Skull Stripping Using Traditional and Soft-Computing Approaches for Magnetic Resonance Images: A Semi-Systematic Meta-Analysis

Humera Azam, Humera Tariq
Department of Computer Science, UBIT, University of Karachi, Karachi, Pakistan
humera.azam@uok.edu.pk

Abstract. MRI scanner captures the skull along with the brain and the skull needs to be removed for enhanced reliability and validity of medical diagnostic practices. Skull Stripping from Brain MR Images is significantly a core area in medical applications. It is a complicated task to segment an image for skull stripping manually. It is not only time consuming but expensive as well. An automated skull stripping method with good efficiency and effectiveness is required. Currently, a number of skull stripping methods are used in practice. In this review paper, many soft-computing segmentation techniques have been discussed. The purpose of this research study is to review the existing literature to compare the existing traditional and modern methods used for skull stripping from Brain MR images along with their merits and demerits. The semi-systematic review of existing literature has been carried out using the meta-synthesis approach. Broadly, analyses are bifurcated into traditional and modern, i.e. soft-computing methods proposed, experimented with, or applied in practice for effective skull stripping. Popular databases with desired data of Brain MR Images have also been identified, categorized and discussed. Moreover, CPU and GPU based computer systems and their specifications used by different researchers for skull stripping have also been discussed. In the end, the research gap has been identified along with the proposed lead for future research work.

Key words: skull stripping, brain MR Images, soft computing, meta-analysis.

1. Introduction

The rich advancement in computing world has made it easier for medical experts to diagnose a particular disease or abnormality in living bodies. There are numerous computer aided diagnostic techniques which are helping doctors, bio-scientists and other medical investigators to understand the novel issues and their proposed solution. Image processing is the backbone of any computer aided mechanism and there are numerous techniques being used in medical field to investigate the human body out of which common techniques are X-rays, Computed Tomography, Magneto Encephalography, Positron Emission Tomography, and the most common and popular technique is the Magnetic Resonance Imaging (MRI) [29, 54].

Primary competitive advantages of using MRI over other types include its quality of being non-invasive and the fact that it provides more detailed, deep and comprehensive

Machine GRAPHICS & VISION 29(1/4):33–53, 2020. DOI: 10.22630/MGV.2020.29.1.3 .

images of organs [3] than the majority of other methods. There are four common modalities of MR images, including Longitudinal Relaxation Time (T1), Transverse Relaxation Time (T2), Proton Density (PD) and Fluid Attenuated Inversion Recovery (FLAIR) [62].

Scanners of MRI scan the body and create numerous images from multiple rotated axes, due to which, different views are reported for diagnosis. The 3D nature of MRI helps taking the view of the body from left to right, top to down, and from front to back [3,4,56]. The common types of anatomical orientation are Coronal plane from front to back; Sagittal plane from left to right; and Transversal plane from top to down [56].

The brain is a very sensitive part of the human body as it is made of soft tissues which are a combination of cerebrospinal fluid and fats. Such a complex system is fully covered with the strongest bone of the body called the skull [44]. MRI scanners capture the skull which needs to be removed for clearer understanding of the actual brain tissues [50]. The process of removing the skull from the brain images is called skull stripping. The more precise and efficient skull stripping ensures better help for clinical diagnosis.

This research study consists in the review of existing methods available for skull stripping from brain MR images along with their merits and demerits. Moreover, identifying the research gap in order to understand the current status and get lead for future research work also belongs to the scope of the present study.

1.1. Significance of the study

This research study provides understanding of the existing research gap and provides an abstraction of the experimental framework for future experiments generally in the field of *digital image processing* and most specifically in the domain of *brain MR images* for removing skull and other non-brain cells, in order to enhance the readability and understanding of brain MR images by medical experts for diagnostic purposes.

1.2. Methodology

This research study is carried out using semi-systematic review of literature pertaining to skull stripping methods. Fully systematic review requires extensive resources as well as at least 18 months to complete. Both said constraints provided the rationale to opt for the semi-systematic approach instead of the fully systematic one. Reviewed research studies are available in the respective cited journals for analyses using the meta-synthesis approach. Thematic convergence of different skull stripping methods has been assessed as the outcome of meta-synthesis on the basis of shared properties or architectural similarity between them.

2. Thematic convergence of skull stripping methods

Thematic convergence of developed, reviewed and discussed methods of skull stripping and image processing by different authors in latest research studies has been discussed in temporal order, i.e. older to newer.

2.1. Traditional methods

The reign of traditional methods have been popular in the field of image processing until the invention of neural networks. The convergence of traditional methods has been synthesized in the following subsections.

