A Synergic Deep Learning Approach for Glioma Grading from Brain Tumor MRI Images

Saransh Chopra[+], Harshvir Sandhu[+], Ishan Bansal[+], Nirmal Yadav[*]
[+,*]Cluster Innovation Centre, University of Delhi
[*]Email: nyadav@cic.du.ac.in, nirmaliitr25@gmail.com

Highlights and Graphical Abstract

- A Synergic Deep Learning (SDL) based model for the classification of glioma into different grades i.e. grade II, grade III, and grade IV, has been proposed.
- Experimental study performed over a small subset of slices from the REMBRANDT dataset for brain tumor.
- Demonstrating that introducing a Gaussian filter during the preprocessing step drops the performance of the SDL model.
- The proposed model trained on a smaller dataset performs better or equally as compared to other state-of-the-art approaches trained on larger datasets.

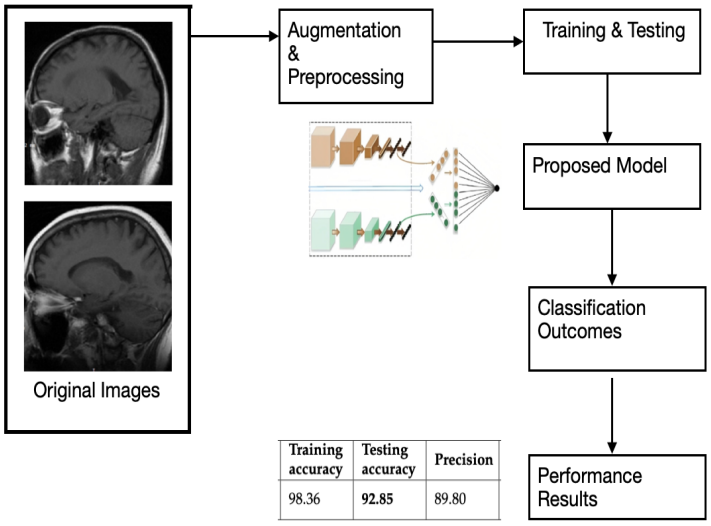


Figure 1: Graphical Abstract

Abstract

Computer-aided diagnosis using deep learning approaches has made tremendous improvements in medical imaging for automatically detecting tumor area, tumor type, and grade. These advancements, however, have been limited due to the fact that 1) medical images are often less in quantity, leading to overfitting, and 2) significant inter-class similarity and intra-class variation between the images. This study proposes a Synergic Deep Learning model[26] [27] with an AlexNet[15] as a backbone for the automatic grading of glioma tumors. The Synergic Deep Learning architecture enables two pre-trained models to mutually learn from each other, allowing them to perform better than vanilla pre-trained models. Our study uses 417 T1-weighted sagittal tumor Magnetic Resonance Imaging (MRI) slices obtained from the REMBRANDT[21] dataset. These 417 slices are pre-processed and augmented before they are fed into the model, which then classifies the tumor into one of the three grades: oligodendro glioma, anaplastic glioma, and glioblastoma multiforme. The proposed architecture achieves accuracy of 98.36%, showing that the model achieves excellent performance metrics even after being trained on an extremely small dataset. Finally, the proposed SDL model trained on less number of MRI images performs wither better or equally with other models trained on larger datasets in the literature.

Keyword Glioma Tumor Grading; Synergic Deep Learning; Transfer Learning; AlexNet; REMBRANDT

1 Introduction

Glioma is the second most prevalent brain tumor in adults after meningiomas. The earliest documented accounts of gliomas were published by Berns in British scientific journals in 1800. Percival Bailey and Harvey Cushing established the foundation for the contemporary classification of gliomas in 1926[22]. As per a National Library of Medicine journal, there are six cases of gliomas diagnosed annually per 100,000 individuals in the United States[17].

Gliomas are tumors that infiltrate the surrounding brain tissue diffusely. They are classified by the World Health Organization's malignancy scale into grades I to IV[12]. Pilocytic astrocytoma, the most common glioma in children, is a benign grade I lesion that has a slow proliferation rate and is relatively well-defined. Grade II tumors, such as astrocytoma, oligodendroglioma, and oligoastrocytoma, have a slow growth rate, are highly differentiated, and infiltrate the normal brain parenchyma diffusely, making them prone to malignant progression. Anaplastic astrocytoma, anaplastic oligodendroglioma, and anaplastic oligoastrocytoma are grade III tumors that exhibit higher cellular density and have atypia and mitotic cells. Glioblastoma and gliosarcoma are the most malignant grade IV tumors and are also the most common gliomas. They exhibit microvascular proliferation and pseudo palisading necrosis in addition to grade III features..

Imaging is crucial for the diagnosis, surveillance, characterization, and therapeutic monitoring of intracranial tumors. The American Cancer Society recommends Magnetic resonance imaging (MRI) and computed tomography (CT) scans as the most commonly used techniques to detect brain tumors. For gliomas, MRI is particularly valuable, and conventional MRI protocols using T1-weighted, T2-weighted, and gadolinium-enhanced

sequences play a significant clinical role. These protocols provide high-resolution structural information in multiple planes, enabling better tissue characterization compared to CT[24].

