International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 6 June 2023, pp: 952-958 www.ijaem.net ISSN: 2395-5252

Brain Tumor Segmentation using RCNN

G Naga Sujini
Assistant Professor, CSE, Mahatma Gandhi Institute of Technology, Hyderabad, India

Date of Submission: 20-06-2023 Date of Acceptance: 29-06-2023

ABSTRACT—Brain tumor is an anomalous growth of tissue in the brain. Tumors are primarily classified as malignant and benign when they develop. It can lead to death; hence it is important to recognize and identify the presence of tumors. Brain tumors can be classified into primary and secondary tumors. The former represents about 70% of all brain tumors, while secondary tumors are the remaining 30%.
The classification is determined according to tumor origin just as tumors that first originate in the brain are called primary tumors. On the other hand, tumors that first arise in any other part of the body and then transferred to the brain are called secondary tumors, and most of them are malignant. Various imaging techniques can be used to detect and classify brain tumors. However, MRI is one of the most common techniques. MRI popularity comes from the fact of using no ionizing radiation during the scan and superior soft-tissue resolution with the ability to acquire different images using various imaging parameters or by employing contrast enhanced agents.
So we are using Mask RCNN to detect the brain tumor in early stages, so that they can be cured in early stages. Mask RCNN is easy to find the tumor and can find multiple tumors in the brain.
Keywords: Tumors,MRI Images and RCNN

I. INTRODUCTION

Brain tumor is an unnatural and uncontrolled growth in brain cells. Since the human skull is a rigid and volume-limited body, consequently, unexpected growth may affect a human function according to the involved part of the brain, moreover, it may spread into other body organs and affect human functions. According to the world cancer report published by the World Health Organization (WHO), brain cancer accounts for less than 2% of human cancer, however, severe morbidity and complications are produced. Cancer research- corporations in the United Kingdom mentioned that there are about 5,250 deaths annually by the act of brain, other Central Nervous systems (CNS) and intracranial tumors in the UK. The different types of brain tumours are given below.

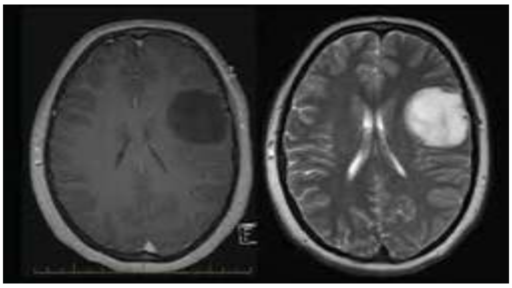

Fig.1 MRI image

Brain tumors can be classified as primary, secondary tumors etc. The former represents about 70% of all the brain tumors and the rest of the 30% of them are secondary tumors and they are malignant. This classification is determined according to tumor origin just as tumors that first originate in the brain are called primary tumors. On the other side, tumors that first arise in any other part of the body and then transferred to the brain are called secondary tumors, and most of them are malignant. Numerous imaging techniques can be used to determine the type of tumor. Gliomas originate in the glial cells of the brain. Gliomas include 30% of all brain tumors and CNS, and 80% of all malignant brain tumors. Gliomas are classified into four grades according to the WHO starting from type I to IV.

Grade I tumors are benign and have a much similar texture to normal glial cells, Grade II is slightly different in texture, Grade III tumors are malignant with abnormal tissue appearance while grade IV is the most severe stage of gliomas and tissue abnormalities that can be visualized by naked eye. Meningioma is a tumor that forms on the membrane that covers the spinal cord and brain. Most meningioma tumors are benign. However, pituitary tumors start from the pituitary glands that control hormones and regulate functions in the body. It can be benign, benign that expands to bones, and malignant.

The classification stage may be a confusing and tedious task for physicians or radiologists in some complicated cases. These cases need experts to work on, localize the tumor, compare tumor tissues with adjacent regions, apply filters on the image if necessary.

