© MAY 2023 | IRE Journals | Volume 6 Issue 11 | ISSN: 2456-8880

A Survey on Medical Image Segmentation in Deep Learning

KARTHIK PANICKER[1], SHUBHRANSHU GUPTA[2], JAYANT BHOWMICK[3]
[1, 2, 3] *Student, Department of Electronics &Telecommunication, Pune Institute of Computer Technology, Pune, Maharashtra, India*

Abstract— Lung cancer is the leading cause of death in most countries around the world. Radiologists usually segment images manually, which takes time and results in less precise results. Since the early discovery of tumors can assist radiologists in determining their nature, sort, and mode of therapy, tumor detection and segmentation from CT Scan images is an important area of study. For experts, automated segmentation facilitates faster data analysis. Deep learning is a machine learning algorithm technique. In deep learning, data transformation takes the form of layers. This paper proposes deep learning, which is based on convolutional neural networks and is mostly used for image processing. The recommended Lung tumor segmentation is being developed utilizing cutting-edge technology such as deep learning algorithms, and CNN, to segment and identify tumors and decide which deep learning approach is most suited for segmenting. Two examples of deep learning techniques are UNET and UNETR. These techniques are frequently applied to segmentation-related problems. The dataset for the segmentation of lung tumors is made available through the NSCLC and Decathlon challenge and consists of lung CT scans.

Indexed Terms— Convolutional Neural Network, Deep Learning, Segmentation, Transformer.

I. INTRODUCTION

Lung cancer is a devastating disease that is responsible for a significant number of cancer-related deaths worldwide. The lungs, which are essential components of the respiratory system, are responsible for oxygenating the blood and removing carbon dioxide from the body. Unfortunately, exposure to harmful substances, such as tobacco smoke or radon gas, can cause mutations in the lung cells that result in the growth of malignant tumors. Due to the lack of symptoms or the similarity of symptoms to respiratory illnesses, lung cancer is typically discovered at an advanced stage, making it difficult to treat effectively. Therefore, the development of a treatment strategy that can diagnose cancer at an early stage is crucial. One way to improve the diagnosis and treatment of lung cancer is through the segmentation of lung tumors. Various studies have been conducted in this area, with researchers utilizing different techniques to segment lung tumors accurately. Manual segmentation by radiologists can occasionally lead to incorrect results due to interobserver variability. On the other hand, semiautomatic segmentation has been shown to have lower interobserver variability than manual segmentation. To address this issue, automatic segmentation of lung cancer using deep learning networks has been explored. These networks have shown promising results in accurately segmenting lung tumors and can be useful in developing treatment strategies that can diagnose lung cancer early and improve patient outcomes.

II. ALGORITHM

Medical image segmentation is a critical component of many clinical applications, such as disease diagnosis, treatment planning, and image-guided interventions. It involves separating the target organ or tissue from the background in medical images, which is essential for accurate and efficient analysis of the image data. While manual segmentation by a radiologist is the gold standard, it can be time-consuming and prone to inter- and intra-observer variability. Automated segmentation algorithms based on machine learning, deep learning, and other computational techniques have shown great promise in improving segmentation accuracy and reducing the time required for segmentation. However, there are still challenges to be addressed, such as handling different imaging modalities, dealing with noise and artifacts, and achieving robust performance across different patient populations.

Training an AI model for lung tumour segmentation typically involves the following steps:
2.1 Image acquisition: This involves capturing medical images using various imaging modalities such as X-ray, CT, MRI, ultrasound, etc. The quality of the acquired images is critical for accurate segmentation.

2.2 Pre-processing: This step involves applying various image processing techniques to the acquired images to enhance their quality, reduce noise, and improve contrast. Common pre-processing techniques include filtering, smoothing, and histogram equalization.

2.3 Segmentation: This is the main step in which the target organ or tissue is identified and separated from the background using various techniques such as thresholding, edge detection, and region growing. The choice of technique depends on the characteristics of the image and the type of organ or tissue being segmented.

