Automatic segmentation of brain tumours from MRI

X. Beltran Urbano[1] and A. Lalande[2]

[1] *University of Burgundy, Erasmus Mundus Joint Master Degree in Medical Imaging and Applications (MAIA)*
[2] *University of Burgundy, ImViA laboratories, Dijon, France*

Abstract— In the field of medical image processing, brain tumour segmentation is an essential task in order to improve the diagnosis, treatment, and follow-up of patients. Early detection of brain tumours has a crucial role in enhancing treatment options and increasing the patient survival rate. Manual segmentation of brain tumours for cancer diagnosis from a large number of MRI images is a complex and time-consuming operation; therefore, automation approaches are necessary to improve the efficiency of this task. The goal of this paper is to provide a review of four MRI-based approaches for brain tumour segmentation. First, an explanation of both machine learning and deep learning methods is provided. Afterwards, quantitative and visual results are presented for each paper. Then, the different algorithms previously mentioned in this work are discussed with an emphasis on the benefits and drawbacks of deep learning approaches. Finally, a conclusion focused on the difference between machine learning and deep learning methods in the field of brain tumour segmentation is addressed.

Keywords— Brain tumour, MRI, Segmentation, CNN, Machine Learning, Deep Learning, Support Vector Machine (SVM), U-NET

I. INTRODUCTION

The ultimate objective of brain tumour imaging analysis is to extract key patient-specific clinical information and diagnostic features. This information, incorporated within multidimensional image data, can guide and monitor interventions once the disease has been diagnosed and localised, ultimately leading to better knowledge for clinical diagnosis, disease staging, and treatment. As a virtue of its high resolution while monitoring tissue, magnetic resonance imaging (MRI) is one of the most commonly used techniques in the field of central nervous system disorders. Due to the multimodal imaging provided by this technology, we are able to segment brain tumours, one of the most difficult segmentation challenges encountered in the medical industry.

In 2012, The Brain tumour Segmentation Challenge (BraTS) was established as an international challenge. This competition, organised by the Radiological Association of North America (RSNA), the American Society of Neuroradiology (ASNR), and the Medical Image Computing and Computer Assisted Interventions (MICCAI) society [1], has considerably boosted the number of segmentation algorithms created for brain tumours. Due to the massive dataset supplied by this challenge, researchers are able to create and train artificial intelligence algorithms capable of segmenting brain tumours with extraordinary precision.

In recent years, numerous researchers in the fields of medical imaging and soft computing have achieved substantial strides in the segmentation of brain tumours. There have been both semi-automatic and completely automatic approaches proposed. However, due to the imperfections of the created algorithms, clinical acceptance has been compromised, as erroneous interpretations are never acceptable.

To demonstrate and analyse the range of approaches for performing the aforementioned objective, we have offered a review of four separate algorithms, with a focus on the distinctions between those that employ Deep Learning methods and those that do not.

II. METHODS

Then, the development methodologies employed in the selected papers for this project will be discussed in detail. In order to accomplish this task, we have divided them into two distinct groups: those that use machine learning (ML) and those that use deep learning (DL).

I. Machine Learning

Machine learning is a subfield of artificial intelligence, which is defined broadly as the capacity of a machine to replicate intelligent human behaviour. In order to improve the efficiency of time-consuming tasks that are often performed by humans in the medical area, these kinds of algorithms are developed.

The first paper we will discuss is titled *'**Fully Automatic Segmentation of Brain tumour Images Using Support Vector Machine Classification in Combination with Hierarchical Conditional Random Field Regularization**'* [2] . In this article, the author has subdivided the task of the provided approach into multiple parts in order to clarify the various algorithm duties. The steps are as follows:

· *Pre-processing*: Initially, the images undergo a preprocessing pipeline. In the first step, the four modalities are registered, using a rigorous registration and mutual information metric. Subsequently, the brain region is extracted from the images using a fully-automatic, customised skull-stripping algorithm. Finally, the noise is eliminated using an edge-preserving smoothing filter and the bias field is corrected.

