GLOBAL ACADEMY OF TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Subject Name	**Project Work Phase II**	**Subject Code**	**17CSP85**
Student Name	**C.P YASHWANTH**	**USN**	**1GA17CS035**
	MADDALI SOWMYA		**1GA17CS080**
	RAKSHITHA MURTHY		**1GA17CS118**
	SWARAJ PARIDA		**1GA17CS163**
Domain	**DEEP LEARNING**	**Group No:**	**11**
Project Title	**“LANDMARK RECOGNITION USING CONVOLUTIONAL NEURAL NETWORKS”**		
Under taken at	**GLOBAL ACADEMY OF TECHNOLOGY**		
Guide Name	**Prof. KAMLESHWAR KUMAR YADAV**		

Agenda

- Introduction
- Problem Statement
- Objectives
- Existing System
- Proposed System
- System Design
- High Level Design
- Implementation Modules
- Conclusion
- References

Introduction

- Computer Vision (CV) is a field of Machine Learning that deals with how computers can gain high-level understanding from digital images or videos.
- Some of its tasks include methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information.
- In the field of Artificial Intelligence (AI) and Computer Vision recognition of objects has become very common, feasible and realistic.

Introduction

- Looking ahead, there will come a time where instance-specific recognition will become a trend and be an everyday problem.
- Artificial Intelligence especially Convolutional Neural Networks (CNN) concept can be used to ease up the life of others.
- CNN is a class of Deep Neural Networks (DNN), most often used to analyze visual imagery.

Problem Statement

- Landmark recognition on Google landmark dataset using various algorithms. The goal is to efficiently recognize objects in an image at an instance level, just not at the base level.

Objectives

- To make a model which recognizes landmarks from an image using different algorithms such as Visual Geometry Group (VGG) and Deep Local Feature (DeLF)
- To improve our model such that it performs better than the present primitive models/techniques.

