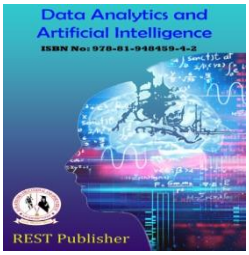
Data Analytics and Artificial Intelligence

Vol: 3(3), 2023

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/



Fake Material Detection Using Artificial Intelligence

* **A. Harishankar, G. Sathish, S. Sheik Hassain Sadiq, S. Udaya Balaji**

Adhiyamaan College of Engineering (Autonomous), Hosur

Krishnagiri,Tamil Nadu, India.

*Corresponding Author Email: haririder.shankar6@gmail.com

Abstract: *The availability of fake products in the market is one of the biggest challenges of online retailing. These products appear to be genuine, but they are imitations of original branded products. Almost 20% of products sold on the Internet are fake. The problem of recognizing counterfeit (fake) products is a tedious task in certain cases and can be dangerous when it comes to medical products (life will be dangerous). It is easier to manufacture and sell fakes if the individual does not properly check the details of the product. This project proposes a better solution using artificial intelligence for non-tech-savvy customers to determine whether a received product is fake or genuine. In this project, the images of the original products are stored in the database (local server), the future scope will be cloud server. If an input image to be analyzed is given, the image is compared to the original product image that is stored in the database. The comparison is made based on the Sift (Scale-invariant feature transform) algorithm and the product is detected as fake or genuine. Automatic image and text recognition and classification of product information enable end customers to accurately and quickly detect counterfeits by comparing them with trained models. The goal of this project is to create an easy-to-use application in which the end user identifies a counterfeit product and contributes to the fight against product piracy.*

Keywords: *Counterfeit Products, Artificial Intelligence, Non-Technical, Databases and Sift Algorithm, image recognition, object detection.*

1. INTRODUCTION

Product detection is a sub problem of object detection that has been extensively researched due to its applications in many fields such as commercial advertising, recommendation retrieval, intelligent transportation systems, and so on. Previously, most traditional product detection methods used hand-crafted features to generate the expression. Where Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. Although we call regression problems, they are best suited for classification. The goal of the SVM algorithm is to find a hyperplane in N-dimensional space that distinctly classifies the data points. It can also handle classification and regression on both linear and non-linear data. A support vector machine works comparatively well when there is an understandable degree of dissociation between classes.

2. OBJECTIVE

- Detect fake material, even if the complex is with a small error
- Product detection technology enables tracking of product variants on products listed on marketplaces and social networks
- It most often uses end-to-end deep learning algorithms in artificial neural networks. It allows product detection in different scales, qualities, colors and even with variations.

3. LITERATURE REVIEW

Encounters in designing a logo detection system include creating a reliable data model representing the random shapes of the logo, the font style and color of the logo, the spelling of the text contained in the logo, and finding ways to mesh the models with accuracy in reality. -time. Other challenges include managing rotation and reducing variations in the original

image. They require historical data in the sense of misinformation as training data to determine which models appear to be a fake product. Based on knowledge from training data, such tools can estimate false product modules. A recent study on bug prediction models shows that an AI-based counterfeit product tracking system can detect 70% of all bugs, while manual code review can detect 35-60% of failures. It is used for a binary classification problem that has only two classes to predict, i.e., whether a given input logo is fake or genuine. Its performance is unrivaled in the linearly separable class. This is based on the probability that a sample will belong to a class. In decision making, which is called the sigmoid or logistic function, depends on the threshold function. A product logo is often recognized as one of the most important elements of a product. Perhaps this is one of the key reasons why the logos of reputable organizations are fake. The counterfeit trade is becoming a growing threat in emerging markets, and many successful products have fallen victim to it. Counterfeit logos not only cost customers who want to get the original product a financial loss, but can also have an adverse effect on the sales of the original. Advice for designing a logo detection system includes creating a reliable data model representing random logo shapes, font style, logo color, spelling of the text contained in the logo, and finding ways to link the models with real-time accuracy. Other challenges include managing rotation and reducing variations in the original image.

