A Customized VGG19 Network with Concatenation of Deep and Handcrafted Features for Brain Tumor Detection

Venkatesan Rajinikanth¹, Alex Noel Joseph Raj², Krishnan Palani Thanaraj¹ and Ganesh R. Naik³,*

¹ Department of Electronics and Instrumentation Engineering, St. Joseph’s College of Engineering, Chennai 600119, India; v.rajinikanth@ieee.org (V.R.); palanithanaraj.k@gmail.com (K.P.T.)
² Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Department of Electronic Engineering, College of Engineering, Shantou University, Shantou 515063, China; jalexnoel@stu.edu.cn
³ MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Milperra, NSW 2560, Australia
* Correspondence: Ganesh.Naik@westernsydney.edu.au

Received: 22 March 2020; Accepted: 12 May 2020; Published: 15 May 2020

Abstract: Brain tumor (BT) is one of the brain abnormalities which arises due to various reasons. The unrecognized and untreated BT will increase the morbidity and mortality rates. The clinical level assessment of BT is normally performed using the bio-imaging technique, and MRI-assisted brain screening is one of the universal techniques. The proposed work aims to develop a deep learning architecture (DLA) to support the automated detection of BT using two-dimensional MRI slices. This work proposes the following DLAs to detect the BT: (i) implementing the pre-trained DLAs, such as AlexNet, VGG16, VGG19, ResNet50 and ResNet101 with the deep-features-based SoftMax classifier; (ii) pre-trained DLAs with deep-features-based classification using decision tree (DT), k nearest neighbor (KNN), SVM-linear and SVM-RBF; and (iii) a customized VGG19 network with serially-fused deep-features and handcrafted-features to improve the BT detection accuracy. The experimental investigation was separately executed using Flair, T2 and T1C modality MRI slices, and a ten-fold cross validation was implemented to substantiate the performance of proposed DLA. The results of this work confirm that the VGG19 with SVM-RBF helped to attain better classification accuracy with Flair (>99%), T2 (>98%), T1C (>97%) and clinical images (>98%).

Keywords: brain tumor; brain MRI slices; VGG19; deep features; handcrafted-features; features concatenation; SVM-RBF

1. Introduction

The brain is one of the primary organs in humans, and it assesses the complete physiological gestures coming from other sensory parts and takes the necessary control measures. The normal operations of brain are badly affected if any infection or disease arises, and an unnoticed and untreated abnormality may lead to various difficulties, including death [1,2].

The regular state of the brain may be exaggerated due to different reasons, such as birth defects, a head injury due to an accident or uncontrolled cell growth (UCG) in a central brain section [3,4]. An irregularity will cause various problems in the physiological system and the untreated brain abnormality also will lead to several major illnesses. A brain abnormality due to UCG is a major threat, and the untreated growth will lead to brain-cancer, which is one of the rapidly increasing cancer burdens globally. The work of Louis et al. [5] clearly discusses the ranking/classification of brain tumors (BTs) as per the 2016 report of World Health Organization (WHO).
Recently, a substantial number of awareness programs have been initiated to protect people from such abnormalities. However, because of different unavoidable causes, such as contemporary lifestyles, food behavior, heredity factors and age, most individuals are suffering due to developed BT [6,7]. If the BT is detected at a premature stage; a promising treatment can be employed to heal/manage the cell growth. The clinical level detection of BT is performed with; (i) single/multi-channel EEG signals and (ii) brain imaging techniques. The image-assisted technique provides more meaningful information compared to the signal assisted technique. Hence, in most clinical-level detection, the imaging procedures are widely preferred, and the image recording procedures, such as computed tomography (CT) and magnetic resonance imaging (MRI), are widely considered to record and check the brain abnormalities using three-dimensional (3D) and 2D images. Compared to CT, MRI is widely preferred due to its varied modalities, and the visibility of the BT in a brain MRI is very clear compared to the CT. Hence, MRIs are largely preferred to evaluate the various brain abnormalities, including the BT. BT evaluation with the modalities such as Flair, T2 and T1C has enhanced tumor visibility compared to the T1 and diffused weight (DW) modality [8–12].

