

PAPER • OPEN ACCESS

Design and Development of Scene Recognition and Classification Model Based on Human Pre-attentive Visual Attention

To cite this article: A R Kudus and C S Teh 2021 *J. Phys.: Conf. Ser.* **1755** 012012

View the [article online](#) for updates and enhancements.



240th ECS Meeting ORLANDO, FL

Orange County Convention Center **Oct 10-14, 2021**

Abstract submission deadline extended: April 23rd

SUBMIT NOW

Design and Development of Scene Recognition and Classification Model Based on Human Pre-attentive Visual Attention

A R Kudus and C S Teh

Faculty of Cognitive Sciences and Human Development, University Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

E-mail: 16020136@siswa.unimas.my

Abstract. Recent works on scene classification still utilize the advantages of generic feature of Convolutional Neural Network while applying object-ontology technique that generates limited amount of object regions. Human can successfully recognize and classify scene effortlessly within short period of time. By utilizing this idea, we present a novel approach of scene classification model that built based on human pre-attentive visual attention. We firstly utilize saliency model to generate a set of high-quality regions that potentially contain salient objects. Then we apply a pre-trained Convolutional Neural Network model on these regions to extract deep features. Extracted features of every region are then concatenated to a final features vector and feed into one-vs-all linear Support Vector Machines. We evaluate our model on MIT Indoor 67 dataset. The result proved that saliency model used in this work is capable to generate high-quality informative salient regions that lead to good classification output. Our model achieves a better average accuracy rate than a standard approach that classifies as one whole image.

1. Introduction

Scene classification problem has become more popular among researchers in the field of image recognition and computer vision as general [1]. Recognizing, classifying and understanding scene is a basic task in computer vision [2][3]. Scene recognition or classification is a process of organizing images and predicting the class category of the image environment [1][2]. This can be done through the process of supplying attributes information of the scene.

As for human, this is such an easy task. Human can successfully recognize and classify scene effortlessly within short period of time. Human recognizes scene's category by observing objects present in the scene environment [3]. The existent objects would be the differentiation factor between two different scenes. Unfortunately, for a complex natural indoor scene which contains massive amount of different objects [3], humans are not able to observe and pay their attention on every object within short period of time [4], yet still able to recognize its category correctly. Indoor scene is consider to be hard to recognize compare to outdoor scene due to its different layouts, decorations and objects [3][5]. Observation process made by human visual attention and their eyes movement are very selective [6],[7]. In a glance, human only fixate and focus on some important objects/areas instead of fixating on the whole visual input entirely [6]. These objects are usually called salient proto-objects.