2.4 Postprocessing: This step involves refining the segmented image to remove any remaining noise or artifacts, and to smooth the boundaries of the segmented region. This is important to ensure that the final segmented image is accurate and visually appealing. Common postprocessing techniques include morphological operations, smoothing, and contour refinement.

1. CNN Based Model

1.1 U-Net:
The U-Net is a convolutional neural network architecture that was proposed by Ronneberger et al. in 2015 for biomedical image segmentation. The U-Net architecture consists of an encoder path that captures the context of the image and a decoder path that enables precise localization of the segmented object. The encoder path of the U-Net consists of a series of convolutional and max-pooling layers that progressively reduce the spatial dimensions of the input image while increasing the number of channels. The decoder path of the U-Net consists of a series of deconvolutional and up sampling layers that increase the spatial dimensions of the output while reducing the number of channels. Skip connections are added between the corresponding layers of the encoder and decoder paths to preserve the spatial information and enable precise localization of the segmented object. The architecture consists of a contracting path and an expansive path, where the contracting path follows the typical architecture of a convolutional network, consisting of repeated applications of two 3x3 convolutions, each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for down sampling. The expansive path, on the other hand, consists of up sampling layers followed by 2x2 convolutions that reduce the number of feature channels, concatenation with the corresponding cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The performance of the U-Net architecture has been evaluated in various studies using different medical imaging datasets. For example, a study by Çiçek et al. (2016) evaluated the U-Net architecture for prostate segmentation using the publicly available PROMISE12 dataset. The U-Net achieved a mean Dice similarity coefficient of 0.863, which outperformed several other state-of-the-art methods. In another study by Isensee et al. (2018), the U-Net architecture was evaluated for brain tumour segmentation using the BraTS 2017 dataset. The U-Net achieved a mean Dice similarity coefficient of 0.858 for whole tumor,0.775 for tumour core and 0.647 for the contrast enhancing tumour, which was among the top-performing methods in the challenge. The author trained and tested the U-Net model on the ISBI cell tracking challenge dataset, which consisted of fluorescence microscopy images of cells. They compared the results of U-Net with other state-of-the-art methods and found that U-Net achieved the best performance with an accuracy of 0.9203 IOU in PhC-U373 and 0.7756 in DIC-HeLa.

1.2 ResUNet++:
The ResUNet++ paper proposes a novel deep-learning architecture for automatic polyp segmentation and detection in colonoscopy images. Polyps are precursors to colorectal cancer, and early detection is crucial for effective treatment. The ResUNet++

architecture is designed to help clinicians identify polyps that may be missed during endoscopic examinations. The ResUNet++ architecture is based on the Deep Residual U-Net (ResUNet) architecture, which combines the strengths of deep residual learning and U-Net. Residual blocks are used to allow the network to learn residual functions, which can make it easier for the network to learn complex mappings. U-Net is a popular architecture for medical image segmentation that uses a contracting path and an expansive path to capture context and localization information, respectively. The ResUNet++ architecture adds several improvements to the ResUNet architecture. Squeeze and excitation blocks are added to allow the network to learn channel-wise dependencies and focus on important features. Atrous Spatial Pyramidal Pooling (ASPP) is used to capture multi-scale information, which can help the network identify polyps of varying sizes. Attention blocks are used to allow the network to focus on relevant regions of the image. The authors evaluate the ResUNet++ architecture on two publicly available datasets and compare its performance to the popular U-Net and ResUNet architectures. They find that the ResUNet++ architecture improves segmentation results significantly for colorectal polyps compared to other state-of-the-art methods. The proposed architecture also works well with a smaller number of images. The performance of the ResUNet++ architecture has been evaluated on two datasets. It was evaluated on the Kvasir-SEG dataset where it gives a dice score of 0.8133 and for the CVC-612 dataset it gives a dice score of 0.7955. The ResUNet++ paper proposes a novel deep-learning architecture for automatic polyp segmentation and detection in colonoscopy images. The architecture incorporates several improvements over previous architectures, including residual blocks, squeeze, and excitation blocks, Atrous Spatial Pyramidal Pooling (ASPP), and attention blocks.

