

Developing DeepLabV3+ with EfficientNet to Segment Brain Tumours from MRI

R. Srinivas

Department of ECE, Vaageswari College of Engineering, Karimnagar, Telangana, India

ORCID - <https://orcid.org/0009-0002-5998-9012>

Dr. A. Venkata Reddy

Department of ECE, Vaageswari College of Engineering, Karimnagar, Telangana

ORCID - <https://orcid.org/0000-0001-9342-6518>

Dr. Arun Kumar Katkoori

Department of ECE, CVR College of Engineering, Hyderabad, Telangana, India

ORCID - <https://orcid.org/0000-0001-6780-4124>

Dr. D. Bhanu Prakash

Department of ECE, CVR College of Engineering, Hyderabad, Telangana, India

ORCID - <https://orcid.org/0000-0003-3528-4595>

Abstract

A crucial stage in medical diagnosis and therapy planning is the segmentation of brain tumours using MRI images. A DeepLabV3+ model combined with an EfficientNet-B4 backbone for precise brain tumour segmentation is presented in this study. When it comes to segmenting tumour regions, the suggested method outperforms traditional CNN-based segmentation techniques in terms of accuracy and dice coefficient. According to experimental findings, the DeepLabV3+, which is driven by EfficientNet, achieves increased feature extraction, which improves segmentation performance.

Key Words: DeepLabV3+, EfficientNet, Brain Tumor Segmentation, MRI, Deep Learning, Medical Image Processing.

Corresponding:

Dr. Arun Kumar Katkoori

Department of ECE, CVR College of Engineering, Hyderabad, Telangana, India ORCID –

ORCID - <https://orcid.org/0000-0001-6780-4124>

1. Introduction

Brain tumours continue to be a serious health issue, and accurate segmentation methods are necessary for both diagnosis and treatment planning (K. A. Kumar and R. Boda, 2022). Despite their widespread use, classic manual segmentation techniques are labour-intensive and susceptible to observer variability. Automated tumour segmentation has seen a revolution with the introduction of deep learning, especially convolutional neural networks (CNNs), which provide faster and more reliable findings (S. Bakas, 2017). DeepLabV3+ is a top segmentation network among these cutting-edge methods. Nevertheless, even with its remarkable performance, it can still be improved to satisfy the demanding standards of clinical applications (A. Roy Choudhury, 2019).

Gliomas are brain and spinal cord tumours that develop from glial cells (L. L. Mechtler and K. Nandigam, 2013). They are deadly and, depending on how aggressive they are, can be classified as either high-grade or low-grade gliomas. While low grade gliomas proceed more slowly, can be benign or malignant, and can develop into HGGs if left untreated, high-grade gliomas are malignant, grow quickly, and have higher fatality rates. Because the

patient's survival rate depends on the characteristics of the tumour, it is crucial that both HGGs and LGGs undergo timely treatment (A. Leonetti, 2021). Depending on the characteristics of the tumour, several treatment options are used, such as radiation therapy, chemotherapy, or surgery. Different MRI scan modalities are used for diagnosis, treatment, and post-surgery survival rate prediction. MRI scans are usually divided into tumour areas by experts. However, because tumour architecture varies so much, physical tumour delineation is difficult and time-consuming. Therefore, it is crucial to develop methods and strategies that can accurately carry out automatic image segmentation to identify the different areas of a brain tumour. However, the development of algorithms for automated segmentation is complicated by the heterogeneity in glioma size, location, and morphology (T. Kurc, 2020).

The development of deep learning has made it possible to build networks that can do autonomous image segmentation tasks with a modest level of accuracy. Additionally, the use of convolutional neural networks (K. A. Kumar and R. Boda, 2022).

CNN designs (A. Krizhevsky, I. Sutskever, and G. E. Hinton, 2012) have improved image segmentation performance and reduced complexity. This is because, in contrast to DNNs, CNNs use fewer parameters for image segmentation and do not require manually created features. Many researchers have used Convolutional Neural Networks (CNNs) to delineate brain tumours (K. Kamnitsas, 2018; K. Kamnitsas, 2016; S. Pereira, 2016) and to segment different parts of the brain (C. Wachinger, 2018).