2.1.1. Traditional methods recently studied

Histogram Analysis and Deformable Model methods comprising the Thresholding and Simplex Mesh respectively have offered significantly positive results on the scale of Jaccard Index = 0.904, Dice Similarity Coefficient (DSC) = 0.95, Specificity = 0.985 [17]. Researchers experimented with Multi Atlas method [11] and Atlas model [20] with significant results of DSC = 0.9802, Specificity = 0.9908, Sensitivity = 0.9802, Average Distance = 0.66 and Hausdorff Distance = 7.72. Binarization method [40] including the irrational filter has provided significant results on the scale of DSC = 0.942, Sensitivity = 0.912, Specificity = 0.971, Overlap Fraction = 0.958 and Extra Fraction = 0.092. The said method remained competitive to the Otsu's method [53]. Another traditional method named as S3 [48] based upon brain anatomy and image intensity has also provided significant results on the scale of Jaccard Similarity $>$ 0.99 and 0.95 for data-sets taken from BrainWeb [5, 6] and IBSR [59] databases respectively; moreover, three measures of DSC, Sensitivity and Specificity $>$ 0.99 for both data-sets. Mathematical Morphology [2] based upon erosion and dilation have also provided better results for skull stripping.

The summary of above discussed traditional methods recently experimented with is presented in Table 1.

2.1.2. Competitive methods in comparison with traditional methods

Common state of the art competitive methods in comparison with traditional methods include Brain Extraction Tool (BET) [11,17,48], Brain Surface Extractor (BSE) [17,48], Robust Brain Extraction (ROBEX) [11,48]. Afore-cited research studies have offered better results in terms of performance measures such as Precision, Accuracy, Effectiveness and Efficiency (PAEE) while comparing with aforementioned state of the art methods.

2.1.3. Data and systems used for traditional methods

Most common data-sets taken for experimenting with the most recently tested traditional methods include Internet Brain Segmentation Repository (IBSR) [2, 11, 17, 40, 48, 59],

Tab. 1. Summary of traditional methods

Author & year	**Methods studied**	**Backbone architecture**	**Measures calculated**	**Methods compared**	**Data type**
Galdames et al. (2012) [17]	Histogram Analyses and Deformable Model	Thresholding and Simplex Mesh	Jaccard Index .904; DSC .9500; Specificity .985; Sensitivity .9900	HWA; BET and BSE	T1 from BrainWeb and IBSR
Doshi et al. (2013) [11]	Multi Atlas Model	Single Atlas and Multi Atlas	DSC .9802; Specificity .9908; Sensitivity .9802; Average Distance .66; Hausdorff Distance 7.72	BET and ROBEX	T1 from ADNI; IBSR and OASIS
Huang and Parra (2015) [20]	Atlas Model	Unified Segmentation Algorithm	Tissue Correlation Map	Intra-method	T1 from BrainWeb and Marom Bikson
Moldovanu et al. (2015) [40]	Binarization Mehtod	Irrational Filter	DSC .942; Senstivity .912; Specificity .971; Overlap Fraction .958; Extra Fraction .092	Otsu [42]; Sauvola [51]; Niblack [41]; Bernsens [1] methods	T1; T2; GAD and PD from WBA; T2 from IBSR
Roy and Maji (2015) [48]	S3	Brain Anatomy and Image Intensity	Jaccard Similarity .99 for BrainWeb and .95 for IBSR; DSC .99; Senstivity .99; Specificity .99	BET; BSE and ROBEX	T1 from BrainWeb; T1 from IBSR
Bhadauria et al. (2020) [2]	Mathematical Morphology	Erosion and Dilation	N/A	Intra-method	WBA and IBSR

BrainWeb [5, 17, 20, 48], and Open Access Series of Imaging Studies (OASIS) [11, 27, 28]. Only T1 weighted brain MR images both simulated and real have been used for the purpose. CPU based systems with 8 GB RAM have been used by the number of researchers for experimenting with traditional methods.

2.2. Deep Learning Neural Network based methods

Deep Learning Neural Network (DLNN) based methods took over the reign of traditional methods because of their enhanced sophistication with their own strengths and weaknesses. The convergence of recently studied DLNN based methods has been synthesized in the following subsections.

2.2.1. Recently developed DLNN methods

Through numerous experiments, the robustness of DLNN based architectures including U-Net, Rectified Linear Unit (ReLU), ConvNet, ResNet, and ConsNet has been proved.

Intensive review has suggested that the most common architectures include U-Net [7,12, 14,21,22,23,30,36,37,55].