Medical imaging techniques have greatly benefited from the use of Artificial Intelligence (AI). In the field of diagnostic imaging, there has been a significant increase in the number of publications on AI, from approximately 100-150 per year in 2007-2008 to 1000-1100 per year in 2017-2018[23]. Researchers have utilized AI to automatically identify complex patterns in imaging data and provide quantitative assessments of radiographic features. AI has also been applied in radiation oncology to various image modalities used in different stages of treatment, such as tumor delineation and treatment assessment. However, AI advancements in biomedical images have been limited due to two primary reasons. Firstly, medical images are often limited in quantity, leading to overfitting of models or high variance. Secondly, there is significant inter-class similarity and intra-class variation between images, making it difficult for the model to classify them accurately.

It is known that the procurement of biomedical image data is chalenging due to several reasons. The limited quantity of datasets in medical image analysis, along with the significant intra-class variation and inter-class similarity, present an even greater challenge in classifying medical images. The Synergic Deep Learning (SDL) model allows for the simultaneous learning of multiple image pairs and utilizes multiple DCNN components without sharing parameters, enabling the model to benefit from an ensemble of multiple networks. The model can be trained end-to-end using classification errors from DCNNs and synergic errors from each pair of DCNNs. If one DCNN correctly classifies an image, the synergic error generated by the other DCNN serves as an additional force to update the model.

Therefore we propose a novel Synergic Deep Learning model with an AlexNet backbone for classifying the grade of a glioma tumor to MRI slices. The Synergic Deep Learning architecture allows our model to grade glioma tumor with high accuracy even with a small training dataset. The pre-trained AlexNet models aim to prevent overfitting caused due to the low number of MRI slices and the overall Synergic Deep Learning architecture allows two AlexNets to mutually learn from each other. Each DCNN is trained separately and the extracted feature vector for an image is fed into the synergic layer, which predicts if the two images passed through two DCNNs belong to the same class. The error produced in the synergic layer is back propagated to the DCNNs, allowing them to learn from each other.

Moreover, this paper makes the following contributions -

- A Synergic Deep Learning model with an AlexNet backbone fine-tuned on an extremely small subset of REMBRANDT dataset (T1-weighted sagittal - assorted scanning views).
- The best SDL model (at $\lambda = 3$) with a training accuracy of 98.36%, a testing accuracy of 92.85%, an average precision of 89.80%, an average recall of 92.3%, and an average F1 score of 90.91%.
- Comparison with models in the literature with results showing that the proposed

model either outperforms or performs equally, even after being trained on a smaller dataset.

- A first-of-its-kind open-source implementation of the Synergic Deep Learning model.
- Discussions on the REMBRANDT dataset, stability interval of the synergic hyperparameter, the use of Gaussian filters and AlexNet as a backbone for Synergic Deep Learning.

2 Related Work

The Synergic Deep Learning model was first proposed for the task of Skin Lesion Classification in Dermoscopy Images. The proposed SDL model used ResNet50 as its backbone and achieved an accuracy of 85.75%, average precision of 66.4%, and AUC of 82.6% on the ISIC 2016[8] skin classification dataset. The accuracy achieved by the model was greater than the accuracy of ResNet50, ResNet502, and the other top 5 performing models on the leaderboard. The SDL model was even able to beat a joint segmentation-classification model, depicting how strong it was. The same authors presented another paper on the SDL model, and generalized the model even further to an SDLn model. Their study used the ImageCLEF-2015[6], ImageCLEF-2016[7], ISIC-2016, and ISIC-2017[5] datasets and outperformed the existing state-of-the-art models, including ResNet50 and ResNet502. However, the paper emphasized the need to use Synergic Deep Learning Models for medical image classification and analysis.

The SDL model has been employed for several diverse datasets, including a study on Diabetic Retinopathy[14]. The paper focused on the classification of DR fundus images on the basis of severity level using a deep learning model. They coupled the SDL model with histogram-based segmentation (to extract useful regions from the image) on the MESSIDOR dataset and achieved an accuracy of 99.28%, a sensitivity of 98%, and a specificity of 99%. To ensure the goodness of the proposed model, they also performed a CT analysis, where their model required a minimum CT of 15.21 seconds to classify the fundus images, followed by AlexNet and then VGG19 models.

A recent research on Grading of glioma tumors using Deep Learning[9] trained a joint segmentation-classification pipeline using Convolutional Neural Networks and achieved an average accuracy of 87% on a dataset of 237 patients. Their model outperformed the methods considering radiomic features alone and also shows the highest performing model (gradient boosting).[25] involved integrating radiomics features with high-level deep learning features to construct a more comprehensive representation of medical images. The features were extracted using a fine tuned VGG-16 model, and the BRATS 2018[2][3] dataset was used, which includes 285 subjects from multiple institutions. The extracted features were then utilized to train three classifiers, namely Logistic Regression (LR), Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA). Moreover the best results were achieved by combining radiomics feature extraction with deep learning feature extraction. Recursive feature elimination (RFE) algorithm was used for feature selection. This method achieved an average accuracy of 89.1% and an average AUC of 93.4%. Another study[1] used the Rembrandt dataset for the detection of brain tumors, specifically for tumor grading. To recognize tumors from brain MRI images,

the Tetrolet Transform (TT)[16] and Support Vector Machine (SVM) classifier were employed. The Tetrolet transform was used to decompose the MRI brain tumor image at a predefined level, and the resulting image features were then classified using SVM. The fifth level of decomposition with SVM-based classification resulted in an accuracy level of 98.8%. In addition, machine learning based models have provide good performance architectures for the classification and segmentation of the brain tumor[18] [10] [20].