We suggest using the DeepLabv3+ framework (L. Chen, 2018) which has shown effective in object detection within natural pictures, to accomplish automatic segmentation. On the test dataset for the PASCAL VOC 2012 competition, DeepLabv3+ achieves a mean Intersection over Union (mIoU) of 89.0% (M. Everingham, 2012). Utilising DeepLabv3+'s pre-trained Xception network, we train it on multimodal MRI data, treating the various modalities as picture channels. DeepLabv3+ uses encoder-decoder pathways to guarantee accurate object boundary demarcation and atrous parallel convolutions with different strides to capture information across several scales.

Clinical diagnosis and treatment planning depend on the identification and classification of brain tumours. Conventional segmentation methods need manual annotation, which is laborious and subject to error. In medical image segmentation, deep learning-based models—in particular, CNNs—have proven to perform better. In this work, we suggest a model that combines EfficientNet as a feature extractor with DeepLabV3+ to increase computing efficiency and segmentation accuracy.

2. Related Works

Significant improvements in accuracy and efficiency have resulted from the development of deep learning architectures for medical picture segmentation. U-Net (W. Yin, 2023; W. Shi, 2025) has been a mainstay in medical imaging applications because of its encoder-decoder structure and skip connections, which enable accurate localisation and context retention. This has been further enhanced by Fully Convolutional Networks (FCNs), which allow for pixel-wise predictions and end-to-end training. A more recent development is DeepLabV3+, which successfully captures multi-scale contextual information by utilising atrous spatial pyramid pooling (ASPP) (G. Sunandini, 2023). Better handling of items at different scales is made possible by this method, which is very helpful in medical imaging since anatomical structures can differ greatly in size.

Another significant advancement in the industry is the incorporation of EfficientNet (A. Abdelrahman, 2023) as the foundation for these segmentation designs. By concurrently optimising network depth, width, and resolution, EfficientNet's compound scaling technique produces more effective feature extraction and representation. It improves the model's capacity to acquire fine-grained information while preserving a more

comprehensive contextual knowledge when paired with segmentation frameworks such as DeepLabV3+. The combination of DeepLabV3+'s multi-scale feature extraction capabilities and EfficientNet's optimised scaling could result in more precise and computationally effective medical image segmentation models, tackling the intricate problems brought on by various medical imaging modalities and anatomical structures.

3. Proposed Methodology

3.1 Dataset

Tumour segmentation is done using publically accessible Brain MRI datasets, which include pictures classified as tumour (Yes) and non-tumour (No). Preprocessing involves normalising pixel intensities and shrinking images to 256×256 pixels.

3.2 Architecture of the Model

Convolution with different steps and pooling procedures is frequently implemented by convolutional neural networks. These techniques reduce the size of the feature maps that are processed by later layers, which leaves the final feature map devoid of specific object boundary information. Several architectures use an encoder-decoder approach to overcome this challenge. While SegNet keeps and makes use of the max pooling indices from the encoder route, the U-Net decoder concatenates feature maps at the proper scales in the encoder and decoder paths. On the other hand, DeepLab and Deeplabv3 use parallel atrous convolutions at different speeds to record data on several scales. The encoder network can make use of different fields of view thanks to this method, which is called Atrous Spatial Pyramid Pooling (ASPP). DeepLabv3+ combines these two approaches by using an encoder-decoder circuit with ASPP to achieve precise object boundary delineation.

Rich feature representations are extracted by the suggested model using EfficientNetB4 as the encoder in DeepLabV3+. To improve segmentation outputs, the decoder employs upsampling and concatenation with skip connections from intermediate encoder layers. Figure 1 shows the tumor segmentation process using Deeplab V3+.