U-Net architectures of both 2D and 3D types have successfully produced significant results for different performance measures of PAEE in different research studies. In an experimental research study, DSC = 0.71 has been achieved while utilizing the following hyperparameters: Epochs = 4, Discount Rate = 0.5 and 0.2, and Learning Rate = 0.0004 [14]. In another study, researchers have achieved DSC = 0.965 with False Negative Rate (FNR) = 0.2 and False Positive Rate (FPR) = 0.8 by implementing three layers of Convolutional Neural Network (CNN) with one steroid in the first and two steroids in the second layer [55]. Simultaneous Truth and Performance Level Estimation (STAPLE) constituted over 2D FCN U-Net has achieved DSC = 0.9575, 0.8887 and 0.8932 for three different data-sets of T1 weighted MR images with Learning Rate = 0.0001; while the measures of Sensitivity, Specificity, Hausdorff and Mean Distance were also significant [36]. The version of 2D U-Net has been extended for establishing 3D U-Net through max-pooling and batch normalization, which has achieved DSC = 0.9903, Sensitivity = 0.9853 and Specificity = 0.9953 on the data-set of T1 weighted MR images [21]. Researchers have experimented with the method HD-BET which is primarily comprised of U-Net CNN with remarkable results for the measures of DSC = 0.976 and Hausdorff Distance = 3.3 using T1, T2 and FLAIR images from databases of European Organization for Research and Treatment of Cancer (EORTC), LONI Probabilistic Brain Atlas (LPBA) and Neurofeedback Skull-stripped (NFBS) [22]. Researchers experimented with 3D U-Net based method comprised of Transfer Learning (TL) and Multi Output Net which performed exceptionally with DSC = 0.785 and 0.843 on the data-set of Multi-Atlas Labeling Challenge (MALC) and Hammers Adult Atlases (HAA), respectively [7]. Researchers experimented with another 2D U-Net based method of STAPLE which offered high rates of DSC = 0.9718 and Symmetric Surface-to-Surface Mean Distance (SSSMD) = 0.037 on T1 weighted images taken from databases of Calgary-Campinas, LPBA and OASIS [37]. The score of other scales like Sensitivity = 0.9891, Specificity = 0.9946 and Hausdorff Distance = 9.713 have also been remarkable but could not outperform other state of the art methods in comparison. Different hyperparameters have been used for the experiment including Learning Rate = 0.001, Exponential Decay = 0.995 after each epoch, and Fixed Kernel Size = 3×3 [37]. Time Distributed U-Net based CNN method has been tested with Model Accuracy = 0.583 in intra-method comparison with T1 weighted images taken from the database of MICCAI Brain Tumor Segmentation (BraTS) [12]. Researchers experimented with the method of Cascade 3D U-Net based CNN while using hyperparameters of Learning Rate = $10-5$, Weight Decay = 0.0005, Momentum = 0.9 (in Adam optimizer), and Epochs = 300 [23]. The method offered considerably good results and achieved Root Mean Square (RMS) = 0.86 on 90 MR images of kidney. In another research study, an experiment with the method of U-Net based CNN named as ACEnet has been carried out with hyperparameters like Epochs

= 100, Dropout Rate = 0.1, Momentum = 0.9 and Weight Decay = 0.0001 [30]. The studied method offered remarkable results as DSC $\geq$ 0.8 and Average Time to Segment $\approx$ 10 s on T1 weighted MR images taken from databases of MALC, Alzheimer's Disease Neuro-imaging Initiative (ADNI), Mindboggle, and SchizBull (see [30] for references).

ReLU architectures have also successfully produced significant results for different measures of PAEE in different research studies. An experiment has been run with ReLU architecture and achieved significant results as DSC = 0.965, FNR = 0.2 and FPR = 0.8 using T1 weighted images taken from NFBS [55]. Apart from this, an experiment has been carried out with ReLU based CNN which provided remarkable results for the measure of Sensitivity $>$ 0.87, Specificity $>$ 0.94 and Accuracy $>$ 0.918 on T1 weighted images taken from OASIS [52]. Another ReLU based CNN named as DeepMedic performed outstanding using hyperparameters of Learning Rate = 0.0005 and Epochs = 35 on T1 weighted MR images taken from different data-sets of OASIS, LPBA, and St. Olavs Hospital [13]. ReLU has also been included in an experiment along with U-Net features and achieved significant results [21]. An experiment has been carried out on ReLU based CNN named as DeepICE using hyper-parameter of Epochs = 20 with significant results of DSC = 0.9889 on T1 weighted MR images taken from IXI, OASIS, and BSTP [38]. CNN based methods of Focal Loss and RetinaNet based upon multiple architectures like ReLU, ConvNet, and ResNet have been experimented with using hyperparameters of Learning Rate = 0.01×0.1 after 60 K and then after 80 K iterations, Momentum = 0.9 and Weight Decay = 0.0001 [31]. The method tested increased the mean Average Precision 3-4 points on each T1 weighted MR image taken from Common Objects in Context (COCO) [33].

