UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingeniería

Medical Image Segmentation for the Detection of Neck and Head Tumors

Juan Ignacio Arguello Herrera
Isabela González Custode
Sebastián Jácome Mejía

Ingeniería Industrial

Trabajo de fin de carrera presentado como requisito
para la obtención del título de
Ingeniero Industrial

Quito, 11 de diciembre de 2024

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias e Ingeniería

HOJA DE CALIFICACIÓN
DE TRABAJO DE FIN DE CARRERA

Medical Image Segmentation for the Detection of Neck and Head Tumors

Juan Ignacio Arguello Herrera

Isabela González Custode

Sebastián Jácome Mejía

Nombre del profesor, Título académico **María Gabriela Baldeón, PhD**

Quito, 11 de diciembre de 2024

© DERECHOS DE AUTOR

Nombres y apellidos: Juan Ignacio Arguello Herrera, Isabela Gonzalez Custode, Sebastián Jácome Mejía

Código: 00214586, 00320503, 00320492

Cédula de identidad: 1720099918, 1721159869, 1719898411

Lugar y fecha: Quito, 11 de diciembre de 2024

ACLARACIÓN PARA PUBLICACIÓN

UNPUBLISHED DOCUMENT

RESUMEN

Los tumores de cabeza y cuello pueden ser difíciles de detectar y tratar, especialmente si se tiene en cuenta que la detección temprana es crucial para el resultado del paciente. En esta investigación, se presenta un enfoque que utiliza una red neuronal convolucional U-Net para detectar tumores de cabeza y cuello. Este estudio se realizó utilizando el conjunto de datos HNTSMRG 2024 Challenge. Además, se aplicó la optimización bayesiana para optimizar los hiperparámetros de la tasa de aprendizaje y el tamaño del lote, lo que resultó en una mejora significativa en las métricas del modelo: intersección sobre unión y coeficiente de similitud de dados. La U-Net propuesta logró un IoU de 0,23 y 0,35 para las clases 1 y 2, respectivamente; y una puntuación de dados de 0,37 a 0,51. El modelo optimizado con Optimización Bayesiana obtuvo mejores resultados: 0,31 para la clase 1 y 0,36 para la clase 2 en términos de IoU, y 0,48 y 0,52 para el Coeficiente de Dice. Estos resultados demuestran el potencial del aprendizaje profundo para la tarea de segmentación de imágenes y los beneficios de aplicar la optimización de hiperparámetros para mejorar los resultados.

Palabras clave: Tumores de cabeza y cuello, Detección de tumores, U-Net, Deep Learning, Redes Neuronales Convolucionales.

ABSTRACT

Head and neck tumors can be challenging to detect and treat, especially considering that early detection is crucial for the patient's outcome. This paper presents an approach using a U-Net Convolutional Neural Network to detect head and neck tumors. This study was made using the HNTSMRG 2024 Challenge dataset. Furthermore, Bayesian Optimization was applied to optimize the hyperparameters of learning rate and batch size, resulting in a significant improvement in the model's metrics: Intersection Over Union and Dice Similarity Coefficient. The proposed U-Net achieved an IoU of 0.23 and 0.35 for classes 1 and 2, respectively; and a Dice Score of 0.37 an 0.51. The model optimized with Bayesian Optimization achieved better results: 0.31 for class 1 and 0.36 for class 2 in terms of IoU, and 0.48 and 0.52 regarding the Dice Coefficient. These results demonstrate the potential of deep learning for image segmentation task and the benefits of applying hyperparameter optimization to improve results.

Key words: Head and neck tumor, Tumor detection, U-Net, Deep Learning, Convolutional Neural Networks.

TABLE OF CONENTS

TABLES INDEX

FIGURES INDEX

INTRODUCTION

Head and neck cancer is one of the most common types of cancer, resulting in 660,000 cases and 325,000 deaths in 2020 (Sung et al., 2021). These types of cancer can vary, depending on their stage. Stages I and II usually have a small tumor and almost no involvement of lymph nodes; whereas stage III and IV tumors invade surrounding structures and may have more lymph nodes (Chow, 2020). If detected in early stages, the five-year overall survival rate is 90%, while in later stages the rate drops to 40% (Mody et al., 2021).

Currently, manual segmentation is done by a radiologist, or an oncologist specialized in these types of cancer. This task can be challenging and time consuming, as the doctor must examine every patient's image; and expensive for the patient, since they need to pay for a trained professional to analyze their case. Furthermore, this method is prone to mistakes as fatigue or human error can impact the outcome of a doctor. Lastly, this way of detecting tumors has limited scalability, because of the manual and time-consuming approach, which does not allow for every patient to receive their diagnosis in the optimal time.

On the other hand, automated systems can process thousands of images in a fraction of the time required by medical professionals, which can help scaling the number of images segmented. Furthermore, the use of computers allows for similar diagnosis, rather than being subject to the doctor's bias or human error. Lastly, automated segmentation can help doctors in their decision-making process, by acting as a guide for tumor detection.

With the improvement of computational capacity and efficiency, deep learning algorithms became more popular for the task of medical image segmentation, thanks to their ability to segment images; specifically, convolutional neural networks (Ronneberger et al., 2015). In recent years, many studies have used convolutional neural networks to achieve segmentation for medical purposes. For example, Zhao, et al., (2019) and Tang, et al., (2021) implemented U-Net arquitectures with convolutions and attention mechanisms, respectively, to segment GTV tumors, obtaining Dice Scores of 0.64 and 0.0.67.

