

ABSTRACT

Scene recognition is considered as one of the most important functionalities of human vision. In the field of computer vision, scene recognition problem is very significant and important. Scene recognition or classification is a process of organizing images and predicting the class category of a scene image. Human can accurately classify scene effortlessly within short period of time. Using this concept, a novel approach of scene classification model which built based on human pre-attentive visual attention has been proposed in this study by utilizing one of the earliest saliency model to generate a set of high-quality regions potentially contain salient objects. An experimental study was performed to investigate the efficiency of Saliency Toolbox on natural indoor scene images when its parameters are manipulated. At the end of this experiment, an acceptable parameter scales have been finalized for the use of Saliency Toolbox in the proposed scene classification model. The proposed model is developed with three main operations; (i) salient region proposals generation, (ii) feature extraction and concatenation, and (iii) classification. The proposed model has been trained and tested on MIT Indoor 67 dataset. An experiment and a benchmarking testing have been conducted on the proposed model. The results of the experiment have clearly shown providing more salient regions means providing more meaningful details of an input image. For the benchmarking testing, the result has proved that saliency model used in this study is capable to generate high-quality informative salient regions that lead to good classification accuracy. The proposed model achieves a higher average accuracy percentage than a standard approach model, which classifies based on one whole image. This indicates the advantages of using deep features of local salient objects over global deep features. Two experiments have been conducted in this study to test and evaluate human performance on scene classification for various visual input conditions. The

accuracy of human classification on complete scene images for a brief period of time in Experiment 1 is compared to the accuracy obtained by the proposed scene classification model. Furthermore, the accuracy of human classification in Experiment 1 is also compared to the accuracy obtained by human in Experiment 2, where their classification performance is tested on cropped salient regions. Evaluation of results from these experiments have shown that the proposed model has not achieved the same standard as human. Using only object features to differentiate between two different scenes is not enough to achieve the best classification accuracy as human. The scene background and layout, relationship between objects and human memory are the other features that affect human classification performance. These other attributes of scene need to be taken in the process of recognition and classification of scene images in further study.

Keywords: Convolutional neural network, pre-attentive visual attention, saliency model, scene classification.

Model Pengenalpastian dan Pengelasan Persekitaran Berdasarkan Visual Pra-perhatian Manusia

ABSTRAK

Mengenal pasti sesuatu persekitaran dianggap sebagai salah satu fungsi terpenting dalam penglihatan manusia. Di dalam bidang visi komputer, masalah mengenal pasti persekitaran adalah sangat signifikan dan penting. Mengenal pasti atau pengelasan persekitaran adalah proses penyusunan imej-imej dan meramal kategori sesuatu imej persekitaran. Manusia mampu untuk mengklasifikasikan persekitaran dengan mudah dan tepat dalam jangka masa yang singkat. Dengan mengambil kira konsep ini, satu pendekatan yang berlainan untuk model pengelasan persekitaran yang dibina berdasarkan visual pra-perhatian manusia telah dicadangkan dalam kajian ini dengan menggunakan salah satu model ‘saliency’ yang terawal untuk menghasilkan usul bahagian yang berkualiti yang berpotensi mengandungi objek-objek yang menyerlah. Satu eksperimen telah dijalankan untuk mengkaji kecekapan Saliency Toolbox ke atas imej persekitaran dalaman apabila parameternya di manipulasi. Di penghujung eksperimen ini, parameter yang sesuai telah dimuktamadkan untuk kegunaan Saliency Toolbox di dalam proses model pengelasan persekitaran yang dicadangkan. Model ini telah dihasilkan dengan tiga operasi utama; (i) penghasilan kawasan yang menyerlah, (ii) pengekstrakan dan penggabungan ciri-ciri, dan (iii) pengelasan. Model tersebut telah dilatih dan diuji ke atas set data MIT Indoor 67. Eksperimen dan ujian penanda aras telah dijalankan ke atas model tersebut. Keputusan daripada eksperimen menunjukkan dengan menyediakan lebih banyak bahagian yang menyerlah bererti membekalkan lebih banyak perincian berkenaan gambar. Keputusan untuk ujian penanda aras membuktikan bahawa model ‘saliency’ yang digunakan dalam kajian ini mampu untuk menghasilkan bahagian yang menyerlah yang berkualiti dan berinformasi lalu membawa kepada ketepatan

pengelasan yang bagus. Model tersebut mencapai purata peratusan ketepatan yang lebih tinggi berbanding pendekatan model yang standard yang mengklasifikasi berdasarkan keseluruhan imej. Hal ini membuktikan kelebihan menggunakan ciri mendalam object menyerlah tempatan berbanding ciri mendalam global.

Dua eksperimen telah dilakukan untuk menguji dan menilai kemampuan manusia untuk mengklasifikasikan persekitaran dalam keadaan visual input yang berbeza. Ketepatan manusia mengklasifikasi imej persekitaran yang penuh pada masa yang singkat di dalam Eksperimen 1 dibandingkan dengan ketepatan yang dicapai oleh model yang dicadangkan. Selain itu, ketepatan manusia mengklasifikasi di dalam Eksperimen 1 juga dibandingkan dengan ketepatan yang dicapai manusia di dalam Eksperimen 2 yang mengklasifikasikan potongan kawasan yang menyerlah. Penilaian keputusan eksperimen tersebut telah menunjukkan bahawa model yang dicadangkan masih belum mencapai standard yang sama seperti manusia. Dengan hanya menggunakan objek-objek untuk membezakan di antara dua persekitaran adalah tidak memadai untuk mencapai ketepatan mengklasifikasi yang terbaik. Latar belakang dan susun atur persekitaran, hubungan di antara objek, dan memori manusia adalah merupakan faktor-faktor lain yang mempengaruhi prestasi manusia untuk mengklasifikasikan imej persekitaran. Ciri-ciri persekitaran ini perlu disertakan sekali di dalam proses mengenal pasti dan mengklasifikasi gambar persekitaran pada masa akan datang.

Kata kunci: *Rangkaian neural konvolusional, visual pra-perhatian, model saliency, pengelasan persekitaran.*