The summary of the above listed DLNN methods is presented in Table 2.

2.2.2. The rise of masking technique in DLNN methods

Along with the success of U-Net and ReLU based DLNN, another great architecture ResNet jointly with Region CNN R-CNN and in the latest cases with Faster R-CNN methods [45, 46] has provided significant results in numerous experiments. The state of the art method of Mask R-CNN [32] has been tested which is primarily based upon the architecture of Faster R-CNN, Feature Pyramid Network (FPN), ResNet, and ResNeXt, and is using hyperparameters of Learning Rate = 0.02, Weight Decay = 0.0001, and Momentum = 0.9 on T1 weighted MR images taken from COCO [18]. Before this, the FPN has been studied which has later been induced to postulate and experiment the revolutionary method of Mask R-CNN [32]. The developed FPN is based upon Faster R-CNN and two versions of ResNet50 and ResNet101 with hyperparameters of Learning Rate = 0.02×0.1 after 60 K and 80 K iterations on T1 weighted MR images from COCO [33] and PASCAL [15]. In continuation of their own work, researchers experimented with RetinaNet which actually received the contribution from their own FPN [31]. Transfer

Learning in Mask R-CNN has successfully been induced with hyperparameters of Learning Rate = 0.02×0.1 after 60 K and then 80 K iterations on T1 weighted MR images from COCO and Visual Genome [19]. Non-local Neural Network functionally comprising Mask R-CNN and ResNet architectures has been tested with hyperparameters of Learning Rate = 0.01×0.1 after every 150 K iterations, Momentum = 0.9, and Weight Decay = 0.0001 [58]. Apart from the novelty of the method, the experiment is unique because the video data has been taken into experiment for segmenting moving objects.

2.2.3. Competitive methods in comparison with DLNN methods

DLNN methods have outperformed traditional methods [16] out of which prominent DLNN methods include Bayesian Evolutionary Analysis by Sampling Trees BEaST [24, 37, 38, 47, 49], ROBEX [21, 22, 24, 37, 47, 49, 55], BET [22, 24, 37, 49], Hybrid Watershed Algorithm (HWA) [24, 37], BSE [22, 24, 37, 49, 55], FMRIB Software Library (FSL) [55], Analysis of Functional NeuroImages (AFNI) [47, 55], Advanced Normalization Tools (ANTs) [22,55], CompNet [10], Spectre [47], Kleesiek's method [21], 3dSkullStripping [22, 24], SLAN [7], Marker based Watershed Scalper (MBWSS), STAPLE and Optimized Brain Extraction Tool (OptiBET) [37], FreeSurfer [57], NICE [38], G-RMI [31, 32], and AttractioNet [32].

2.2.4. Data and systems used for DLNN methods

Experimental studies conducted to test different DLNN methods of skull stripping has taken data from different databases out of which some are publicly available and for the rest of them the prior permission is needed to access the database and to use data. Leading databases provided different types of brain MR images like T1 weighted, T2 weighted, FLAIR etc. and such databases include OASIS [10, 13, 24, 36, 37, 38, 52], IBSR [24, 57], LPBA [13,22,24], MALC [7,30], ADNI [30], PASCAL [19,32], COCO [18,19,31,32], Hammers [7], NAMIC [49], MPRAGE [49], UKBB [7], BraTS [12, 14], Visual Genome [19], NFBS [21, 22, 55], and Calgary-Campinas, [22, 36].

In addition to databases, different GPU based computer systems have been utilized by researchers for image processing; out of which, NVIDIA Tesla M40 [18,19], NVIDIA GTX 1050 TI [12, 23] NVIDIA GTX 970 [55], and NVIDIA GTX Titan [13, 22, 30, 37, 38, 47] are common.

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	Methods studied	Backbone architecture	Hyper-parameters	Measures calculated	Methods compared	Data type	System used
Prasad et al. (2014) [43]	Deformable Model	Intensity Analyses and Motor Control	N/A	Jaccard Index .8478; DSC .9175; Hausdorff 36.35; FPE .0253; FNE .1328	HWA; BET and BSE	T1 from ADNI and manually segmented	N/A
Dai et al. (2015) [8]	Mask R-CNN	CFM based on R-CNN; Spatial Pyarmid Pooling and RPN	N/A	Mean IoU for non-CFM 44 and CFM 50 and 50.9 for design A and B respectively	Inter-dataset comparison	PASCAL Context	Nvidia GTX Titan GPU based on the Caffe Library
Ren et al. (2015) [45]	Region Proposal Networks (RPN)	Faster R-CNN	N/A	Mean Average Precision mAP = .788 and .759 for PASCAL VOC 2007 and 2012 respectively	VGG-16 architecture	T1PASCAL VOC 2007; 2012 and COCO	NVIDIA Tesla K40
Kleesiek et al. (2016) [24]	3d CNN	CNN	N/A	Public Data DSC .9530 and Specificity .9936; Brain Tumor Data DSC .9519 and Specificity .9924	BEaST; Robex; BET; 3dSkullStrip; HWA and BSE	T1; T2; FLAIR IBSR; LPBA; OASIS and Non-enhanced Images	N/A