· *Feature Extraction*: This process consists of the feature extraction of each image, followed by its subsequent classification. To accomplish this task, the author has focused his methods on the use of intensity and first-order texture features (mean, variance, skewness, kurtois, energy). Once the features have been extracted, they are stored in a 28-dimensional feature vector x, which consists of the voxel-wise concatenation of the multimodal intensities I and multimodal textures T at each voxel i as shown in equation (1).

$$x(i) = [I_{T_1}(i), I_{T_{1c}}(i), I_{T_2}(i), I_{T_{2f}}(i), T_{T_1}(i), T_{T_{1c}}(i), T_{T_2}(i), T_{T_{2f}}(i)] \quad (1)$$

· *Classification*: The goal of this procedure is to classify data between healthy and tumour regions in order to partition the healthy region into cerebrospinal fluid (CSF), grey matter (GM), and white matter (WM) and the tumour area into necrotic part, active part, and edema part. To accomplish this task, the author has used a soft-margin SVM classifier, which is based on the LibSVM implementation described in [3] .

· *Regularization:* The final step of this approach is to calibrate the models in order to minimise the adjusted loss function and avoid over-fitting or under-fitting. In this paper, the author uses Conditional Random Fields (CRF) techniques to solve the SVM classifier's spatial connection deficiency. This type of classifier assumes that the data is independent and uniformly distributed, which is obviously not the case for image voxels, since the majority of voxel labels are highly dependent on their neighbours.

The title of the second and final paper (employing machine learning techniques) we will discuss is ***'Context-sensitive Classification Forests for Segmentation of Brain tumour Tissues'*** [4] . As previously stated, the international competition known as BraTS has contributed to increasing the number and quality of the techniques developed in this field. This article is a perfect demonstration of this, as it was the 2012 challenge winner.

The major objective of this study is to segment each multimodal image into three classes `C = {B, T, E}` for background (B), tumour (T), and edema (E), with the highest precision. To perform the mentioned task, the author has presented a discriminative multiclass classification approach consisting of a standard classification forest (CF), based on spatially non-local features, which is combined with initial probabilities estimates for the individual tissue classes. These probabilities are used as additional input channels for the forest, together with the MRI image data.

In order to clarify the numerous algorithmic responsibilities, the author has broken the approach's duties into multiple components. The procedure is as follows:

· *Pre-processing*: Before starting the algorithm's several steps, the author has chosen to pre-process the data. His technique consists of first applying bias-field normalisation using the ITK N3 implementation from [5]. Then, the mean intensities of the images within each channel are aligned by a global multiplicative factor.

· *Estimating Initial Tissue Probabilities*: As a first step of our approach, the author estimates the initial class probabilities for a given patient based on the intensity representation in the MRI input data. These probabilities are computed based on the likelihoods obtained by training a set of Gaussian Mixture Models (GMM) on the training data. For each class, a single GMM is trained to represent the probability of multidimensional intensities. Once these intensities have been collected, we utilise them as input for the classification forest. The input vector classifier's final shape includes both the multichannel MRI intensities and the posterior intensities $pGMM(c|p)$. In the equation (2), we can observe the channels of a single patient as an example.

$$C = (I_{T1-gad}, I_{T1}, I_{T2}, I_{FLAIR}, p_{AC}^{GMM}, p_{NC}^{GMM}, p_{E}^{GMM}, p_{B}^{GMM}) \quad (2)$$

· *Classification Forest*: The goal of this process consists of classifying the input data mentioned in the previous step, into background (B), tumour (T) or edema (E). To achieve this task the author has presented a classification forest (CF) which is based on spatial non-local information from the channels C.