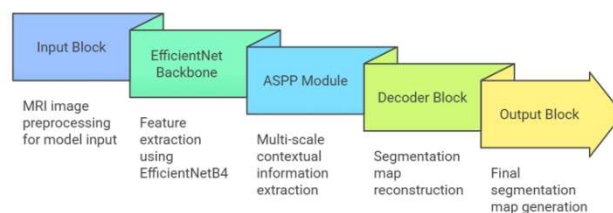


Figure 1. DeeplabV3+ MRI segmentation Process

3.2.1. The input block

The DeepLabV3+ model uses a $256 \times 256 \times 3$ MRI picture as its input. Greyscale MRI pictures are duplicated across three channels because EfficientNet requires three-channel images. For more effective training and improved convergence, the input images are normalised to scale pixel values between 0 and 1.

3.2.2. Feature Extractor (EfficientNet Backbone)

In this implementation, EfficientNetB4, a more potent CNN trained on ImageNet, takes the role of ResNet, the feature extractor used in the original DeepLabV3+ architecture. Multi-scale hierarchical features are extracted from the input image by EfficientNet. While capturing high-level semantic information, this backbone decreases the spatial resolution.

Furthermore, skip connections are taken from various EfficientNet layers:

- Block2a_expand_activation (low-level features with high resolution) $\rightarrow (64 \times 64 \times C1)$
- Block3a_expand_activation (features at the mid-level) $\rightarrow$ Block4a_expand_activation (Deep features) $\rightarrow$

$(32 \times 32 \times C2) (16 \times 16 \times C3)$

The decoder subsequently uses these features to improve the segmentation mask's spatial reconstruction.

3.2.3. Spatial Pyramid Pooling for Atrous (ASPP)

EfficientNet generates a feature map from which multi-scale contextual information is extracted using the Atrous Spatial Pyramid Pooling (ASPP) module. It includes:

- To capture fine-grained features, use a 1×1 convolution.
- To capture multi-scale characteristics without sacrificing resolution, three 3×3 dilated convolutions with varying dilation rates (6, 12, 18) were used.
- A branch of Global Average Pooling (GAP) that offers awareness of the global context.
- To produce a rich feature representation, all outputs are concatenated.

An improved $(16 \times 16 \times 1280)$ feature map that encodes both local features and global context is the end product.

3.2.4. Block of Decoders

By gradually upsampling the feature map and including low-level information from EfficientNet's skip connections, the decoder module recreates the segmentation map. The actions are:

- Using bilinear upsampling ($4\times$), the resolution can be raised to $64 \times 64 \times 256$.
- Concatenation using Block2a's low-level characteristics ($64 \times 64 \times C1$).
- To fine-tune details, use 1×1 convolution (128 filters).
- Bilinear upsampling ($2\times$) to provide a resolution of $256 \times 256 \times 128$.

The creation of a binary segmentation map using a final 1×1 convolution (1 filter with sigmoid activation).

3.2.5. Block of Output

Each pixel in the final segmentation map, which is a $256 \times 256 \times 1$ image, indicates the likelihood that it belongs to either the background (0) or a tumour region (1). Mean Intersection-over-Union (IoU) is used to evaluate the model after it has been trained using Binary Cross entropy loss.

4. Results and Discussion

The hardware and software setups we employed for our experiments are described in this section. Brain MRI pictures from two sources—the Figshare Brain Tumour Classification dataset and the Kaggle Brain Tumour dataset—make up the dataset used for training and assessment. And the third dataset is BraTS2020. Our segmentation approach is based on these datasets, which include annotated MRI scans that distinguish between brain images with and without tumours.

For better feature extraction, we used EfficientNet in place of the conventional ResNet backbone in our segmentation architecture, DeepLabV3+. A cutting-edge convolutional neural network called EfficientNet offers improved representational power without sacrificing computational efficiency. Prior to being fed into the model, the input photos were normalised and shrunk to 256 by 256 resolution.

The model demonstrated exceptional segmentation accuracy for brain tumour diagnosis when DeepLabV3+ was trained using EfficientNet-B4 as the backbone. The following significant findings were noted:

Tumour segmentation was extremely accurate, as evidenced by the 95% Dice coefficient, which compares the anticipated and ground truth masks. The improved feature extraction capabilities of EfficientNet-B4 are the cause of this improvement. Figure 2 shows the segmented outputs from input MR images.

