

Guest Editorial Preface

Special Issue of Deep Structured Learning Approaches for Multimedia Modelling and Content Analysis

Rashmi Agrawal, Faculty of Computer Applications, Manav Rachna International University, India

D. Ganesh Gopal, School of Computing Science and Engineering, Galgotias University, India

R. Lakshmana Kumar, Department of Computer Applications, Hindusthan College of Engineering and Technology, India

Multimedia has exponentially creating the significant size of data with different modalities like video, text, audio, and image. Multimedia applications need a lot of resources and utilities to support the real-time audio and video applications. Deep structured learning has gained a lot of research interest in the multimedia applications very recently. Deep structure learning has been utilized for video transmission, retrieval content analysis.

Deep structured learning been applicable to fields like image demosaicing and denoising natural language processing, social network filtering, bioinformatics, design of drugs, medical image analysis, audio/video processing, board game programs, computer vision, material inspection and speech recognition, object detection, feature extraction, visual object classification, scene understanding, speech recognition and activity prediction.

With the advancements in smart phones, semiconductor, digitalized cameras, workstations, portable devices and consumer electronics are employing different variety of powerful processors. To generate process, analyze, evaluate, and support the content analysis of image and video various architectures like SoC, MLSoC, etc. have been proposed. Variety of machine learning algorithms namely supervised, unsupervised and reinforcement learning algorithms were used in applications of content analysis of multimedia. But the biggest limitation is that not meeting the real time requirement.

The deep learning-based image recognition system has seen a variety of deep learning models for the media quality. Even though lot of research has been made with the deep structured models, but still it has limitations like encoding the quality related attributes to deep structured learning model is a complex one, relatively smaller data sets to be trained in deep structured learning model is difficult.

Deep structured learning methods study depictions to train the multimodal data for cross media transmissions. It can attain a new way of research in demonstrating the connections between the multimodal data, multimedia content recommendation systems, multimedia content analysis. Multimedia quality estimators are accessed through deep structured machine learning has gained a lot of attraction very recently. This combined approach will overcome the existing limitations of the present methodology which experiences a couple of constraints namely identification with the subjective perspectives of preparing and testing information, the utilization of ill-advised example measures for factual testing, potentially subordinate example perceptions, and an absence of spotlight on measuring the learning capacity of the machine learning based target quality indicator.

This special issue of the *International Journal of Gaming and Computer-Mediated Simulations* (IJGCMS) had received 9 papers. After going through 6 months of review process with several levels of blind review finally 3 papers have been selected for publication.

The first paper that was selected with the title “Sentiment Enhanced Content-Based System for Online Recommendations and Rating Prediction” written by the authors Dr. Akshi Kumar, Ms. Simran Seth, Mr. Shivam Gupta and Mr. Shubham. The main aim of this paper is to design a sentiment enhanced content based recommender system (SEC-Rec). Model has four modules namely key feature extraction module, feature sentiment analysis module, recommendation module and rating prediction module. Key feature extraction module uses hybrid of RAKE and TextRank to uncover key product features. They propose a hybridized model HSVADER (Hybrid SVM and VADER) for feature sentiment evaluation. The recommendation module combines sentiment and similarity for robust product ranking strategy. The practical benefits of SEC-Rec are demonstrated using Amazon Camera dataset and the results are compared to the state-of-the-art. The rating prediction module uses key feature sentiment score to estimate the overall user-rating resolving the multi-criteria decision-making issue.

The second paper selected was titled “Behaviour and Emotions of Working Professionals Towards Online Learning Systems: Sentiment Analysis” written by Dr. Venkata Ramana Attili, Dr. Sreenivasa Rao Annaluri, Prof. Suresh Reddy Gali and Mr. Ramasubbareddy Somula. The main aim of this paper is to find the student behavior in the classroom depends on various influential factors (such as family, friends, locality, habits, etc.). Once a student enters into professional life after completing the graduation, it finds difficult to get back to the learning process due to a variety of issues. In such situations, most of the students go for online courses to improve their skills or to get a promotion at work by upgrading their academic degrees. The tendency of working professionals attending online courses is increasing rapidly due to the vast development in technology in recent times and due to the demand for innovative technologies. In this paper, a detailed study on a variety of participants from different work domains was carried out to study the sentiments of working professionals by analyzing their behavior and emotions using Hadoop, big data and R-Language.

The third paper that was selected is “A Deep Structured Model for Video Captioning” written by the authors Ms. V Vinodhini, Prof. B Sathiyabhama, S Sankar and Mr. Ramasubbareddy Somula. The main aim of this paper is to provide a deep structured model for video captioning. Video subtitles are the description that is automatically generated in natural language as sentences to describe the scene. Captioning covers a wide range from helping the visually impaired to robot interaction. Many advanced areas like Computer vision and Human computer interaction play a vital role as there is a successful growth of deep learning technique. Numerous surveys on deep learning models are evolved with different methods and metrics. Working with the video subtitles is still challenging in term of activity recognition in video. To cope with this challenge a self-gated activation function “swish” is used. The 3D CNN model is used to generate the subtitle for a given video without the audio information in a pop on style. The video is divided into 6:4 ratios for training and testing, the first 50 frames are chosen for segmentation to identify the characters in the image, this include labelling the image based on gestures and other semantic content. The video captioning is given by natural language.

After a rigorous review we have finalized 3 articles that were suitable and were ready to be within the scope of the special issue. Since this is a special issue, we didn’t accommodate any papers that were away from the scope of the theme and after further scrutinizing and several levels of blind review this has been done. Hope this special issue will gather more citation and improve the impact factor of the journal. We greet the editor in chief and manager publications for giving us an opportunity to do a special issue and further we expect to do more special issue with them. Thank You.

*Rashmi Agrawal
D. Ganesh Gopal
R. Lakshmana Kumar
Guest Editors
IJGCMs*