Neural Networks have hyperparameters that can be set to improve the algorithm's performance. However, setting these hyperparameters can be hard because of the "missing guidance on the choice and configuration of hyperparameter optimization methods for the problem at hand; difficulty to define the search space of a hyperparameter optimization process appropriately" (Bischl, et al., 2023). Optimizing hyperparameters can significantly increase the model's performance, as Borgli, et al., (2019) showed in his study, where through the use of Bayesian Hyperparameter Optimization, the Dice Score increased by around 10%, compared to the base model on the ImageNet Dataset; or the case of Borgli, (2018) where the F1 score was improved by 13% when optimizing hyperparameters on the Kvasir dataset. Both of these studies focused on using Bayesian Optimization in a Convolutional Neural Network (CNN), which is a type of neural network that uses convolutional layers to detect patterns and features, with Transfer Learning, on gastrointestinal cancer.

In light of the above, the aim of this work is to propose the use of a U-Net model for medical image segmentation in head and neck tumors and implement Bayesian Optimization for hyperparameter tuning. In this work we use the HNTSMRG 2024 Challenge dataset (Wahid et. al, 2024) to segment GVTp and GTVn, the evaluation metrics selected are IOU (Intersection Over Union) and Dice Coefficient. After 200 epochs with a batch size of 4 and a learning rate of 1e-3, the model achieved IoU scores of 0.23 and 0.34 for classes 1 and 2, respectively, with corresponding Dice coefficients of 0.51 and 0.34.

TOPIC DEVELOPMENT

LITERATURE REVIEW

Deep learning has become a crucial tool in medical image segmentation, especially in the detection of head and neck tumors. This literature review describes models utilized in this field, followed by different methods for hyperparameter optimization in medical images.

i. Models for the detection of head and neck tumors

Zhao, et al., (2019) designed a U-Net with dilated convolutions to capture contextual information for the image without increasing the filter size or the number of parameters. It replaces the standard convolutional layers with dilated convolutions increasing the receptive field (the amount of image area that a particular neuron can process). Their model was able to achieve a Dice Score of 0.644. Moreover, Ali, et al., (2024) designed a flexible SAM model, that uses encoder and decoder components. It can apply previous knowledge to unseen data making it very useful for image segmentation tasks such as MRI brain tumor detection. Their study obtained a Dice Score of 0.676 and an IOU of 0.689. Tang, et al., (2021) improves the U-Net architecture by adding two types of attention mechanisms, Position attention module (PAM) and Channel attention module (CAM). The first type of attention (PAM) helps the network focus on relevant regions such as the tumors. The second attention type (CAM) adjusts the importance of the feature channels capturing precise tumor features. They obtained a Dice Score of 0.785 (2021). Furthermore, Ghimire, et al., (2021) did a Patch Based 3D U-net model that uses patches (smaller pieces of the image) so that the model focuses on a small area at the time. They obtained a Dice Score of 0.691. Finally, Mohammad, et al., (2021) made a U-net with transformers. This model works by using a U-net and then transformers, which pay attention to the relationships between far away parts of the image allowing it to focus on global context and learning how all the parts of the image fit together. They were able to get a Dice Score of 0.680.

ii. Methods for hyperparameter optimization:

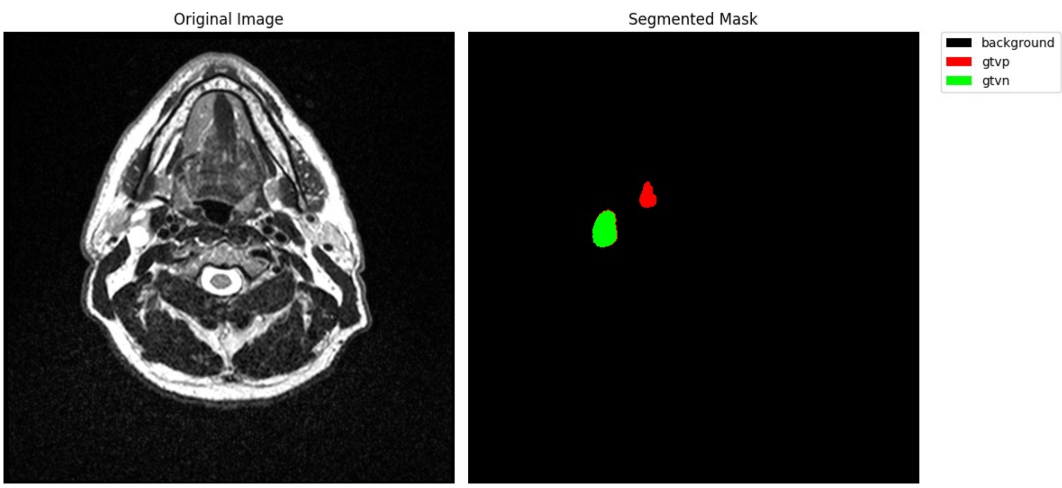
There are several techniques used for hyperparameter optimization such as grid search or random search. Grid search is an exhaustive method that tests all possible combinations of predefined values of hyperparameters, making it very computationally expensive. On the other hand, random search instead of testing all possible combinations, it randomly chooses a subset of these predefined values and tests them, which means it can omit finding the best set of optimal hyperparameters (Alibrahim & Ludwig, 2021). Deepa, et al., (2021) applied a stochastic gradient descent (SGD) in a support vector machine. SGD is an optimization method that updates the model's parameters in smaller steps based on gradients, and it can incorporate hyperparameters such as learning rate, weight decay, and momentum to minimize the loss function. Unlike the Bayesian method, SGD doesn't use probabilities to make future combinations. Applying this method they obtained a 3% increase in accuracy. Additionally, Borgli, et al., (2019) applied Bayesian Optimization in a Convolutional Neural Network (CNN) in Keras in two gastrointestinal datasets and obtained a 10% increase in Dice Score. Also, Borgli, (2018) did another Bayesian Optimization on transfer learning for medical image classification and obtained an increase in F1 Score of 13% over the base model. Lastly, Gao and Ding, (2020) performed a Bayesian Optimization in different classifiers, such as XGBoost, GBDT, LightGBM and obtained an increase of approximately 2% overall on the F1 Score.