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	Methods studied	Backbone architecture	Hyper-parameters	Measures calculated	Methods compared	Data type	System used
Erden et al. (2017) [14]	3D Fully CNN	3D U-Net and Average Pooling	Epochs = 4, Dropout Rate = 0.5, 0.2 Learning Rate = 0.0004 ”	DSC = .71	Intra Algorithm Comparison	285×T1 BraTS Multimodal Brain Tumor Segmentation	N/A
He et al. (2017) [18]	Mask R-CNN	Faster R-CNN; Feature Pyramid NetworkFPN; ResNet-50-C4; ResNet-101-RPN; ResNeXt-101-FPN	Learning Rate = 0.02; Weight Decay 0.0001; Momentum = 0.9	mAP; Average Time = 195 ms/image	Inter-dataset comparison	COCO	Nvidia Tesla M40 GPU x8
Lin et al. (2017a) [32]	Feature Pyramid Networks	Faster R-CNN; ResNet-50; ResNet-101; VGG-16	Learning Rate = 0.02 for first 60k batches; Learning Rate = 0.002 for next 20k batches; RoIs per image = 2000	AR increased by 8 points; AP on COCO and PASCAL increased by 8 and 3 points respectively; Average Spead 0.165 img/sec for ResNet 50; Average Spead 0.19 img/sec for ResNet 101	G-RMI; AttractioNet; Faster R-CNN; Multi-path	COCO-2016; PASCAL	NVIDIA M40 GPU

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	**Methods studied**	**Backbone architecture**	**Hyper-parameters**	**Measures calculated**	**Methods compared**	**Data type**	**System used**
Lin et al. (2017b) [31]	Focal Loss and RetinaNet	FPN; ResNet; ConvNet; ReLU	Learning Rate = 0.01; divided by 10 after 60k and then at 80k iterations; Momentum = 0.9; Weight Decay = 0.0001	AP increased by 3 to 4 points; Average Speed of RetinaNet-ResNet-101-600 is 50 ms/per image faster	Two Staged Methods; Faster R-CNN+++; Faster R-CNN – FPN; Faster R-CNN – GRMI; Faster R-CNN – TDM; One Stage Method; YOLOv2; SSD-513; DSSD-513	COCO	NVIDIA M40 GPU
Mehta et al. (2017) [39]	BrainSeg-Net	CNN	Epochs = 30; Momentum= 0.75; Learning Rate = 0.05	mDC = 0.844 for IBSR; 0.824 for LONI-LPBA; 0.840 for Hammers67; 0.808 for Hammers83	MALP; Patch based; Classification based	T1 from LONI-LPBA; Hammers-67; Hammers-83; IBSR; MICCAI	NVIDIA Tesla K40
Dey and Hong (2018) [10]	CompNet	CNN	Learning Rate = 0.001; Epochs = 10	DSC = .9827; Sensitivity = .9826; Specificity = .9980	Plain U-Net; Dense U-Net; Probability CompNet; Plain CompNet	T1 from OASIS	N/A

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	**Methods studied**	**Backbone architecture**	**Hyper-parameters**	**Measures calculated**	**Methods compared**	**Data type**	**System used**
Eide (2018) [13]	3D U-Net; Deep-Medic; 3D CNN by Kleesiek et al. [24]	ReLU	Learning Rate = 0.0005; iEpochs = 35	DSC; Senstivity; Specificity	Inter-method and Inter-data comparison	T1 from OASIS; LPBA; Data provided by St. Olavs Hospital and LiTS for liver images	Nvidia TitanX
Hu et al. (2018) [19]	Transfer Learning in Mask R-CNN	Faster R-CNN; ResNet-50-FPN; ResNet-101-FPN	Learning Rate = 0.02x 0.1 after 60k and 80k iterations; Weight Decay = 0.0001; Momentum = 0.9	mAP; IoU threshold = 0.5 - 0.95	Inter-dataset comparison	COCO; PASCAL VOC; Visual Genome	Nvidia Tesla M40 GPU x8
Lucena et al. (2018) [36]	Simultaneous Truth and Performance Level Estimation	2D FCN U-Net; Loss Function = Negative of DSC	Learning Rate = 0.00001	DSC = .9579; Sensitivity = .9771; Specificity = .9887; Hausdorff = 1.49; Mean = 0.075	Gold Standard only	LONI-LPBA; CC-359; and OASIS	N/A
Roy and Maji (2018) [49]	Anatomic Region of Surgical Interest ARoSI	Rough Fuzzy Connectedness	N/A	DSC = .9625 and .9553; Senstivity = .9555 and .9520; Specificity = .9965 and .9985 for BrainWeb and NAMIC respectively	AFNI; BSE; BET; ROBEX; BEaST; CNN	T1 from BrainWeb; NAMIC with Normal Control; and NAMIC with Chizophrenic	N/A