4. EXISTING SYSTEM

The database is very large and complex to implement. The algorithm that used SIFT because it has many disadvantages, it is difficult to understand. Which is a long-term high cost and not effective for low-power devices. The RANSAC algorithm has no upper limit on the time required to calculate these parameters. Previously, most traditional product detection methods used hand-crafted features to generate the expression. For example, SIFT key points have been used to perform feature representation and product image recognition. The image of the product is fully visible in the image, is not damaged by noise and is not subject to changes. So, they cannot be applied to a real-world image that can be corrupted by noise

Disadvantages:

- Difficulty understanding the complexity of the algorithm. Bugs can't be fixed that soon and the output takes a long time to implement.
- The cost of this application will be very high and it will not be easy to use

5. PROPOSED SYSTEM

Image Preprocessor, which helps improve the image of the image data, which suppresses unwanted distortion or improves some image properties important for further processing. A segmentation process that helps improve the image detection image. In this process, matching is done by dividing the product image into rows and columns. Once this process is complete, the pairing will be very accurate. The solution has proven to be highly effective and meets the requirements of product detection and recognition in real-world images. Clever edge detection that fully frames the product. A Sobel filter that removes product error.The SVM algorithm is a supervised machine learning algorithm used for both classification and regression.Products can appear anywhere in an image with any orientation and scale, and multiple products can coexist in the same image, different object designs are generated for each image, and these areas are more likely to contain a product.These designs are then cropped to a common size to match the input dimensions and propagated through an SVM specially trained for product recognition.

Advantage:

- We use SVM algorithm for more flexible and accurate output
- The database requires less space, so it will be easy to train the entire algorithm.
- Reliable usage and cost will be very less after implementation.

6. ARCHITECTURE DESIGN

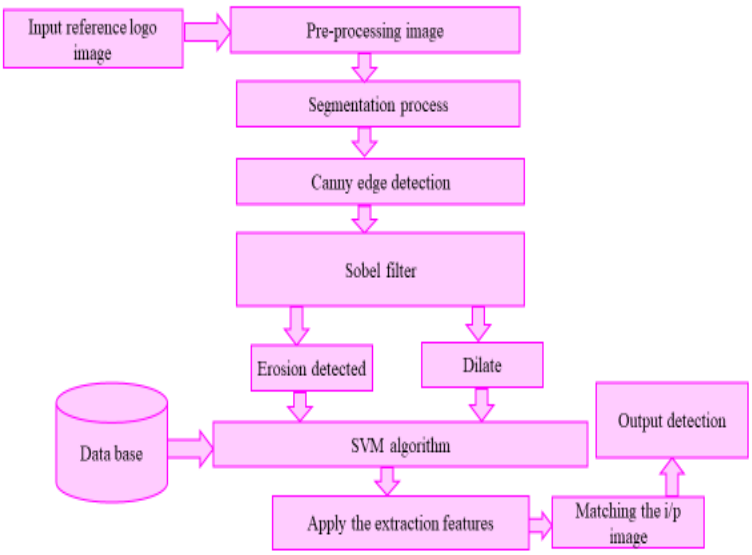


FIGURE 1. Block diagram

7. MODULE

There are 5 modules used in this project

- Input video processing
- Segmentation process
- Canny edge detection
- SIFT encode module
- Data base module

8. MODULE DESCRIPTION

Input video processing:
In this module, the real time camera is used to get input image, is the camera will capture live video streaming and convert the video into image frames by extracting the image from video. It is called input video processing
Segmentation process:
Image segmentation involves converting an image into a collection of pixel regions that are represented by a mask or labeled image. By dividing the image into segments, you can process only the important segments of the image instead of processing the entire image.
Canny edge detection:
Smart edge detection first removes noise from the image by smoothing. It then finds the gradient of the image to highlight areas with high spatial derivatives. The algorithm then follows these areas and suppresses any pixels that are not at their maximum.
SIFT encode module:
Scale-Invariant Feature Transform (SIFT) - SIFT is an algorithm in computer vision for detecting and describing local features in images. It is a feature extraction is widely used in image processing.
Data base module:

They collect information about people, and project information image collected to store in dataset. This information is collected in one place so it can be tracked and analyzed. A database can be thought of as a collection of information and data.