Existing System

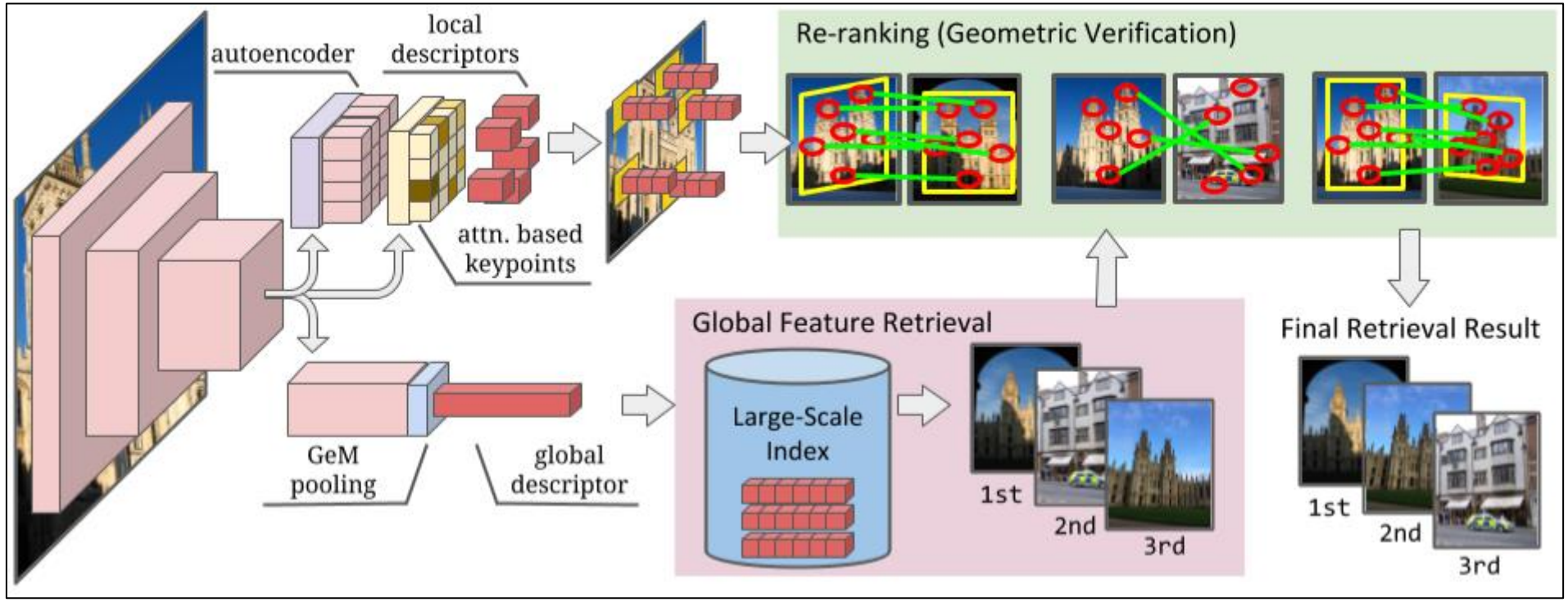

Figure 1: DeLF Architecture

Existing System

DeLF architecture as shown in figure 1 can be decomposed into four main blocks :