3 Proposed Architecture

The architecture of proposed work includes preprocessing of dataset and synergic deep learning model. A detailed description of these steps are given below.

3.1 Pre-processing

The T1-weighted sagittal MRI slices were selected as the tumor was visible in these scans. Figure 2 shows some sample images from the dataset that has been used while training the classification model. This was done to address the fact that a large number of axial images contained no information about the tumor, but we could see the tumor in almost all the sagittal images.

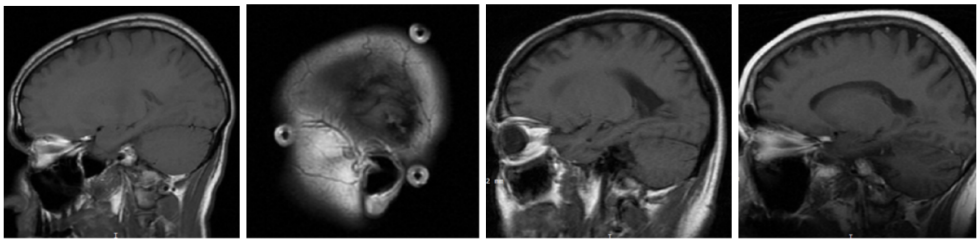


Figure 2: First 2 images (both grade 3 tumor) depicting intra-class variation and last 2 images (from left to right: grade 4 and grade 3) depicting inter-class similarity.

The Synergic Deep Learning model inputs a pair of images as its input with 3 labels: 1) the label of the first image, 2) the label of the second image, 3) the synergic label conveying if both the images belong to the same class (1) or not (0). Images belonging to each grade were randomly shuffled and the first 50% of images belonging to a particular grade were paired with each other, giving them a synergic label of 1. The remaining 50% of the images were paired randomly with an image of another grade, giving them a synergic label of 0. This split also resulted in a loss of images; the last set of images could not be paired up with an image belonging to a different grade, as all the remaining images had the same grading. Finally, the testing images were not paired because both the DCNNs were tested independently of each other. Given that the MRI slices have only 1 channel and the pre-trained models accept images with 3 channels, every image was stacked depthwise to make the resulting image have 3 channels where each channel held the same information. After pairing images, the dataset was split in a ratio of 7:3 for the training and testing phase. and every MRI slice was resized to 128X128 to allow the model to converge faster. Furthermore, the images were augmented without increasing the size of the dataset. The augmentation process included randomly rotating the images in the range of $\pm$ 10 degrees and applying Gaussian filters of kernel sizes 3x3, 5x5, and 7x7.

3.2 Synergic Deep Learning model

The Synergic Deep Learning model constitutes of a pair images as its input with 3 labels, 2 independent DCNNs and a synergic layer. Figure 3 shows a diagrammatic description of the Synergic Deep Learning model. The pair of input images are then fed independently to the two DCNN components with their respective labels. During forward propagation, the feature vector of the pair of images is obtained from the second last layer of each DCNN and is concatenated for the synergic layer. This concatenated feature vector is fed to the synergic layer, which then predicts if the images belong to the same class or not.

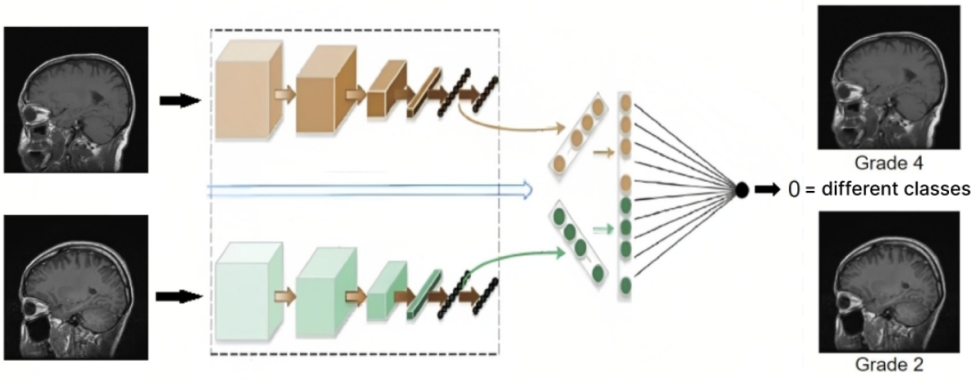


Figure 3: Diagrammatic representation of the proposed Synergic Deep Learning model.

For a data pair Z_A,Z_B with labels y_A,y_B where $A \neq B$, the synergic label y_s is defined as:

$$y_s = \begin{cases} 1 & y_A = y_B \\ 0 & y_A \neq y_B \end{cases}$$

In this work, we use 2 AlexNets as the DCNN components with their classification layer replaced with a single layer of neurons followed by the prediction layer of 3 neurons. We also define a single neuron layer as the synergic layer with 1 output neuron and sigmoid activation, to predict the synergic label. The synergic layer of the SDL model minimizes the following binary cross entropy loss:

$$Loss_{SDL} = y_s \log \hat{y_s} + (1 - y_s) \log (1 - \hat{y_s})$$

and the DCNN components are trained to minimize the following Cross Entropy (CE):

$$Loss_{DCNN} = \sum_{i=1}^{n} \sum_{c=1}^{3} y_{ic} \log p_{ic}$$

where, n is the batch size, c is the number of classes, and p is the predicted probability of the DCNN component. Figure 4 shows the architecture of our data pipeline and model with the double headed arrows in the model section depicting both forward and back propagation.