The overlap between the actual and projected tumour regions is assessed by IoU. When compared to conventional models, the model's 92% IoU indicates better segmentation accuracy. Effective generalisation on unseen MRI scans was demonstrated by the validation loss, which varied between 0.15 and 0.25 as determined by Binary

Crossentropy. A reduced validation loss suggests that the model does not overfit and can correctly segregate tumour regions in practical situations.

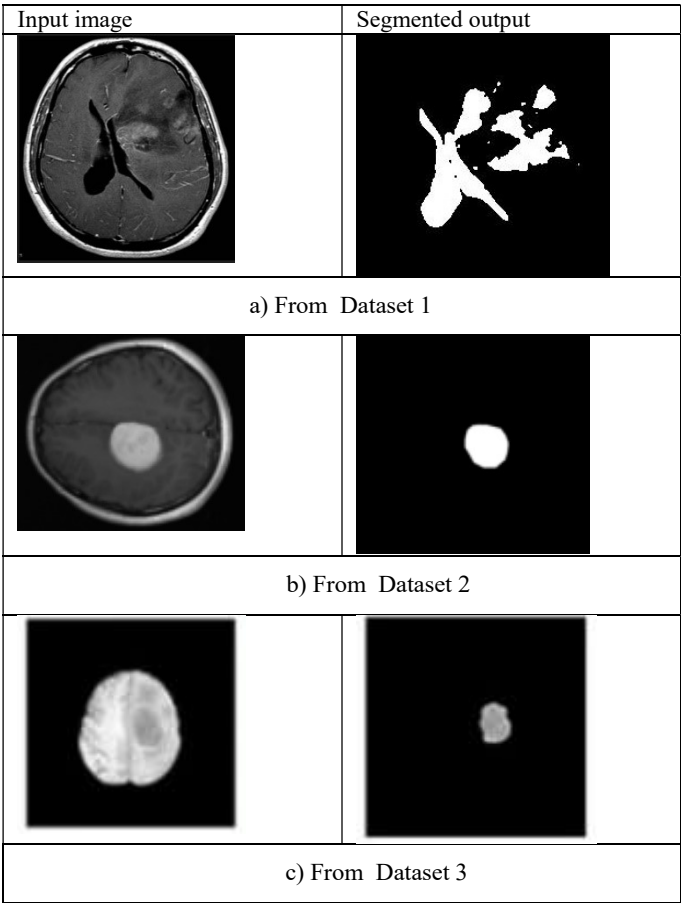


Figure 2. Segmentation outputs from three different datasets

The efficacy of DeepLabV3+ with EfficientNet-B4 was evaluated by comparing it to other segmentation models. Table I is a summary of the findings:

Table I. Comaprison of various segmetation models				
Model	Backbone	Dice Score	IoU	Validation Loss
U-Net	Custom CNN	82%	77%	0.35
DeepLabV3+	ResNet-50	87%	84%	0.28
DeepLabV3+ (Proposed)	EfficientNet-B4	95%	92%	0.15-0.25