· *Context-sensitive Feature Types*: As previously stated, the classification features are intensity and textures based. In particular, the author employs three context-sensitive feature types that describe a point to be classified based on its non-local neighbourhood. The first two of these feature categories are quite generic, as they quantify the difference between intensity and intensity meaning. The third feature, however, is not generic because it is designed with the purpose of detecting structural changes.

II. Deep Learning

Deep learning (DL) is one of the AI subfields that has experienced exponential growth in recent years. The scientific community has focused its attention on DL due to its versatility, high performance, high generalisation capacity, and multidisciplinary uses, among many other qualities [6].

The majority of tough tasks in neuroimaging, such as brain tumour segmentation, are performed using DL methods. In particular, since 2015, one of the most used architectures for medical imaging segmentation is U-NET, which was proposed in the International Symposium on Biomedical Imaging (ISBI) challenge for segmentation of neuronal structures in electron microscopic stacks [8]. Due to the excellent performance that was demonstrated, this work now has over 40.000 citations and is considered the gold standard in this field.

Then, two papers utilising the U-NET architecture will be described. The initial article to be discussed is titled ***'nnU-Net for Brain tumour Segmentation' [7]***. This research was developed for the segmentation task of the aforementioned BraTS international competition. Specifically, this article was submitted for the 2020 challenge, and due to the accuracy obtained, it was deemed the winner. The method utilised was based on nn-UNet (see the architecture in the Fig. 1). Moreover, in order to obtain the highest accuracy in the segmentations, the author performed some modifications to the nnU-Net configuration. Those modifications are as described below:

- Region-Based Training: The provided labels for training are 'edema', 'non-enhancing tumour and necrosis'

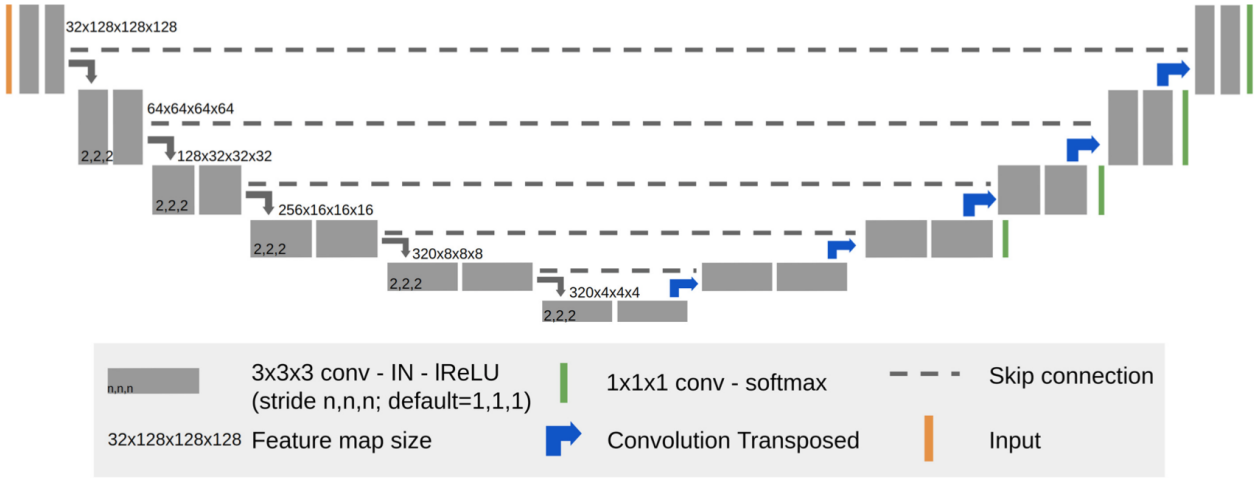

Fig. 1: nnU-Net architecture proposed in the paper [7].

and 'enhancing tumour'. The evaluation of the segmentations is, however, performed on three partially overlapping regions: whole tumour (consisting of all 3 classes), tumour core (non-enhancing and necrosis + enhancing tumour) and enhancing tumour. So, since it has been shown previously that directly optimising these regions instead of the individual classes can improve the performance of the segmentations, the author has employed this configuration. Moreover, the cross-entropy loss term is replaced by a binary cross-entropy, which optimizes each of the regions independently.

