Chapter 5
Multi-Scene Recognition in Single Aerial Images Using CNN

G. Ananthi
https://orcid.org/0000-0003-3227-2134
Mepco Schlenk Engineering College, India

S. Mehala Shevani
Mepco Schlenk Engineering College, India

P. Priyadharshini Devi
Mepco Schlenk Engineering College, India

ABSTRACT

Scene recognition from aerial images is a primary application in utilizing high resolution satellite images. In the recent past, many studies were carried out in this field to classify an image to one scene category. But the real scenarios are different, they often contain more than one scene. Therefore, we explore an approach to recognize multiple scenes from a single aerial image. To do this, three different datasets namely UCM dataset, AID dataset and MAI dataset consisting of 2100, 10000, 3923 images respectively were used. Out of the three datasets, UCM and AID dataset

DOI: 10.4018/979-8-3693-7525-9.ch005

are Single scene datasets and the MAI dataset is a multi-scene dataset. These datasets are utilized in two configurations namely UCM2MAI and AID2MAI. Prototype based memory network using different CNN baseline models is used as the backbone. First, the single scene datasets were used to extensively train different baseline models and store them as prototypes in the external memory. Experiments were conducted to analyse the performance of different baseline models.

INTRODUCTION

In recent years, the scopes and applications of remote sensing techniques have been phenomenal, ranging from urban cartography to traffic monitoring, terrain surface analysis to ecological scrutiny, and determining the damage to the landscape caused by natural calamities such as earthquakes, tsunamis, hurricanes, and so on. These are some of the notable developments aided by aerial image classification. Because of the vast and vital scopes of aerial image classification, it has become a hot topic among researchers.

To begin, aerial image classification refers to providing images with labels associated with the scene visible on the images based on a single semantic category. Because satellites capture aerial images from great distances, they contain not only a single scene but also many objects in the landscape. This macroscopic view of single object classification may sometimes result in ignoring the major issue to be addressed by rendering the model infeasible in providing a comprehensive view of scenes in images for mission-critical applications.

A large number of algorithms have been proposed in the literature to identify each and every pixel of an image, localize objects with bounding boxes, and recognize the ground truth. However, these methods often require precise annotations, high computational power, and access to high-quality datasets—resources that are not always readily available. Furthermore, acquiring and curating relevant remote sensing datasets can be expensive and time-consuming due to the need for manual annotation and domain expertise.