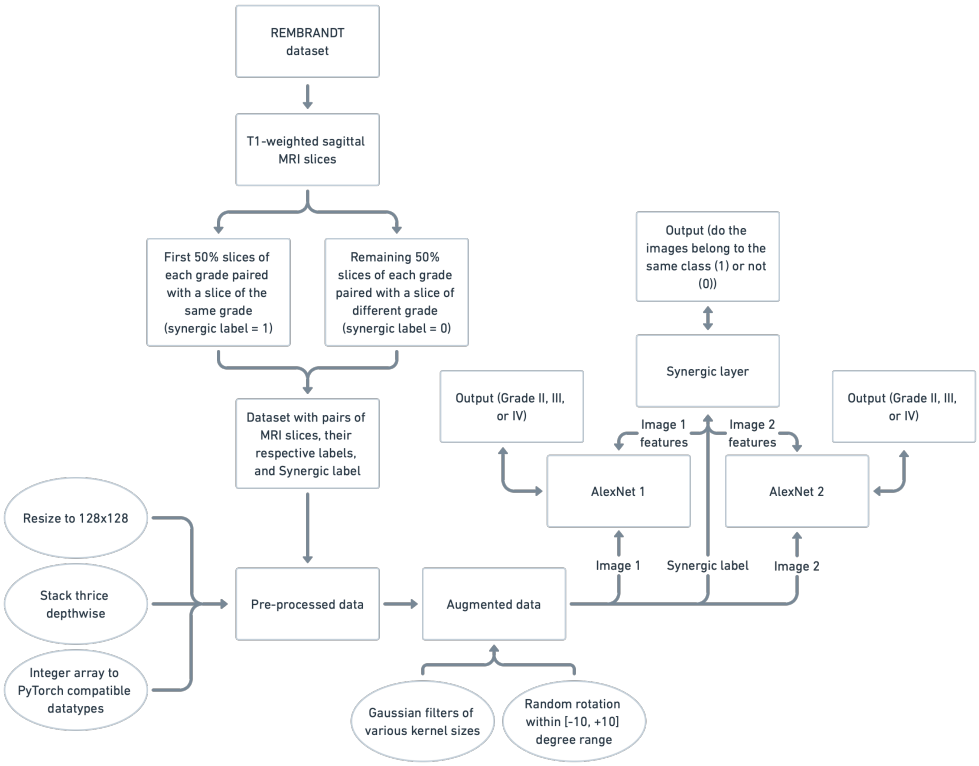


Figure 4: Architecture of our data pipeline and model.

3.2.1 Forward propagation

For forward propagation, we consider f_i as the features extracted by $DCNN_i$ from X_i where X_i is a batch of images, with labels y_i. P_1,P_2 are the class probability distributions predicted by $DCNN_1$, $DCNN_2$. Similarly P_{SDL} is the probability predicted by the synergic layer. The loss for each DCNN and the synergic layer is computed [27] as follows:

$$f_{SDL} \leftarrow f_1 \cdot f_2$$

$$P_{SDL} \leftarrow SDL(f_{SDL})$$

$$loss_{DCNN_1} \leftarrow CE(P_1, y_1)$$

$$loss_{DCNN_2} \leftarrow CE(P_2, y_2)$$

$$loss_{SDL} \leftarrow BCE(P_{SDL}, y_{SDL})$$

3.2.2 Back propagation

Here θ_1, θ_2 are defined as the parameters of $DCNN_1$, $DCNN_2$ respectively and θ_{SDL} as the parameters of the synergic layer. λ or the synergic hyper-parameter indicates the factor by which the synergic error is back-propagated to the individual DCNNs and η represents the learning rate. The model parameters are updated [27] as follows:

$$\theta_1 \leftarrow \theta_1 - \eta * (\frac{\partial loss_{DCNN_1}}{\partial \theta_1} + \lambda * \frac{\partial loss_{SDL}}{\partial \theta_{SDL}})$$

$$\theta_2 \leftarrow \theta_2 - \eta * (\frac{\partial loss_{DCNN_2}}{\partial \theta_2} + \lambda * \frac{\partial loss_{SDL}}{\partial \theta_{SDL}})$$

$$\theta_{SDL} \leftarrow \theta_{SDL} - \eta * \frac{\partial loss_{SDL}}{\partial \theta_{SDL}}$$

4 Experiments

4.1 Platform

The training and testing were performed on a machine with 8GB RAM, Intel-i5-10300H CPU @ 2.5 GHz, and a dedicated NVIDIA GTX 1650ti GPU. Training of the model took 1281 seconds for 184 pairs of images, making each epoch last for an average of 5.124 seconds.

4.2 Hyperparameters

The model was trained for 250 epochs with the data divided into a batch size of 8 and it was trained using Stochastic Gradient Descent optimizer with initial learning rate of 0.0001. The learning rate scheduler is of the form -

$$\eta = \frac{\eta}{1 + (10^{-4} * epoch)}$$

Additionally, we varied the synergic hyper-parameter and recorded the maximum training and testing accuracies obtained from $\lambda = 3$ to $\lambda = 8$. We also experimented by adding a Gaussian filter to the SDL model and varied the size of the Gaussian kernel (from 3x3 to 7x7). The other pre-trained models were independently fine-tuned and trained using the mini-batch SGD optimizer.