- Increased Batch Size: In machine learning, batch size refers to the number of training samples used in a single iteration. In a conventional nnU-Net, the small batch size leads to noisy gradients, which may prevent overfitting but limits the model's ability to accurately fit the training data. In the configuration of this paper, due to large datasets (as it is BraTS dataset) it may be beneficial to increase the batch size. Because of this, the mentioned term has been increased from 2 to 5 in an effort to improve the model's accuracy.

- More Data Augmentation: In an effort to increase the robustness of the models, the author has recommended more aggressive data augmentation strategies, despite the fact that nnU-Net currently employs a wide variety of aggressive techniques. Using the *batchgenerators* framework, all augmentations to nnU-Net for brain tumour segmentation are applied dynamically during training.

- Batch normalization: Since dice BraTS scores for test cases are frequently lower than the values reported on the training and validation datasets, the author has proposed a batch normalisation technique to address this problem.

- Batch Dice: Instead of addressing the samples in a mini-batch independently, the author implements a new dice loss computation consisting of computing the dice loss over all samples in the batch (pretending they are just a single large sample). This essentially normalises this score, as samples with few annotated voxels (these samples can cause large gradients and dominate the parameter updates during training) are now eclipsed by other samples within the same batch.

The title of the second and final paper we will discuss is ***'Cascade multiscale residual attention CNNs with adaptive ROI for automatic brain tumour segmentation'***[9]. In this paper, the author has proposed a novel method that follows a hybrid input strategy and includes MRA-UNet architectures employed in a cascade fashion to stage-wise segment the whole, enhanced, and core tumour regions.

In order to provide a clearer explanation of this method, the author has separated the article into the developed methods, which are as follows:

· *Pre-processing*: The pre-processing of this method consists of three different steps: Firstly, the images are refined in order to remove the background, and they are normalized to 160×160. Secondly, to enhance the image quality, a histogram equalization (HE) technique and normalization of intensity values have been employed. Finally, in order to utilize the 3D sequential information, we merged the two adjacent slices into a three-channel image. The overall pre-processing has been depicted in Fig. 2 .

· *Cascade multiscale residual attention UNet*: To achieve the segmentation process, the author has proposed a multiscale residual attention UNet (MRA-Unet) architecture in a cascade fashion (see Fig. 3). The primary characteristics of the aforementioned architecture are as follows:

- Residual Block: The major role of the residual block is to propagate information across layers in order to create a deeper neural network capable of overcoming the degradation difficulties of every encoder. It minimises computational costs while also enhancing channel interdependencies.

- Atrous spatial pyramidal pooling (ASPP): ASPP is based on the concept of spatial pyramidal pooling, which proved beneficial for resampling features at different scales. In the proposed architecture, the ASPP