1.3 nnU-Net:

nnU-Net, short for "No New-Net", is a deep learning-based framework designed for medical image segmentation. The U-Net architecture has been widely used for medical image segmentation tasks due to its ability to handle small datasets and achieve high accuracy. However, U-Net has some limitations, such as its fixed architecture, lack of adaptability to different datasets, and sensitivity to initialization. To address these limitations, nnU-Net introduces a self-adapting U-Net-based framework that can adjust itself to any given dataset. The framework is based on a modular design, with a core U-Net architecture and several modules that adapt the model to different datasets. To adapt the model to different datasets, nnU-Net introduces several modules, each of which is responsible for a specific adaptation task. These modules can be combined and configured in different ways to form a customized framework for a particular dataset.

Some of the modules introduced in nnU-Net include:
1.3.1 Data pre-processing module: This module applies various pre-processing techniques to the input data, such as intensity normalization and histogram equalization, to improve the performance of the model.
1.3.2 Data augmentation module: This module generates synthetic data by applying various transformations to the input data, such as rotation and scaling, to increase the size of the training set and improve the robustness of the model.

1.3.3 Training module: This module provides various options for training the model, such as different loss functions and optimization algorithms, to fine-tune the model to the given dataset.

1.3.4 Inference module: This module applies post-processing techniques, such as connected component analysis and conditional random fields, to the output of the model to improve the segmentation results.

The authors demonstrated the effectiveness of the nnU-Net framework by participating in the medical segmentation Decathlon challenge. They used a five-fold cross-validation approach with three different automatically configured U-Net models for each dataset and selected the model with the highest mean foreground dice score for final submission. The nnU-Net performed competitively on the held-out test sets of seven highly distinct medical datasets and achieved the highest mean dice scores for all classes of all tasks, except for class 1 in the Brain Tumour dataset. The dice scores of the test sets of lung sections are 0.692. nnU-Net also introduces a novel self-adaptation mechanism that can adjust the framework to a given dataset automatically. The self-adaptation mechanism consists of a self-adaptive layer that learns to adjust the number of filters and the size of the convolutional kernels based on the input data. This mechanism allows nnU-Net to handle datasets of different sizes and resolutions and adapt the model architecture to the specific characteristics of each dataset.

2. Transformer Based Models.

2.1 UNETR:

UNETR is a recent paper that proposed a new architecture for 3D medical image segmentation using transformers. Transformers are a popular architecture for natural language processing tasks, but they have shown promising results in other domains such as computer vision. The authors of UNETR aimed to leverage the power of transformers for 3D medical image segmentation tasks, which is a challenging problem due to the high variability and complexity of medical images. To address this limitation, researchers have combined self-attention modules with convolutional layers to improve non-local modeling capabilities. Additionally, transformer-based models have achieved state-of-the-art benchmarks in various tasks in natural language processing (NLP) and computer vision due to their great capability of modeling long-range dependencies and capturing global context. In particular, the Vision Transformer (ViT) and its variants have shown excellent capabilities in learning pre-text tasks that can be transferred to downstream applications. Building on this, the authors propose a novel architecture called UNET Transformers (UNETR) for volumetric medical image segmentation. The architecture reformulates the task of 3D segmentation as a 1D sequence-to-sequence prediction problem and uses a transformer as the encoder to learn contextual information from the embedded input patches. After the embedding layer, they a stack of transformer blocks comprising of multi-head self-attention (MSA) and multilayer perceptron (MLP). The extracted representations from the transformer encoder are merged with the CNN-based decoder via skip connections at multiple resolutions to predict the segmentation outputs. Instead of using transformers in the decoder, the proposed framework uses a CNN-based decoder since transformers are unable to properly capture localized information, despite their great capability of learning global information. The architecture utilizes a transformer encoder to capture global context, which has been shown to be beneficial in various computer vision tasks. A combination of soft dice loss and cross-entropy loss was used. The proposed model is validated on two public datasets: BTCV and MSD, where UNETR achieves new state-of-the-art performance on the BTCV leaderboard and outperforms competing techniques on the MSD dataset. The model average dice score for BTCV and MSD was 0.891 and 0.711, respectively.