4.3 Dataset

The REMBRANDT dataset used in this study obtained through The Cancer Imaging Archive (TCIA)[4]. The dataset houses MRI slices of 127 patients stored in the DICOM format and the typical folder structure used for storing DCM MRI slices. It contains image scans of 89 patients of distinct ages, races, maladies and contains glioma grading information. From this dataset, we selected all the T1-weighted sagittal MRI slices. The selected 417 slices are described in table 1. These selected images are divided into two sets for training and testing purpose. Training set consists of 295 slices and testing set with 122 slices.

Table 1: Selected dataset description

Grade	**Patients**	**Selected Slices**
II	41	128
III	25	108
IV	23	181

4.4 Evaluation Metrics

For comparison with the previous models, we used different parameters such as precision, recall, F1 score, accuracy and specificity, which can be defined as

$$Precision = \frac{True\ positive}{True\ positive + False\ positive}$$

$$Recall = \frac{True\ positive}{True\ positive + False\ negative}$$

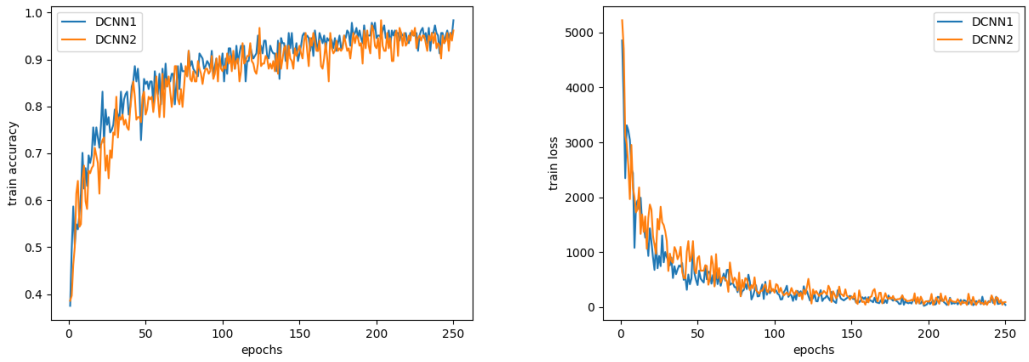
$$F1\ score = \frac{2 * precision * recall}{precision + recall}$$

$$Specifity = \frac{True\ negative}{True\ positive + False\ positive}$$

$$Accuracy = \frac{True\ positive + True\ negative}{True\ positive + False\ positive + True\ negative + False\ negative}$$

4.5 Results

The model achieved the highest testing accuracy with the synergic hyper-parameter λ set to 3. We varied the value of the synergic hyperparameter λ within the stability interval described in[27] 8a. The testing accuracy of the SDL model went down as we increased the value of λ. The best model achieved a training accuracy of 98.36% and a testing accuracy of 92.85%. Figure 5a shows the training accuracies of both the DCNNs from the SDL model with $\lambda = 3$. It can be observed that the accuracies of both the DCNNs converge. Figure 5b shows the loss values for the same training phase, and it can be observed that both the loss functions converge as well. Furthermore, figure 6 shows the loss and accuracy of the synergic layer, the layer that was responsible for classifying the pair of images passed to the DCNNs as 1 (same class) or 0 (different class). It can be observed that the synergic layer quickly attains an accuracy of 100% and the loss function first fluctuates but soon converges exactly to 0.



(a) Training accuracy vs epochs for two DCNNs.

(b) Training loss vs epoch for two DCNNs.

Figure 5: Training phase of the DCNNs.

The testing accuracies and losses of the DCNN components of the SDL model fluctuate but converges at the same time, as shown in figure 7a and figure 7b. The DCNN components from the best SDL model achieved a testing accuracy of 92.85%. It can also be seen the that the best testing accuracy was not achieved on the last epoch, rather it was achieved in an intermediate epoch.

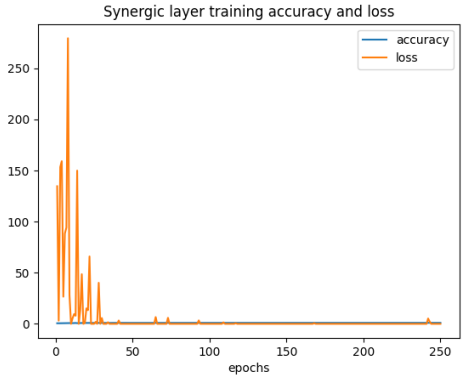
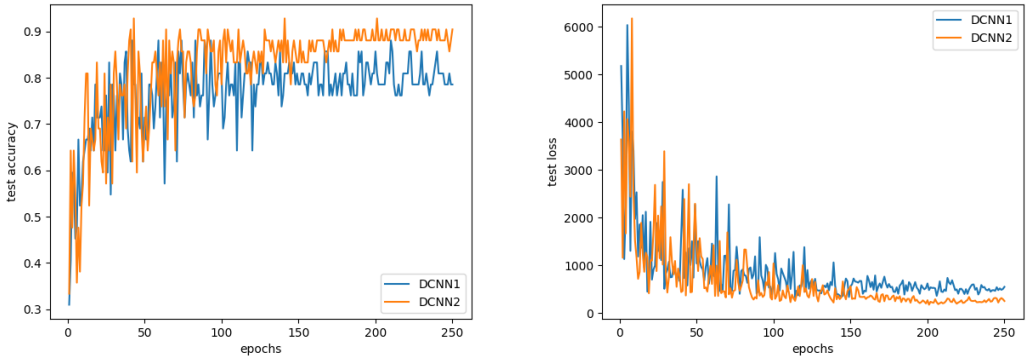


Figure 6: Training phase of the synergic layer.