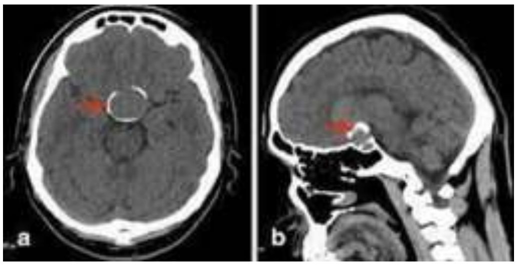

Fig 2. Glioma

MRI Magnetic resonance imaging (MRI) is a medical imaging technique used in radiology to form pictures of the anatomy and the physiological processes of the body.MRI images are used for brain tumor detection because it is a non-invasive and painless procedure. It also does not use any harmful ionizing radiation. MRI does not involve X-rays or the use of ionizing radiation, which distinguishes it from CT (Computed Tomography) and PET (Positron Emission Tomography) scans. MRI is a medical application of nuclear magnetic resonance (NMR). NMR can also be used for imaging in other NMR applications, such as NMR spectroscopy.While the hazards of ionizing radiation are now well controlled in most medical contexts, an MRI may still be seen as a better choice than a CT scan. MRI is widely used in hospitals and clinics for medical diagnosis and staging and follow-up of disease without exposing the body to radiation. An MRI may yield different information compared with CT.


Fig 3. Pituitary Tumor

Fig 3. Use OF MRI in classification For the classification of brain tumor, MRI imaging technique is used and this is because of the non-invasive technique of the MRI. To identify a particular brain tumor, the radiologist or doctor needs a good image and where the tumor is located as well as how large the tumor is so as to diagnose it correctly. If the brain tumor is not detected correctly and classified into their respective classes, the life of the patient is at risk. So a good imaging technique is required which is less harmful for the patient and gives exact information. For this purpose, we opt for MR imaging technique.

II. REVIEW OF RELATED LITERATURE

It presents a machine-based approach for segmentation of brain images using thresholding segmentation technique followed by the identification and classification of tumor into its types using CNN classification approach. An accuracy of 97.05% for the T1-weights for FLAIR weighted[1] MR images is obtained.

An extensive preprocessing is done to eliminate the noise, focusing on the preprocessing of the particular brain image using morphological operations[2].At last, the Naive classifier is used to classify the pre-processing images. This model has obtained an accuracy of 84%. Algorithms start with a group of random generated solutions and the optimal solution is investigated iteratively..

The segmentation is done using Fuzzy clustering means(FCM) and classification is done using CNN with an accuracy of 97.87% . Various segmentation methodologies are explained[3].In spite of huge research, there is no universally accepted method for image segmentation. All methods are equally good for a particular type of image.

The MRI images enhanced using contrast improvement technique, abnormal cells are localized and tumor region is segmented by machine learning classifiers like . fuzzy C-mean, KNN, K-mean[4] with an segmentation accuracy of 98.97, 89.96 and 79.95 respectively .Features extraction is done using GLCM[5].

A hybrid approach which classifies MRI tumor images[5] into a benign tumor and malignant tumor was proposed by Sanjeev Kumar et al. The approach includes discrete wavelet transform (DWT) to be used for extraction of features, support vector machine for brain tumor classification[6] .This approach reduces the time taken for classification and avoids human error.

Image net database is used for classification[7]. It is one of the pre-trained models. So the training is performed for only the final layer. Raw pixel values with depth, width and height feature values are extracted from CNN[8]. Finally, the Gradient descent-based loss function is applied to achieve high accuracy. The training accuracy, validation accuracy and validation loss are calculated[9]. The training accuracy is 97.5%.

The complete model is divided into two sub parts ,first part involves image suitable for pre-processing, second part involves classification using svm[10]. This model got an accuracy of 97.1% for 70 images. The tumor affected area for symmetrical analysis. They showed its application with several data sets with different tumor sizes. MR Images give better results compared to other techniques like CT images and X-rays.

III. DESIGN METHODOLOGY

Medical image classification and segmentation is a field where deep learning can make a huge impact with promising results. It facilitates the automation of non-invasive imaging-based diagnosis. In architecture the user tries to find and collect the dataset from different research papers. Then the user does some pre-processing techniques on the dataset. By splitting the data into train data and test data.

MRI images are taken as input from the user . These images are pre-processed using different methods so that it can be easier for the model to learn and also be used for testing purposes. Later these pre-processed images are sent to the model for training the model .Some parts of these images are kept aside for the testing purposes .After the training is completed these images are classified using the model whether it has the tumor or not . If the image has a tumor then the image is further segmented and the tumor part is carefully analysed and the tumor is marked .