2.2 Swin-UNETR:

Swin-UNETR is a novel deep-learning architecture for medical image segmentation. This architecture is built on top of two state-of-the-art models in the field of computer vision: the Swin Transformer and the UNETR. The Swin Transformer is a recent advancement in the field of computer vision that has achieved impressive results in image classification and object detection. It is a hierarchical transformer network that breaks down the image into smaller patches and processes them independently before assembling them back into the final output. This approach is much more efficient than traditional convolutional neural networks (CNNs) as it reduces the amount of computation needed for each layer. The architecture of the Swin Transformer consists of four key components: a patch embedding layer, a hierarchical transformer layer, a Swin Transformer block, and a classification head. The Swin Transformer block is the core building block of the Swin Transformer architecture. It consists of several key components, including a shifted window self-attention mechanism, a shifted window feed-forward network, and a residual connection. The shifted window self-attention mechanism allows the model to capture spatial dependencies within the image by attending to neighboring patches in a shifted window pattern, rather than a traditional sliding window pattern. The researchers evaluated Swin-UnetR on several medical image datasets, including liver and prostate segmentation. The results showed that Swin-UnetR outperformed several state-of-the-art segmentation models in terms of accuracy and efficiency. The researchers also showed that Swin-UnetR was able to generalize well to unseen data, indicating its potential for use in real-world applications. The model was trained and evaluated on the BraTS dataset where it gives an average dice score of 0.913.

III. RESULTS AND DATASETS

SN.	Model	Dataset	Dice Score
1	U-Net	PhC-U373	0.9203
2	U-Net	DIC-HeLa	0.7756
3	ResUNet++	Kvasir-SEG	0.8133
4	ResUNet++	CVC-ClinicDB	0.7955
5	nnU-Net	Decathlon Lung Dataset	0.6920
6	nnU-Net	Decathlon Brain Dataset	0.612
7	UNETR	BTCV	0.891
8	UNETR	MSD	0.711
9	Swin-UNETR	BraTS	0.913

Table 1. Model and Dice Score for dataset

Summary of datasets used by researchers for Medical Segmentation:

1. PhC-U373:
The PhC-U373 dataset is a collection of phase contrast images of human glioblastoma cells (U373MG cell line) acquired with a 20x objective. It was created by the Biomedical Imaging Group of the École Polytechnique Fédérale de Lausanne (EPFL) and is commonly used as a benchmark dataset for evaluating image segmentation algorithms. The dataset contains 327 fluorescence images of cell nuclei with accompanying ground truth annotations of the nuclei boundaries. The images were captured with a Zeiss Axiovert microscope and are available in 8-bit TIFF format with a resolution of 696x520 pixels. The PhC-U373 dataset is freely available for research purposes and can be downloaded from the EPFL website. It has been widely used in the development and evaluation of various computer vision and machine learning techniques for image segmentation and cell tracking.

2. DIC-HeLa:
The DIC-HeLa dataset is a collection of differential interference contrast (DIC) microscopy images of HeLa cells. It was created by the Cell Biology and Biophysics Unit of the European Molecular Biology Laboratory (EMBL) and is often used as a benchmark dataset for evaluating image segmentation and tracking algorithms. The dataset contains 469 DIC images of HeLa cells with accompanying ground truth annotations of the cell boundaries and nuclei positions. The images were acquired with a 100x objective and are available in 16-bit TIFF format with a resolution of 512x512 pixels. The DIC-HeLa dataset is freely available for research purposes and can be downloaded from the EMBL website. It has been widely used in the development and evaluation of various computer vision and machine learning techniques for cell segmentation, tracking, and analysis.