METHODOLOGY:

This section describes the characteristics of the dataset used from HNTSMRG 2024 Challenge. Additionally, it provides an in-depth overview of the U-Net architecture and the Bayesian optimization method for hyperparameter tuning. Finally, it explains the IoU and Dice Coefficient metrics used in the study.

- **Database and preprocessing**

The Training Dataset for the HNTSMRG 2024 Challenge (Wahid et al., 2024) includes 5,331 head and neck medical images from 150 anonymous patients, originally in .nii.gz format. These were converted to .png and standardized to a 512x512 resolution to ensure consistency. Images were normalized to a grayscale range of 0-255, converted to three channels, and zero-padded as needed. Masks underwent similar preprocessing with nearest-neighbor interpolation to preserve class labels, as shown in Figure I.



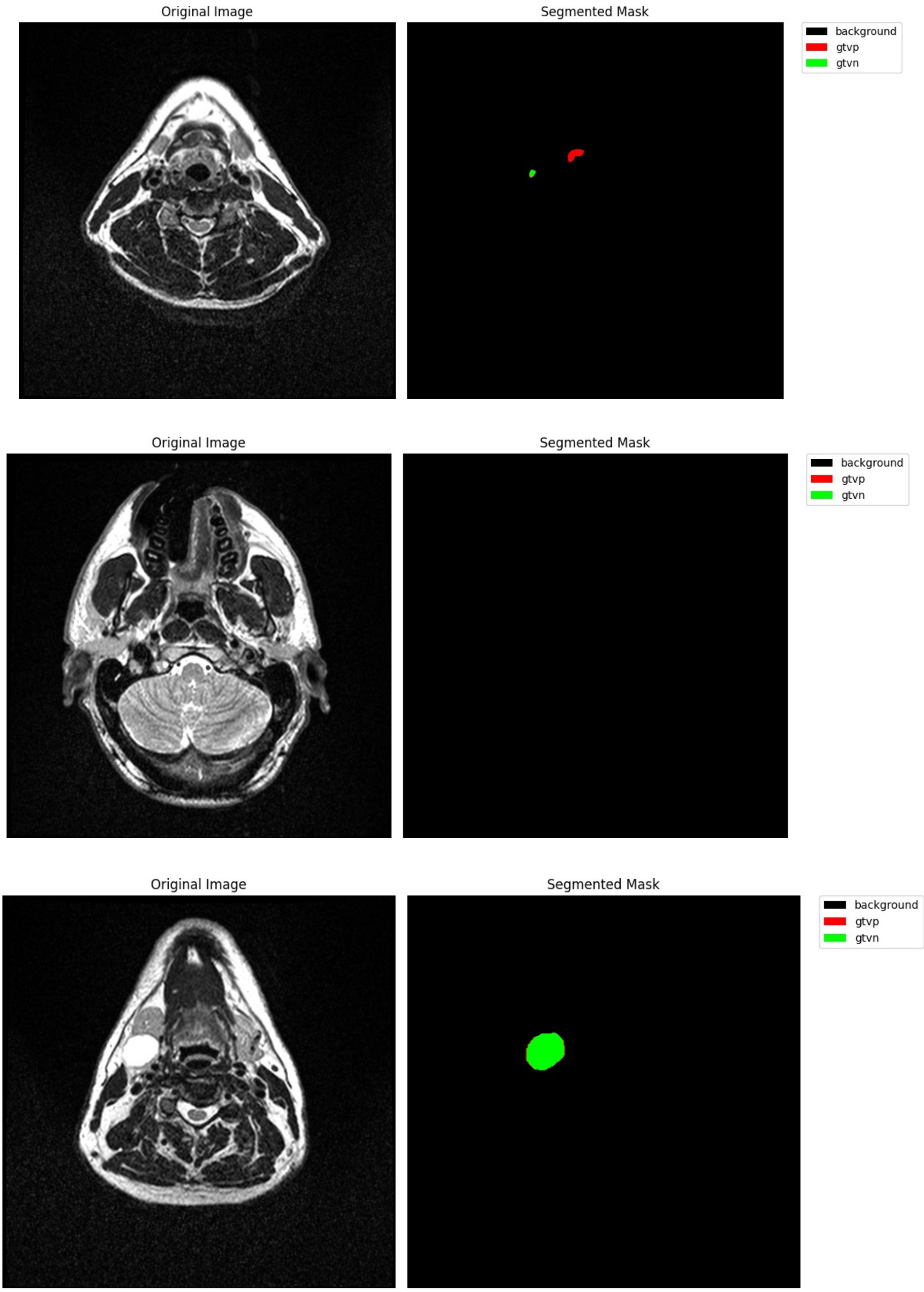


Figure I: Processed Image and Mask

The original dataset exhibited a significant class imbalance, with a large proportion of images lacking tumors. This imbalance caused the model to primarily predict the background class (class 0). To address this, a balancing strategy was applied to ensure that 80% of the images contained either class 1 (GVTp), class 2 (GVTn), or both, while only 20% contained no tumors. This process reduced the dataset size to 2,386 images and masks, divided into training, validation, and test sets with a 70-15-15 split. Before balancing, only 11.44% of the training images, 11.50% of the validation images, and 8.32% of the test images contained all classes. After balancing, these proportions increased to 21.13%, 23.85%, and 17.89% for the training, validation, and test sets, respectively.

- **Convolutional Neural Networks:**

The model used for the detection of tumors is a specific type of convolutional neural network (CNN). This network is especially useful in learning spatial features of images by applying several layers to the data (Gu, et al., 2018). The key components of the network are:

1. Convolutional layer: This layer receives as input the image and applies a filter (set of weights) to extract important features of the input. This filter, also known as kernel, passes through all the image detecting patterns. Afterwards it computes the dot product between the pixels and the weight giving as an output a feature map (Wu, 2017).
2. Activation function: After the convolution layer is applied then it passes through an activation function which basically makes different transformations, depending on the function used, to the values in the feature map. It helps introduce nonlinearity to the model making it better at capturing complex patterns in the data (Wu, 2017).
3. Pooling layer: After the transformations are applied the next step is the pooling layer. This layer computes a down sampling operation. This means it reduces the dimensions of the image by applying some methods such as max pooling or average pooling. It works like the convolutional layer as it also applies a filter to make the down sampling. However, instead of making the dot product between the feature map and the filter it

takes the average or the max value giving as an output a reduced feature map (Wu, 2017).