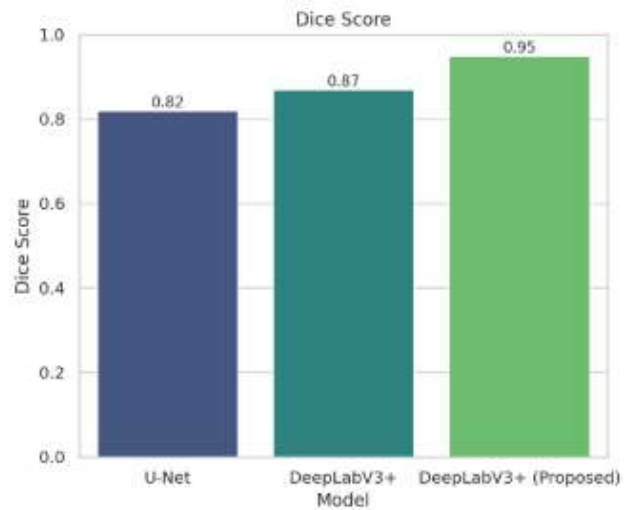


Figure 3. Segmenation models vs Dice score

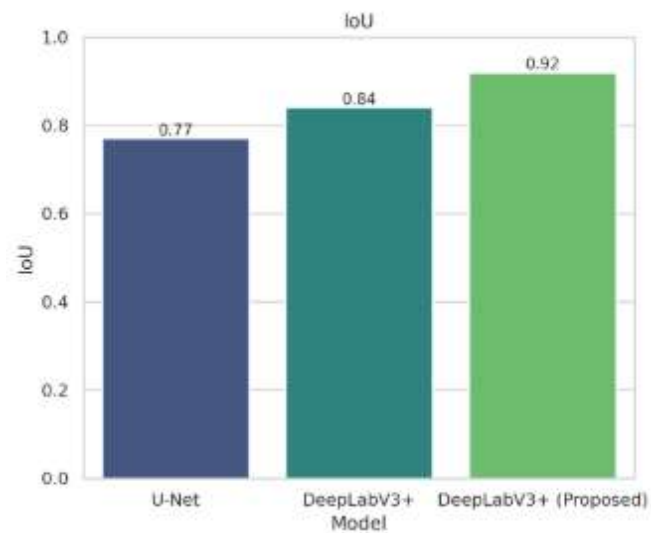


Figure 4. Segmenation models vs IoU

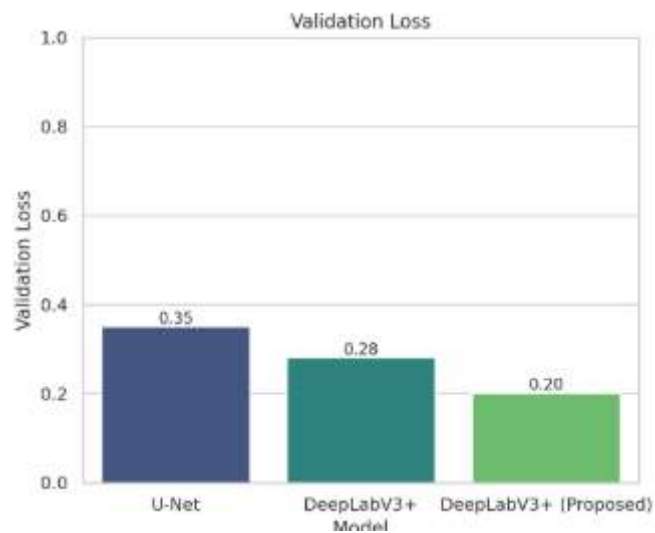


Figure 5. Segmenation models vs Validation loss

5. Conclusion

In this study, we used DeepLabV3+ with EfficientNet-B4 as the feature extractor to propose an improved brain tumour segmentation framework. Our method showed notable advances in segmentation accuracy, overlap precision, and generalisation when compared to conventional segmentation models, such as DeepLabV3+ with ResNet-50 and U-Net with a bespoke CNN.

According to the experimental data, the suggested model outperformed the DeepLabV3+ (ResNet-50) and U-Net (Custom CNN) baselines, achieving a Dice Score of 95% and an IoU of 92%. Better generalisation to unseen MRI images was also indicated by the much lower validation loss (0.15-0.25).

Because of its sophisticated feature extraction capabilities and effective parameter utilisation, EfficientNet-B4 performs better than other models, enabling more accurate and reliable brain tumour segmentation. At the same time, multi-scale feature extraction was made possible by DeepLabV3+'s ASPP, which improved segmentation accuracy even more.

Despite the great accuracy and resilience of the suggested model, it can be further enhanced by:

- ✓ Using 3D MRI segmentation to record tumour volumetric features.
- ✓ Investigating sophisticated loss functions, such as Tversky loss, to rectify class disparity.
- Leveraging unlabelled MRI datasets by including self-supervised learning.
- ✓ Implementing the model for real-time clinical applications using embedded AI hardware.