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	Methods studied	Backbone architecture	Hyper-parameters	Measures calculated	Methods compared	Data type	System used
Roy et al. (2018) [47]	CNN	AlexNet; Gaussian Filter to the binary segmentation	N/A	DSC = .9719; Jaccard Index = .9454; PPV = .9963; Volume Difference = .0477	BEaST; Spectre; Robex; manual method	T1 from MPRAGE	Nvidia Titan X GPU
Selvathi and Vanmathi (2018) [52]	CNN	ReLU; NonLocal Mean for noise reduction	N/A	Senstivity ¿ .87; Specificity ¿ .94; Accuracy ¿ .918	Intra-method and Inter-image comparison	OASIS	N/A
Valvano et al. (2018) [55]	CNN	U-Net – CNN; ReLU – CNN	2x CNN layers of 3x3; 1 stroid in first and 2 stroids in second layer	DSC = 0.965; FNR = 0.2; FPR = 0.8	Robex; FSL; BSE; AFNI; ANTS	T1 from NFBS	Intel Xeon E5-2620 v4 CPU; GPU NVIDIA GTX 970;10 GHz; with 16 cores and 32 threads
Wang et al. (2018) [58]	Non-local Neural Networks	Mask R-CNN; ResNet; Batch-Norm; ImageNet	Learning Rate = 0.01; Momentum = 0.9; Wieght Decay = 0.0001	Average Precision	Inter-dataset comparison	Kinetics dataset; 246k and 20k videos; Charades dataset; 8k and 8k	N/A

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	Methods studied	Backbone architecture	Hyper-parameters	Measures calculated	Methods compared	Data type	System used
Yilmaz et al. (2018) [61]	Multi-stable Cellular Neural Network – MCNN	Contrast Enhancement Using Linear Image Combinations Algorithm CEULICA	N/A	Jaccard = 0.838; DSC = 0.899; True Positive = 0.151; True Negative = 0.013	BET; BSE	T1 from BrainWeb; MIDAS-NAMIC; Healthy PeopleHospital	Intel Core TM i7-382060GHZ processor; 16 GB RAM and 64 bit OS
Dai et al. (2019) [7]	Transfer Learning and Multi Output Net MO-Net	3D U-Net; Two staged training	N/A	DSC on MALC = .785; DSC on HAA = 0.843	U-Net FS; U-Net FT; SLAN T8; SLAN T27	UKBB; HAA; MICCAI; MALC	NVIDIA GPU
Dalca et al. (2019) [9]	Atlas Based CNN	CNN	N/A	DSC = .835	Inter-dataset comparison	T1 from OASIS; ABIDE; ADHD; MCIC; PPMI; HABS; and Harvard GSP. PD and manuualy annotated images	NVIDIA Titan Xp GPU.

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	**Methods studied**	**Backbone architecture**	**Hyper-parameters**	**Measures calculated**	**Methods compared**	**Data type**	**System used**
Hwang et al. (2019) [21]	3D U-Net	Extension of 2D U-Net; ReLU; Max Pooling; Batch Normalization	N/A	DSC = .9903; Senstivity = .9853; Specificity = .9953	BSC; ROBEX; Kleesiek's DLNN	T1 from NFBS	2x NVIDIA 1080Ti GPU
Isensee et al. (2019) [22]	HD-BET	U-Net – CNN	N/A	DSC 0.976; Hausdorff Distance 3	Robex; BET; BSE; 3dSkull-Stripping; ANTS	EORTC; LPBA NFBS; Calgary-Compinas	NVIDIA TITAN Xp GPU
Lucena et al. (2019) [37]	Simultaneous Truth and Performance Level Estimation	2D U-Net; reffered as CONSNet	Learning Rate = 0.001; Exponential Decay = 0.995 after each epoch; Fixed Kernel Size = 3 x 3	DSC = .9718; Senstivity = .9891; Specificity = .9946; Hausdorff Distance = 713; Symmetric Surface to Surface Mean Distance SSSMD = .037	ANT's; BEaST; BET; BSE; HWA; MBWSS; OPTIBET; ROBEX; STAPLE-12	T1 from Calgary-Campinas; LPBA; OASIS	NVIDIA had 12 Gbyte; CPU Xeon E3-1220 v3; 4x 10 GHz Intel; GeForce Titan X