3. Kvasir-SEG:
The Kvasir-SEG dataset is a collection of gastrointestinal endoscopic images and accompanying annotations, designed for research and development of computer vision and machine learning algorithms in the medical domain. The dataset was created by the Norwegian University of Science and Technology (NTNU) and SINTEF Digital, and it is freely available for research purposes. The dataset contains a total of 1,000 annotated images from the gastrointestinal tract, captured using an Olympus GIF-H180J endoscope with a 4K UHD camera. The images are classified into eight different classes: normal-z-line, normal-pylorus, normal-cecum, esophagitis-a, esophagitis-b-d, polyps, ulcerative-colitis, and dyed-resection-margins. Each image is manually annotated by at least two medical experts with pixel-level segmentation masks, indicating the location of the various classes of gastrointestinal tract diseases and conditions in the image. The Kvasir-SEG dataset has been widely used in the development and evaluation of various computer vision and machine learning techniques for image segmentation, classification, and object detection in the medical domain.

4. CVC-ClinicDB:
The CVC-ClinicDB dataset is a collection of endoscopic images and videos of the gastrointestinal tract. It was created by the Computer Vision Center (CVC) of the Autonomous University of Barcelona and the Hospital Clinic of Barcelona, Spain, and it is often used as a benchmark dataset for evaluating computer vision and machine learning algorithms in the medical domain. The dataset contains 612 images and 612 videos from 13 different patients, with a variety of gastrointestinal diseases such as ulcerative

colitis, Crohn's disease, and colon cancer. The images and videos were acquired using a Pentax endoscope with a resolution of 768x576 pixels and 25 frames per second. Each image and video are manually annotated with pixel-level segmentation masks, indicating the location of various classes of gastrointestinal diseases and conditions in the image or video. The annotations include labels for normal tissue, polyps, ulcers, and other lesions. The CVC-ClinicDB dataset is freely available for research purposes and has been widely used in the development and evaluation of various computer vision and machine learning techniques for image and video analysis in the medical domain.

5. Decathlon Dataset
The Decathlon dataset is a collection of medical imaging datasets designed for benchmarking computer vision and machine learning algorithms in the medical domain. It was created by the Computational Radiology Laboratory of the Children's Hospital of Philadelphia and the Department of Radiology of the Perelman School of Medicine at the University of Pennsylvania. The Decathlon dataset consists of 10 different medical imaging datasets, each with its own unique characteristics and challenges, including:

5.1 Brain Tumour Segmentation (BraTS)
5.2 Cardiac MRI Segmentation (ACDC)
5.3 Lung CT Segmentation (LUNA16)
5.4 Prostate MRI Segmentation (PROMISE12)
5.5 Liver CT Segmentation (LiTS)
5.6 Hippocampus Segmentation (ADNI)
5.7 Colonography (CTC)
5.8 Retinal Vessel Segmentation (DRIVE)
5.9 Chest X-Ray Abnormalities Detection (ChestXray14)
5.10 Skin Lesion Segmentation (ISIC)
Each dataset is accompanied by detailed annotations, including ground-truth segmentation masks and clinical information about the patients. The Decathlon dataset is freely available for research purposes and has been widely used in the development and evaluation of various computer vision and machine learning techniques in the medical domain, such as image segmentation, classification, and registration.