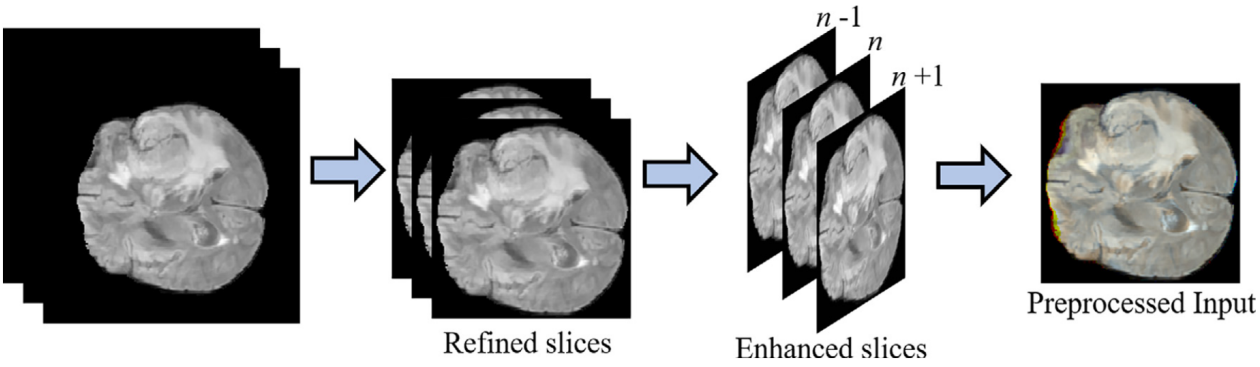

Fig. 2: Pre-processing stage including scan refinement, image enhancement and slice concatenation [9].

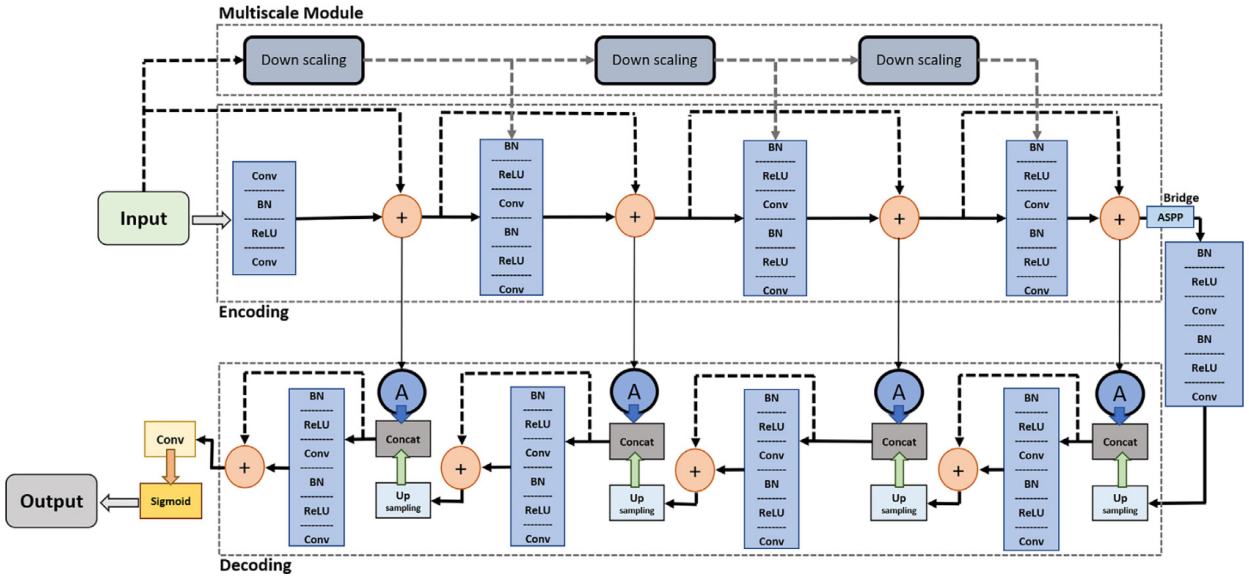

Fig. 3: Block diagram of the proposed MRA-UNet architecture [9].

operates as a bridge between the encoder and the decoder, as depicted in Fig. 3.

- Attention gates: The attention gate is a mechanism that allows the network to focus its attention on significant regions while inhibiting feature activation in unrelated areas. In this project, this mechanism is used to focus the network only on the whole tumour region while discarding feature responses in irrelevant background regions.
- Multiscale residual learning: It is known that features at multiple scales are advantageous because they help encode both local and global contexts. Segmentation of brain tumours is greatly aided by the use of multiscale characteristics due to the tumour's significant inter-and intra-size fluctuations. As shown in Fig. 3, the author has therefore presented a method in which several scale features are extracted and concatenated with network layers at various levels of the encoder network. The input image size is kept to 128×128 and then downsampled to half at three levels and subsequently fed to the proposed network's encoder block.
- Loss Function: The loss function measures how effectively your algorithm models your dataset. The author has utilised the Dice similarity coefficient (DSC) loss to complete this work. This function is commonly used in the segmentation of medical images to calculate the similarity between two samples.