4. Fully connected layer: This layer is the final layer, and it acts as the classifier. This stage of the network connected all the neurons of the previous layers to consolidate and extract the most important features learned to make a final prediction. It calculated the sum of all the weights and biases of all the neurons in the network. Afterwards, another activation function is applied such as Relu, Softmax or Sigmoid, depending on the case of the problem. The output of this layer is a final vector with values equal to the number of classes, where each value represents the probability for each class (Wu, 2017).

- **U-net arquitecture**

In this study U-net, which is a type of convolutional neural network, was chosen. This model is especially useful in biomedical predictions such as tumor detection. It is composed of two main parts: the encoder (contracting stage) and the decoder (reconstructing stage). The encoder reduces the dimension of the input image by applying convolutional layers, activation functions and a pooling layer. In the center of the architecture lies a bottleneck, which works as a bridge for the encoder and decoder, where the image is in its smallest dimension and the model can focus on the most important features. Additionally, the decoder reconstructs the image, returning it to the initial dimensions, while maintaining all the characteristics learned from the first stage. It applies up-sampling techniques and skip connections to concatenate what the model learned from the encoder to the decoder. Finally, a convolution generates a final output where each pixel is designated a specific class prediction (Zahra, et al., 2020).

Structure	**Type**

Objective function	Cross-Entropy Loss
Optimizer	Adam
Encoder	Convolution: DoubleConv
	Activation function: Relu
	Pooling: MaxPool2d
Decoder	Up sampling: ConvTranspose2d
	Skip connections
Output	Output convolution

Table I: Parameters U-net model

The U-net model implemented in this study integrates the components shown in Table I. It uses a cross-entropy loss objective function which measures the difference between the predicted probability and the actual probability of the mask. The goal of this function is to highly penalize incorrect predictions. Moreover, to update the weights of the network the Adam optimizer is used. This optimizer updates the models' parameters considering two aspects: the moving average of the gradients (momentum) and their squared values. In the first stage (encoder), the model incorporates a consecutive double convolution (DoubleConv) to extract complex features from the input image. Afterwards, an activation function (Relu) is applied to the feature map generated from the first layer to introduce nonlinearity. This enables the model to generalize and learn complex patterns. To reduce the dimension of the image and capture the most important features a Maxpool2d layer is applied. Furthermore, in stage 2 (decoder) the model uses ConvTranspose2d up sampling technique to reconstruct the image while maintaining what the model learned from the encoder stage while applying skip connections. These connections concatenate the information captured in the encoder to the decoder. Finally, a final convolution is applied to generate an output image where each pixel represents a class prediction.