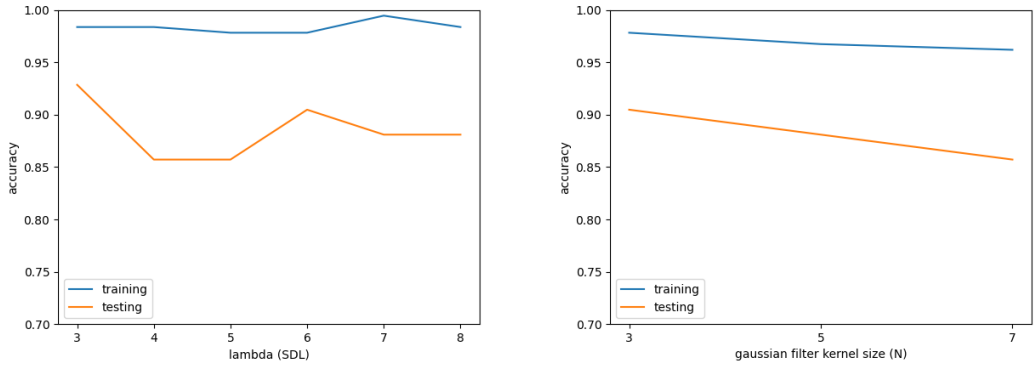
(a) Accuracy vs epochs

(b) Loss vs epoch

Figure 7: Performance Analysis for the proposed model.

Further, we varied the value of the synergic hyper-parameter λ from 3 to 8 and recorded the highest testing and training accuracies achieved by the proposed model. We observed (figure 8a) that as λ increases, the training accuracy varies only by a small factor, whereas the variation in the testing accuracy is noticeable. The given model achieved the training accuracy of 99.45% and the testing accuracy of 92.85%. Upon adding Gaussian Filters to our model, the accuracy went down. More specifically, as the kernel size of the Gaussian Filter increased, the accuracy went down, as shown in figure 8b. Finally, figure 9 depicts the images for the classification results obtained from input MRI image dataset to the proposed classification model. The figure is divided into two horizontal halves, with the upper half containing vertical pair of images with synergic label 1 and the lower half containing vertical pair of images with synergic label 0. Each image includes the original label of the slice and the label predicted by the two DCNNs.

Table 2 compares the SDL model with other pre-trained models and models available in the literature. The implemented model shows comparable accuracy of the proposed model with other state of art models. Table 3 shows the outcomes of various performance metrics obtained for all three grades after the classification and can be observed that model performed constantly in terms of accuracy and other metrics.



(a) Accuracy vs lambda for proposed model. (b) Accuracy vs Gaussian filter size for proposedmodel.

Figure 8: Varying synergic parameter and Gaussian filter size.

Table 2: Comparing the proposed SDL model with popular pre-trained models.

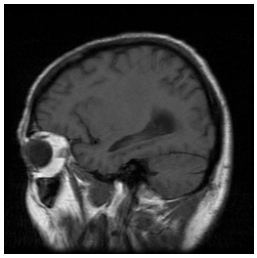
Model	Accuracy	Precision	Specificity	Sensitivity
Proposed model	97.36	**96.05**	**97.19**	**96.07**
CNN model[11]	**98.14**	85.19	92.66	81.11
Sajjad[20]	**90.67**	81.84	80.88	78.89
Kabir[13]	**90.90**	**92.92**	87.04	87.04
Naser 2020[19]	**89.00**	N/A	92.02	87.19

Table 3: Classwise classification report.

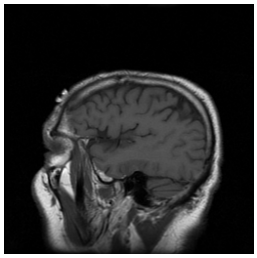
Grade	Accuracy	Precision	Recall	Specificity	Sensitivity	F1 score
II	90.90	90.91	90.91	96.77	90.91	90.91
III	**100**	88.89	**100**	**97.06**	**100**	**94.12**
IV	91.30	**95.45**	91.30	94.74	91.30	93.33

5 Discussion

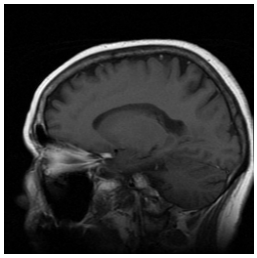
We chose AlexNet over the ResNet models for the backbone of the proposed model to avoid overfitting. The ResNet models are comparatively bigger and have a more complicated architecture than the AlexNet model, which led to the SDL model having a high variance while training. Other smaller models like VGG19 led to the SDL model having a high bias; hence, we decided to use the AlexNet model as the SDL backbone, minimizing the bias-variance tradeoff. An independently fine-tuned AlexNet performed poorly on the testing dataset with an accuracy of 78.57%, average precision of 77.15%, average recall of 74.44%, and average F1 score of 75.10%. But, the same pre-trained model when fine-tuned within the SDL architecture performs exceptionally well and gave good results on both the training and testing data. This can be credited to the fact that the SDL architecture enables 2 AlexNets to mutually learn from each other, allowing them to achieve better metrics than the ones achieved by fine-tuning a single



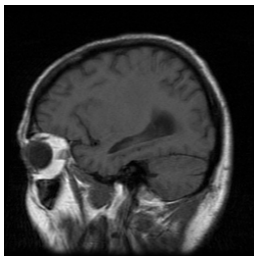
(a1) Label = 4, predictions = (4, 2).