In the literature, a significant amount of conventional and modern BT detection procedures are proposed and implemented by researchers with a chosen machine learning (ML) or deep learning (DL) technique [13–18]. The chief aim of the existing automated and semi-automated disease evaluation procedure is to develop an accurate disease detection system to assist the doctor during the diagnosis and treatment-planning process. Most of the new disease diagnosis systems implement the DL technique due to its superiority and detection accuracy. The work of Talo et al. [19] implemented a transfer-learning-based deep learning architecture (DLA) to detect tumors using 2D MRI slices and achieved a classification accuracy of >98%. Further, Talo et al. [20] presented a detailed analysis on the existing DLA in the literatures and confirmed that the ResNet50 offers a better classification accuracy (>95%) during the brain tumor detection process. Amin et al. [21] implemented a brain tumor evaluation procedure using BRATS2013, 2015 and clinical database and achieved an accuracy of >98%. The work of Sharif et al. [22] implemented an enhanced binominal thresholding and multi-features selection-based technique to classify the brain tumor and attained enhanced result. The work of Fabelo et al. [23] implemented a DLA to detect the glioblastoma using hyperspectral 3D and 2D brain images. Sajid et al. [24] implemented a DL based brain tumor detection procedure and attained better values of sensitivity and specificity. The higher order spectra feature-based detection and classification of the abnormal section in a brain MRI is discussed by Acharya et al. [25]. Further, a considerable number of approaches are proposed and implemented by a considerable number of researchers to improve the detection accuracy on a class of brain MRI images ranging from the benchmark datasets and clinical images [26–31].

The work in this paper aimed to evaluate the performances of the existing DLAs, such as AlexNet, VGG16, VGG19, ResNet50 and ResNet101 [20] for detecting the BT using the 2D brain MRI slices. Initial detection was performed with a transfer learning procedure using the attained deep-features and the SoftMax classifier. After finding the best suitable DLA for the considered task, the concatenation of the deep and hand-crafted features was performed to enhance the classification accuracy with the SoftMax classifier. Further, a detailed comparative study with other classifiers, such as random forest (RF), decision tree (DT), k-nearest neighbor (KNN), SVM-linear and SVM-RBF was also performed to attain better classification accuracy.

For the experimental investigation, the 2D brain MRIs recorded using Flair, T2 and T1C images of the dimensions 227 × 227 × 1 were considered, and the essential images for this task were collected using the benchmark datasets, such as BRATS (without skull section) [32,33] and TCIA (with skull section) [34–37] to train, test and validate the DLAs. Further, a clinical level dataset [38] with the skull was considered to validate the DLA. This clinical level dataset was already used to test the machine-learning systems [2–4]. In this work, the performance of the proposed system was confirmed by computing the accuracy, precision, sensitivity, specificity, F1-Score and negative predictive value
(NPV). This work also implemented a ten-fold cross validation, and based on the average value, the performance of the proposed DLA was confirmed.

The remaining parts of this work are set out as follows: Section 2 presents the materials and methods. Section 3 presents the details of experimental investigation, and conclusion of this work is discussed in Section 4.

2. Materials and Methods

In the literature, a number of DLAs are proposed to detect the abnormalities in medical images using conventional and customized DLAs [17,39–41]. Development of a new DLA from the scratch is complex and requires complex work to build, train, test and validate the architecture for a chosen problem. Hence, most of the earlier works adapt the proven DLAs existing in the literature to solve a disease detection problem. Furthermore, selecting and implementing a particular architecture requires prior knowledge about its structure, complexity in implementation, initial tuning and validation procedures [17,39–42].

This work initially considers the existing DLAs discussed in [20] to detect the brain tumors from the considered MRI database. This work employs the transfer learning concept to train, test and validate the adopted DLAs using the SoftMax classifier. The DLA which offered the enhanced classification accuracy is then selected and its performance is further enhanced using the proposed method. The initial experimental outcome of this research is confirmed: the VGG19 offers better classification accuracy compared to the alternatives, and hence conventional and customized VGG19 are then considered in this research to attain the better tumor detection accuracy.

After selecting the VGG19 to solve the considered image examination problem, its performance enhancement is tried using the following approaches: (i) replacing the SoftMax classifier with DT, KNN, SVM-Linear and SVM-RBF classifiers, and (ii) enhancing the outcome of VGG19 using a new feature vector obtained by fusing the handcrafted and deep features. The earlier research work confirms that if the feature vector is enhanced, then the pre-trained DLA will offer better detection accuracy compared to the conventional DLA [1,29,30]. Figure 1 depicts the customized VGG19 proposed and implemented in this research. The pre-trained DLA will provide a one-dimensional deep-feature vector, and the handcrafted texture features are attained using counterlet transform (COT) (38 features), curvelet transform (CUT) (121 features) and discrete wavelet transform (DWT) (40 features) [43]. The existing deep and handcrafted features are sorted and serially combined based on principal component analysis (PCA), and these features are then used to train, test and validate the classifier unit, which separates the given image into normal/tumor classes [1].

