faulty / fake, or the probability of a module being original, given its traits. Then, it allocates a module to the faulty / fake class, if its posterior probability is greater than a pre-defined threshold i.e (0.5). Otherwise, the module is classified as the original.

Naïve Bayes classifier is in use for numerous reasons. It is a broadly used, modest, and strong machine learning technique used in various applications such as pattern recognition, medical diagnosis, and fault prediction. Eg, in fault prediction model with a Naïve Bayes, the classifier delivers the best prediction accuracy on public datasets compared with models with other classifiers. One of the reasons for the success of the Naïve Bayes classifier over other methods is that it combines signals coming from multiple sources.

Logistic Regression

Logistic regression is the most broadly used machine algorithm for classification problems. It is used for binary classification problem which has only two classes to predict i.e if the given input logo is fake or original. Its performance is unmatched in linearly separable class. This is grounded on the probability for a sample to belong to a class. It is reliant on a threshold function to make a decision that is called the Sigmoid or Logistic function [8].

HOG Based Logo Detection

Histogram of Oriented Gradients (HOG) method comprises training and test phases. In the training phase, a Support Vector Machine (SVM) -based logo classifier is trained for a precise class individually. In the test phase, multi-scale sliding window scanning with SVM for a precise class is used individually to detect logos of that class in scene images [9].

Shape Detection

Shape-based technique is used to detect logos in real-world images. The shape features are more outstanding as related to local features like Scale - Invariant feature transform (SIFT) because most object categories are better defined by their shape than texture. As the local features contain a lot of noise, it is always better to go for shape detection [10].

III. PROPOSED SYSTEM

We propose a mobile application for logo detection which can find the difference between a counterfeit product and an original product based on various shapes, text, font, color features using an Artificial Intelligence algorithm.

This proposed system works under 2 phases. The first phase is logo detection using spelling detection and color recognition and the second phase is training the ML model and then detecting fake or original logo with the help of the Feature Extraction method

a. Spelling Detection

A normal logo detector or scanner only captures images and thereby medical products like medicines specifically do not have a logo and so a spelling detector is added as an extra feature within the scanner [11]. The dataset consisting of various medicines with their correct spelling is fed to the machine and the input image or text is captured with the help of a scanner and the mobile application gives the output whether the given input is fake or original. There are two approaches for spelling detection as edit distance measure and hamming distance measure [11].

Edit distance is a method of scheming how unlike

ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 11, Iss 10, 2022

two strings (e.g., words) are to one another by counting the least number of processes obligatory to convert one string into the other [12].

The Minimum Edit distance amongst two strings, str1 (string 1) and str2 (string 2) is demarcated as the least number of insert/delete/substitute processes obligatory to transform str1 into str2.

The Hamming distance amongst two strings of equivalent length is the number of location at which the corresponding symbols are dissimilar [13]. In additional words, it measures the least number of replacements required to alter one string into the other or the least number of faults that could have transformed one string into the other. The symbols could be letters, bits, or decimal digits, among other likelihoods.

Figure 1 illustrates the example of Edit distance [13]

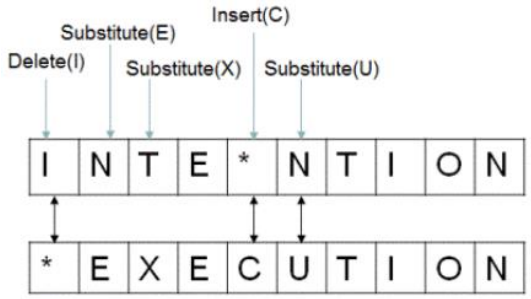

Fig 1: Example of Edit Distance [12]

The Examples of hamming distance have been used in medicine names as follows [14]:

"Nebicard" and "Nevisard" is 2.

"Tozaar-H" and "Tosaar-H" is 1.

"Glycomet" and "Glikomet" is 2.

b. Color Recognition

Color moments are mostly used for color indexing resolutions as features in image retrieval applications in order to compare how alike two images are grounded on color. Basically, one image is compared to a database of digital images with pre-computed features in order to find and retrieve a similar image [15].

For an RGB image, four-color moments are figured for each channel leading to a total of 12 color moments.

These are as explained below:

Mean: The first color moment construed as the average value of an image channel and given by:

$$A_i = \frac{1}{N}\sum_{j=1}^{N} P_{i,j|}$$

Here N is the number of pixels in the image and Pi, is the value of the j-th pixel of the i-th color channel [16].