  In the equation (3), we can observe an example of the DSC_{loss} function for one single sample.

$$DSC_{loss} = 1 - \frac{2 \times |A \cup B|}{|A \cap B|} \tag{3}$$

· *Post-processing*: In order to enhance the brain tumour segmentation performance and its classification, two post-processing methods have been proposed: conditional random field (CRF) and test time augmentation (TTA).

- Conditional random field (CRF): CRF is a highly effective statistical modelling method used to refine the network's output semantic segmentation map. In this approach, the author employed this method to enhance segmentation efficiency and achieve a more precise outcome.

TABLE 1: QUANTITATIVE RESULTS USING DICE SIMILARITY COEFFICIENT (DSC).

Method	Dataset	WT [1]	N [2]	E [3]	A [4]
Bauer et al. [2]	10 patients from the ContraCancrum brain tumour database	0.84 ± 0.03	0.61 ± 0.24	0.73 ± 0.04	0.71 ± 0.09
Method	**Dataset**	**WT [1]**	**TC [5]**	**ET [6]**	**A [4]**
Zikic et al. [4]	BraTS2012 dataset	0.71 ± 0.24	-	0.7 ± 0.09	-
Isensee et al. [7]	BraTS2020 dataset	0.89 ± 0.13	0.85 ± 0.24	0.82 ± 0.19	-
Ullah et al. [9]	BraTS2020 dataset	0.9 ± 0.08	0.87 ± 0.11	0.86 ± 0.21	-

WT [1]= Whole tumour. N [2]= Necrotic. E [3]= Edema. A [4]= Active. TC [5]= Tumour core. ET [6]= Enhancing tumour.