![Image](image1.png)

**Figure 1.** Structure of the customized VGG19 network.

2.1. Image Collection and Processing

In every medical evaluation procedure, the performance of the developed diagnosis system depends mainly on the database considered based on the problem to be solved. To solve the brain tumor detection problem, the most commonly considered images are attained from the well-known benchmark images of the Multimodal Brain Tumor Segmentation Challenge (BRATS) [32,33]. Figure 2 depicts the image dataset considered in this work along with the available MRI modalities. BRATS and TCIA will not have the MRI modalities, such as DW and T1C respectively, and hence, it is denoted as not available (NA) in Figure 2. In this work, the 2D MRI slices of Flair, T2 and T1C are considered for the assessment [2–4,6].
not available (NA) in Figure 2. In this work, the 2D MRI slices of Flair, T2 and T1C are considered for the assessment [2–4,6].

The Cancer Imaging Archive (TCIA) also provides clinical grade medical images for research purposes [34–37]. In this work, the glioma images associated with the skull section are chosen for the assessment. Furthermore, the clinical grade MRIs collected from the Proscons Ltd are considered to validate the proposed DLA [2–4,38]. In this work, the data augmentation is implemented to increase the BRATS and TCIA images with the help of image-flip and image-rotate (90° left/right) operations. This procedure helped to achieve the considerable number of test images for both the normal and tumor classes.

Table 1 presents the details of the test image datasets and its modalities considered in this work. This table also depicts the images considered for training and testing the classifier unit.

<table>
<thead>
<tr>
<th>Image Class</th>
<th>Modality</th>
<th>Number of Images for Training</th>
<th>Number of Images for Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Mixed (Flair+T2)</td>
<td>1000</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>1500</td>
<td>600</td>
</tr>
<tr>
<td>Abnormal</td>
<td>T2</td>
<td>1500</td>
<td>600</td>
</tr>
<tr>
<td>(BRATS)</td>
<td>T1C</td>
<td>1500</td>
<td>600</td>
</tr>
<tr>
<td>Abnormal</td>
<td>T2</td>
<td>1000</td>
<td>400</td>
</tr>
<tr>
<td>(TCIA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abnormal</td>
<td>T2</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>(Clinical)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.2. Handcrafted Feature Extraction

In ML and DL techniques, feature extraction is the principal procedure which helps to extract the meaningful information from the image based on its shape and the texture values. Based on these features, the implemented classifier units are trained, tested and validated. In the literature, a substantial number of feature extraction techniques are implemented for a class of RGB/gray scaled pictures [44–49]. The implemented CUT, COT and DWT helped to get a sum of 199 features, which were then fused with the deep-feature vector to improve its feature dimension. The features extracted with the conventional scheme are called the handcrafted features and every approach provides a 1D feature vector. The feature vectors of chosen procedures are represented as: $FV_1 = 1 \times 1 \times 38$, $FV_2 = 1 \times 1 \times 121$, ...
$FV_3 = 1 \times 1 \times 40$ and the handcrafted-feature vector $FV_h = FV_1 + FV_2 = FV_3 = 1 \times 1 \times 199$. The other details on the feature extraction and selection can be found in [1,30].

2.3. Feature Selection and Concatenation

The major objective of this research was to achieve a solitary feature vector by fusing the $FV_h$ with deep features (DF) of the considered DLA. In the literature, features are fused with the serial and parallel process and in this work; serial concatenation is adopted due to its simplicity and the $FV_h$ to be combined with the DF selected and sorted based on the PCA. Normally, PCA transfer $n$-vectors $(p_1, p_2, \ldots, p_n)$ of $D$-dimensional space into $D'$ space with values $(p'_1, p'_2, \ldots, p'_n)$ where $D$ and $D'$ are positive integers with size $D' \leq D$.

The new feature of PCA can be represented as:

$$f_n' = \sum_{k=1}^{\nu} R_{k,i} S_k$$  \hspace{1cm} (1)

where $S_k = $ eigenvectors and $R_{k,i} = $ primary components [1,26,29,30]. After implementing the feature concatenation, the final feature vector will be $(1 \times 1 \times 199) + (1 \times 1 \times 1024) = 1 \times 1 \times 1223$.