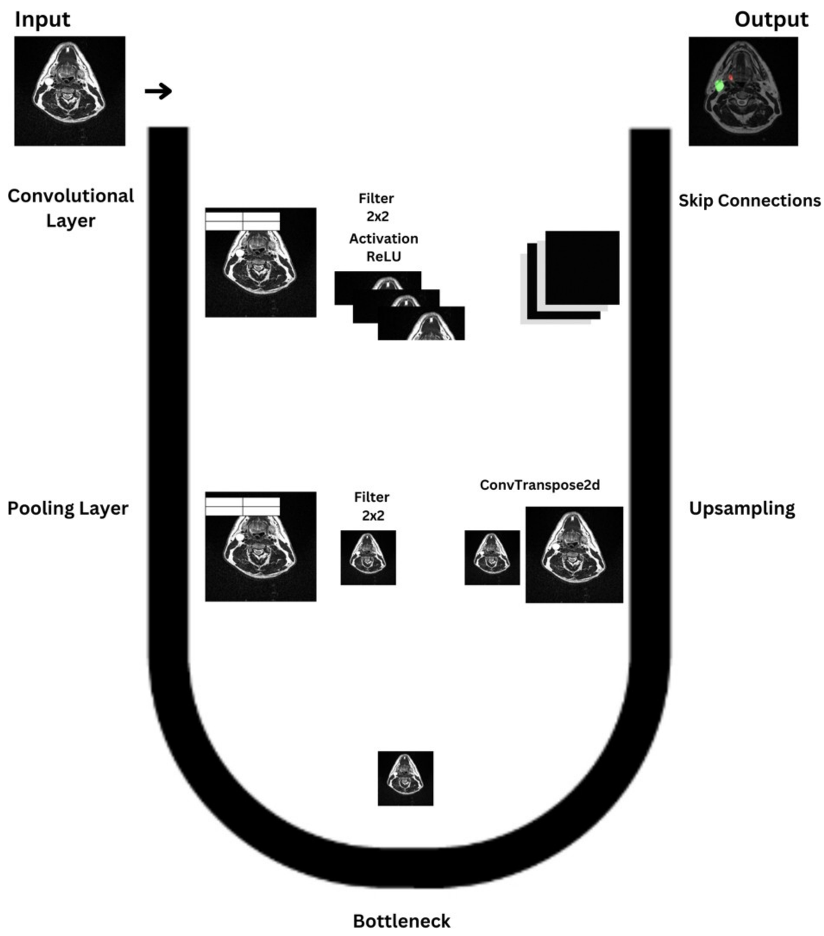


Figure II: U-net Structure

- **Bayesian optimization**

The U-net model can improve its final prediction by adding a hyperparameter optimization method, which includes finding an optimal combination of hyperparameters that will assure a better performance of the model.

Given that we have a large and complex dataset a Bayesian Optimization is applied. This technique provides an efficient way to find the optimal configurations by minimizing the number of evaluations needed using a probabilistic map to guide the search for the best hyperparameters (Gao et al., 2023). It uses prior information to find the configuration that obtains the minimal loss function and the highest accuracy (Victoria & Maragatham, 2021).

The hyperparameters chosen for optimization were learning rate and batch size (table II) that were defined prior to the training of the dataset. The algorithm ran the model with Bayesian optimization for 6 trails and 20 epochs. This means that the model was trained for a total of 120 epochs. For every trail the model defined a different combination of hyperparameters learned from the probabilistic map generated. The selected combinations are shown in table 3. Afterwards, the best combination of hyperparameters is stored in a path using the following libraries: Optuna, Pandas, OS, Torch, TQDM, and Albumentations. This path, that stores the best model, is then passed for another set of training of 200 epochs.

