Notice for the PhD Viva Voce Examination

Ms Sagaya Mary J (Reg. No. 1540073), PhD scholar at the School of Sciences, CHRIST (Deemed to be University), will defend her PhD thesis at the public viva voce examination on Thursday, 18 August, 2022 at 3.00 pm in the Syndicate Room, (Room No. 802), Ground Floor, Auditorium Block, CHRIST (Deemed to be University), Bengaluru - 560029.

Title of the Thesis	:	**Computer Vision based Indian Sign Language Recognition using Deep Learning**
Discipline	:	**Computer Science**
External Examiner (Outside Karnataka)	:	**Mary Metilda** Associate Professor Department of Computer Applications Queen Mary's College, Mylapore Chennai – 600004 Tamil Nadu
External Examiner (Within Karnataka)	:	**Dr Krishnan R** Associate Professor Department of Computer Science St Francis College, Koramangala Bengaluru - 560034 Karnataka
Supervisor	:	**Dr Nachamai M** Research Scientist Siemens Healthcare Private Limited Bengaluru - 560100 Karnataka

The faculty members of the Department and the School, interested experts and research scholars of all the branches of research are cordially invited to attend this open viva.

Registrar

Place: Bengaluru
Date: 16 August 2022

ABSTRACT

Speech is a human default and unique modality for language development and communication which is essential for memory and overall cognitive development. Excellency in language permits a child to be extrovert enriching the development of cognitive and psychosocial skills; whereas, for auditory deprived children, the misalignment of the brain and ear makes them impotent to communicate with the society which creates a central dogma that hearing-loss is a disability which further ignores their psycho-social identity. To fill such gaps and make their community more freewheeling in India, Indian Sign Language (ISL) - a complete language with its own linguistic and verbal elements was framed. Though ISL is appropriate and absolute in every linguistic approach, lack of prerequisite and proficiency enforces dedicated teachers to teach the curriculum through contrived signs for the sake of convenience that not only diminishes the distinctiveness of ISL but also dislodges the idea of learning their mother tongue. This creates an imbalance in the analogous learning of communication and curriculum language. In order to balance the level in learning, effective vision-based days of the week ISL model is developed through Convolution Neural Network (CNN) architecture which boasts independent learning of ISL. The proposed model comprises of six stages: dataset creation, preprocessing, splitting dataset into train, validation and test, applying various types of image augmentation techniques according to split, constructing CNN model for feature extraction and classification and finally evaluating the result through evaluation measures.

Initially, an image dataset is created as there is a scarcity of standard ISL datasets in internet sources. The images are created on vision-based technique to avoid of carrying additional superfluous hardware gadgets for human computer interaction. Days of the week image dataset is created by capturing the static sign gesture on Olympus PEN Mini E-PM2 camera from *100* different persons in *5* different views, and with various challenges that real time data includes. Acquired images are preprocessed in the following manner: resizing the original image into *100×100* and adding horizontally flipped images with original image to support the people who have the habit of sinistrality. Once the preprocessing is over, the dataset is manually partitioned into a training set, a validation set, and a test set. In the next stage, various types of image augmentation techniques are applied according to split for increasing the generalisation ability and number of dataset. The custom model is built on CNN architecture for feature learning and classification. Feature learning structure encompasses *5* layer groups from which each group consists of convolution, activation, maxpooling and dropout layers. Initially, *100×100* input image is fed into 1st convolution layer of layer group-1. Even though CNN achieves state-of-the-art result in various aspects, producing too many number of parameters while learning make the researchers to shift from CPU based system to GPU configuration. Thus the proposed model is constructed as a light-weight in terms of number of parameters and number of layers. The other subsequent layers (activation and pooling) are auxiliary layers enhances the learning process. To improve the accuracy, types of feature elevating functions (recently used activation functions like ReLU, Leaky ReLU (LReLU) and Parametric ReLU (PReLU)) are examined in the proposed model of which PReLU procured good accuracy. In classification, the challenge of drastic increment of parameters is overcome in the proposed work by replacing the Fully Connected (FC) layer with the Global Average Pooling (GAP) layer that reduces the size of the parameters without affecting the accuracy. Finally, the performance of the model is evaluated through various metrics. The custom model built in the proposed work procured *98.037%* accuracy in training and *94.686%* in validation and *95.25%* accuracy in test dataset with *9* epochs and *35* as its batch size and the entire model learned only *1,16,175* parameters which utilized approximately *47-48%* of CPU memory space. This makes the proposed model light-weight and cost effective. The portability of the model is also high as it is developed only with the types of open source software packages. ISL learning through proposed model delivers proficient knowledge in language which permits successful social interaction and further leads to higher order cognitive skill development. Hence the new attitude bloom as the hearing-loss is mere a difference but not a self-made fault or mistake.

Keywords: *Indian Sign Language (ISL), Deep Learning, Convolution Neural Network (CNN), Primate Visual Cortex, Image Augmentation, Convolution Layer, Feature Map, Representation learning, Parametric Rectified Linear Unit (PReLU), Dropout Layer, Global Average Pooling (GAP), Parameters, Hyperparameters, Multi-Layer Perceptron (MLP), Mathew's Correlation Coefficient (MCC), Fβ Score.*