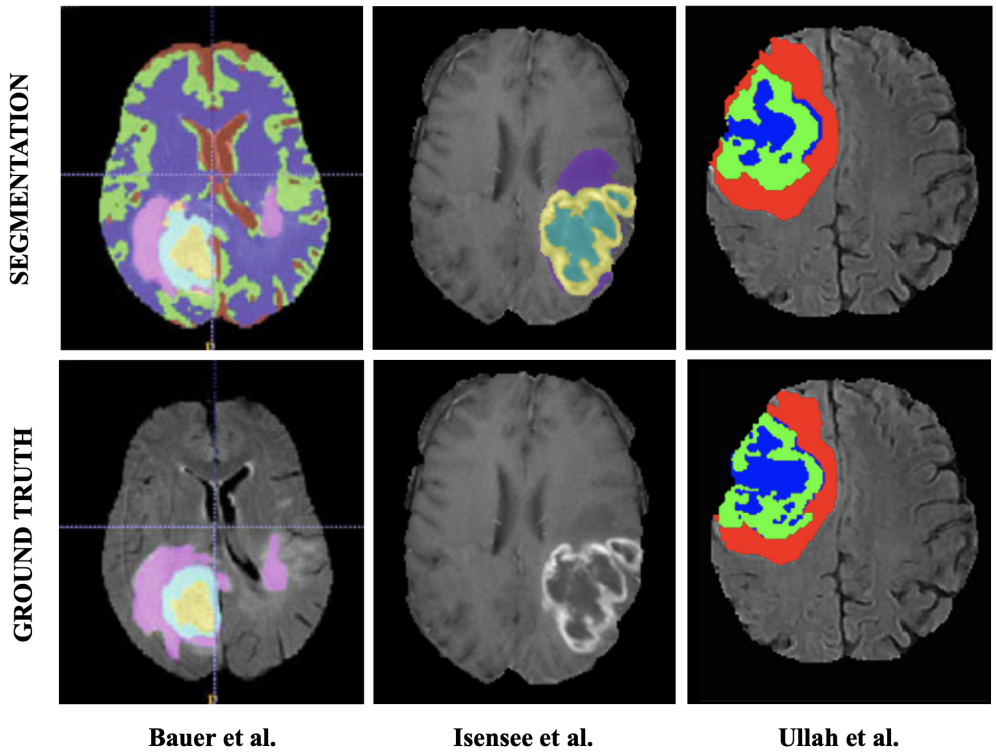

Fig. 4: One axial slice segmentation result for each of the previously stated papers. All segmentations have been submitted with the appropriate ground truth, with the exception of the approach performed by *Isensee et al.* [7], for which we have submitted the T1c image of the corresponding segmentation due to the lack of ground truth in the original paper.

- Test time augmentation (TTA): The TTA proposed by the author in this article consists of four steps: augmentation, prediction, dis-augmentation, and merging. Firstly, the augmentation is initially applied to the test image. Then, we predict for both the augmented and the original images, and afterwards, using the obtained prediction, we revert the transformation, which is known as dis-augmentation. Finally, the last step, called merging, is employed using an extended merging method inspired by one of the Data Science Bowl (DSB) [10].

III. RESULTS

Then, the quantitative results of each paper will be presented using the previously mentioned metric, the Dice similarity coefficient (DSC).

In Table 1 we can observe the quantitative results of the different methods presented in this paper. Since the papers segment the tumour into different regions, the table submitted in this paper has been divided according to this. Moreover, in Fig. 4, we can observe the result of an axial slice segmentation, as well as the corresponding ground truth, performed for all of the aforementioned methods with the exception of the method performed by Zikic et al. [4], for which it was not possible to display the segmentation result due to the lack of a segmentation image.

IV. DISCUSSION

This paper presented a thorough survey of techniques used in brain tumour segmentation. The survey encompasses two traditional machine learning and two deep learning-based

methods with their quantitative performance, which can be observed in the Table 1. The first method presented was based on an unsupervised machine learning approach such as the soft-margins SVM classifier [2]. This method is mainly affected by outliers and noise. To overcome this challenge, a hierarchical conditional random field regularization has been proposed by the author. However, even with the regularisation mentioned, the accuracy obtained for some regions still has to increase considerably in order to use this information for the diagnosis and treatment of real patients. The second and last method based on machine learning presented in this review consists of the context-sensitive classification forest method [4]. The major limitation of this approach is the over-segmentation of edema, since the classifier estimates more FP than FN. These limitations are reflected in the quantitative results, since with this approach the accuracy obtained is lower than the accuracy obtained with the previous method.

Then, the next and last two papers presented were based on deep learning approaches. The first one was based on a nnU-Net. This approach obtained high accuracy for the three regions segmented. However, due to the limitations of this method, the accuracy obtained was lower than those obtains with the approach performed by *Ullah et al.* [9]. These limitations are mainly based on the configuration of the neural network and its hyperparameters. A more thorough optimization of these values could result in further performance gains. Finally, the last paper presented was based on a multiscale residual attention UNet (MRA-UNet). This approach achieved competitive performance for whole-brain tumours and significantly outperformed for enhanced and core tumour segmentation. Even though it has obtained high accuracy of the segmented regions, the performance of this method could be improved including the extension of MRA-UNet with deep supervised learning to enable the end-to-end learning to avoid the cascade mechanism without degrading the performance for enhanced and core tumour regions. This method, as can be observed in the Table 1, has obtained the highest accuracy of all the presented methods. However, since the different methods presented in these reviews employ different datasets to perform the brain tumour segmentation task and, the segmentation developed is performed in different tumour regions, we can not claim that this is the best method among all papers. Nevertheless, since the accuracy obtained with deep learning techniques is higher than that obtained with machine learning techniques in all the regions segmented, we can assert that this segmentation methodology is the most suitable for the task we wish to complete.

