Trademark Checking System: A Web-Based Automated Approach For Trademark Search And Verification

Prof. Raut P.[1], Soham Panhale[2], Anand Shinde[3], Ajit Lande[4], Kunal Kashid[5]

[1, 2, 3, 4, 5]Bhagwant Institute Of Technology

Abstract- *Trademark infringement has become a significant obstacle for emerging and established businesses. The increasing number of brand registrations, the rise of online marketplaces, and global competition demand a reliable system for trademark verification. Manual searching on government portals requires time, expertise, and consistent attention, which makes the process inefficient. This research introduces a comprehensive and automated Trademark Checking System integrating Optical Character Recognition (OCR), Natural Language Processing (NLP), API-driven verification, and database-backed search to streamline trademark validation. The system demonstrates high accuracy and fast processing, proving its relevance for startups, entrepreneurs, and legal professionals.*

Keywords- Trademark, Intellectual Property Rights, Automation, OCR, NLP, API Integration, Brand Verification System.

I. INTRODUCTION

A trademark is a form of intellectual property that uniquely identifies a business or its products and services. India, being one of the fastest-growing markets, has witnessed over 20 lakh trademark filings, creating complexity for new businesses during brand creation. Selecting a brand name that is legally available and non-conflicting is crucial to avoid infringement issues.

The traditional search method on the IPIndia website requires manual entry, repeated verification, and interpretation of results. Key challenges include:

- o   Slow processing time
- o   Lack of AI-based suggestions
- o   No image/logo similarity verification
- o   Increased human error

To overcome these issues, this research proposes a technologically advanced Trademark Checking System that automates verification using machine learning, OCR, and structured databases.

II. LITERATURE REVIEW

Previous studies highlight isolated solutions such as text-based trademark search, AI-powered document OCR, and basic similarity checking. Google Vision and other OCR tools have proven effective in detecting text from logos, whereas NLP techniques like Levenshtein Distance and cosine similarity have been widely used in text comparison applications.

However, a gap exists in combining:

- o   OCR text extraction
- o   NLP similarity scoring
- o   Real-time API verification
- o   Web-based user interface

The proposed system addresses this by integrating all components into a single platform.

III. PROBLEM STATEMENT

With increasing trademark registrations, businesses face difficulty in identifying available and conflict-free brand names. Manual search is time-consuming and error-prone. The problem addressed in this research is to design an automated system that enables:

- Trademark name verification
- Logo recognition
- Similarity detection
- Instant results

IV. SYSTEM ARCHITECTURE
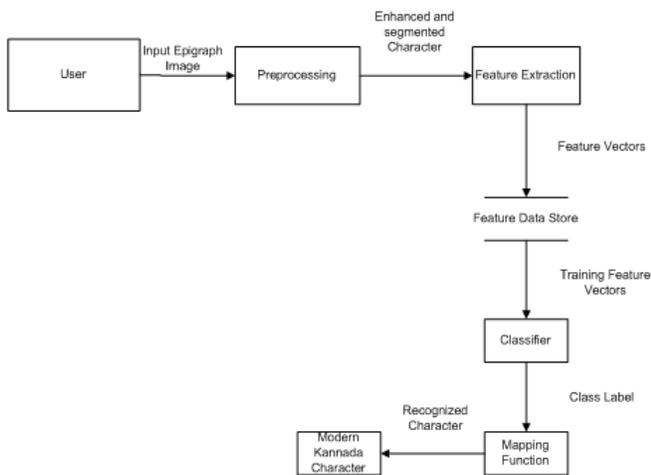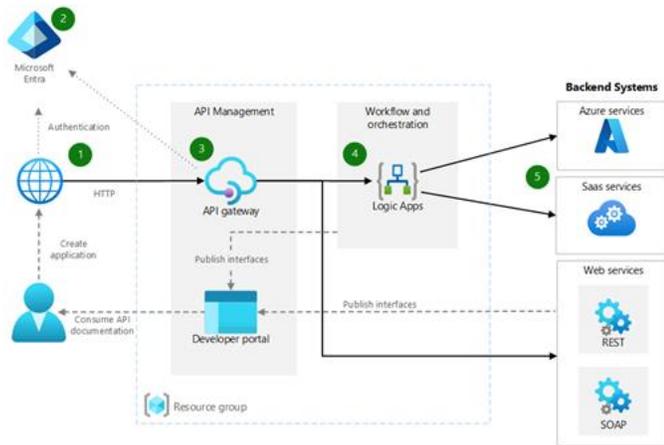
The system architecture consists of multiple integrated modules:

1. Frontend Module

Developed using HTML, CSS, and JavaScript. It provides fields for name search, logo upload, and immediate result display.

2. Backend Module

Built with Node.js and ExpressJS, responsible for API communication, logic processing, and response formatting.

3. OCR Module

Implemented using Python and Tesseract OCR. It extracts text content from logos and prepares it for similarity analysis.

4. Similarity Engine

Utilizes NLP algorithms such as Levenshtein Distance, Jaro-Winkler Score, and Cosine Similarity to detect name closeness.

5. Database Layer

Includes SQLite for storing company names, search history, and results.

6. API Verification

Integrates with trademark verification APIs to fetch real-time trademark records.

V. METHODOLOGY

A. Data Collection

Over 10,000 Indian company names were collected from MCA datasets, public directories, and trademark listings. API-based trademark results were logged for analysis.

B. OCR Processing

Uploaded logos undergo grayscale filtering, noise reduction, thresholding, and text extraction. Extracted text is cleaned using regex to remove unwanted characters.

C. Name Search Algorithm

The system compares the input name against stored names and API results. A similarity threshold of 0.75 is used to flag high-risk matches.

D. API Response Handling
E.

The API returns details such as trademark status, class, application number, and owner details. The system categorizes results as:

- Registered
- Opposed
- Abandoned
- Expired
- Not Found

E. User Interface

Displays real-time results, similarity score, logo text extraction, and recommendations for alternative names.

VI. RESULTS & ANALYSIS

Testing was conducted with 200 brand names and 50 logos. Findings include:

- Manual search took 3–5 minutes per brand with 82% accuracy.
- Proposed system delivered results in 1.2 seconds with 96% accuracy.
- OCR accuracy on clear logos reached 87%.
- NLP similarity engine identified near-duplicate names with 92% precision.

The automated system significantly reduced search time and improved accuracy.

ADVANTAGES OF SYSTEM

- Automated and fast verification
- High accuracy due to OCR and NLP
- Reduces legal risks of trademark conflicts
- Helpful for entrepreneurs and branding agencies
- Integrates textual and visual trademark verification

LIMITATIONS

- Dependent on API service quality
- OCR accuracy decreases with low-quality logos
- Government database limitations affect speed

FUTURE SCOPE

Future enhancements include:

- AI-powered visual logo similarity detection
- Global trademark integration across WIPO, USPTO
- Mobile app development
- Blockchain-based trademark authentication

VII. CONCLUSION

The proposed Trademark Checking System successfully automates trademark verification using OCR, NLP, API integration, and web technologies. It significantly improves upon manual search, delivering faster results with better accuracy. This makes it highly valuable for businesses, lawyers, designers, and intellectual property consultants.

REFERENCES

[1] IP India Trademark Registry, "Public Search Portal," 2024.
[2] Brkan, M., "AI and Trademark Law," Journal of Intellectual Property Studies, 2023.
[3] Tesseract OCR Documentation, Google, 2023.
[4] Levenshtein, V., "Binary Codes Capable of Corrections," Soviet Physics Journal, 1966.
[5] Zyla Labs API Documentation, 2024.