1. Dense localized feature extraction
2. Keypoint selection
3. Dimensionality reduction and
4. Indexing and retrieval

Proposed System

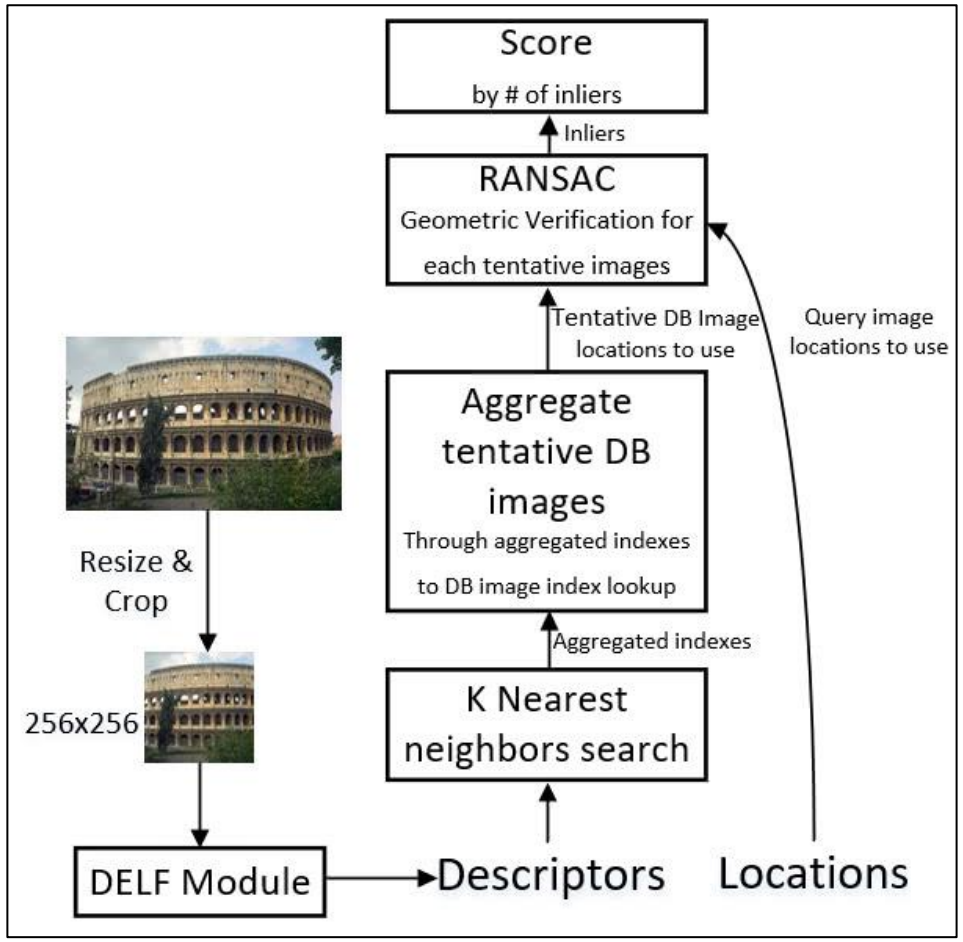

Figure 2: Proposed System

System Architecture

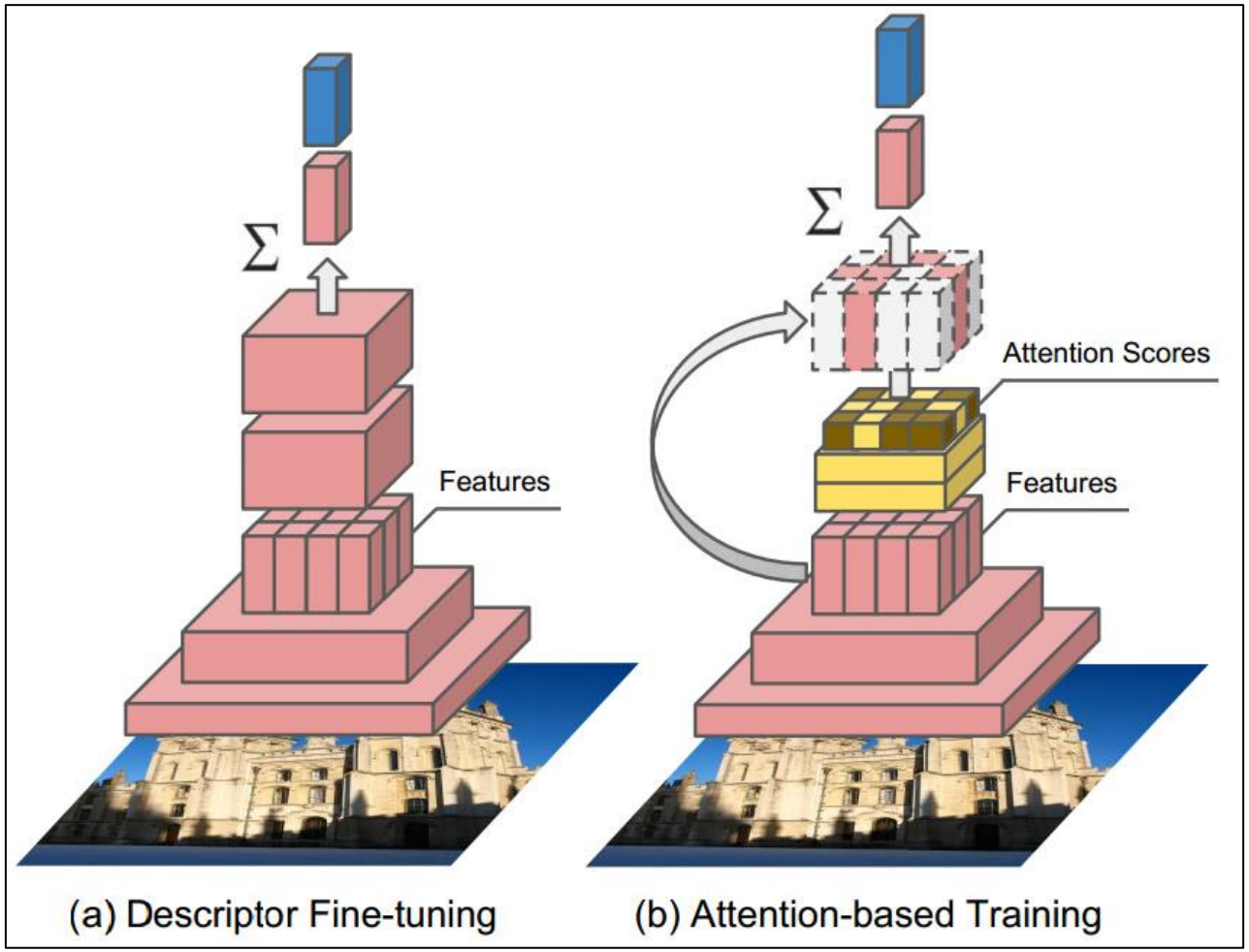

Figure 3: System Architecture

High Level Design