Hyperparameters	**Range**
Batch Size	2-16
Learning rate	1e-5 – 1e-1

Table II: Ranges chosen for hyperparameters

Trial	**Batch Size**
1	13
2	4
3	4
4	8
5	5
6	3

Table III: Trials and Batch Size chosen by Bayesian Optimization

- **Data Augmentation**

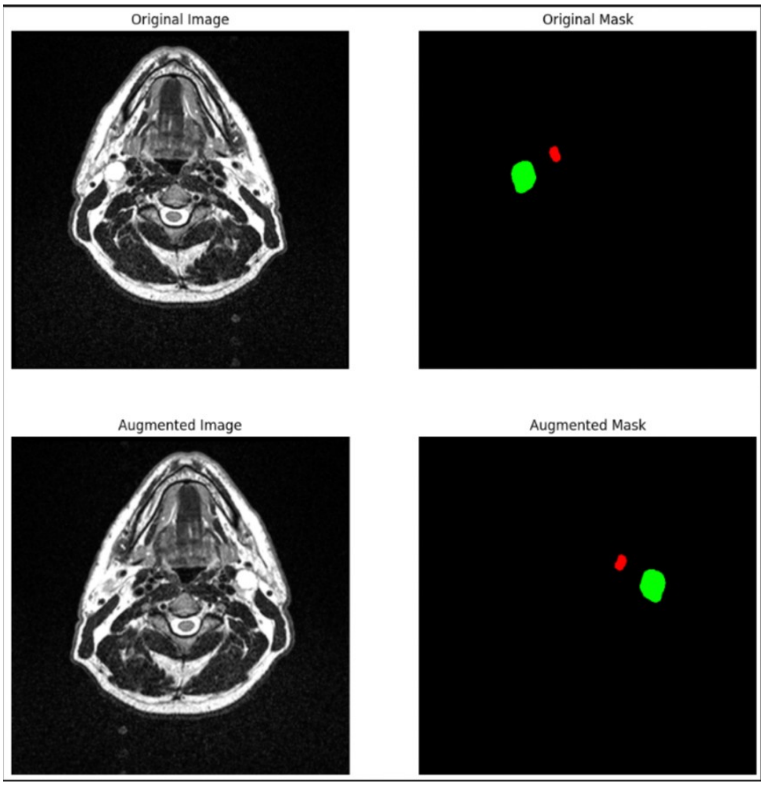


Figure III: Image and Mask after augmentation

Augmentation techniques were applied to the training set, including horizontal and vertical flipping, random 90-degree rotations, brightness and contrast adjustments, scaling, padding, random cropping, noise addition, and elastic deformations, as shown in Figure III. These augmentations improved model robustness by simulating various conditions and ensuring masks aligned with their corresponding images.

RESULTS

Trial	Batch Size	Learning Rate	Validation Loss
1	13	0.01979	0.04457
2	4	0.00027	0.03625
3	4	0.00019	0.04175
4	8	0.00007	0.03652
5	5	0.00849	0.04394
6	3	0.01548	0.04292

Table IV: Bayesian Optimization results based on validation loss

Trials with lower learning rates (e.g., 0.00027 and 0.00007) achieved the lowest validation losses, demonstrating the effectiveness of Bayesian optimization in identifying optimal hyperparameters, as shown in Table IV. The variation in batch sizes did not show a clear correlation with performance, suggesting that learning rate optimization plays a more crucial role in enhancing model performance. The model from Trial 2, which achieved the lowest validation loss, was then trained for 200 epochs to compare its performance with the original U-Net model.

Model/Metrics	Test Loss	IoU		Dice Coefficient	
		Class 1	*Class 2*	*Class 1*	*Class 2*
UNET	0.03724	0.2298	0.3449	0.3737	0.5129
UNET with Bayesian Optimization	0.01507	0.31091	0.35559	0.4744	0.5246

Table V: Performance metrics on the test set for U-Net with and without Bayesian optimization

Bayesian optimization significantly improves U-Net's performance across all metrics. The test loss decreased from 0.03724 to 0.01507, indicating better performance. Both

Intersection over Union (IoU) and Dice Coefficient values increased for classes 1 and 2, with the most notable improvement in IoU for class 1 (from 0.2298 to 0.31091). Overall, these results demonstrate the effectiveness of Bayesian optimization, even by changing only one hyperparameter in comparison to the original UNET model, in enhancing segmentation performance by fine-tuning hyperparameters.