A. Mask R-CNN architecture

It is very similar to R-CNN except there is another layer to predict segmentation. The stage of region proposal generation is the same in both the architecture and the second stage which works in parallel, generates a bounding box as well as outputs a binary mask for each RoI.

It comprises –

- Backbone Network
- Region Proposal Network
- Mask Representation
- RoI Align

Backbone Network:

The authors of Mask R-CNN experimented on two kinds of backbone networks. The first is standard ResNet architecture (ResNet-C4) and another is ResNet with a feature-pyramid network.

Region Proposal Network:

All the convolution feature map that is generated by the previous layer is passed through a 3*3 convolution layers. The output of this is then passed into two parallel branches that determine the objectness score and regress the bounding box coordinates.

Mask Representation:

A mask contains spatial information about the object. Thus, unlike the classification and bounding box regression layers, we could not collapse the output to a fully connected layer to improve since it requires pixel-to-pixel correspondence from the above layer. Mask R-CNN uses a fully connected network to predict the mask.

RoI Align:

RoI align has the same motive as the RoI pool, to generate the fixed size regions of interest from region proposals.

B. Regional Convolutional Neural Network

Convolutional Neural Network (CNN) is a class of deep neural networks and is mostly applied for visual imagery. There are various layers in CNN arranged in hierarchical order placed one after the other. The various layers include image input layer, convolutional layer, pooling layers (like max-pooling, average-pooling),normalization layer, ReLU layer, dropout layer, fully connected layer, SoftMax layer, classification layer.

Re-LU Layer:

Re-LU means Rectified Linear Unit. It is an activation function. The Rectified Linear Unit layer is the most commonly developed activation function defined as the positive part of the argument. Mathematically, it is represented as f (x) = max (0, x). In a neural network, the activation functions are responsible for transforming the summed weighted inputs from the node to activation of the node or output for that input. The main purpose of applying an activation function is to increase the non- linearity of the image or the input. Images are naturally non-linear i.e. an image contains several nonlinear features such as the transition between the pixels, the colors, borders etc.

When this image comes through the convolution layer the results are as such. If observed carefully, the image details i.e. the features of the image are clearly seen as it is composed of pixels from white to black with many shades of gray in between. The rectifier i.e. the ReLU layer serves to break up the linearity in the image. It removes all the black elements from it and keeps only the ones that are positive i.e. the ones that have gray and white. The essential difference between the rectified image and the non- a rectified image is the progression of the colors. It classifies the weights of the image into only non-negative values. This indicates that the linearity is disposed.

Normalization layer:

Then, a cross-channel normalization layer is used to normalize the input layer by scaling and

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 6 June 2023, pp: 952-958 www.ijaem.net ISSN: 2395-5252

adjusting the related activations. It makes a local response normalization layer based on channel-wise with a window of a particular size (it has been arbitrarily chosen to be 5). Normalization can be used in back propagation and network training acceleration.

Pooling layer:

After the convolution we usually perform a pooling operation to minimize the dimensionality. Pooling enables us to reduce the number of parameters which shortens the training time and also combats over-fitting. By down sampling each feature map independently, pooling layers reduce the height and width by keeping the depth intact.

Pooling is used to down-sample the detection of features in feature maps. Pooling helps to make the representation become approximately equivalent to small translations of the input. Invariance to translate the input by a small amount, the values of the most of the pooled outputs do not change. Pooling reduces the amount of information in each feature obtained in the convolutional layer while maintaining the most important information.

The most common type of pooling is max-pooling which takes the max value in the pooling window. Unlike convolution, pooling has no parameters. It takes a window and slides it over the input and takes the maximum value in the window. Just like in convolution, we mention the size and stride of the window or filter or kernel.

Below is the example of the max-pooling with a window of size 2x2 and a stride of 2.

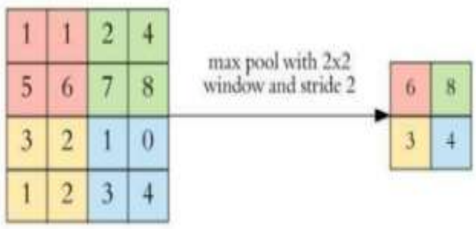

Fig 4. Max-pooling

Here each color denotes a different window and as the sizes of the window and stride are the same, the windows are not overlapping each other. This down samples the feature map while keeping the important details of the image or the input.