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	**Methods studied**	**Backbone architecture**	**Hyper-parameters**	**Measures calculated**	**Methods compared**	**Data type**	**System used**
Wang and Li (2019) [57]	Label Fusion Method	Pixel Grayscale Probability Information; Sparse Representation; Weighted Voting	N/A	DSC; Recall; Precision; Hausdorff Distance HD	FreeSurfer; Lin et al. method [34]; Kushibar et al. method [26]; Xu et al. method [60]; Liu et al. method [35]	T1 from IBSR	PC Intel Core 70 GHz CPU
Dutta et al. (2020) [12]	Time-Distributed U-Net based CNN TD-U Net-CNN	U-Net – CNN; and ReLU	N/A	Model Accuracy = 5 3%	Intra-method	T1w; T2w; FLAIR from MICCAI BraTS; MICCAI BraTS	NVidia GTX 1050Ti GPU; 4 GB of VRAM; 100 GB of DRAM; an Intel i5-8600 K CPU overclocked to 10 GHz
Kim et al. (2020) [23]	Cascade 3D U-Net	U-Net; Active Learning; CNN	Learning Rate = 10 - 5; Weight Decay = 0.0005; Momentum = 0.9; Epochs = 300	Root Mean Squre RMS = 0.86	CNN; Manually Segemeted	MR images of kidney	NVIDIA GTX 1080 TI

to be continued in the next page

Tab. 2: Summary of DLNN based skull stripping methods

Author & year	**Methods studied**	**Backbone architecture**	**Hyper-parameters**	**Measures calculated**	**Methods compared**	**Data type**	**System used**
Kuang et al. (2020) [25]	Skull R-CNN	Faster R-CNN; Skeleton-based region proposal method	RMSprop with = 0.9; Learning Rate = 0.0001	Average Precision AP = 0.62	Faster R-CNN + FPN; Skull R-CNN + FPN	N/A	NVIDIA GTX 1080 TI
Li et al. (2020) [30]	Anatomical Context-Encoding Network – ACENet	2D CNN; U-Net; Spatial Squeeze and Excitations	Weight Decay = 0.0001; Momentum = 0.9; Dropout Rate = 0.1; Epochs = 100	Average Time to Segment 10 seconds; i DSC 0.8	Intra-dataset Comparison	T1 from MALC; ADNI; Mindboggle-101; SchizBull	NVIDIA TITAN XP GPU; 12GB of RAM
Manjón et al. (2020) [38]	Deep Intracranial Cavity Extraction – DeepICE	ReLU	Epochs = 20	DSC = .9889	Non-Local Intracranial Cavity Extraction – NICE; BEaST; VBM8	T1 from OASIS; BSTP	Nvidia TitanX 12 GB RAM

3. Research Gap

In the light of intensive literature review, we have come to the conclusion that the most recent development has been made in the domain of DLNN and the scientific progress has led the experts of digital image processing to successfully experiment with the latest and robust CNN variant named as Mask R-CNN [18] for image segmentation. The comprehensive literature audit did not provide sufficient empirical evidence pertaining to the use of Mask R-CNN for skull stripping. The availability of deep learning weights for hundreds of objects and classes and non-availability of the same for the skull stripping in giant public digital libraries like COCO etc. are also empirical evidences addressing the dearth of research stated above in the realm of image segmentation. The research gap identified and discussed above needs prompt attention of researchers. Therefore, the scientific research study may be carried out to experiment skull stripping using Mask R-CNN along with its underlying structure and auxiliaries to ultimately bridge the existing research gap.

References

[1] J. Bernsen. Dynamic thresholding of gray level images. In *Proc. 8th Int. Conf. on Pattern Recognition ICPR*, page 1251–1255, Paris, France, 27–31 Oct 1986.

[2] A. S. Bhadauria, V. Bhateja, M. Nigam, and A. Arya. Skull stripping of brain MRI using mathematical morphology. In S. Satapathy et al., editors, *Smart Intelligent Computing and Applications, Proc. 3rd Int. Conf. Smart Computing and Informatics SCI 2018-19 (Vol. 1)*, volume 159 of *Smart Innovation, Systems and Technologies*, pages 775–780, Bhubaneswar, India, 21–22 Dec 2018. Springer, Singapore 2020. doi:10.1007/978-981-13-9282-5_75.

