

CHAPTER I

INTRODUCTION

This chapter discusses the background, research problem formulation, research objectives, research benefits, and research limitations. The background explains the urgency of the researched topic related to building asset management in Bojonegoro Regency, which requires an approach based on deep learning technology and Google Street View imagery. The problem formulation outlines the main issues to be discussed, including the implementation of ResNet50 as a feature extractor, the application of cosine similarity for image similarity measurement, redundant image elimination, and the integration of Google Maps API for building location address identification. The research objectives state the outcomes to be achieved in accordance with the established problem formulation. The research benefits explain the contributions of the research both academically and practically for the research object. The research limitations establish the boundaries of the research scope to keep the discussion focused on the researched topic.

1.1 Background

Population growth and urban area development have an impact on the increasing need for infrastructure, asset management, and regional spatial planning [1]. The development of artificial intelligence and deep learning technology in recent decades has opened new opportunities for automating spatial data management more efficiently [2]. Bojonegoro Regency, as one of the regions in East Java Province, has also experienced population growth that has led to an increase in the demand for housing and other supporting facilities. Based on data from the Central Bureau of Statistics (BPS), the population of Bojonegoro Regency in 2020 reached 1,344,038 people, an increase compared to the previous year [3]. In the context of continuously developing regional management, the ability of computer systems to understand and analyze visual imagery becomes an important component in supporting data-driven decision-making [4]. Pattern recognition and machine learning-based approaches enable visual data analysis to be conducted more systematically and measurably [5]. This condition makes the need for a more accurate and integrated building data management system increasingly important.

Along with the increasing economic activity and regional growth, local governments require spatial and visual building data capable of supporting effective asset management. However, the available building data is generally still limited to administrative information without adequate visual documentation. This limitation causes the process of building identification and monitoring of building function changes to be less than optimal. Manual field surveys are still frequently used to obtain visual building data, but this method requires considerable cost, time, and labor. One alternative that can be utilized to obtain visual building data efficiently is street view imagery. Google Street View was first introduced as a service capable of recording panoramic street views using cameras mounted on vehicles, thereby providing large-scale visual data automatically. Street view imagery provides street-level panoramic views capable of displaying building facade details more clearly compared to satellite imagery. The utilization of street view imagery has also been proven effective in research on urban environment perception analysis, including for assessing the visual condition of residential areas. The utilization of Google Street View has also been widely used in computer vision and urban analysis research because it can provide large-scale visual data with integrated spatial information [6][7].

Nevertheless, the use of street view imagery has challenges in the form of high image redundancy due to image capture at adjacent locations. This condition causes many images to have nearly identical visual appearances, thereby increasing data storage requirements and computational load. Digital image processing, which includes the processes of acquisition, representation, and image analysis, is an important foundation in addressing this problem [8]. Therefore, a method capable of measuring the degree of similarity between images is needed to select the most representative imagery. In the field of computer vision, image similarity measurement can be performed using the cosine similarity method, which compares the proximity between feature vectors in multidimensional space [9]. The application of deep learning methods in image similarity measurement has shown better results compared to traditional approaches, primarily due to their ability to extract features hierarchically and automatically [10]. In order for visual features to be well represented, this research uses the ResNet50 architecture as a feature extractor. Various deep convolutional neural network architectures have developed

rapidly and have been proven capable of producing strong feature representations for various visual pattern recognition tasks [11]. ResNet was introduced as a deep learning architecture capable of overcoming the vanishing gradient problem in deep neural networks through the residual learning mechanism [12]. This capability makes ResNet50 widely used in image feature extraction because it can produce more discriminative visual representations.

After the redundant image elimination process is performed, the remaining images will be integrated with Google Maps API to obtain address information based on geographic coordinates [13]. Evaluation of the system results is conducted using appropriate evaluation metrics to measure the effectiveness of the redundant image elimination process [14]. The implementation of a deep learning-based system such as this requires an in-depth understanding of model architecture, transfer learning techniques, and appropriate training strategies so that the system can work optimally on real-world data [15]. The integration between image processing and spatial data is expected to produce a more efficient, representative, and structured building dataset to support regional asset management in Bojonegoro Regency.

1.2 Problem Statement

Based on the problems presented in the background section, the problem formulation can be composed as follows:

1. How to implement ResNet50 as a feature extractor on Google Street View imagery?
2. How to apply cosine similarity to measure the degree of similarity between building images?
3. How to eliminate redundant images based on the results of similarity measurement?
4. How to integrate Google Maps API to obtain addresses based on geographic coordinates?

1.3 Research Objectives

Based on the problem formulation above, the objectives of this research are as follows:

1. To implement the ResNet50 method as a feature extractor on Google Street View imagery.
2. To apply cosine similarity to measure the degree of similarity between building images.
3. To eliminate redundant images based on the results of similarity measurement.
4. To integrate Google Maps API to obtain address information based on the geographic coordinates of images.

1.4 Research Benefits

Based on the research objectives mentioned above, the results of this research are expected to provide the following benefits:

1. For Academics

- a. This research can serve as a reference for students, lecturers, and researchers in the fields of informatics engineering, information technology, and computer engineering, particularly those interested in the development of deep learning models and computer vision applications in the context of spatial data.
- b. The results of this research can broaden the understanding of the implementation of the ResNet50 algorithm in detecting image similarity, as well as how the method is adapted for the case of redundant image elimination from street view data.
- c. To add literature and academic insight regarding the application of Google Maps API as a location data retrieval tool integrated with an image processing system, so that this research can serve as a reference for the development of similar projects in the future.

2. For the Research Object

- a. To provide hands-on experience in implementing a deep learning-based system for managing building image data in a specific administrative region, particularly in Bojonegoro Regency.

- b. To improve efficiency in building data updating and regional asset management because the image collection process can be performed automatically through the system, without having to conduct manual field surveys.
- c. To reduce computational load and data storage by eliminating images with a high degree of similarity, so that the stored data becomes lighter, more organized, and more efficiently used.
- d. To facilitate the process of mapping building addresses and coordinates through the utilization of Google Maps API, which is capable of providing accurate geographic information based on the results of street view image identification.
- e. To provide a foundation for the development of visual and spatial data-based applications in the future, both for spatial planning purposes, business building data collection, and digital technology-based regional monitoring programs.

1.5 Scope and Limitations

The research limitations in this study are formulated to clarify the scope of discussion and to avoid the expansion of issues that are not relevant to the research objectives. The research limitations in this study are as follows:

1. The system developed uses the Python programming language, with the utilization of supporting libraries such as TensorFlow, Keras, and OpenCV for image processing,
2. The dataset used originates from Google Street View imagery obtained from certain areas in Bojonegoro Regency, East Java. This area was selected because it has a sufficiently diverse variation of residential and commercial buildings,
3. Image analysis focuses on measuring the degree of similarity between images,
4. The deep learning architecture used is ResNet50,