2.4. Classification

The general performance of the DLA relies mainly on the classifiers employed to classify the considered images into normal/tumor class. The traditional DLA considers the SoftMax, which provides a reasonable accuracy with using transfer learning approach. Further, the performance of the DLA can be enhanced by employing appropriate classifiers [50–53].

In this work, the SoftMax classifier is replaced with other classifiers, discussed below:

- Decision tree: DT is one of the more famous methodologies used to categorize the linear and non-linear information with a sequence of testing methods, which expands like a tree. The DT utilizes a quality exploration situation as the root and internal nodes, and the class label forms terminal nodes. Once a DT has been shaped, categorization is accomplished by the conclusions taken in each branch of the tree. Other particulars of DT can be found in [43–46].

- K-nearest neighbor: KNN is a well-known technique often considered to classify medical images based on an existing feature set. In this work, KNN is considered to classify the brain MRIs of varied modality. During the classification task, the KNN evaluates the space among new features to each training feature and discovers the best neighbor. The earlier works on the KNN can be found in [43–46].

- Support vector machine: SVM categorizer uses a hyperplane for labeling of dataset based on features gathered throughout the training stage. SVM is one of the most frequently used to categorize MRI images. Radial basis function-based SVM (SVM-RBF) is used to sort the 2D MRI with the elected features. In SVM-RBF, the kernel value is controlled by a scaling parameter “$\sigma$”; and this value is varied from 0.2 to 1.9 with a step size of 0.1. Furthermore, the SVM with linear polynomial kernel (SVM-Linear) is also adopted to grade the MRI database [43].

2.5. Performance Measures and Validation

The evaluation of the performance of classifiers is normally carried by computing the essential performance values. In this work, the classifier performance is to be assessed based on a chosen performance values. The initial assessment computes the essential measures, such as true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) values [43–45,54].

$$\text{Accuracy} = ACC = \frac{TP + TN}{TP + TN + FP + FN}$$  \hspace{1cm} (2)
Precision \( = \text{PRE} = \frac{TP}{TP + FP} \)  
(3)

Sensitivity \( = \text{SEN} = \frac{TP}{TP + FN} \)  
(4)

Specificity \( = \text{SPE} = \frac{TN}{TN + FP} \)  
(5)

F1-score \( = F1S = \frac{2TP}{2TP + FN + FP} \)  
(6)

Negative Predictive Value \( = \text{NPV} = \frac{TN}{TN + FN} \)  
(7)

3. Experimental Outcome and Discussions

This part of the work presents the experimental results and discussions. This work is executed using the workstation i5 processor, 8 GB RAM and 2 GB VRAM within a Matlab environment. During this work, the following initial values are assigned for every DLA: epoch size = 55, iteration size = 1200, iteration per epoch = 110, updating frequency = five iterations, learning error rate = 1e-5, stopping criteria = best validation or maximum iteration.

Initially, the DLAs such as AlexNet, VGG16, VGG19, ResNet50 and ResNet101 are considered to examine and classify the dataset into normal/tumor classes using SoftMax classifier trained and tested with the deep-features. Initially, the AlexNet is implemented to solve the classification problem and the attained results are depicted in Figure 3. Figure 3a, b presents the accuracy and the loss value attained during the training and testing, respectively. A similar procedure is repeated with the other DLAs and the corresponding results attained are depicted in Table 2. The results available in this table confirm that the performance values attained with the VGG19 DLA are better for all the MRI modalities, such as Flair, T2 and T1C. This result confirms that, for the chosen dataset, the VGG19 offered better results compared to the alternatives, and hence, the VGG19 is then considered in this work for further enhancement to attain a better accuracy using the proposed methodology.

![Figure 3. Convergence of the training and validation process of AlexNet. (a) Accuracy. (b) Loss value.](image_url)
Table 2. Performance measures attained with conventional DLA with SoftMax classifier.