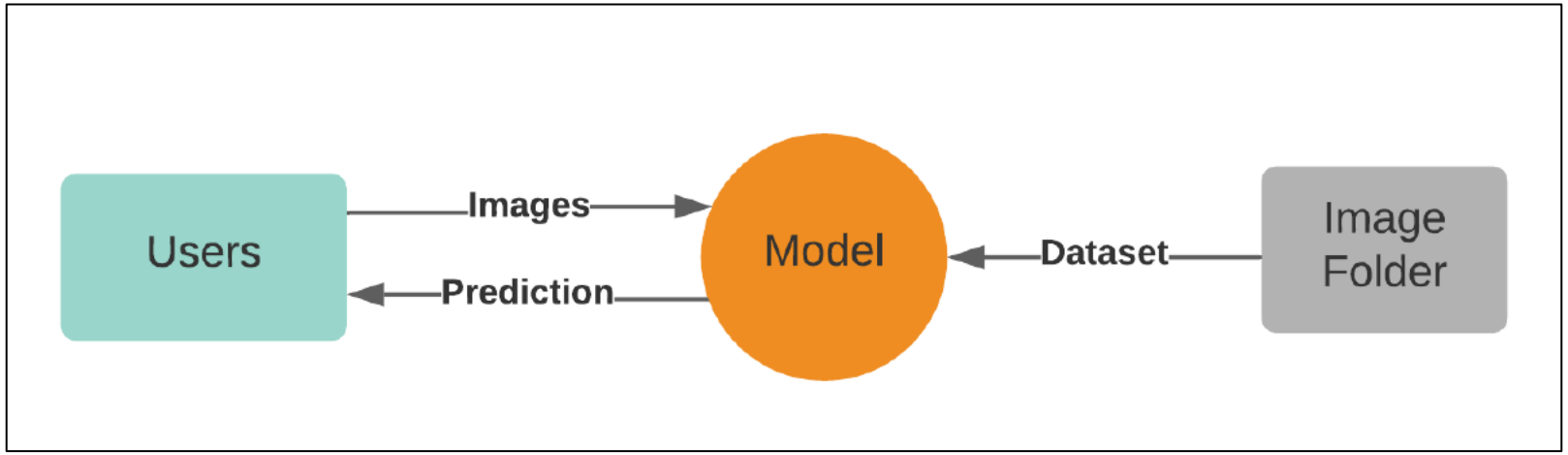

Figure 4: Data Flow Diagram Level 0

In figure 4 -

- User provides the query image which is fed to the model.
- The model compares the query image with the already existing images in the Image Folder.
- It performs necessary functions and outputs a prediction back to the user.