6. BraTS:
The BraTS dataset is a widely used benchmark dataset for brain tumour segmentation and diagnosis in medical imaging. It stands for "Multimodal Brain Tumour Segmentation Challenge" and is organized by the Medical Image Computing and Computer Assisted Intervention Society (MICCAI) and the BraTS organizing committee. The BraTS dataset includes Magnetic Resonance Imaging (MRI) scans of the brain, including T1-weighted, T1-weighted with contrast enhancement, T2-weighted, and Fluid Attenuated Inversion Recovery (FLAIR) sequences. The dataset also includes ground-truth segmentation labels for three types of brain tumours: glioma, meningioma, and pituitary adenoma. The BraTS dataset has been used to evaluate various algorithms for brain tumour segmentation, including deep learning-based approaches. It is widely used in research and development of brain tumour diagnosis and treatment planning. The dataset is freely available for research purposes and can be downloaded from the official BraTS website.

3. Medical Segmentation Decathlon (MSD):
The Medical Segmentation Decathlon (MSD) is a dataset collection designed for benchmarking algorithms in medical image analysis. The MSD dataset consists of ten different medical imaging tasks, including brain tumour segmentation, liver segmentation, and cardiac segmentation, each with a corresponding dataset. The goal of the MSD challenge is to compare the performance of different segmentation algorithms on these tasks. The MSD dataset includes various imaging modalities, such as MRI, CT, and ultrasound, and is annotated with ground truth segmentations by medical experts. The datasets are publicly available and can be used for research purposes.

The tasks included in the MSD dataset are:
6.1 Brain tumour segmentation
6.2 Hippocampus segmentation
6.3 Liver segmentation
6.4 Lung segmentation
6.5 Pancreas segmentation
6.6 Prostate segmentation
6.7 Cardiac segmentation
6.8 Spleen segmentation
6.9 Colon segmentation
6.10 Head and neck CT segmentation
The MSD dataset has been widely used in medical image analysis research, and several algorithms have

been developed and evaluated on this dataset. The MSD challenge is organized by the Computational Radiology Laboratory at Boston Children's Hospital and the Brigham and Women's Hospital.

7. Beyond The Cranial Vault (BTCV)

The Beyond the Cranial Vault (BTCV) Segmentation Challenge dataset is a collection of medical images of various anatomical structures, including the head and neck, thorax, and abdomen. The dataset was created for the BTCV Segmentation Challenge, which was organized in conjunction with the MICCAI 2019 conference.

The BTCV Segmentation Challenge dataset consists of three sub-datasets:

7.1 The head and neck dataset, which includes computed tomography (CT) scans of the head and neck region, and is designed for segmentation of the larynx, pharynx, and various other structures.

7.2 The thorax dataset, which includes CT scans of the thorax, and is designed for segmentation of the lungs, heart, and other structures.

7.3 The abdomen dataset, which includes CT scans of the abdomen, and is designed for segmentation of the liver, spleen, and various other structures. Each sub-dataset includes training and validation sets with corresponding ground truth annotations, as well as a test set for evaluation. The annotations were created by medical experts, and the dataset is publicly available for research purposes. The BTCV Segmentation Challenge dataset has been used to evaluate and compare the performance of various segmentation algorithms and has led to the development of new methods for medical image analysis.

CONCLUSION

In this article, we have explored existing medical image segmentation model techniques based on various descriptor methods combined with convolution neural networks such as U-Net, ResU-Net++, and transformer-based such as UNETR, Swin-UNETR. We also listed common datasets used for segmentation. A limitation of existing medical image segmentation is ambiguity and complexity of the image, inter-patient variabilities such as age, gender, etc., lack of interpretability, and class imbalance. Our study provides insight into existing medical image segmentation models for researchers wishing to conduct research in this area. A challenge for future research is to develop a robust medical image segmentation model that can overcome the existing limitations.

REFERENCES

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. ArXiv. https://doi.org/10.48550/arXiv.1505.04597.

[2] Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., de Lange, T., Halvorsen, P., & Johansen, H. D. (2019). ResUNet++: An Advanced Architecture for Medical Image Segmentation. ArXiv. https://doi.org/10.48550/arXiv.1911.07067.