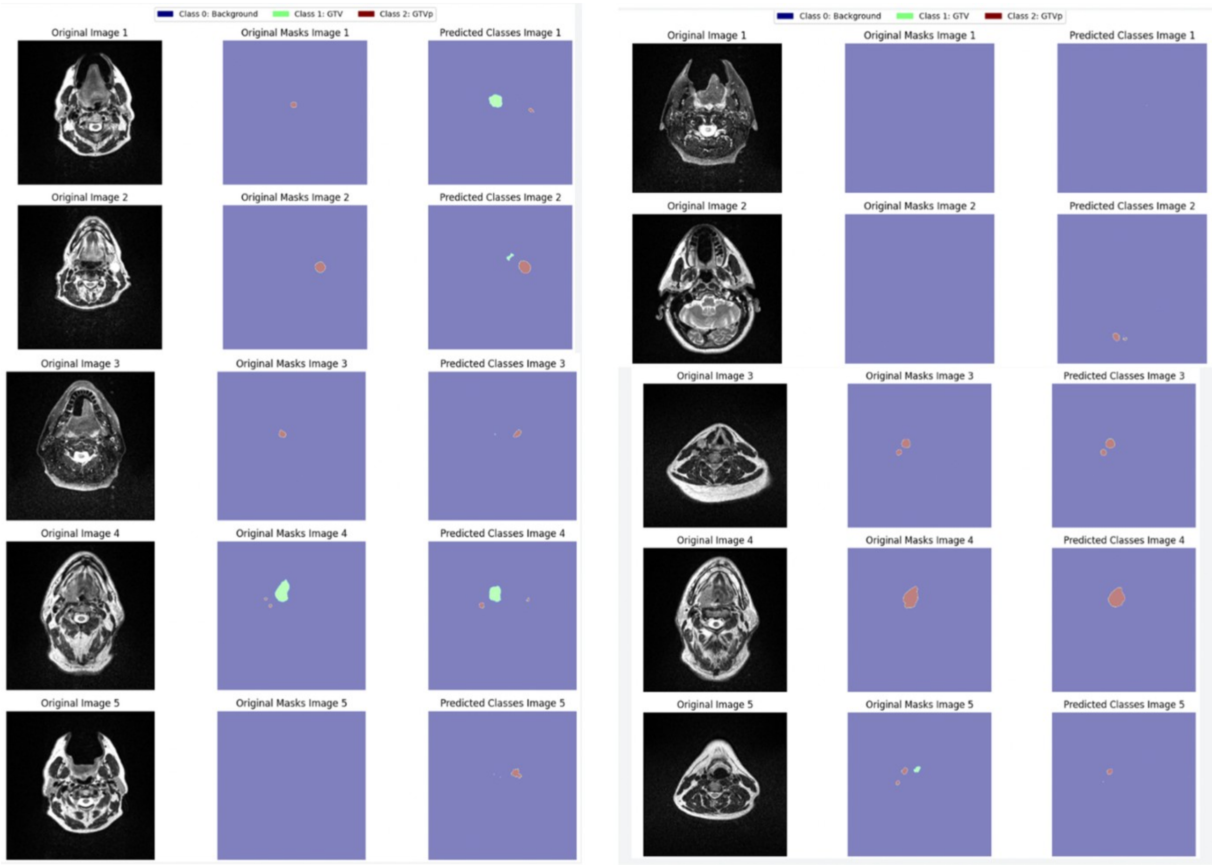


Figure IV: Examples of segmentation results for UNET model (left) and UNET with Bayesian Optimization (right)

Figure IV shows examples of segmentation results for both models. Both models perform a decent segmentation by identifying the presence of GTVp and GTVn tumors. However, there is an improvement when using Bayesian Optimization, especially at the segmentation of the primary tumor (GTVp).

DISCUSSION:

Our results demonstrated that using Bayesian Optimization, IoU and Dice Coefficient can have a better performance, when compared to a base model. Nevertheless, the Dice Similarity Coefficient and IoU achieved by our base model indicate that there is room for improvement. This performance level is not sufficient for clinical applications, where segmentation accuracy must meet high standards to ensure the patient's safety. However, what should be noted is that our Bayesian Hyperparameter optimization did improve the model's performance significantly.

Comparing our results to previous studies, we notice that the dice coefficient metric for the U-Net did not perform as good as the studies mentioned above. Zhao, et al., (2019), Ali, et al., (2024), Tang, et al., (2021) and Mohammad, et al., were able to reach scores closer to 0.6 on their Dice Score; and (2021) Ghimire, et al., (2021) was able to obtain a score that was close to 0.8. This can be because of the modifications in the architectures that they used, like using transformers or adding attention mechanisms to their models, which is something our U-Net did not include.

However, when comparing the results obtained from the Bayesian Optimization with the base model, we can see that we got an increase of 35% for class 1 and 3.5% for class 2, in terms of IoU; 27% for class 1 and 2.3% for class 2, in terms of Dice Coefficient. When comparing our Dice Coefficient increase to the improvement made by the studies of Deep, et al., (2021), Borgli, et al., (2019), Borgli, (2018) and Gao and Ding, (2020), we can see that class 2 got similar results to them, having a one-digit increase. However, for class 1, through the optimization of hyperparameters, we were able to have a significant increase in both the Dice Coefficient and the IoU; showing that just by changing a hyperparameter, we can see a notable improvement in the model's performance.

A limitation that was presented to us was the time constraint. If we had more time, we could run more trials with different combinations for the optimization or increase the number of epochs. A solution for this could also be implementing early stopping techniques for neural networks (Bai, et al., 2021), where the model can stop on its own once it realizes that metrics are not improving. For future works, architectural enhancements can be made. As shown before, previous studies have achieved notable results by modifying the U-Net architecture; future work could explore incorporating techniques like transformers to improve the model´s performance.

CONCLUSIONS

In conclusion, this study shows an approach to segment images to detect head and neck tumors and the effectiveness of optimizing hyperparameters for deep learning models. Using a U-Net Convolutional Neural Network, we were able to segment images based on not only if they had a tumor, but also the type of tumor. By applying Bayesian Optimization to optimize the hyperparameters of learning rate and batch size, significant improvements were achieved in the model's performance metrics. The results demonstrate the potential of deep learning models to help patients with their diagnosis, reducing the work needed from medical professionals and costs for the patient. These findings emphasize the importance of both neural network architectures and hyperparameter optimization in enhancing the performance of medical image segmentation models. Furthermore, other metrics can be analyzed, such as the Hausdorff distance.

REFERENCES

Aljabri, M., & AlGhamdi, M. (2022). A review on the use of deep learning for medical images segmentation. In Neurocomputing (Vol. 506, pp. 311–335). Elsevier B.V. https://doi.org/10.1016/j.neucom.2022.07.070

Ali, L., Alnajjar, F., Swavaf, M., Elharrouss, O., Abd-Alrazaq, A., & Damseh, R. (2024). Evaluating segment anything model (SAM) on MRI scans of brain tumors. Scientific Reports, 14(1), 21659.

Bai, Y., Yang, E., Han, B., Yang, Y., Li, J., Mao, Y., ... & Liu, T. (2021). Understanding and improving early stopping for learning with noisy labels. Advances in Neural Information Processing Systems, 34, 24392-24403.

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., ... & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13(2), e1484.

Borgli, Rune Johan, et al. "Automatic hyperparameter optimization for transfer learning on medical image datasets using bayesian optimization." 2019 13th international symposium on medical information and communication technology (ISMICT). IEEE, 2019.

Borgli, Rune Johan. (2018) " Hyperparameter optimization using Bayesian optimization on transfer learning for medical image classification." University of Oslo, 2018.

Caeiro, J., Dirección, C., Mosqueira, E., María, R. E., & Pereira, H. (n.d.). Modelo de aprendizaje profundo para la segmentación y clasificación de imágenes médicas Estudante.