<table>
<thead>
<tr>
<th>Network</th>
<th>Modality</th>
<th>TP</th>
<th>FN</th>
<th>TN</th>
<th>FP</th>
<th>ACC</th>
<th>PRE</th>
<th>SEN</th>
<th>SPE</th>
<th>F1S</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>AlexNet</td>
<td>Flair</td>
<td>377</td>
<td>23</td>
<td>581</td>
<td>19</td>
<td>95.80</td>
<td>95.20</td>
<td>94.25</td>
<td>96.83</td>
<td>94.72</td>
<td>96.19</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>368</td>
<td>32</td>
<td>573</td>
<td>27</td>
<td>94.10</td>
<td>93.16</td>
<td>92.00</td>
<td>95.50</td>
<td>92.58</td>
<td>94.71</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>361</td>
<td>39</td>
<td>568</td>
<td>32</td>
<td>92.90</td>
<td>91.86</td>
<td>90.25</td>
<td>94.67</td>
<td>91.05</td>
<td>93.57</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>379</td>
<td>21</td>
<td>585</td>
<td>15</td>
<td>96.40</td>
<td>96.19</td>
<td>94.75</td>
<td>97.50</td>
<td>95.47</td>
<td>96.53</td>
</tr>
<tr>
<td>VGG-16</td>
<td>T2</td>
<td>364</td>
<td>36</td>
<td>570</td>
<td>30</td>
<td>93.40</td>
<td>92.39</td>
<td>91.00</td>
<td>96.83</td>
<td>94.68</td>
<td>94.06</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>368</td>
<td>32</td>
<td>564</td>
<td>36</td>
<td>93.20</td>
<td>91.09</td>
<td>92.00</td>
<td>95.50</td>
<td>92.58</td>
<td>94.71</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>381</td>
<td>19</td>
<td>584</td>
<td>16</td>
<td>96.50</td>
<td>95.97</td>
<td>95.25</td>
<td>97.33</td>
<td>96.61</td>
<td>96.85</td>
</tr>
<tr>
<td>VGG-19</td>
<td>T2</td>
<td>378</td>
<td>22</td>
<td>581</td>
<td>19</td>
<td>95.90</td>
<td>95.21</td>
<td>94.50</td>
<td>96.83</td>
<td>94.86</td>
<td>96.35</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>370</td>
<td>30</td>
<td>574</td>
<td>26</td>
<td>94.40</td>
<td>93.43</td>
<td>92.50</td>
<td>95.67</td>
<td>92.96</td>
<td>95.03</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>371</td>
<td>29</td>
<td>567</td>
<td>33</td>
<td>93.80</td>
<td>91.83</td>
<td>92.75</td>
<td>94.50</td>
<td>92.29</td>
<td>95.13</td>
</tr>
<tr>
<td>ResNet-50</td>
<td>T2</td>
<td>358</td>
<td>42</td>
<td>565</td>
<td>35</td>
<td>92.30</td>
<td>91.09</td>
<td>89.50</td>
<td>94.17</td>
<td>92.49</td>
<td>94.08</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>362</td>
<td>38</td>
<td>566</td>
<td>34</td>
<td>92.80</td>
<td>91.41</td>
<td>90.50</td>
<td>94.33</td>
<td>92.95</td>
<td>94.71</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>374</td>
<td>26</td>
<td>583</td>
<td>17</td>
<td>95.70</td>
<td>95.65</td>
<td>93.50</td>
<td>97.17</td>
<td>95.46</td>
<td>96.73</td>
</tr>
<tr>
<td>ResNet-101</td>
<td>T2</td>
<td>361</td>
<td>39</td>
<td>568</td>
<td>32</td>
<td>92.90</td>
<td>91.85</td>
<td>90.25</td>
<td>94.67</td>
<td>91.05</td>
<td>93.57</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>366</td>
<td>34</td>
<td>565</td>
<td>35</td>
<td>93.10</td>
<td>91.27</td>
<td>91.50</td>
<td>94.17</td>
<td>91.39</td>
<td>94.32</td>
</tr>
</tbody>
</table>

The performance of the chosen VGG-19 is then verified by replacing the SoftMax classifier with other approaches, such as DT, KNN, SVM-Linear and SVM-RBF, and the obtained results are depicted in Table 3. This table confirms that the pre-trained VGG-19 offered a classification accuracy of 96.70% for Flair modality (SVM-RBF), 96.10% for T2 modality (DT) and 94.60% for the T1C modality MRI slices. These results confirm that, if the SoftMax is replaced with a chosen classifier unit, the detection accuracy can be improved.