If we clearly observe the dimensions of the feature map before and after the pooling, we can see that the dimensions have changed. If the input of the pooling layer has a dimensionality of 16x16x10 using the same pooling parameters as above, we would get a feature map of 8x8x10. Both the height and width are halved but the depth remains unchanged. This is because the pooling works on each depth slice of the input. The below image illustrates how the depth remains unchanged while pooling. By halving the heights and widths we reduce the number of weights to 1⁄4 of the input. As the number of inputs increases, we change the dimensions of the pooling window. Usually in CNN a 2x2 dimensional window with stride 2 and zero padding is used.

Dropout layer:

Dropout layer is a neural network corresponding to dropping out a neuron in a neural network. Dropout is commonly used to make the neural networks regularized. Most types of layers are used along with it, such as fully connected layers (dense), convolutional layers and recurrent layers like long short-term memory network layers. Dropout layer is applied on all the hidden and visible layers of the network. Dropout layer is not applied on the output layer. Application of

The dropout layer on the fully connected layers and the convolutional layers is a very different operation. But, in the community of deep learning using deep neural networks that dropout layer has limited benefits. Over fitting can be prevented from the model using a dropout layer. By setting the outgoing edges to the hidden units of 0 at each update of the training phase dropouts work in a randomized manner.

Applying the dropout in the fully-connected networks and taking out a zero-out column from the matrix that represents the fully-connected layer. In other words it means removing a neuron from the neural network. To drop a neuron is accordingly reasonable because it quantitatively helps the promotion of the redundancy in the weight matrix i.e.,the operation can be robustly performed by the sub-networks. The information is lost if we apply the drop out in the first layer which is nothing but the convolutional layer, so there has to be extra care taken when dropout is applied to the convolutional layer. So, it is recommended to use low dropout in the first layer at the beginning and further increase it gradually. The major problem when applying the dropout to the convolutional layer is the network complexity. To completely avoid overfitting, network complexity must be decreased.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 6 June 2023, pp: 952-958 www.ijaem.net ISSN: 2395-5252

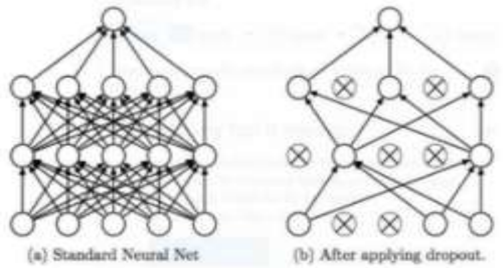

Fig 5. Before and after applying the Dropout layer

Fully Connected Layer:

In the fully connected layer, all the inputs from one of the convolutional layers to every activation unit of the next layer. To detect these high-level features from various layers there is a need for an additional layer which connects all the previous layers and also the fully connected layer to end the network. The output is of the layers preceding it such as the ReLu layer or convolutional layers. Input volume which gives an N dimensional output vector.(N is the number of classes that the program has to choose from).

For example, if we want a digit classification program, there would be 10 digits. Hence N is 10. The probability of a certain class is represented by each number in this N dimensional vector. In other example, if the resulting vector for a digit classification program is [0.1 .2 .70 0 0 0 0 0 0 0.05], then it represents a 10% probability that the image is a 1, a 20% probability that the image is a 2, a 70% probability that the image is a 3, and a 30% probability that the image is a 9.(Though there are other ways to represent, we chose this as an example of soft max representation). This fully connected layer looks at the output of the previous layer (which as we remember should represent the activation maps of high-level features) and determines the features which correlate the most to a particular class For instance, consider a program which is used to predict any image of a dog, the values of the activation maps will be high that represents the high level predicting features like 4 legs, paws, tail etc. Consider another program that predicts birds, then the high-valued activation maps will consider features like two legs and wings as the high-value parameters.

The basic functionality of the fully connected layer is to look for the high-level features which are used to correlate to particular classes easily and strongly and assign particular weightage to every feature so that when we compute the products between the weights and the previous layer, and get the correct results for the different classes.