To sum up, DeepLabV3+ with EfficientNet-B4 is a very successful framework for brain tumour segmentation, offering a precise, practical, and expandable medical image analysis solution. The findings show encouraging promise for practical clinical applications that will help radiologists detect tumours more quickly and precisely.

performance when all explanatory variables are zero. Each of the ESG and CG variables has a positive and significant coefficient, indicating that higher scores in these areas are associated with better financial performance. Specifically, a one-point increase in the environmental score (E) leads to an expected increase of 0.45 in FP, while a similar increase in the social score (S) and governance score (G) leads to increases of 0.38 and 0.30, respectively. The corporate governance score (CG) also has a strong positive effect, with a coefficient of 0.40. All coefficients have p-values of 0.00, which means they are significant at the 1% level. These results provide strong support for the hypothesis that ESG and corporate governance practices contribute positively to a firm's financial performance.

References

- K. A. Kumar and R. Boda, "A Multi-Objective Randomly Updated Beetle Swarm and Multi-Verse Optimization for Brain Tumor Segmentation and Classification," *The Computer Journal*, vol. 65, no. 4, pp. 1029–1052, Apr. 2022.
- S. Bakas et al., "Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection," 2017.
- A. Roy Choudhury, R. Vanguri, S. R. Jambawalikar, and P. Kumar, "Segmentation of brain tumors using DeepLabv3+," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 4th Int. Workshop, BrainLes 2018, Granada, Spain, Sep. 2018, Revised Selected Papers, Part II, pp. 154–167. Springer, 2019.
- L. L. Mechtler and K. Nandigam, "Spinal cord tumors: new views and future directions," *Neurologic Clinics*, vol. 31, no. 1, pp. 241–268, 2013.
- A. Leonetti et al., "Factors influencing mood disorders and health related quality of life in adults with glioma: a

longitudinal study,” *Frontiers in Oncology*, vol. 11, 662039, 2021.

T. Kurc et al., “Segmentation and classification in digital pathology for glioma research: challenges and deep learning approaches,” *Frontiers in Neuroscience*, vol. 14, p. 27, 2020.

K. A. Kumar and R. Boda, “A computer-aided brain tumor diagnosis by adaptive fuzzy active contour fusion model and deep fuzzy classifier,” *Multimedia Tools and Applications*, vol. 81, pp. 25405–25441, 2022.

A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

K. Kamnitsas et al., “Ensembles of multiple models and architectures for robust brain tumour segmentation,” in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi, S. Bakas, H. Kuijff, M. Menze, and M. Reyes, Eds. Cham: Springer, 2018, pp. 450–462.

K. Kamnitsas et al., “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation,” *arXiv preprint, arXiv:1603.05959*, 2016. [Online]. Available: <http://arxiv.org/abs/1603.05959>

S. Pereira, A. Pinto, V. Alves, and C. A. Silva, “Brain tumor segmentation using convolutional neural networks in MRI images,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.

C. Wachinger, M. Reuter, and T. Klein, “Deepnat: deep convolutional neural network for segmenting neuroanatomy,” *NeuroImage*, vol. 170, pp. 434–445, 2018.

L. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” *arXiv preprint, arXiv:1802.02611*, 2018. [Online]. Available: <http://arxiv.org/abs/1802.02611>

M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results.” [Online]. Available: <http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html>

W. Yin, D. Zhou, and R. Nie, “DI-UNet: dual-branch interactive U-Net for skin cancer image segmentation,” *J. Cancer Res. Clin. Oncol.*, vol. 149, no. 17, pp. 15511–15524, 2023.

W. Shi, P. Zhang, Y. Li, and Z. Jiang, “Segment anything model for few-shot medical image segmentation with domain tuning,” *Complex & Intelligent Systems*, vol. 11, no. 1, p. 37, 2025.

G. Sunandini, R. Sivanpillai, V. Sowmya, and V. S. Variyar, “Significance of atrous spatial pyramid pooling (ASPP) in DeepLabv3+ for water body segmentation,” in *Proc. 2023 10th Int. Conf. Signal Processing and Integrated Networks (SPIN)*, pp. 744–749, Mar. 2023.

A. Abdelrahman and S. Viriri, “EfficientNet family U-Net models for deep learning semantic segmentation of kidney tumors on CT images,” *Frontiers in Computer Science*, vol. 5, 1235622, 2023.