Table 3. Performance values obtained with VGG-19 with chosen classifiers.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Modality</th>
<th>TP</th>
<th>FN</th>
<th>TN</th>
<th>FP</th>
<th>ACC</th>
<th>PRE</th>
<th>SEN</th>
<th>SPE</th>
<th>F1S</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>DT</td>
<td>Flair</td>
<td>379</td>
<td>21</td>
<td>587</td>
<td>13</td>
<td>96.60</td>
<td>96.68</td>
<td>94.75</td>
<td>97.83</td>
<td>95.71</td>
<td>96.55</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>377</td>
<td>23</td>
<td>584</td>
<td>16</td>
<td>96.10</td>
<td>95.93</td>
<td>94.25</td>
<td>97.33</td>
<td>95.08</td>
<td>96.21</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>368</td>
<td>32</td>
<td>571</td>
<td>29</td>
<td>93.90</td>
<td>92.69</td>
<td>92.00</td>
<td>95.17</td>
<td>92.35</td>
<td>94.69</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>369</td>
<td>31</td>
<td>581</td>
<td>19</td>
<td>95.00</td>
<td>95.10</td>
<td>92.25</td>
<td>96.83</td>
<td>93.65</td>
<td>94.93</td>
</tr>
<tr>
<td>KNN</td>
<td>T2</td>
<td>371</td>
<td>29</td>
<td>580</td>
<td>20</td>
<td>95.10</td>
<td>94.88</td>
<td>92.75</td>
<td>96.67</td>
<td>93.80</td>
<td>95.24</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>368</td>
<td>32</td>
<td>578</td>
<td>22</td>
<td>94.60</td>
<td>94.36</td>
<td>92.00</td>
<td>96.33</td>
<td>93.16</td>
<td>94.75</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>370</td>
<td>30</td>
<td>583</td>
<td>17</td>
<td>95.30</td>
<td>95.61</td>
<td>92.50</td>
<td>97.17</td>
<td>94.03</td>
<td>95.11</td>
</tr>
<tr>
<td>SVM-Linear</td>
<td>T2</td>
<td>374</td>
<td>26</td>
<td>579</td>
<td>21</td>
<td>95.30</td>
<td>94.68</td>
<td>93.50</td>
<td>96.50</td>
<td>94.09</td>
<td>95.70</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>372</td>
<td>28</td>
<td>571</td>
<td>29</td>
<td>94.30</td>
<td>92.77</td>
<td>93.00</td>
<td>95.17</td>
<td>92.88</td>
<td>95.33</td>
</tr>
<tr>
<td></td>
<td>Flair</td>
<td>378</td>
<td>22</td>
<td>589</td>
<td>11</td>
<td>96.70</td>
<td>97.17</td>
<td>94.50</td>
<td>98.17</td>
<td>95.82</td>
<td>96.40</td>
</tr>
<tr>
<td>SVM-RBF</td>
<td>T2</td>
<td>374</td>
<td>26</td>
<td>582</td>
<td>18</td>
<td>95.60</td>
<td>95.41</td>
<td>93.50</td>
<td>97.00</td>
<td>94.44</td>
<td>95.72</td>
</tr>
<tr>
<td></td>
<td>T1C</td>
<td>369</td>
<td>31</td>
<td>574</td>
<td>26</td>
<td>94.30</td>
<td>93.42</td>
<td>92.25</td>
<td>95.67</td>
<td>92.83</td>
<td>94.88</td>
</tr>
</tbody>
</table>

The performance of the traditional VGG-19 is further improved by combining the deep-features with the handcrafted features ($F_{V_h}$). During this process, a serial concatenation procedure is implemented to combine the deep-features of dimension $1 \times 1 \times 1024$ with the $F_{V_h}$ of dimension $1 \times 1 \times 199$ to attain a new feature vector size of $1 \times 1 \times 1223$. This feature set is then considered to train, test and validate the classifier implemented in the VGG-19 network. The sample results attained with various layers of the VGG-19 are presented in Figure 4 and the area under the curve (AUC) of 98.5% attained using the VGG-19 with SVM-RBF classifier is presented in Figure 5.
Table 4 shows the overall performance measures achieved after a 10-fold cross validation for the BRATS, TCIA and clinical brain MRI datasets, and averages of the performance measures are considered for the validation. The results shown in this table confirm that the proposed network helped to achieve better classification accuracy for all the considered datasets irrespective of the modalities of the test images. Figure 6 depicts the sample classification result attained with the proposed VGG19 with SVM-RBF classifier for the TCIA database. The TCIA database consists of the brain MRI slices with the skull section, and hence, the average detection accuracy attained for the TCIA database is less compared with the results attained with the BRATS database. Figure 7 depicts the overall results attained with the customized VGG19 for the considered test images, and these results confirm that the VGG19 works well on the considered image dataset. Further, this architecture helped to attain classification accuracy of 98.17%. This outcome confirmed that the proposed DLA is clinically significant, and in future, it can be considered to detect tumors in clinical-grade brain MRI slices. During the real time implementation, this DLA can be used as an assisting tool for the doctor to make the possible decisions during brain tumor detection and treatment-planning processes.