9. SYSTEM FUNCTION

Segmentation process: Image segmentation is a major domain of computer vision supported by a large body of research involving both image processing-based algorithms and learning-based techniques.Along with being one of the most important areas of computer vision, image segmentation is also one of the oldest problem statements researchers have thought about, with the first works involving the primitive region being developed as early as 1970-72.Image segmentation is a subdomain of computer vision and digital image processing that aims to group similar regions or segments of an image under appropriate class labels.Since the entire process is digital, a representation of the analog image in the form of pixels is available, so the task of creating segments is equivalent to the task of grouping pixels.Image segmentation is an extension of image classification the image , wanted and unwanted image to perform a localization. Thus, image segmentation is a superset of image classification with a model that pinpoints where the corresponding object is present using the object's boundary contour.

Figure 2. Segmentation processing

Pre-Processing Image: As a machine learning engineer, data preprocessing or data cleaning is a crucial step, and most ML engineers spend a significant amount of time preprocessing data before building a model. Some examples of data preprocessing include detecting outliers, treating missing values, and removing unwanted or noisy data. Similarly, image preprocessing is a term for image operations at the lowest level of abstraction. These operations do not increase the image information content, but decrease it if entropy is an information measure. The aim of pre-processing is to improve the image data, which suppresses unwanted distortion or improves some image properties important for further processing and analysis.

Database Management: A DBMS manages incoming data, organizes it, and provides ways for users or other programs to modify or extract the data. Some examples of DBMS include MySQL, PostgreSQL, Microsoft Access, SQL Server, FileMaker, Oracle, RDBMS, dBase, Clipper, and FoxPro. Oracle makes software, called database management systems (DBMS), for creating and managing databases. RDBMS is a relational database management system. An Oracle database (or Oracle RDBMS) is a collection of data organized by type, with relationships maintained between different types. Databases are used to store the image and videos, to proper maintain and access to kind of data information store in dataset. They collect information about people, and project information image collected to store in dataset. This information is collected in one place so it can be tracked and analyzed. A database can be thought of as a collection of information and data.

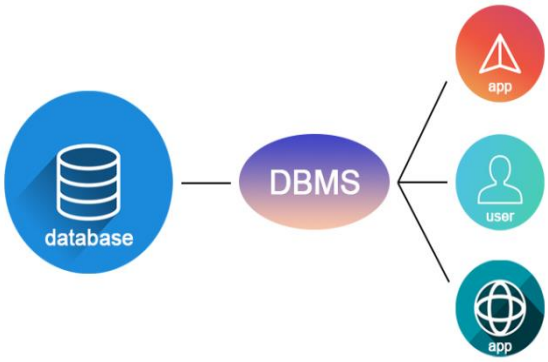


FIGURE 3. Database Management

Sobel Filter: The Sobel operator, sometimes called the Sobel–Feldman operator or the Sobel filter, is used in image processing and computer vision, particularly in edge detection algorithms, where it produces an edge-enhancing image. It is named after Irwin Sobel and Gary Feldman, colleagues at the Stanford Artificial Intelligence Laboratory (SAIL). At each point in the figure, the result of the Sobel-Feldman operatories either the corresponding gradient vector. The Sobel-Feldman operator is based on the convolution of the image using a small, separable, integer filter in the horizontal and vertical directions, and is therefore relatively inexpensive in terms of computation. On the other hand, the gradient approximation it produces is quite rough, especially for high-frequency variations in the image.

Support Vector Machine Algorithm: Support Vector Machine or SVM is one of the most popular learning algorithms used for both classification and regression problems. Easily remove SVM algorithm support. However, it is primarily used for classification in machine learning. The goal of the SVM algorithm is to create a line or decision boundary that can segregate the n-dimensional space into classes so that we can easily place the creation of a new data point with the correct information in the future. The best decision line of this SVM algorithm is called the hyperplane. The SVM algorithm selects the extreme points/vectors that help in creating the hyperplane. These extreme cases are called support vectors, which is why the algorithm is called a support vector machine.

