K. J. Somaiya School of Engineering, Mumbai - 400 077
(A Constituent College of Somaiya Vidyavihar University)
Dept. of INFORMATION TECHNOLOGY
L.Y. B. Tech. Semester - VII (2025 - 26)
LY Project Poster

Automated Description Generation for Jewellery Images using Deep Learning

Introduction

Jewellery businesses depend on accurate tagging for inventory and search, but the current process is slow and often inaccurate. Existing systems can detect jewellery but fail to capture important details such as metal colour or gemstone type. Because of this limited understanding, maintaining a consistent and searchable digital catalogue becomes difficult.

Problem Statement

Develop an automated system that detects necklaces and earrings from a person's image, generates captions describing their metal colour, gemstones and classifies them based on presence of gemstones, metal colour, etc.

Scope

Functional

- Accept **input image** of necklace or earrings
- Detect and **identify necklaces and earrings** in the image
- Generate captions detailing **metal colour and gemstones** if present (based on the color)
- Output **structured descriptions and classifications** for each detected item
- Provide a **user interface** for image upload and result display

Non - Functional

Accuracy
94.62% - Jewellery Classification (Earring or Necklace)
90.44% - Description Generation

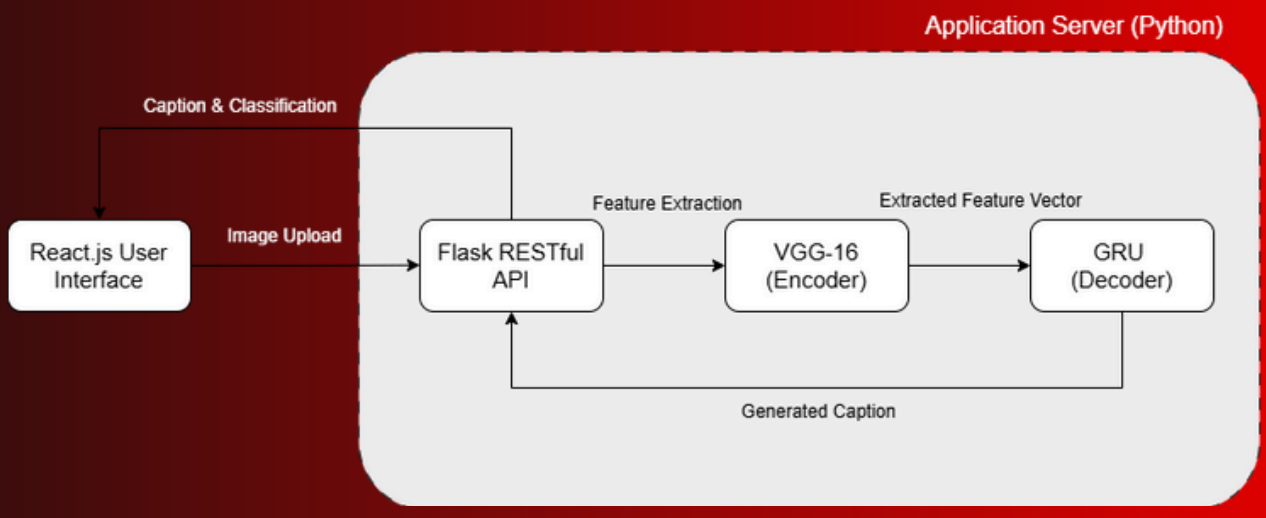
Response-Time
1-2 seconds

Usability

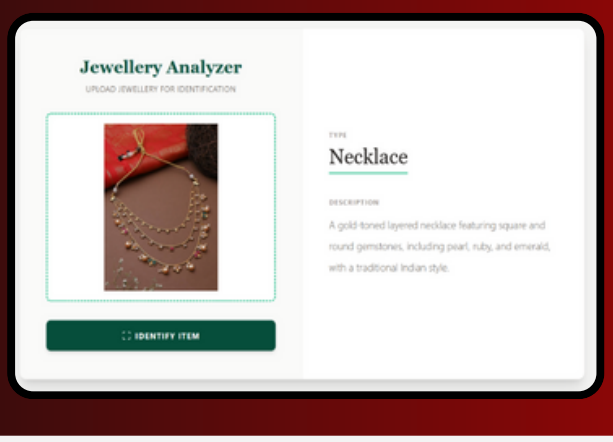
References

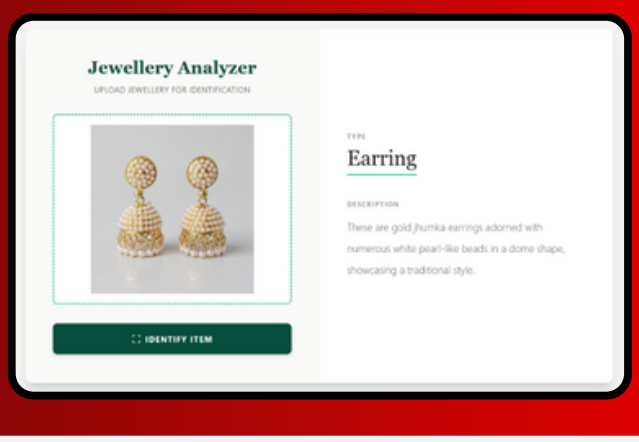
1. S. S. A. Zaidi et al., "A Survey of Modern Deep Learning based Object Detection Models," arXiv, 2021.
2. J. M. Alcalde-Llergo et al., "Jewellery Recognition via Encoder-Decoder Models," arXiv, 2024.
3. P. Khalilian et al., "Jewellery rock discrimination... using... deep learning," Sci. Rep., vol. 14, 2024.
4. J. M. Alcalde-Llergo et al., "Automatic Identification and Description of Jewellery...," Appl. Sci., vol. 15, 2025.

Methodology



- The system uses a **client–server** architecture combining a web interface with deep learning models.
- Users upload jewellery images through the **React.js** UI.
- The images are sent to the backend via a **Flask REST API**.
- Flask connects the UI with the ML pipeline.
- The **VGG-16 encoder** extracts visual features like shape, texture, and jewellery patterns.
- The **GRU decoder** uses these features to generate jewellery descriptions (metal type, colour, gemstones).
- The backend returns the final classification and caption to the frontend for display.





Results and Discussions

Classification Results

Metric	Value
Accuracy	94.62%
Precision	94.62%
Recall	94.62%
F1-Score	94.58%

Description Results

Metric	Value
Jewellery Type Accuracy	0.8113
Strict Pass Rate	0.7925
Semantic Accuracy (BLEU-4)	0.407
Max Validation Accuracy (word-level)	~91%
Total Images Tested	53

Conclusion

The project delivers a fast, accurate VGG-16 + GRU system that identifies jewellery and generates captions through a smooth React–Flask pipeline.

Group - 4
Guide : Prof. Avani Sakhapara
16010422200 : Tanisha Mangaonkar
16010422233 : Prachi Gandhi
16010422234 : Chandana Galgali
16010422235 : Mahek Thakkar