Table 4. Performance values obtained with the proposed VGG19 architecture.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Classifier</th>
<th>Modality</th>
<th>TP</th>
<th>FN</th>
<th>TN</th>
<th>FP</th>
<th>ACC</th>
<th>PRE</th>
<th>SEN</th>
<th>SPE</th>
<th>F1S</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRATS</td>
<td>DT</td>
<td>Flair</td>
<td>388</td>
<td>12</td>
<td>589</td>
<td>11</td>
<td>97.70</td>
<td>97.24</td>
<td>97.00</td>
<td>98.17</td>
<td>97.12</td>
<td>98.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2</td>
<td>385</td>
<td>15</td>
<td>586</td>
<td>14</td>
<td>97.10</td>
<td>96.49</td>
<td>96.25</td>
<td>97.67</td>
<td>96.37</td>
<td>97.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1C</td>
<td>383</td>
<td>17</td>
<td>581</td>
<td>19</td>
<td>96.40</td>
<td>95.27</td>
<td>95.75</td>
<td>96.83</td>
<td>95.51</td>
<td>97.16</td>
</tr>
<tr>
<td></td>
<td>KNN</td>
<td>Flair</td>
<td>387</td>
<td>13</td>
<td>591</td>
<td>9</td>
<td>97.80</td>
<td>97.73</td>
<td>96.75</td>
<td>98.50</td>
<td>97.24</td>
<td>97.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2</td>
<td>384</td>
<td>16</td>
<td>588</td>
<td>12</td>
<td>97.20</td>
<td>96.96</td>
<td>96.00</td>
<td>98.00</td>
<td>96.48</td>
<td>97.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1C</td>
<td>385</td>
<td>15</td>
<td>589</td>
<td>11</td>
<td>97.40</td>
<td>97.22</td>
<td>96.25</td>
<td>98.17</td>
<td>96.73</td>
<td>97.52</td>
</tr>
<tr>
<td></td>
<td>SVM-Linear</td>
<td>Flair</td>
<td>392</td>
<td>8</td>
<td>592</td>
<td>8</td>
<td>98.40</td>
<td>98.00</td>
<td>98.00</td>
<td>98.67</td>
<td>98.00</td>
<td>98.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2</td>
<td>391</td>
<td>9</td>
<td>590</td>
<td>10</td>
<td>98.10</td>
<td>97.50</td>
<td>97.75</td>
<td>98.33</td>
<td>97.63</td>
<td>98.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1C</td>
<td>384</td>
<td>16</td>
<td>587</td>
<td>13</td>
<td>97.10</td>
<td>96.72</td>
<td>96.00</td>
<td>97.83</td>
<td>96.36</td>
<td>97.34</td>
</tr>
<tr>
<td></td>
<td>SVM-RBF</td>
<td>Flair</td>
<td>395</td>
<td>5</td>
<td>596</td>
<td>4</td>
<td>99.10</td>
<td>99.00</td>
<td>98.75</td>
<td>99.33</td>
<td>98.87</td>
<td>99.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2</td>
<td>394</td>
<td>6</td>
<td>595</td>
<td>5</td>
<td>98.90</td>
<td>98.74</td>
<td>98.50</td>
<td>99.17</td>
<td>98.62</td>
<td>99.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1C</td>
<td>389</td>
<td>11</td>
<td>588</td>
<td>12</td>
<td>97.70</td>
<td>97.01</td>
<td>97.25</td>
<td>98.00</td>
<td>97.13</td>
<td>98.16</td>
</tr>
<tr>
<td></td>
<td>TCIA</td>
<td>SVM-RBF</td>
<td>393</td>
<td>7</td>
<td>391</td>
<td>9</td>
<td>98.00</td>
<td>97.76</td>
<td>97.25</td>
<td>97.75</td>
<td>98.01</td>
<td>98.24</td>
</tr>
<tr>
<td></td>
<td>Clinical</td>
<td>SVM-RBF</td>
<td>395</td>
<td>5</td>
<td>194</td>
<td>6</td>
<td>98.17</td>
<td>98.50</td>
<td>98.75</td>
<td>97.00</td>
<td>98.63</td>
<td>97.49</td>
</tr>
</tbody>
</table>
The performance of the proposed DLA is then validated with other methods existing in the literature, and the results are presented in Table 5. These results confirm that customized VGG19 offers enhanced outcomes for the BRATS, TCIA and clinical level images compared to the results of existing methods.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Approach</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>98 (Naïve-Bayes)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>86 (Ensemble)</td>
</tr>
<tr>
<td>Mallick et al. [28]</td>
<td>Deep neural network</td>
<td>97 (DT)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>97 (KNN)</td>
</tr>
<tr>
<td>Sharif et al. [30]</td>
<td>Deep learning with feature fusion</td>
<td>97.8 (SoftMax)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>93.6 (MSVM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>92.2 (KNN)</td>
</tr>
<tr>
<td>Gudigar et al. [17]</td>
<td>Shearlet transform + texture + PSO SVM</td>
<td>94.4 (Ensemble)</td>
</tr>
<tr>
<td>Khawaldeh et al. [42]</td>
<td>Convolutional neural networks</td>
<td>97.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>91.16</td>
</tr>
</tbody>
</table>