Standard Deviation: The second color moment computed by taking the square root of the variance of the color distribution [16].

$$S_i = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(P_{i,j} - A_i)^2}$$

Skewness: The third color moment measures how asymmetric the color distribution is and gives information about the shape of the color distribution [16].

$$W_i = \sqrt[3]{\frac{1}{N}\sum_{j=1}^{N}(P_{i,j} - A_i)^3}$$

Kurtosis: The fourth color moment which in addition to the shape of color distribution provides information about how tall or flat the distribution is compared to the normal distribution [16].

$$K_i = \sqrt[4]{\frac{1}{N}\sum_{j=1}^{N}(P_{i,j} - A_i)^4}$$

ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 11, Iss 10, 2022

IV. ALGORITHM

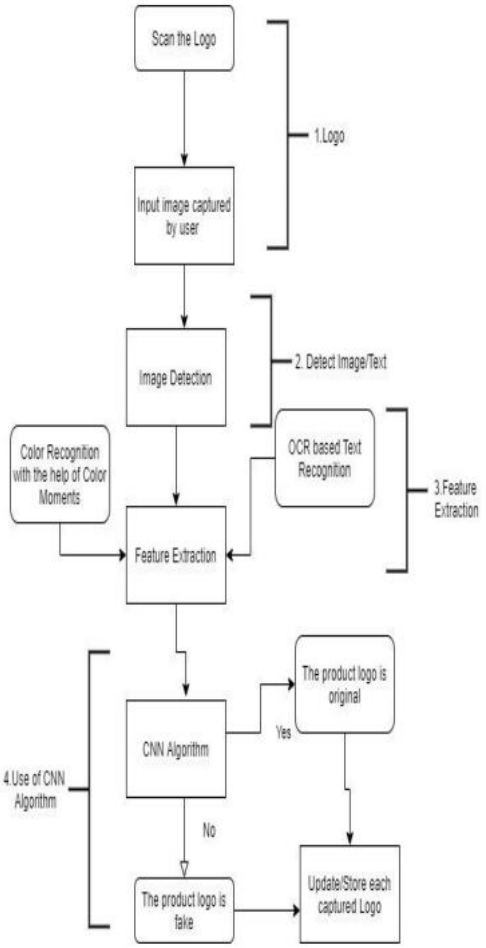

Fig 2: Block diagram of the setup

V. CONCLUSION

There exist many devices for the detection of Counterfeit products, but they just capture the image of the Logo or Barcode. The aim of this research is to propose a way to construct a device that will capture the product logo image and process it with the help of artificial intelligence along with text and color recognition indicating whether a product is fake or original. This application will prove to be portable and easy to use. It will be much helpful for the people who isn't tech-savvy.

REFERENCES

1. Riel, Cees B.m. Van, and Anouschka Van Den Ban. "The Added Value of Corporate Logos - An Empirical Study." European Journal of Marketing, vol. 35, no. 3/4, 2001, pp. 428–440., doi:10.1108/03090560110382093.

2.https://www.inbrief.co.uk/intellectualproperty/intellectual-property-rights/

3. Pathak, Abhishek, et al. "Implicit and Explicit Identification of Counterfeit Brand Logos Based on Logotype Transposition." Journal of Product & Brand Management, vol. 28, no. 6, 2019, pp. 747–757., doi:10.1108/jpbm-06-2018-1921.

4. Sheetala.s.c, U.d. Dixit. "Logo Recognition, Detection and Logo Based Document Image Retrieval: A Review." International Journal of Innovative Research in Science, Engineering and Technology, vol. 4, no. 5, 2015, pp. 2822–2827., doi:10.15680/ijirset.2015.0405024.

5. Li, Zhe, et al. "Fast Logo Detection and Recognition in Document Images." 2010 20th International Conference on Pattern Recognition, 2010, doi:10.1109/icpr.2010.665.

6. Shull, F., Boehm, V.B., Brown, A., Costa, P., Lindvall, M.,Port, D., Rus, I., Tesoriero, R., and Zelkowitz, M. 2002.What We Have Learned About Fighting Defects.In Proceedings of the Eighth International Software Metrics Symposium, 249-258, 2002.

7. Practical Statistics for data scientists by P.Bruce and A.Bruce- Chapter -5 : Classification

8. Dreiseitl, Stephan, and Lucila Ohno-Machado. "Logistic Regression and Artificial Neural Network Classification Models: a Methodology Review." Journal of Biomedical Informatics, vol. 35, no. 5-6, 2002, pp. 352–359., doi:10.1016/s1532-0464(03)00034-0.

9. Li, Kuo-Wei, et al. "Logo Detection with Extendibility and Discrimination." Multimedia

ISSN PRINT 2319 1775 Online 2320 7876
 Journal Volume 11, Iss 10, 2022

Tools and Applications, vol. 72, no. 2, 2013, pp. 1285–1310., doi:10.1007/s11042-013-1449-1
10. Liu R,Wang Y, Baba T, Masumoto D (2010) Shape detection from line drawings with local neighbourhood structure. Pattern Recognition 43(5):1907–1916
11. Hodge, V.J.; Austin, J. (2003). A comparison of standard spell checking algorithms and a novel binary neural approach. IEEE Transactions on Knowledge and Data Engineering, 15(5), 1073–1081. doi:10.1109/TKDE.2003.1232265
12.https://www.ideserve.co.in/learn/img/editDistance_0.gif
13. https://towardsdatascience.com/9-distance-measures-indata-science-918109d069fa
14.https://anandasupermarket.com/pharmacy_product.php
15. Ghosh, Souvik, and Ranjan Parekh. "Automated Color Logo Recognition System Based on Shape and Color Features." International Journal of Computer Applications, vol. 118, no. 12, 2015, pp. 13–20., doi:10.5120/20796-3457.