Softmax Layer:

The layer which is used to squash all the predicted classes between 0 and 1 probability such that the total sum will be equal to 1 i.e., 100%. Softmax layer is similar to the activation functions like sigmoid, tanh and ReLu which is applied to the output of the very last layer and it is defined as

(y)y =ez

Here j -> the probability of the class
k -> number of different classes
The total summation will be equal to 1

It is used to normalize the output of the neural network to fit between the probability 0 and1. In other words, the softmax layer converts the output of the neural network into probability distribution.

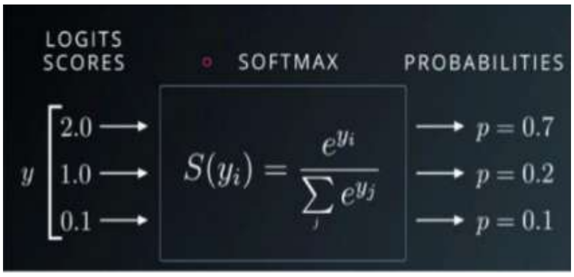

Fig 6.Softmax layer

Softmax function turns logits into probabilities..thetermlogits layer is used as the last neuron layer for neural network for classification task which produces raw prediction values as real numbers ranging from [-infinity,+ infinity]. The accuracy of the image classification depends upon the probability.

Classification Layer

The common last layer of every architecture is the classification layer .We use a classification layer which is based on cross-entropy loss to estimate the classification loss and provide the final predicted categorical label for each input image.

IV. RESULTS

Test case 1:

The brain tumor segmentation model takes the MRI images as input and segments the tumors to show the difference between the G.T and detected tumor with high accuracy and precision.

Input images

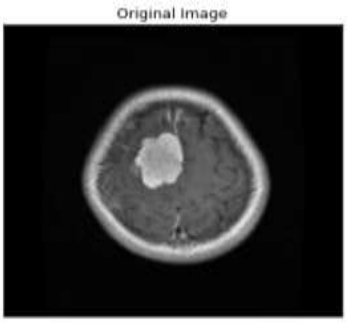

Fig 7. Input image

Output image

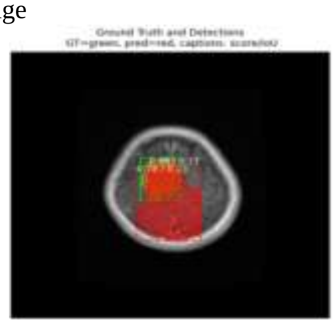

Fig 8. Final output image

The output image represents the ground truth of the tumor and the prediction of the model when run with 25 epochs also presents the accuracy of the tumor predicted.

Testcase 2:
Input image:

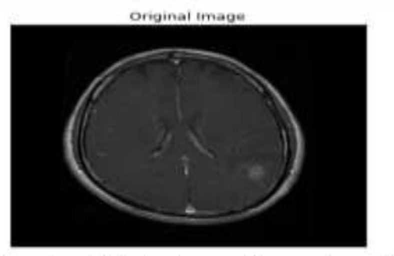

Fig 9. Input image

Output image:

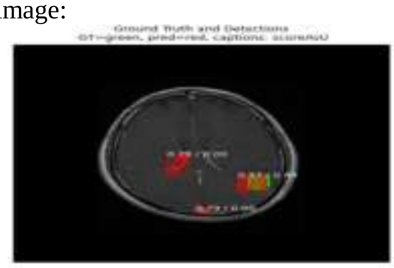

Fig 10. Final output image

The output image represents the ground truth of the tumor and the prediction of the model when run with 25 epochs also presents the accuracy of the tumor predicted.

Test case 3:
Input image:

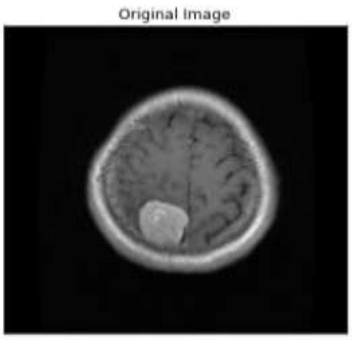

Fig 11. Input Image

Output image:

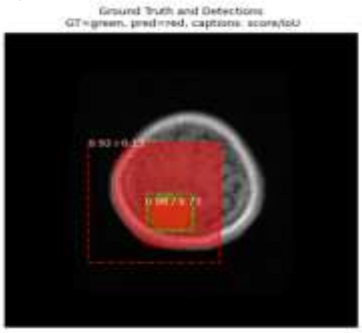

Fig 12. output image

The output image represents the ground truth of the tumor and the prediction of the model when run with 25 epochs also presents the accuracy of the tumor predicted.