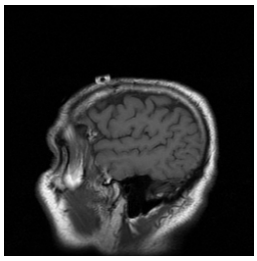
(b1) Label = 2, predictions = (2, 2).



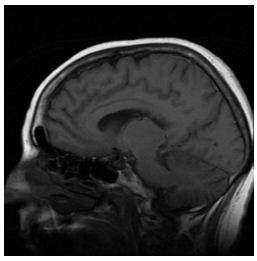
(c1) Label = 3, predictions = (3, 3).



(a2) Label = 4, predictions = (4, 4).

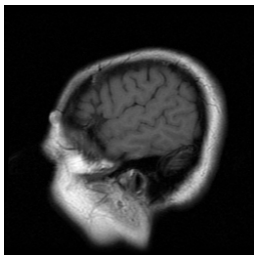


(b2) Label = 2, predictions = (2, 2).

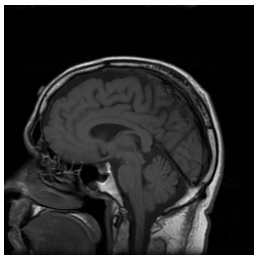


(c2) Label = 3, predictions = (3, 3).

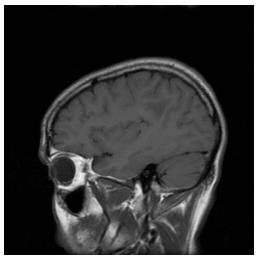
(a) Synergic label = 1 (b) Synergic label = 1 (c) Synergic label = 1



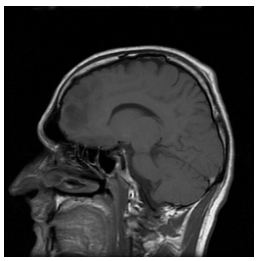
(d1) Label = 2, predictions = (2, 4).



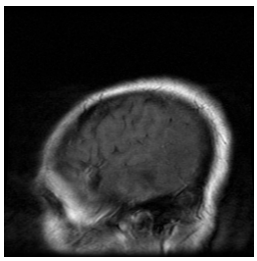
(e1) Label = 2, predictions = (2, 2).



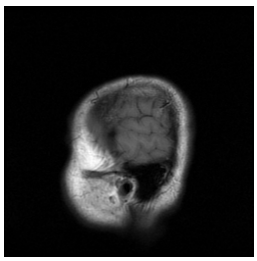
(f1) Label = 3, predictions = (3, 3).



(d2) Label = 3, predictions = (3, 3).



(e2) Label = 4, predictions = (4, 4).



(f2) Label = 4, predictions = (2, 4).

(d) Synergic label = 0 (e) Synergic label = 0 (f) Synergic label = 0

Figure 9: The upper horizontal half contains three vertical pair of images with synergic label 1 and the lower half contains three vertical pair of images with synergic label 0. Each image includes the original label of the slice and the label predicted by the two DCNNs.

AlexNet on the same dataset.

6 Conclusion

This paper presented a Synergic Deep Learning model with an AlexNet backbone for automated grading (II, III, and IV) of glioma tumor MRI scans to overcome the 2 prevalent problems present in medical image classification models - 1. Overfitting, and 2. Significant intra-class variation and inter-class similarity. The work used an extremely small subset of the REMBRANDT dataset, i.e T1-weighted sagittal MRI scans for training the SDL model. The SDL model, even after being trained on a small subset of the REMBRANDT dataset, demonstrated equal proficiency when compared to several other models trained on comparatively larger datasets. Adding a Gaussian filter to the SDL model brought down both, the training and the testing accuracies. Additionally, increasing the kernel size of the Gaussian filter further brought down the accuracy of the SDL model. Further, the best SDL model (at $\lambda = 3$) gave us an average precision of 89.80%, an average recall of 92.3%, and an average F1 score of 90.91%. Finally, the AlexNet model independently performed poorly on the testing dataset (testing accuracy = 78.57%) but performed very well (testing accuracy = 92.85%) when used as the backbone of the SDL model.

7 Future work

This paper restricts the SDL model to include exactly 2 DCNNs in its architecture but we can extend this to the SDL^n model[27], where n is the number of DCNNs used in a single SDL model. Further, The MRI slices can be segmented or cropped such that only the region of interest is visible before they are fed into the model to remove noisy and unnecessary details of the scan. Hence, a joint segmentation classification model can be trained with the classification model being the SDL model. This work can also be extended to a refined dataset of axial MRI scans, allowing the model to learn the features from different angles of an MRI scan. Finally, more sophisticated methods like Reinforcement Learning can be applied to automatically select the best value of the synergic hyper-parameter for classifying the MRI scans.