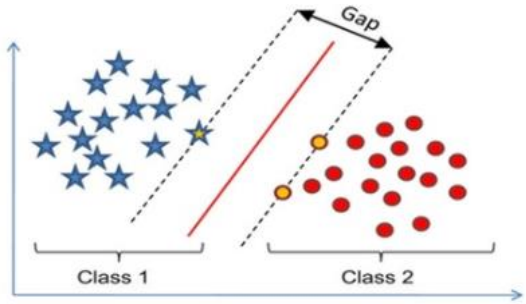


Fig no 4: SVM algorithm

Canny Edge Detection: Canny edge detection uses linear filtering with a Gaussian kernel to smooth the noise and then calculates the edge strength and direction for each pixel in the smoothed image. Candidate edge pixels are identified as those that survive a thinning process called non-maximum suppression. In this process, the edge strength of each candidate edge pixel is set to zero if its edge strength is not greater than the edge strength of two neighboring pixels in the gradient direction. Thresholding is then performed on the thinned edge size image using hysteresis. Two edge strength thresholds are used in hysteresis. All candidate edge pixels below the low threshold are labeled as non-edge, and all pixels above the low threshold that can be connected to any pixel above the high the threshold values in the chain of edge pixels are labeled as edge pixels. The Canny edge detector requires the user to enter three parameters. The first is sigma, the standard deviation of the Gaussian filter specified in pixels. The second parameter low is the low threshold, which is specified as a fraction of the calculated high threshold. The third parameter high is the upper threshold for use in hysteresis and is specified as a percentage point in the distribution of gradient magnitude values for candidate edge pixels.

10.RESULTS

The large volume of counterfeit products available online, the counterfeiting industry is increasing at an exponential rate. As a result, there is a great demand to detect counterfeit products, and AI technology is utilized to detect products fake or real.

Fig no 5: Fake product

Fig no 6: Edge detection

11. 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 project research is to propose a way to construct a device that will capture the product 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]. Li, Zhe, et al. "Fast Logo Detection and Recognition in Document Images." 2020 20th International Conference on Pattern Recognition, 2020.

[2]. Shull, F., Boehm, V.B., Brown, A., Costa, P., Lindvall, M.,Port, D., Rus, I., Tesoriero, R., and Zelkowitz, M. 2021.What We Have Learned About Fighting Defects.In Proceedings of the Eighth International Software Metrics Symposium, 249-258, 2021.
[3]. Dreiseitl, Stephan, and Lucila Ohno-Machado. "Logistic Regression and Artificial Neural Network Classification Models: a Methodology Review", 2022.
[4]. Pathak, Abhishek, et al. "Implicit and Explicit Identification of Counterfeit Brand Logos Based on Logotype Transposition." Journal of Product & Brand Management, 2019.
[5]. Li, Zhe, et al**. "**Fast Logo Detection and Recognition in Document Images**."** 2020.
[6]. EkaDyarWahyuni and ArifDjunaidy, "Fake Review Detection from a Product Review Using Modified Method of Iterative Computation Product", January 2019 Research Gate.
[7]. M. C. Jayaprasanna, V. A. Soundharya, M. Suhana and S. Sujatha, "A Block Chain based Management System for Detecting Counterfeit Product in Supply Chain," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks.
[8]. Singh, Shivam& Choudhary, Gaurav & Kumar, Shishir &Sihag, Vikas & Choudhary, Arjun. (2021). Counterfeited Product Identification in a Supply Chain using Blockchain Technology.
[9]. 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, 2020.
[10]. Hodge, V.J.; Austin, J. (2003). A comparison of standard spell-checking algorithms and a novel binary neural approach, 2019.
[11]. Daoud, Eduard, et al. "Improving Fake Product Detection Using Ai-Based Technology." 2021.
[12]. Cheng, Gong, et al. "When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification via Learning Discriminative CNNs." 2020.
[13]. Karaoglu, Sezer, et al. "Con-Text: Text Detection for Fine-Grained Object Classification." IEEE Transactions on Image Processing, vol. 26, no. 8, 2019.
[14]. 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, 2019.
[15]. Dr. T Ramashri, et al. "Content Based Image Retrieval Using Color Moments and Texture", 2021.
[16]. Daoud, E. &Gaedke, M., 2019. Decentralizing Products Certificates Using Blockchain, 2021.