Testcase 4:
Input images:

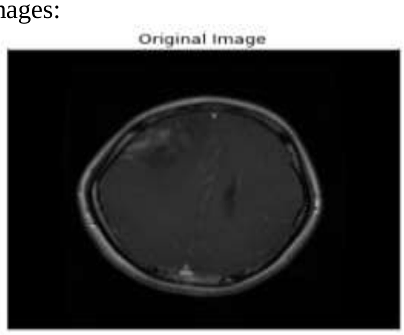

Fig 13. Input image

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 6 June 2023, pp: 952-958 www.ijaem.net ISSN: 2395-5252

Output image

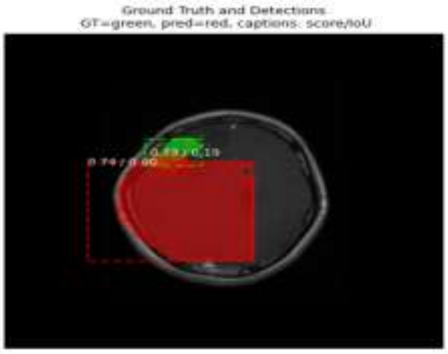

Fig 14. Final output image

The output image represents the ground truth of the tumor and the prediction of the model when run with 25 epochs also presents the accuracy of the tumor predicted.

V. CONCLUSION

Recent research has focused on applying different machine learning and deep learning models over the MRI images to segment the tumors present in it.but, it lacked in accuracy and also couldn't predict small tumors present in them which can a potential threat in the future.Our model has handled these problems by using preprocessing techniques and high capacity RCNN architecture.Most of the models currently present cannot predict multiple tumors even this case is also handled. In overall, Our model can help the radiologist to better predict the tumors,potential tumors and act accordingly to the situation.

REFERENCES

[1]. M. J. Pereira, L. Coheur, P. Fialho, and R. Ribeiro, "Chatbots' greetings to human-computer communication," arXiv preprint arXiv:1609.06479,2016.

[2]. O. Deryugina, "Chatterbots," Scientific and Technical Information Processing,vol. 37, no. 2, pp. 143–147, 2010.

[3]. J. S. Malik, P. Goyal, and A. K. Sharma, "A comprehensive approach towards data preprocessing techniques & association rules," in Proceedings of The4th National Conference, 2010.

[4]. B. A. Shawar and E. Atwell, "Chatbots: are they really useful?" in LDV Forum, vol. 22, no. 1, 2007, pp. 29–49.

[5]. A. Bordes, S. Chopra, and J. Weston, "Question answering with subgraph embeddings," arXiv preprint arXiv:1406.3676, 2014.

[6]. B. Setiaji and F. W. Wibowo, "Chatbot using a knowledge in database: Human-to-machine conversation modeling," in Intelligent Systems, Modeling and Simulation (ISMS), 2016 7th International Conference on.IEEE, 2016, pp. 72–77.

[7]. H. Wang, Z. Lu, H. Li, and E. Chen, "A dataset for research on short-text conversations." in EMNLP, 2013, pp. 935–945.

[8]. D. Britz, "Deep learning for chatbots, part 1–introduction," 2017.

[9]. I. V. Serban, A. Sordoni, Y. Bengio, A. C. Courville, and J. Pineau, "Building end-to-end dialogue systems using generative hierarchical neural network models." in AAAI, 2016, pp. 3776–3784.

[10]. J. Hill, W. R. Ford, and I. G. Farreras, "Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations," Computers in Human Behavior, vol. 49, pp. 245–250, 2015.

[11]. N. Asghar, P. Poupart, J. Xin, and H. Li, "Online sequence-to-sequence reinforcement learning for open-domain conversational agents," arXiv preprint arXiv:1612.03929, 2016. [12] A. Kerly, R. Ellis, and S. Bull, "Calm System: a conversational agent for learner modeling," Knowledge-Based Systems, vol. 21, no. 3, pp. 238–246, 2008.