[3] Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., Wasserthal, J., Koehler, G., Norajitra, T., Wirkert, S., & H., K. (2018). NnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. ArXiv. https://doi.org/10.48550/arXiv.1809.10486.

[4] Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H., & Xu, D. (2021). UNETR: Transformers for 3D Medical Image Segmentation. ArXiv. https://doi.org/10.48550/arXiv.2103.10504.

[5] Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H., & Xu, D. (2022). Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. ArXiv. https://doi.org/10.48550/arXiv.2201.01266.

[6] Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M., Unal, G., Wells, W. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. MICCAI 2016. Lecture Notes in Computer Science (), vol 9901. Springer, Cham. https://doi.org/10.1007/978-3-319-46723-8_49.

[7] Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., & H., K. (2018). Brain Tumor Segmentation and Radiomics Survival

Prediction: Contribution to the BRATS 2017 Challenge. ArXiv. /abs/1802.10508.

[8] D. Jha, P. H. Smedsrud, M. Riegler, P. Halvorsen, T. de Lange, D. Johansen, and H. Johansen, "Kvasir-seg: A segmented polyp dataset," in International Conference on Multimedia Modeling. Springer, 2020. [Online]. Available: https://datasets.simula.no/kvasir-seg/.

[9] J. Bernal, F. J. Sanchez, G. Fern ´andez-Esparrach, D. Gil, C. Rodr iguez, and F. Vilarino, "Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians," Computerized Medical Imaging and Graphics, vol. 43, pp. 99–111, 2015.

[10] Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Landman, B. A., Litjens, G., Menze, B., Ronneberger, O., Summers, R. M., van Ginneken, B., Bilello, M., Bilic, P., Christ, P. F., Do, R. K., Gollub, M. J., Heckers, S. H., Huisman, H., Jarnagin, W. R., McHugo, M. K., . . . Cardoso, M. J. (2021). The Medical Segmentation Decathlon. ArXiv. https://doi.org/10.1038/s41467-022-30695-9.

[11] B Landman, Z Xu, J Igelsias, M Styner, T Langerak, and A Klein. Miccai multi-atlas labeling beyond the cranial vault–workshop and challenge. In Proc. MICCAI Multi-Atlas Labeling Beyond Cranial Vault—Workshop Challenge, 2015.

[12] Amber L Simpson, Michela Antonelli, Spyridon Bakas, Michel Bilello, Keyvan Farahani, Bram Van Ginneken, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv:1902.09063, 2019.

[13] Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E.R., Weber, M.A., Arbel, T., Avants, B.B., Ayache, N., Buendia, P., Collins, D.L., Cordier, N., Corso, J.J., Criminisi, A., Das, T., Delingette, H., Demiralp, C., Durst, C.R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K.M., Jena, R., John, N.M., Konukoglu, E., Lashkari, D., Mariz, J.A., Meier, R., Pereira, S., Precup, D., Price, S.J., Raviv, T.R., Reza, S.M.S., Ryan, M.T., Sarikaya, D., Schwartz, L.H., Shin, H.C., Shotton, J., Silva, C.A., Sousa, N., Subbanna, N.K., Szekely, G., Taylor, T.J., Thomas, O.M., Tustison, N.J., Unal, G.B., Vasseur, F., Wintermark, M., Ye, D.H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., Leemput, K.V.: The multimodal brain tumor image segmentation benchmark (brats). IEEE Trans. Med. Imaging 34(10), 1993–2024 (2015).

[14] Bakas, S., Reyes, M., et Int, Menze, B.: Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge. In: arXiv:1811.02629 (2018).

[15] Bakas, S., Akbari, H., Sotiras, A., Bilello, M., Rozycki, M., Kirby, J., John Freymann, K.F., Davatzikos, C.: Segmentation labels and radiomic features for the pre-operative scans of the tcga-lgg collection. The Cancer Imaging Archive (2017), https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF.