

A fine-tuned Deep Learning model for Medical Image Segmentation

Agughasi Victor Ikechukwu A.¹, Murali S.², Prithvi Raj G.D.³

Research Scholar, Department of CSE, Maharaja Institute of Technology Mysore¹

Professor, Department of CSE, Maharaja Institute of Technology Mysore²

UG Student, Department of CSE, Maharaja Institute of Technology Mysore³

Abstract : In medical imaging, segmentation plays a vital role towards the interpretation of Xrays, CT Scans and MRIs where salient features are detected and extracted with the help of image segmentation. Finding an optimal medical image reconstruction methodology is becoming increasingly difficult as technology advances. As a result, medical imaging has benefited from advancements in analysis and diagnosis. Without undergoing surgery, clinicians and radiologists employ various modalities ranging from X-Rays and CT-Scans to ultrasonography, and other imaging techniques to visualise and examine interior human body organ and structures. The focus of this study is on the segmentation approached applied to chest x-ray images, tumour obtained from CT and MRI images.

Keyword: Pattern Recognition, Image Segmentation on X-rays, Tumour Detection, MRI, Medical imaging.

INTRODUCTION

The proliferation of data and availability of high-end devices has made modern 3D medical imaging techniques generate images that are extremely dependable. 2D images includes but not limited to images captured with a 2D camera, graphical images and text-based images (including artificial image)[1]. X-rays, ultrasounds, MRIs, CT-scans, and satellite images captured from the most recent datasets are all examples of 2-Dimensional medical images. In positron emission tomography (PET) images, attenuation correction (AC) is a critical step in obtaining the correct pixel intensity value. Recently, a PET/MR system for cancer diagnosis without radiation has been developed[2]. Image segmentation techniques divide an entire image into useful sections that can be further analysed. The underlying data and solution space are usually used to drive the segmentation process, with the latter being helpful in cases where the images are distorted or contain artefacts due to image acquisition constraints.[3].

In many Clinical Applications, medical image segmentation has a significant impact on all subsequent research and diagnosis. Manual annotation, on the other hand, is reliant on domain expertise and expert skills. It is boring and time-consuming, and prone to intra and interobserver inconsistencies. As a result, creating automated segmentation algorithms for exact annotation of medical images is both clinically useful and necessary.[4] Abdominal imaging organ segmentation can aid clinical processes in several areas, including diagnostics, patient preparation, and therapy delivery. Organ segmentation is critical for computer-assisted diagnosis and biomarker measurement systems.[5].

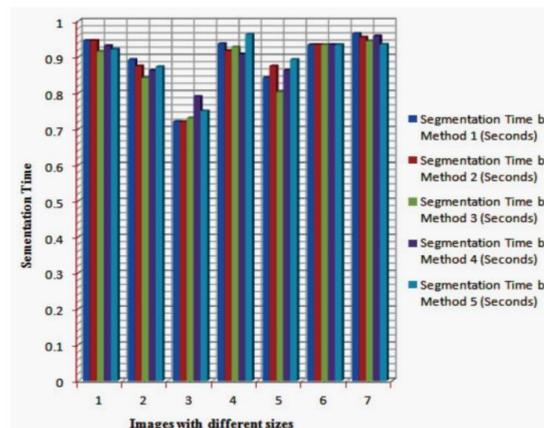


Fig 1. Segmentation Time vs Different image sizes for training.



METHODOLOGY

The DRINet (Dense-Res-Inception Net) has an analysis and synthesis path, much like the FCN (Fully Connected Networks). The research path, which is inspired by DenseNet, is made up of stacks of dense link blocks instead of regular convolution layers. The Res-Inception Net is used to simulate the synthesis course, which consists of residual inception and unpooling blocks. The DRINet does not have any skip links in order to save memory [6].

CNNs for Image Segmentation: FCN and DeepLab are two of the most advanced approaches for natural image segmentation. Effective networks such as U-Net, DCAN, and Nabla-net have been proposed for 2D biomedical image segmentation. patch-based 3D volumes VNet, HighRes3DNet, and 3D profoundly supervised network are three more efficient end-to-end 3D CNNs that have been proposed for segmentation of the brain tumour and pancreas [7].

CNNs in Chest X-Ray Segmentation:

We summarise the three schemes for fusing information from various image modalities as Fusing at feature level, fusing at the classifier level, and Fusing at the decision level, as any supervised learning-based approach is conceptually comprised of three steps: feature extraction, classification, and decision making [8]. Unsupervised segmentation is achieved using K-means clustering. A spatial Kmeans classifies an image by grouping adjacent pixels in the feature space into clusters by incorporating pixel intensity, average, and median pixel intensities of a local window into a feature space [9]. OCT was invented in the 1990s, but it wasn't until the introduction of spectral domain SD-OCT, which became commercially available in 2006 and allowed for quicker signal acquisition, that image quality and resolution improved enough for quantitative image analysis to be accurate. As a result, OCT-based retinal image analysis is a relatively, new field, with automatic retinal layer segmentation being one of the first applications. As a result, OCT retinal image analysis is a relatively new field, with automatic retinal layer segmentation being one of the first applications in stable and moderately diseased retinas, followed by fluid segmentation in retinas with macular edema [10].

Techniques Used	Sensitivity (%)	Accuracy (%)
Geometric Transform Invariant Segmentation and Analysis of Brain Tumor Images	89	90
Brain Tumour MR Image Segmentation and Classification Using Support Vector Machines and Artificial Neural Networks	98	97.37
Levenberg Marquardt Tumor Diagnosis in MRI Brain Images Using Artificial Neural Networks	90.98%	93.74%
Medical Image Segmentation and Classification Using Local Independent Projection	97.3%	98.5%
Brain Tumor Image Segmentation and Classification Using Wavelet and Zernike	79%	77%

Table 1: Performance metrics of various Techniques visa-viz Sensitivity and Accuracy

RELATED WORK

The approach is divided into two steps: intuitionistic image representation and IFCM (Iterative Fuzzy C-Means). Image representation based on intuition due to the imprecise pixel intensity values in images produces ambiguity. The authors

used a tweaked IFCM algorithm to segment medical images[11][12]. By looking for local minima, the proposed a new approach to select clusters from feature vectors[13].

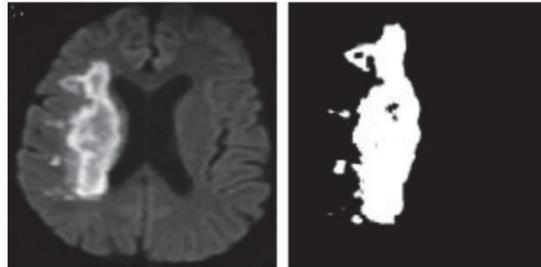


Figure 2: original MRI and segmented image

Another paper suggested segmentations “Without re-initialization”, in which variational LSMs was employed. To avoid the re-initialization method, Haiping Yu et al.,[13] suggested using a variational formulation as a penalty word which changed the level set function during curve evolution in order to overcome the flaws of the re-initialization-based LSM.

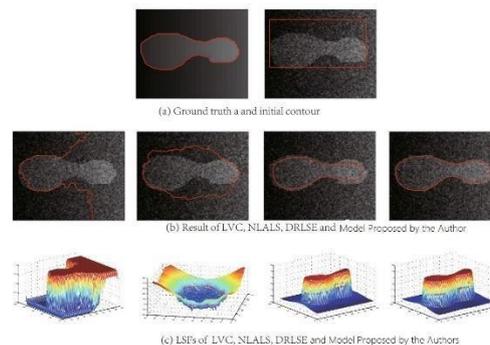


Figure 3: Comparison of Yu's proposed method [13] against LCV, DRLSE based on the ground truth.

Autoencoders (AE) are variants of neural network that tries to mirror a midway representation that can be used to extract the original input. It features a secret layer h on the inside, whose activations indicate the input image, sometimes known as codes. Thus, AE is often programmed to be incomplete to prevent it from copying its output directly, resulting in a code size that is smaller than the input dimension.

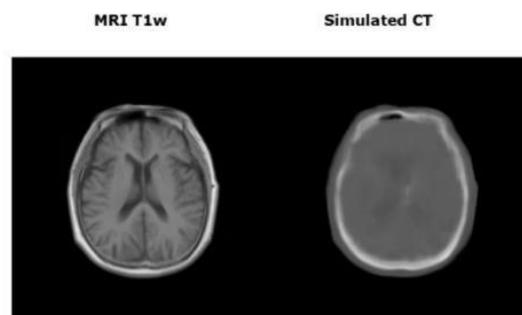
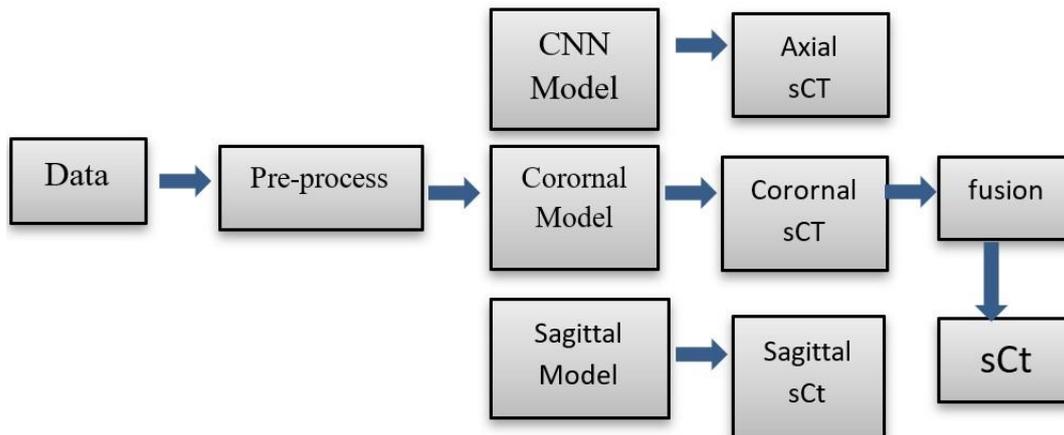


Figure 4: MRI vs Synthetic CT Scan result: The left sub-image is the actual MR image, and the right sub-image is the synthetic CT created from the left MR image using the traditional U-Net method.

Super-Resolution (SR) in Medical

Images: The aim of super-resolution (SR) image generation with CNNs is to retrieve spatial frequency information that is beyond the low resolution image's spatial bandwidth (LR), and observation $x \in RN$ to predict a high resolution (HR) image such that $y \in RM$ ($N \ll M$) and the network model only works with segmentation masks, which restricts their use to SR problems with an intensity image as the model output. To solve this problem and extend the conventional denoising to the T-L regularisation model, Oktay *et al* merged the AE with a prediction network. [3]. Convolutional Neural Network (CNNs) have been found worthy in winning competitions because convolutional layers are used to generate feature maps by convoluting an image with kernels or filters. Often, datasets with less images tend to suffer from overfitting and to deal with this, volumetric constraints was applied by deleting clusters in the segmentation obtained by the CNN that are smaller than a predefined threshold in the segmentation obtained by the CNN[14]. In general, contour detectors do not



promise that they can produce closed contours, and therefore do not guarantee that the image will be partitioned into regions. However, closed contours can still be recovered from regions in the shape of their borders. Early contour detection methods aimed to calculate the presence of a boundary at a specific image position using local measurements. Despite the fact that much of the comprehensive literature on contour detection predates its creation, it is still useful in present scenarios[15]. Fully convolutional networks (FCNs) are robust group of models that can be used to solve a wide range of pixel-level problems. By passing pretrained classifier weights, various neurons can be represented for learning salient features from a given dataset. FCNs for semantic segmentation significantly improve accuracy. Learning and inference are simplified and accelerated for complete pixel-to-pixel processing [16]. They primarily focused on the Histogram Thresholding based technique, which can be considered an easy and effective methodology for segmenting images with light artefacts on dark backgrounds for tumour identification and the Support vector machine (SVM) classifier for determining the type of tumour. Since tumour diagnosis is such a difficult and delicate process, precision and consistency are crucial. As a result, this paper urges further study and development of segmentation and classification methods to achieve more accurate outcomes. In the 3-D implementation of the HMRFM algorithm, a functional problem must be discussed. The MRF neighbourhood system could theoretically be three dimensionally isotropic. The slice thickness of a 3-D volume, on the other hand, is often greater than the interslice voxel dimensions. In this case, an isotropic neighbourhood structure may trigger issues. As a result, an anisotropic 3-D neighbourhood structure with a lower weight across slices is used [17].

PROBLEM STATEMENT

Enhancement of images is the proposed tool's first procedure which can be used to pre-process the test image. This stage improves the picture by grouping similar pixels depending on the threshold amount selected. Researchers have used a variety of image segmentation techniques in the literature to derive the area of interest from medical test photographs, including seed region increasing (RG), distance regularised level range (DRLS), principal component analysis (PCA), and watershed (WS)[18]. Due to the enormous number of weights in a CNN and the scarcity of labelled data, the iterative weight update, which starts with a random weight initialization, may result in an unacceptable local minimum for the cost function.[19]. From single-slice helical CT to 128-slice dual-source CT, there are many different CT technologies available today. CT image conversions such as 3D surface models, can be saved in a variety of file formats.[20].

During training, it would be great to process a conditional generative model from conditional distribution $P(b|a)$, a trained model generated from multiple samples b from each input a . The generator G will generate as many alternative values for each a as there are z values given an input picture a and a random vector z . Images can be captured using various modalities and configurations (e.g., echo time, repeat time), geometry (2D vs. 3D), and hardware (e.g., MRI, fMRI, CT, ultrasound) (e.g., field power, gradient performance) which can all affect the location of body organs and tumour types, especially when it comes to medical image segmentation.[21].

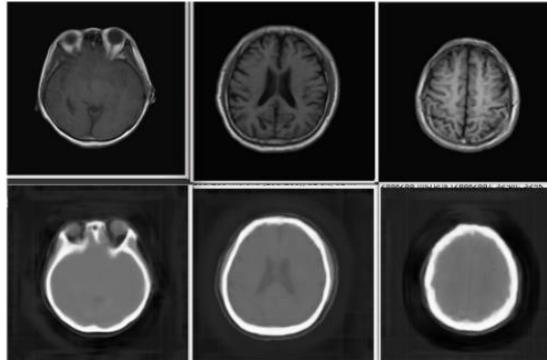


Figure 6: The outcomes of the approach for MR and synthetic CT proposed by G. Li et .al.. The first set of photos are original MR pictures in various slices. By subtracting synthetic CT images from above MR pictures, the second lowest level is reached.

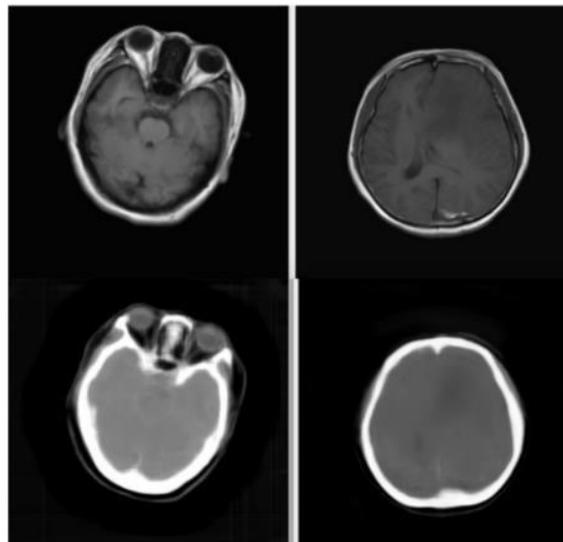


Figure 7: The findings of an MR scan and a synthetic CT scan in the case of a significant rotation or lesion. The original MR image is shown on the left top sub-image. The original MR image with lesion location is shown in the right top sub-picture. Two sub-pictures in the second row are corresponding synthetic CT scans created by the author. [2]

$$PSNR = 10X \log_{10} \left(\frac{(2^N - 1)^2}{MSE} \right) \quad (i)$$

Evaluation was done using Peak Signal to Noise Ratio(PSNR) as shown in table 3 below:

<i>Method</i>	<i>MAE</i>	<i>PSNR</i>
<i>U – Net</i>	~42	~14.6
<i>Proposed (Tuned U – Net)</i>	~31	~18

Table 3: U-Net vs Proposed method

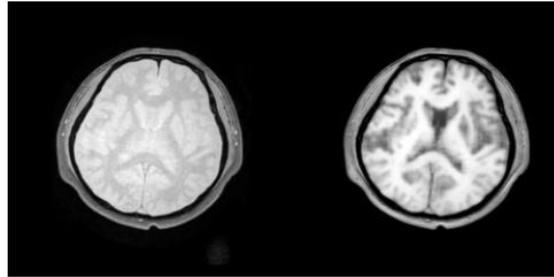


Figure 8. The outcomes of several MR sequences' synthetic pictures. The original PDW MR images on the left. The author's T1W image on the right is synthesized one.

The CNN model was trained on a dataset of MRI prostate scans with a fixed dimension of $128 * 128 * 64$ voxels and a spatial resolution of $1 * 1 * 1.5$ millimetres. Owing to the huge amount of time required by manually annotation by experts in the accurate ground-truth generalization, annotated medical volumes are difficult to come by. To address this and improve on the test dataset accuracy, we felt it important to complement the initial training dataset[22].

DESIGN

Thresholding is a technique for segmenting scalar images that creates a binary partitioning of image intensities that is linked depending on certain predefined parameters known as region growing. The algorithm can be implemented as follows: **Step 1:** Calculate the image's grey levels and sort them in ascending order.

Step 2: Determine the frequency of the image's grey levels.

Step 3: Assign the first pixel to the position of the first seed point. Step

Step 4: Combine the pixels to get a better seed point.

The mathematical morphology method is used for renal segmentation and feature extraction in the renal images, and it is also used to see how the diseased kidney affects the neighbouring organs in which active contours was used. Contours segmentation by grouping the same pixels are used to find the snake type boundary's roots.

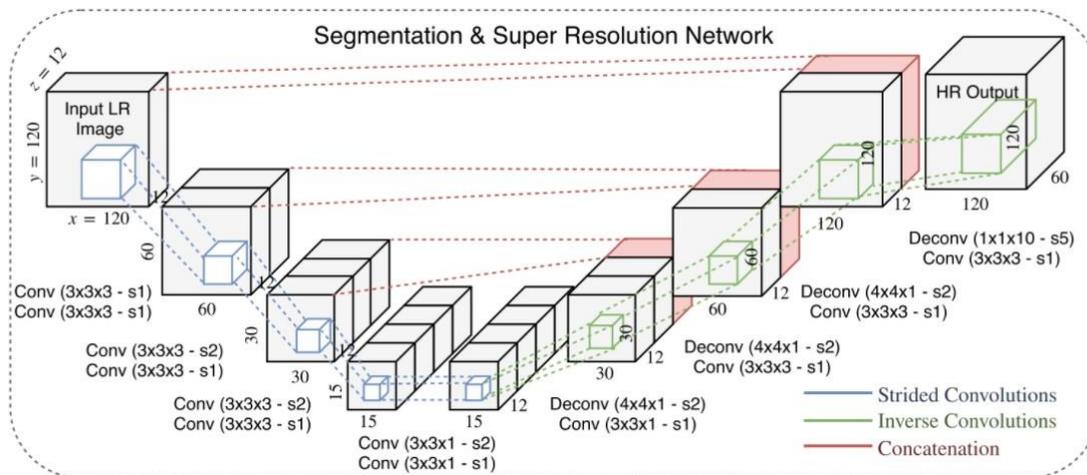


Figure 7: Baseline segmentation model for a super-resolution models with the T-L regularization and SR features in a spatial low-resolution

A similar methodology was implemented by [23] by employing diagnostic and prognostic techniques based on the morphology. Lymphocyte and epithelial nuclei are two groups of nuclei that are often studied. The appearance of nuclei varies depending on a variety of factors, including nuclei type, disease malignancy, and nuclei life cycle. Lymphocytes are white blood cells that play a critical function in the immune response. Lymphocyte nuclei (LN) are inflammatory nuclei that are smaller and more regular in form than epithelial nuclei[23].



RESULTS/DISCUSSION

There have been several segmentation approaches developed till date. In this article, a methodology for region-based segmentation is discussed. Watershed algorithm first implemented in MATLAB 2010a, was used to compare the results. A greater risk of redundancy between the auxiliary dataset and testing dataset was observed. We performed a detailed analysis that validates the degree to which PASCAL test images were found within the ILSVRC 2012 training and validation set, despite the fact that the tasks of object recognition and whole-image classification are significantly different, making such cross-dataset duplication much less concerning.

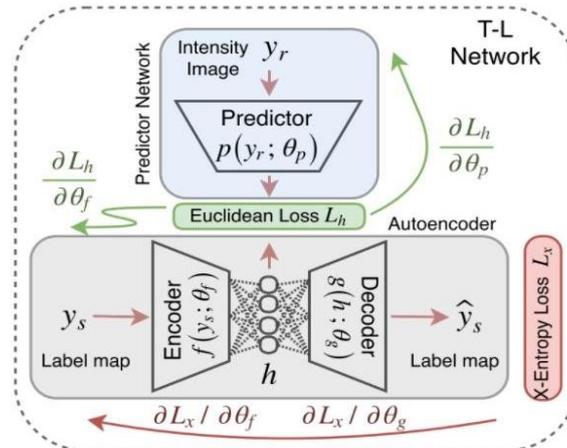


Figure 8: A stacked convolutional autoencoders trained on segmented labels

Both the hyperparameters and architecture decisions were tested following PASCAL VOC best practises. The fine-tuned CNN model was employed on the final results of the VOC 2010-12 datasets[24]. Different forms of area dissimilarities, comparison transitions, and conditions may all be integrated using the proposed structure wherein an ablation tests on MSRA-B, extracting one component at a time was used to assess the efficacy of these components. A one factor at a time test, was investigated leaving the other variables constant, since the system has three main factors: area dissimilarity, contrast transformation, and encoding[25].

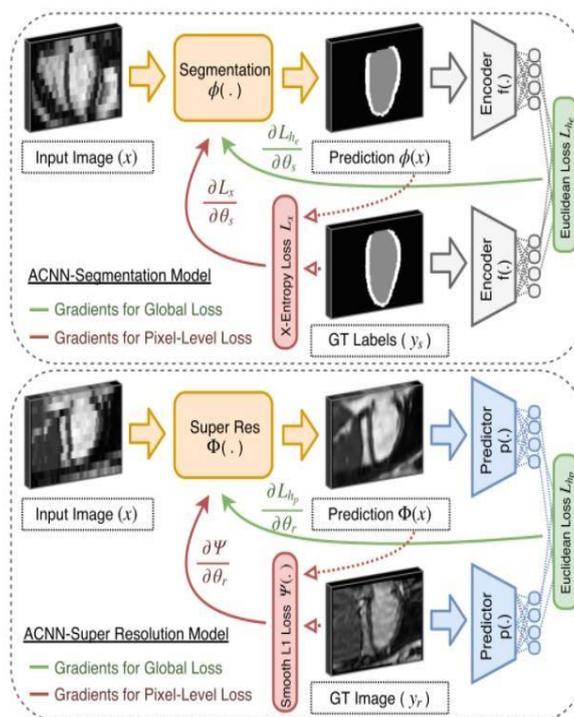


Figure 9: A super-resolution Anatomically Constrained Convolutional Neural Network (ACNN) for image segmentation tasks.



COMPARISON

VGG-16 vs. ResNet-101: As seen in the diagram above, using DeepLab based on ResNet-101 produces better segmentation results along entity boundaries than VGG16. We believe ResNet-101's identity mapping has a similar impact to hypercolumn features, which use intermediate layer features to help localise boundaries. Inside the “trimap”, we quantify this effect further (a narrow band along object boundaries). As seen in the diagram, using ResNet-101 before a CRF achieves nearly the same object boundary precision as using VGG-16 in combination with a CRF. The segmentation result is improved even further when the ResNet-101 result is postprocessed with a CRF[26]. The early convergence of the RGB and depth channels at the input is known as RGB D. HHA is the depth embedding of horizontal and the angle of the local surface normal with the inferred gravity direction as horizontal disparity height above ground. The RGB-HHA model blends RGB and HHA predictions in a late fusion model. [27]. Learning all at once instead of in steps produces almost identical outcomes, while preparation is quicker and less repetitive. However, because of the different aspect sizes, naive training proved vulnerable to deviance from the normal. An efficient approach using a stream equivalent to staged-learning rate was employed. These values were chosen randomly to equalise the mean function across various streams[27]. The accurate identification of structures or lesions is a crucial step in the diagnostic process, and it usually necessitates the localization of minor lesions in medical photographs. A CNN-based method for detecting nodules in x-ray images was introduced in a study. To detect micro-bleeds, a 3D CNN was used to process brain MRI results. Multi-stream CNNs have been used to explore PET and CT data together in a few studies[28].

CONCLUSION

Various approaches to methodologies applied in digital image segmentation was briefly explained in this analysis of medical image imaging research. The work also examines the findings on various image segmentation research methodologies crucial for pattern identification and recognition using edges, images, and points. This study analysed and compared a few image segmentation strategies and their efficiencies. The universal segmentation algorithm has been the primary target of medical image processing, giving image segmentation a bright future. Despite numerous breakthroughs and new discoveries, a widely agreed image segmentation approach that produces more reliable results is yet to be established, as image segmentation is influenced by a variety of factors, including the segmented images' objectives [29]. The shortage of marked databases, which hampered training and task evaluation, is a recurring trend in machine learning. More info, on the other hand, is acknowledged to increase efficiency, as shown on Google dataset of 300 million images using state-of-the-arts variations of CNN with hyperparameter optimizations [30][31].

FUTURE DIRECTIONS

An attempt was made to highlight some research work in medical image segmentation with emphasis on CT scan of the human brain. Future study will focus on a more in-depth examination of other forms of segmentations, such as the Threshold Method, Edge-Based Segmentation, Method based on a watershed and lots more while extending it to various modalities such as MRIs, X-Rays and PETs.

REFERENCES

- [1] K. K. Gupta, N. Dhanda, and U. Kumar, “A comparative study of medical image segmentation techniques for brain tumor detection,” *2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018*, pp. 1–4, 2018.
- [2] G. Li, L. Bai, C. Zhu, E. Wu, and R. Ma, “A Novel Method of Synthetic CT Generation from MR Images Based on Convolutional Neural Networks,” *Proc. - 2018 11th Int. Congr. Image Signal Process. Biomed. Eng. Informatics, CISP-BMEI 2018*, pp. 1–5, 2019.
- [3] O. Oktay *et al.*, “Anatomically constrained neural networks (ACNN): Application to cardiac image enhancement and segmentation,” *arXiv*, vol. 37, no. 2, pp. 384–395, 2017.
- [4] X. Chen *et al.*, “Anatomy-Regularized Representation Learning for CrossModality Medical Image Segmentation,” *IEEE Trans. Med. Imaging*, vol. 40, no. 1, pp. 274–285, 2021.
- [5] E. Gibson *et al.*, “Automatic MultiOrgan Segmentation on Abdominal CT with Dense V-Networks,” *IEEE Trans. Med. Imaging*, vol. 37, no. 8, pp. 1822–1834, 2018.
- [6] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, and D. Rueckert, “DRINet for Medical Image Segmentation,” *IEEE Trans. Med. Imaging*, vol. 37, no. 11, pp. 2453–2462, 2018.
- [7] G. Wang *et al.*, “Interactive Medical Image Segmentation Using Deep Learning with Image-Specific Fine Tuning,” *IEEE Trans. Med. Imaging*, vol. 37, no. 7, pp. 1562–1573, 2018.
- [8] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, “Medical image segmentation based on multi-modal convolutional neural network: Study on image fusion schemes,” *arXiv*, no. Isbi, pp. 903–907, 2017.
- [9] L. Zhang, V. Gopalakrishnan, L. Lu, R. M. Summers, J. Moss, and J. Yao, “Self-learning to detect and segment cysts in lung CT images without manual annotation,” *arXiv*, no. Isbi, pp. 1100–1103, 2018.
- [10] H. Bogunovic *et al.*, “RETOUCH: The Retinal OCT Fluid Detection and Segmentation Benchmark and Challenge,” *IEEE Trans. Med. Imaging*, vol. 38, no. 8, pp. 1858–1874, 2019.