This study was focused on implementing a DLA to segregate MRI slices into normal/tumor classes, and after the segregation, the MRI slices with the tumor were further examined by the doctor. In future,
the outcome of the proposed system along with the clinically collected data can be used to develop a
computerized model to track the ependymal tumor dissemination.

The future scope of the proposed research includes:

(i) Enhancing the handcrafted feature vector by considering the additional texture and shape features.
(ii) Adjusting the fully-connected and drop-out layers to improve the categorization accuracy.
(iii) Improving the feature-concatenation technique to attain better results.
(iv) Implementing the proposed VGG19 DLA to classify the tumors into low/high grade gliomas.
(v) Developing a neural-network model for ependymal tumor dissemination.

4. Conclusions

The main objective of the proposed research was to identify and improve a suitable deep-learning
architecture which would help to achieve a better detection of brain tumors from 2D MRIs. With an
experimental investigation, this work identified that the VGG19 helped to attain better results compared
to AlexNet, VGG16, ResNet50 and ResNet101. The proposed work implemented the following
techniques to improve the detection accuracy of the VGG19; (i) replacing the SoftMax classifier with
well-known classifiers, such as decision tree, k-nearest Neighbor, SVM-linear and SVM-RBF, and (ii)
 improving the performance of the pre-trained VGG19 by implementing a future fusion technique to
help improve the detection accuracy. In this work the customized VGG was developed using the
handcrafted features of dimension 1x1x199 and deep features of dimension 1x1x1024, and then these
features were sorted based on the PCA and fused using the serial concatenation technique. The final
feature vector of size 1x1x1223 is then considered to enhance the classification accuracy. In this work,
the brain images of BRATS, TCIA and clinical datasets were considered for the examination, and the
overall results attained with the proposed VGG19 with SVM-RBF classifier helped to attain better
results on Flair, T2 and T1C modality images. The performance was confirmed with ten-fold cross
validation and classification accuracies of >99%, >98% and >97% for the modalities Flair, T2 and
T1C, respectively.

and G.R.N.; resources, K.P.T., A.N.J.R. and G.R.N.; data curation, V.R. and K.P.T.; writing—original draft preparation,
authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by The Research Start-Up Fund Subsidized Project of Shantou
University, China, grant number NTF17016.

Acknowledgments: The authors of this article would like to acknowledge M/S. Proscans Diagnostics Pvt. Ltd., a
leading scan center in Chennai, for providing the clinical brain MRI for experimental investigation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bhandary, A. Deep-learning framework to detect lung abnormality–A study with chest X-Ray and lung CT
2. Pugalenthi, R.; Rajakumar, M.P.; Ramya, J.; Rajinikanth, V. Evaluation and classification of the brain tumor
3. Fernandes, S.L.; Tanik, U.J.; Rajinikanth, V.; Karthik, K.A. A reliable framework for accurate brain image
examination and treatment planning based on early diagnosis support for clinicians. Neural Comput. Appl.
2019, 1–12. [CrossRef]
4. Dey, N. Social-group-optimization based tumor evaluation tool for clinical brain mri of
5. Louis, D.N. The 2016 world health organization classification of tumors of the central nervous system:


38. Proscans Diagnostics PVT. LTD. Homepage. Available online: [https://proscans.in](https://proscans.in) (accessed on 1 November 2019).


54. Satapathy, S.C.; Rajinikanth, V. Jaya algorithm guided procedure to segment tumor from brain MRI. *J. Optim.* 2018, 2018, 3738049. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).