Funding: This research received no external funding
Conflicts of interest: The author declares no conflict of interest.

References

[1] Babu, B. and Sourirajan, V. (2017). Detection of brain tumour in mri scan images using tetrolet transform and svm classifier. *indian journal of science and technology*, 10:1–10.

[2] Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., and Davatzikos, C. (2017a). Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific Data*, 4.

[3] Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., Freymann, J., Farahani, K., and Davatzikos, C. (2017b). Segmentation labels and radiomic fea-

tures for the pre-operative scans of the TCGA-GBM collection. *The Cancer Imaging Archive*.

[4] Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., and Prior, F. (2013). The cancer imaging archive (TCIA): Maintaining and operating a public information repository. *Journal of Digital Imaging*, 26:1045–1057.

[5] Codella, N. C. F., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., and Halpern, A. (2018). Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC).

[6] García Seco de Herrera, A., Müller, H., and Bromuri, S. (2015). Overview of the ImageCLEF 2015 medical classification task. In *Working Notes of CLEF 2015 (Cross Language Evaluation Forum)*.

[7] García Seco de Herrera, A., Schaer, R., Bromuri, S., and Müller, H. (2016). Overview of the ImageCLEF 2016 medical task. In *Working Notes of CLEF 2016 (Cross Language Evaluation Forum)*.

[8] Gutman, D., Codella, N. C. F., Celebi, E., Helba, B., Marchetti, M., Mishra, N., and Halpern, A. (2016). Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic).

[9] Gutta, S., Acharya, J., Shiroishi, M., Hwang, D., and Nayak, K. (2021). Improved glioma grading using deep convolutional neural networks. *American Journal of Neuroradiology*, 42(2):233–239.

[10] Habib, H., Amin, R., Ahmed, B., and Hannan, A. (2021). Hybrid algorithms for brain tumor segmentation, classification and feature extraction. *Journal of Ambient Intelligence and Humanized Computing*, 13(2).

[11] Irmak, E. (2021). Multi-classification of brain tumor mri images using deep convolutional neural network with fully optimized framework. *Iranian Journal of Science and Technology*, 45:1015–1036.

[12] Jiang, Y. and Uhrbom, L. (2012). On the origin of glioma. *Upsala journal of medical sciences*, 117:113–21.

[13] Kabir, A. A., Ayati, M., and Kazemi, F. (2019). Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. *Biocybern Biomedical Engineering*, 39(1):63–74.

[14] Kathiresan, S., Sait, A. R. W., Gupta, D., S.K, L., Khanna, A., and Pandey, H. M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. *Pattern Recognition Letters (Elsevier)*, 133:210–216.

[15] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C., Bottou, L.,

and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.

[16] Krommweh, J. (2010). Tetrolet transform: A new adaptive haar wavelet algorithm for sparse image representation. *Journal of Visual Communication and Image Representation*, 21:364–374.

[17] Mesfin, F. and Al-Dhahir, M. (2021). *Gliomas*. StatPearls (TreasureIsland, FL: StatPearls publishing).

[18] Nalepa, J., Kotowski, K., Machura, B., Adamski, S., Bozek, O., Eksner, B., Kokoszka, B., Pekala, T., Radom, M., Strzelczak, M., Zarudzki, L., Krason, A., Arcadu, F., and Tessier, J. (2023). Deep learning automates bidimensional and volumetric tumor burden measurement from MRI in pre- and post-operative glioblastoma patients. *Computers in Biology and Medicine*, 154:1–19.

[19] Naser, M. A. and Deen, M. J. (2020). Brain tumor segmentation and grading of lower-grade glioma using deep learning in mri images. *Computers in Biology and Medicine*, 121.

[20] Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., and Baik, S. W. (2019). Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *Journal of Computational Science*, 30:174–182.

[21] Scarpace, L., Flanders, A. E., Jain, R., Mikkelsen, T., and Andrews, D. W. (2019). Data from REMBRANDT [data set]. *The Cancer Imaging Archive*.

[22] Stoyanov, G. and Dzhenkov, D. (2017). On the concepts and history of glioblastoma multiforme-morphology, genetics and epigenetics. *Folia medica*, online:ahead of print.

[23] Tang, X. (2020). The role of artificial intelligence in medical imaging research. *BJR|Open*, 2(1):20190031.

[24] Upadhyay, N. and Waldman, A. (2011). Conventional MRI evaluation of gliomas. *The British journal of radiology*, 84 Spec No 2:S107–11.

[25] Xiao, T., Hua, W., Li, C., and Wang, S. (2019). Glioma grading prediction by exploring radiomics and deep learning features. In *Proceedings of the Third International Symposium on Image Computing and Digital Medicine*, ISICDM 2019, pages 208–213, New York, NY, USA. Association for Computing Machinery.

[26] Zhang, J., Xie, Y., Wu, Q., and Xia, Y. (2018). Skin lesion classification in dermoscopy images using synergic deep learning. In Frangi, A. F., Schnabel, J. A., Davatzikos, C., Alberola-López, C., and Fichtinger, G., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pages 12–20, Cham. Springer International Publishing.

[27] Zhang, J., Xie, Y., Wu, Q., and Xia, Y. (2019). Medical image classification using synergic deep learning. *Medical Image Analysis*, 54:10–19.