Bangladesh University of Business and Technology (BUBT)

Course No. : **CSE - 465**

Course Title: **Machine Learning**

Semester: **Summer 2022 [Tri-Semester]**

Survey Paper Name: A Hybrid Approach combining CNN and LSTM To classify Brain Tumor on 3D MRI Scans Performing Ablation Study.

Submitted by -

Md. Abdul Baten	ID: 18192103230
Emran Hossain	ID: 18192103231
Reza Mia	ID: 18192103249
Md. Alif -ur- Rahman	ID: 18132303234
Khandoker Maliha Fairuz	ID: 17183103002 , 39/1

Submitted to -

Khan Md. Hasib

Faculty Member, Department of CSE, BUBT

Date- 28 SEP 2022

Summary of the Paper:

In this paper honorable “SIDRATUL MONTAHA 1 , SAMI AZAM 2 , A. K. M. RAKIBUL HAQUE RAFID1 , MD. ZAHID HASAN 1 , (Member, IEEE), ASIF KARIM 2 , (Member, IEEE), AND ASHRAFUL ISLAM 3” Considering the challenges of tumor biopsies, three dimensional (3D) Magnetic Resonance Imaging (MRI) are extensively used in analyzing brain tumors using deep learning. As we know that Identification of brain tumors at an early stage is crucial in cancer diagnosis, as a timely diagnosis can increase the chances of survival. In this paper, three BraTS datasets are employed to classify brain tumors into two classes where each of the datasets contains four 3D MRI sequences for a single patient. This research is composed of two approaches. In the first part, they propose a hybrid model named TimeDistributed-CNN-LSTM (TD- CNN-LSTM) combining 3D Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) where each layer is wrapped with a TimeDistributed function. The objective is to consider all the four MRI sequences of each patient as a single input data because every sequence contains necessary information of the tumor. Therefore, the model is developed with optimal configuration performing ablation study for layer architecture and hyper-parameters. In the second part, a 3D CNN model is trained respectively with each of the MRI sequences to compare the performance. Moreover, the datasets are preprocessed to ensure highest performance. Results demonstrate that the TD-CNN-LSTM network outperforms 3D CNN achieving the highest test accuracy of 98.90%. Later, to evaluate the performance consistency, the TD-CNN-LSTM model is evaluated with K-fold cross validation. The approach of putting together all the MRI sequences at a time with good generalization capability can be used in future medical research which can aid radiologists in tumor diagnostics effectively.

Unique Contribution of the Paper:

Our proposed TD-CNN-LSTM model outperforms all these studies with a test accuracy of 98.90%. The idea of using four 3D images of a particular subject as a single input to a deep learning model contributed to the effectiveness of our mode. With this process more details of the brain tumor can be learned by the model than the other approaches described in literature. Training on this combined dataset makes the training process less time consuming and highly detailed 3D images make it easier for the classifier to distinguish between HGG and LGG tumors. Moreover, in this research, vigorous K-fold cross validation experimentations are performed which conclude in consistent performance of the model in all 12 K-fold configurations. A large external test dataset (265 samples) ensures that the model is tested properly for every ablation study case in building the model that adds to sound evaluation of the robustness of the proposed approach. Furthermore, the optimal accuracy achieved from an external test dataset gives a glimpse of the model's absolute interpretation of 3D MRIs in real world application. With this proposed approach TD-CNN-LSTM trained on the combined dataset can outperform all other studies while also requiring fewer epochs for training. This demonstrates the potential of the proposed approach in classifying brain tumors from 3D MRI scans.

How the proposed model works in the paper:

We propose a hybrid model called TimeDistributed-CNN-LSTM (TD-CNN-LSTM) in this survey, which combines 3D Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM), with each layer wrapped in a TimeDistributed function. Because each sequence contains necessary tumor information, the goal is to consider all four MRI sequences of each patient as a single input data set. As a result, the model is created with the best configuration for performing ablation studies on layer architecture and hyper-parameters. In the second part, a 3D CNN model is trained with each of the MRI sequences in order to compare performance. Furthermore, the datasets are preprocessed to ensure maximum performance. The results show that the TD-CNN-LSTM network outperforms 3D CNN, with the highest test accuracy of 98.90%. The TD-CNN-LSTM model is then tested for performance consistency using K-fold cross validation. The method of assembling all of the MRI sequences at once with good generalization capability can be used in future medical research to aid radiologists in tumor diagnostics.

Advantages of the paper:

- To improve overall performance and reduce computational complexity, all 3D images are normalized to a standard scale of 0-1 using the min-max normalization technique and resized to 128 x 128 x 32 during the pre-processing step.
- Instead of using single MRI sequences, all of the sequences and the Region of Interest (ROI) of a specific patient are aimed to be fed into the CNN model as a single input.
- A hybrid TD-CNN-LSTM model is proposed where the layers of CNN are wrapped with a TimeDistributed function, in order to pass all the 3D images of a single subject as one input.
- A 3D CNN model is introduced and trained with single sequences, in order to compare the performance with the TD-CNN-LSTM network in terms of accuracy and performance consistency.
- Both of the models are evaluated using different performance metrics: Accuracy, Precision, Recall, Specificity, F1-score, MAE, RMSE and AUC score.

Disadvantages of the paper:

- The strong, static magnetic field will attract magnetic objects (from small items such as keys and cell phones, to large, heavy items such as oxygen tanks and floor buffers) and may cause damage to the scanner or injury to the patient or medical professionals if those objects become projectiles. Careful screening of people and objects entering the MR environment is critical to ensure nothing enters the magnet area that may become a projectile.
- The magnetic fields that change with time create loud knocking noises which may harm hearing if adequate ear protection is not used. They may also cause peripheral muscle or nerve stimulation that may feel like a twitching sensation.
- The radiofrequency energy used during the MRI scan could lead to heating of the body. The potential for heating is greater during long MRI examinations.
- The strong, static magnetic field of the MRI scanner will pull on magnetic materials and may cause unwanted movement of the medical device.
- The radiofrequency energy and magnetic fields that change with time may cause heating of the implanted medical device and the surrounding tissue, which could lead to burns.
- The magnetic fields and radiofrequency energy produced by an MRI scanner may also cause electrically active medical devices to malfunction, which can result in a failure of the device to deliver the intended therapy

Conclusion:

The goal of this research is to translate a medical expert's diagnosis process of analyzing all of a single patient's MRI sequences to determine cancer into a deep learning approach with the highest accuracy and lowest computational complexity. This paper proposes a completely automated hybrid CNN model called TD-CNN-LSTM that combines CNN and LSTM for classifying brain tumors into HGG and LGG using 3D volumetric MRI. The input layer is wrapped with a TimeDistributed function so that all four MRI sequences of a patient can be passed as single input data. In addition, to compare the performance of our proposed model, a 3D CNN approach is used to train with MRI sequences independently. The results show that the optimally configured TD-CNN-LSTM network outperforms the 3D CNN model, achieving the highest accuracy of 98.90% with optimizer Nadam and learning rate of 0.0001. Furthermore, the K-fold cross validation results validate the robustness and consistency of our model's performance across various training scenarios. Our method of analyzing all MRI sequences together and developing a CNN model using ablation study to achieve the highest accuracy while maintaining the smallest time complexity may be useful in future research and real-world tumor diagnosis systems.