Chow, L. Q. (2020). Head and neck cancer. New England Journal of Medicine, 382(1), 60-72.

Deepa, N., Prabadevi, B., Maddikunta, P. K., Gadekallu, T. R., Baker, T., Khan, M. A., & Tariq, U. (2021). An AI-based intelligent system for healthcare analysis using Ridge-Adaline Stochastic Gradient Descent Classifier. The Journal of Supercomputing, 77, 1998-2017.

Du, G., Cao, X., Liang, J., Chen, X., & Zhan, Y. (2020). Medical Image Segmentation based on U-Net: A Review. Journal of Imaging Science & Technology, 64(2).

Gao, S., Zhou, H., Gao, Y., & Zhuang, X. (2023). BayeSeg: Bayesian Modeling for Medical Image Segmentation with Interpretable Generalizability. http://arxiv.org/abs/2303.01710

Ghimire, K., Chen, Q., & Feng, X. (2021). Patch-based 3D UNet for head and neck tumor segmentation with an ensemble of conventional and dilated convolutions. In Head and Neck Tumor Segmentation: First Challenge, HECKTOR 2020, Held in Conjunction with MICCAI

2020, Lima, Peru, October 4, 2020, Proceedings 1 (pp. 78-84). Springer International Publishing.

Goceri, E. (2023). Medical image data augmentation: techniques, comparisons and interpretations. Artificial Intelligence Review, 56(11), 12561-12605.

Gómez, V., Tutor, J., Nolla, C., Gómez Adrían, F., & Ander Peris Abril, J. (n.d.). Redes neuronales para la clasificación y segmentación de imágenes médicas TRABAJO FIN DE MÁSTER.

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. (2018). Recent advances in convolutional neural networks. Pattern recognition, 77, 354-377.

Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process, 5(2), 1.

Mody, M. D., Rocco, J. W., Yom, S. S., Haddad, R. I., & Saba, N. F. (2021). Head and neck cancer. The Lancet, 398(10318), 2289-2299.

Oreiller, V., Andrearczyk, V., Jreige, M., Boughdad, S., Elhalawani, H., Castelli, J., ... & Depeursinge, A. (2022). Head and neck tumor segmentation in PET/CT: the HECKTOR challenge. Medical image analysis, 77, 102336.

Prakash, O., Sudipta, V., Subhash, R., Pandey, C., & Mittal, M. (n.d.). Advancement of Machine Intelligence in Interactive Medical Image Analysis. http://www.springer.com/series/16171

Rahman, H., Bukht, T. F. N., Imran, A., Tariq, J., Tu, S., & Alzahrani, A. (2022). A Deep Learning Approach for Liver and Tumor Segmentation in CT Images Using ResUNet. Bioengineering, 9(8). https://doi.org/10.3390/bioengineering9080368

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.

Salmanpour, M. R., Hajianfar, G., Rezaeijo, S. M., Ghaemi, M., & Rahmim, A. (2021). Advanced automatic segmentation of tumors and survival prediction in head and neck cancer. In 3D Head and Neck Tumor Segmentation in PET/CT Challenge (pp. 202-210). Cham: Springer International Publishing.

Santos, M. (2022). Bayesian Optimization for Hyperparameter Tuning. Journal of Bioinformatics and Artificial Intelligence, 2(2), 1-13.

Sebastián, W., López, G., Milena, P., & Vaca, I. (n.d.). UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ Cardiac lesions: a comparative study based on Convolutional Neural Networks and Visual Transformers in the segmentation of MR cardiac images.

Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. IEEE Access, 9, 82031–82057. https://doi.org/10.1109/ACCESS.2021.3086020

Sung H, Ferlay J, Siegel R L et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021; 71: 209-249.

Tang, P., Zu, C., Hong, M., Yan, R., Peng, X., Xiao, J., ... & Wang, Y. (2021). DA-DSUnet: dual attention-based dense SU-net for automatic head-and-neck tumor segmentation in MRI images. Neurocomputing, 435, 103-113.

van Beers, F., Lindström, A., Okafor, E., & Wiering, M. (2019, March). Deep neural networks with intersection over union loss for binary image segmentation. In Proceedings of the 8th international conference on pattern recognition applications and methods (pp. 438-445). SciTePress.

Victoria, A. H., & Maragatham, G. (2021). Automatic tuning of hyperparameters using Bayesian optimization. Evolving Systems, 12(1), 217–223. https://doi.org/10.1007/s12530-020-09345-2

Wahid, K., Dede, C., Naser, M., & Fuller, C. (2024). Training Dataset for HNTSMRG 2024 Challenge [Data set]. Zenodo. https://doi.org/10.5281/zenodo.11199559

Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 5(23), 495.

Yu, H., Caldwell, C., Mah, K., Poon, I., Balogh, J., MacKenzie, R., ... & Tirona, R. (2009). Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. International Journal of Radiation Oncology* Biology* Physics, 75(2), 618-625.

Zhao, B., Soraghan, J., Di Caterina, G., & Grose, D. (2019, September). Segmentation of head and neck tumours using modified U-net. In 2019 27th European signal processing conference (EUSIPCO) (pp. 1-4). IEEE.

Zahra, E., Ali, B., & Siddique, W. (2020). Medical image segmentation using a U-Net type of architecture. arXiv preprint arXiv:2005.05218.