Deep Learning methods are recognised as state-of-the-art in the field of brain tumour segmentation. However, as can be observed in this review, they have some limitations. The primary constraint of this method is that it requires a large amount of data to develop accurate models. In recent years, thanks to competitions such as BraTS, the quantity of data available for training neural networks for this field has grown significantly. Nonetheless, as demonstrated by the quantitative results of this review, the models we have developed are able to achieve high accuracy, but they are not flawless, leading us to believe that we must continue expanding the size of our datasets in order to improve the accuracy obtained with them. Another constraint of this technique is that in order to acquire the best possible results, we must determine the architecture and configuration of hyperparameters that best fit the task we wish to complete. A poor configuration of our neural network can have a direct effect on the accuracy of our model, which must be nearly perfect if we are to use them to complete the aforementioned task, as even the smallest error in clinical diagnosis and treatment can be life-threatening for the patient.

V. CONCLUSION

Automating the brain tumour segmentation and classification task has tremendous benefits in improving the diagnosis, treatment planning, and follow-up of patients. In this paper, both machine learning and deep learning methods for brain tumour segmentation have been presented. Even though machine learning has obtained a good performance, it is not as good as the one obtained with deep learning methods. Due to the advantage of automatically learning representative complex features for both healthy brain tissues and tumour tissues directly from the multi-modal MRI images, deep learning methods are capable of obtaining a DSC around 0.9 for all the segmented regions. Future enhancements and modifications to the architectures of these methods, as well as the incorporation of complementary information from other imaging modalities, may lead to the development of more accurate and clinically acceptable automatic glioma segmentation methods for improved patient diagnosis and treatment.

REFERENCES

[1] U. Baid, S. Ghodasara, S. Mohan, M. Bilello, E. Calabrese, E. Colak, K. Farahani, J. Kalpathy-Cramer, F. C. Kitamura, S. Pati, and L. M., "The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification," 7 2021.

[2] S. Bauer, L. P. Nolte, and M. Reyes, "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6893 LNCS, pp. 354–361, 2011. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-23626-6_44

[3] "Libsvm – a library for support vector machines." [Online]. Available: https://www.csie.ntu.edu.tw/~cjlin/libsvm/

[4] D. Zikic, D. Zikic, B. Glocker, E. Konukoglu, J. Shotton, A. Criminisi, D. H. Ye, C. Demiralp, O. M. Thomas, T. Das, R. Jena, and S. J. Price, "Context-sensitive classification forests for segmentation of brain tumor tissues," *IN MICCAI*. [Online]. Available: http://130.203.136.95/viewdoc/summary?doi=10.1.1.457.9629

[5] N. J. Tustison and J. Gee, "N4itk: Nick's n3 itk implementation for mri bias field correction," *The Insight Journal*, 1 2022.

[6] A. Anaya-Isaza, L. Mera-Jiménez, and M. Zequera-Diaz, "An overview of deep learning in medical imaging," *Informatics in Medicine Unlocked*, vol. 26, p. 100723, 1 2021.

[7] F. Isensee, P. F. Jäger, P. M. Full, P. Vollmuth, and K. H. Maier-Hein, "nnu-net for brain tumor segmentation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12659 LNCS, pp. 118–132, 11 2020. [Online]. Available: https://arxiv.org/abs/2011.00848v1

[8] W. Weng and X. Zhu, "U-net: Convolutional networks for biomedical image segmentation," *IEEE Access*, vol. 9, pp. 16 591–16 603, 5 2015. [Online]. Available: https://arxiv.org/abs/1505.04597v1

[9] Z. Ullah, M. Usman, M. Jeon, and J. Gwak, "Cascade multiscale residual attention cnns with adaptive roi for automatic brain tumor segmentation," *Information Sciences*, vol. 608, pp. 1541–1556, 8 2022.

[10] "2018 data science bowl | kaggle." [Online]. Available: https://www.kaggle.com/c/data-science-bowl-2018