- [11] A. Van Opbroek, H. C. Achterberg, M. W. Vernooij, and M. De Bruijne, "Transfer learning for image segmentation by combining image weighting and kernel learning," *IEEE Trans. Med. Imaging*, vol. 38, no. 1, pp. 213–224, 2019.
- [12] S. V. Aruna Kumar and B. S. Harish, "A Modified Intuitionistic Fuzzy Clustering Algorithm for Medical Image Segmentation," *J. Intell. Syst.*, vol. 27, no. 4, pp. 593–607, 2018.
- [13] H. Yu, F. He, and Y. Pan, "A novel segmentation model for medical images with intensity inhomogeneity based on adaptive perturbation," *Multimed. Tools Appl.*, vol. 78, no. 9, pp. 11779–11798, 2019.
- [14] 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.
- [15] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, 2011.
- [16] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *Intas Polivet*, vol. 10, no. 2, pp. 227–228, 2009.
- [17] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectationmaximization algorithm," *IEEE Trans. Med. Imaging*, vol. 20, no. 1, pp. 45–57, 2001.
- [18] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth, "Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation," *J. Ambient Intell. Humaniz. Comput.*, vol. 0, no. 0, pp. 1–12, 2018.
- [19] N. Tajbakhsh *et al.*, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.
- [20] M. van Eijnatten, R. van Dijk, J. Dobbe, G. Streekstra, J. Koivisto, and J. Wolff, "CT image segmentation methods for bone used in medical additive manufacturing," *Med. Eng. Phys.*, vol. 51, pp. 6–16, 2018.
- [21] M. Rezaei, H. Yang, and C. Meinel, "Recurrent generative adversarial network for learning imbalanced medical image semantic segmentation," *Multimed. Tools Appl.*, vol. 79, no. 21–22, pp. 15329–15348, 2020.
- [22] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," *Proc. - 2016 4th Int. Conf. 3D Vision, 3DV 2016*, pp. 565–571, 2016.
- [23] and D. R. Humayun Irshad, Antoine Veillard, Ludovic Roux, "Methods for Nuclei Detection, Segmentation, and Classification in Digital Histopathology: A Review—Current Status and Future Potential Humayun," *IEEE Trans. Med. Imaging*, vol. 9, no. 1, pp. 1–11, 2019.
- [24] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Pattern Recognit.*, pp. 580–587, 2014.
- [25] L. Tang, F. Meng, Q. Wu, N. L. Sowah, K. Tan, and H. Li, "Salient Object Detection and Segmentation via Ultra-Contrast," *IEEE Access*, vol. 4, no. c, 2018.
- [26] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, 2018.
- [27] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, 2017.
- [28] Y. Song, M. N. A. Rana, J. Qu, and C. Liu, "A Survey of Deep Learning Based Methods in Medical Image Processing," *Curr. Signal Transduct. Ther.*, vol. 15, pp. 1–14, 2019.
- [29] L. Subki, N. A. Ali, and R. Alwee, "a Review on Medical Image Segmentation : Techniques and Its a Review on Medical Image Segmentation : Techniques," *Perintis eJournal*, vol. 7, no. March 2019, pp. 59–82, 2017.
- [30] S. Shekkizhar and M. Lababidi, "Tensorflow implementation of Fully Convolutional Networks for Semantic Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, 2017.
- [31] H. P. Chan, R. K. Samala, L. M. Hadjiiski, and C. Zhou, "Deep Learning in Medical Image Analysis," *Adv. Exp. Med. Biol.*, vol. 1213, pp. 3–21, 2020.