High Level Design

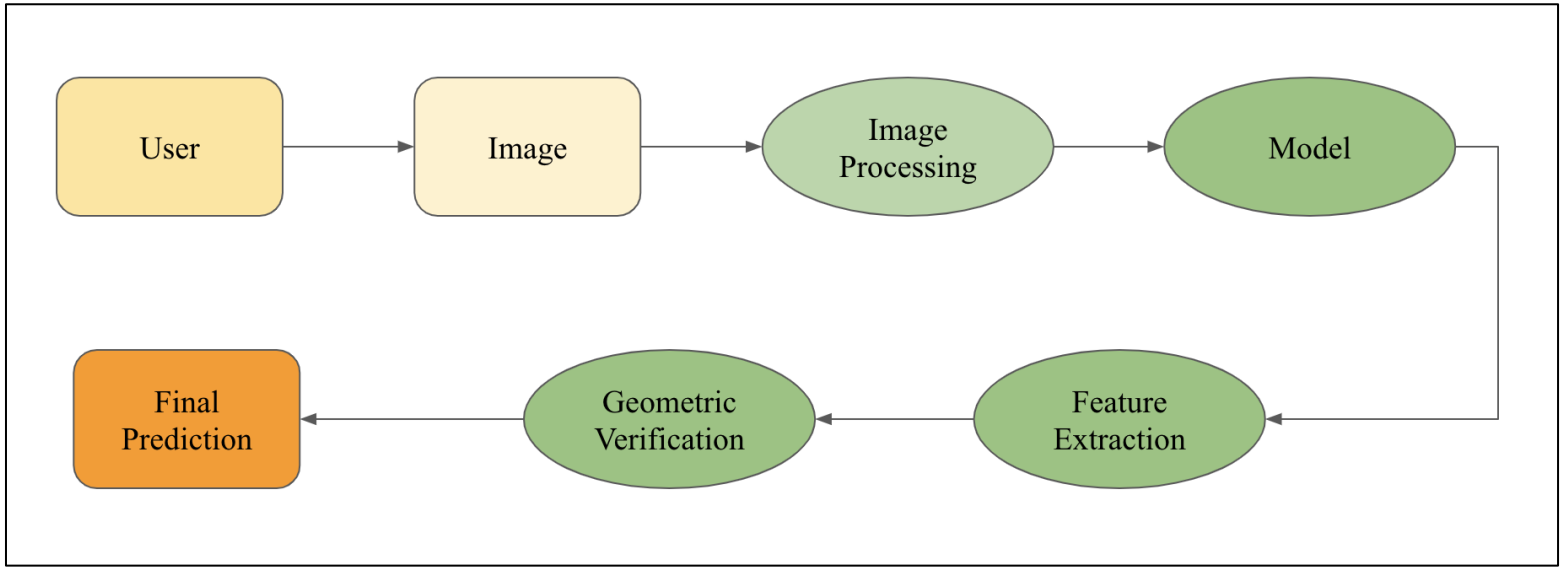

Figure 5: Data Flow Diagram Level 1

In figure 5 -

- The user provides the query image, it is re-sized and converted into NumPy arrays.
- The model obtains the location and feature vectors from the NumPy arrays.
- The feature vectors are used to verify the query image with the database image.
- Finally the geometric verification retrieves the most similar image.

High Level Design

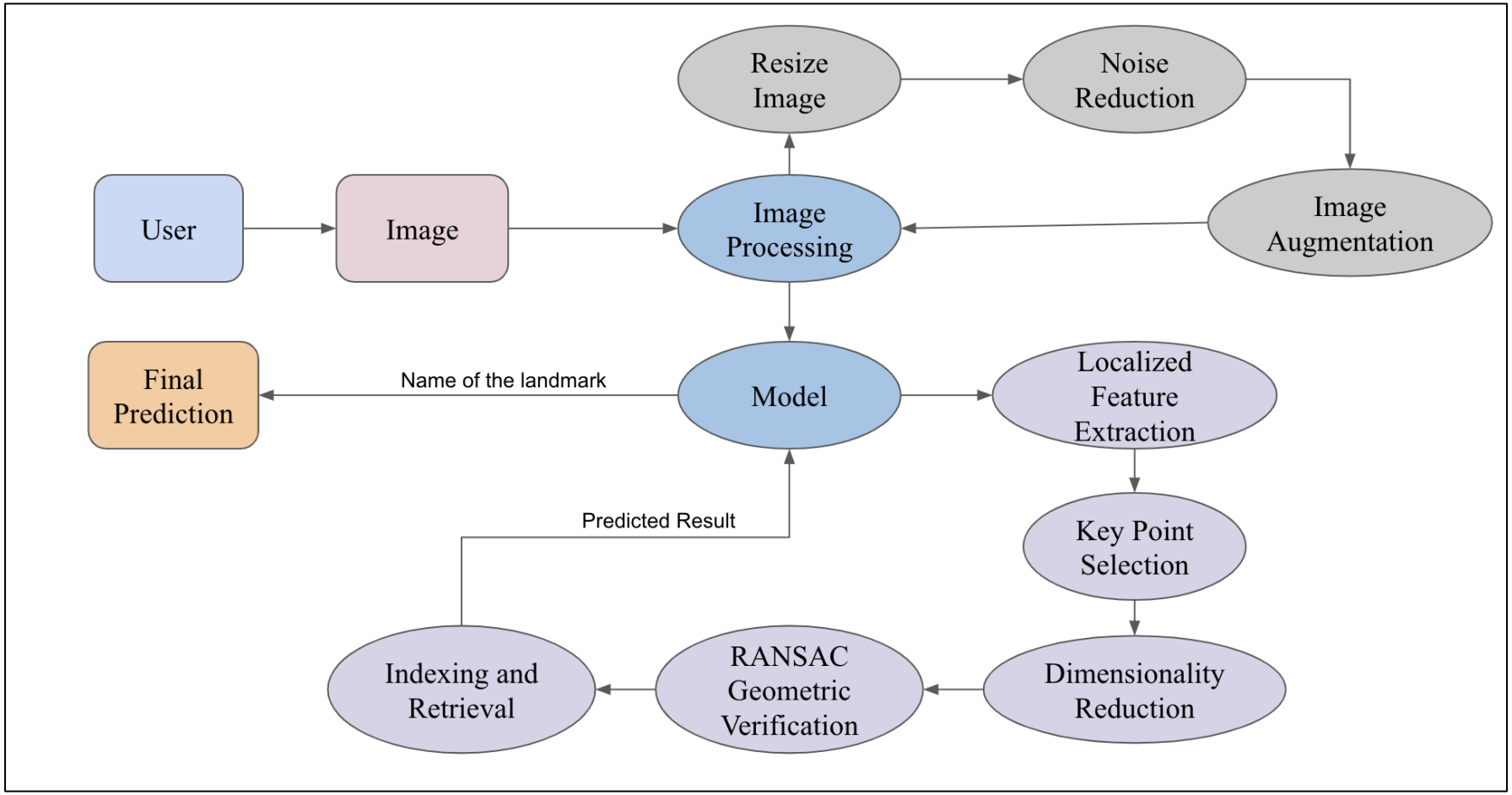

Figure 6: Data Flow Diagram Level 2

High Level Design

In figure 6 -

- Image Processing
  - Resize Image
  - Noise Reduction
  - Image Augmentation
- Building a Model
  - Localized Feature Extraction
  - Dimensionality Reduction
  - RANSAC Geometric Verification

Implementation Modules

1. Image Acquisition
2. Image Pre-processing
3. Feature Extraction
4. Geometric Verification
5. Final Predictions

Image Acquisition

- Functionality:- Fetching images from the given Universal Resource Locators (URLs)
- Input:- Universal Resource Locators (URLs)
- Output:- Customized images

Image Pre-processing

- Functionality:- Re-sizing images and converting images into NumPy arrays
- Input:- Images of different sizes
- Output:- Uniform sized images

Feature Extraction

- Functionality:- Obtaining the location and feature vectors
- Input:- NumPy array
- Output :- Array of location and feature vectors

Geometric Verification

- Functionality:- Verifying query image with database image
- Input:- Query image and database images
- Output:- Number of inliers among the matched images

Final Prediction

- Functionality:- To retrieve the most similar image
- Input:- Number of inliers
- Output:- Image with the highest number of inliers

Conclusion

- Image Processing is a challenging task.
- There is always a trade-off between scalability and accuracy.
- Hence, this technology can be a solution to predict landmark labels directly from image pixels, to help people better understand and organize their photo collections.

References

[1] André Araujo, Bingyi Cao, Jack Sim, Tobias Weyand. (2020). *Google Landmarks Dataset v2 - A Large-Scale Benchmark for Instance-Level Recognition and Retrieval*. IEEE. pp. 2575-2584. DOI: 10.1109/CVPR42600.2020.00265

https://ieeexplore.ieee.org/document/9157053

[2] Christof Henkel, Philipp Singer. (2020). *Supporting Large-scale Image Recognition with out-of-domain Samples*. eprint arXiv:2010.01650

https://arxiv.org/abs/2010.01650

[3] Andre Araujo, Bohyung Han, Hyeonwoo Noh, Jack Sim, Tobias Weyand. (2017). *A Large-Scale Image Retrieval With Attentative Deep Local Features.* IEEE. pp. 3456-3465. DOI: 10.1109/ICCV.2017.374

https://ieeexplore.ieee.org/document/8237636

[4] Jeff Donahue, Jitendra Malik, Ross Girshick, Trevor Darrell. (2014). *Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation tech report (v5).* IEEE

https://openaccess.thecvf.com/content_cvpr_2014/html/Girshick_Rich_Feature_Hierarchies_2014_CVPR_paper.html

- [5] Ross Girshick. (2015). *Fast R-CNN*. IEEE. pp. 1440-1448*. DOI: 10.1109/ICCV.2015.169*

https://ieeexplore.ieee.org/document/7410526

[6] SeungKee Jeon. (2020). *1st Place Solution to Google Landmark Retrieval.* eprintarXiv:2009.05132

https://arxiv.org/abs/2009.05132#:~:text=The%20solution%20is%20based%20on,enhance%20the%20model's%20performance%20further.

Thank You

Q & A