[3] P.-F. Chen, R. G. Steen, A. Yezzi, and H. Krim. Brain MRI T1-map and T1-weighted image segmentation in a variational framework. In *Proc. 2009 IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pages 417–420, Taipei, Taiwan, 19-24 Apr 2009. doi:10.1109/ICASSP.2009.4959609.

[4] M. Cheour. Advantages of brain MRI. *RadiologyInfo.org*, 2010.

[5] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans. *BrainWeb*: Simulated Brain Database, 1998. `https://brainweb.bic.mni.mcgill.ca`. [Accessed 10 Oct 2020].

[6] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998. doi:10.1109/42.712135.

[7] C. Dai, Y. Mo, E. Angelini, Y. Guo, and W. Bai. Transfer learning from partial annotations for whole brain segmentation. In *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data. Proc. MICCAI Workshop on Domain Adaptation and Representation Transfer DART 2019*, volume 11795 of *Lecture Notes in Computer Science*, pages 199–206, Shenzen, China, 13 Oct 2019. doi:10.1007/978-3-030-33391-1_23.

[8] J. Dai, K. He, and J. Sun. Convolutional feature masking for joint object and stuff segmentation. In *Proc. 2015 IEEE Conf. Computer Vision and Pattern Recognition CVPR*, pages 3992–4000, Boston, MA, USA, 7-12 Jun 2015. doi:10.1109/CVPR.2015.7299025.

[9] A. V. Dalca, E. Yu, P. Golland, et al. Unsupervised deep learning for bayesian brain MRI segmentation. In D. Shen, T. Liu, T. M. Peters, et al., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, volume 11766 of *Lecture Notes in Computer Science*, pages 356–365, 2019. doi:10.1007/978-3-030-32248-9_40.

[10] R. Dey and Y. Hong. CompNet: Complementary segmentation network for brain MRI extraction. In *Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention MICCAI 2018*, volume 11072 of *Lecture Notes in Computer Science*, pages 628–636, Granada, Spain, 16-20 Sep 2018. doi:10.1007/978-3-030-00931-1_72.

[11] J. Doshi, G. Erus, Y. Ou, et al. Multi-atlas skull-stripping. *Academic Radiology*, 20(12):1566–1576, 2013. doi:10.1016/j.acra.2013.09.010.

[12] J. Dutta, D. Chakraborty, and D. Mondal. Multimodal segmentation of brain tumours in volumetric MRI scans of the brain using time-distributed U-Net. In A. K. Das et al., editors, *Proc. Conf. Computational Intelligence in Pattern Recognition CIPR 2019*, volume 999 of *Advances in Intelligent Systems and Computing*, pages 715–725, 2020. doi:10.1007/978-981-13-9042-5_62.

[13] Ø. A. Eide. Skull stripping MRI images of the brain using deep learning. Master's thesis, Norwegian University of Science and Technology, 2018. `https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2566509`.

[14] B. Erden, N. Gamboa, and S. Wood. 3D convolutional neural network for brain tumor segmentation. Technical report, Computer Science, Stanford University, Stanford, USA, 2017. `http://cs231n.stanford.edu/reports/2017/pdfs/526.pdf`.

[15] M. Everingham, L. van Gool, C. Williams, et al. The PASCAL Visual Object Classes homepage, 2012. `http://host.robots.ox.ac.uk/pascal/VOC/`. [Accessed 10 Oct 2020].

[16] A. Fatima, A. R. Shahid, B. Raza, et al. State-of-the-art traditional to the machine-and deep-learning-based skull stripping techniques, models, and algorithms. *Journal of Digital Imaging*, 33(6):1–22, 2020. doi:10.1007/s10278-020-00367-5.

[17] F. J. Galdames, F. Jaillet, and C. A. Perez. An accurate skull stripping method based on simplex meshes and histogram analysis for magnetic resonance images. *Journal of Neuroscience Methods*, 206(2):103–119, 2012. doi:10.1016/j.jneumeth.2012.02.017.

[18] K. He, G. Gkioxari, P. Dollár, and R. Girshick. Mask R-CNN. In *Proc. IEEE Int. Conf. Computer Vision ICCV*, pages 2961–2969, Venice, Italy, 22-29 Oct 2017. doi:10.1109/ICCV.2017.322.

[19] R. Hu, P. Dollár, K. He, et al. Learning to segment every thing. In *Proc. 2018 IEEE Conf. Computer Vision and Pattern Recognition CVPR*, pages 4233–4241, Salt Lake City, USA, 18-23 Jun 2018. doi:10.1109/CVPR.2018.00445.

[20] Y. Huang and L. C. Parra. Fully automated whole-head segmentation with improved smoothness and continuity, with theory reviewed. *PloS ONE*, 10(5):e0125477, 2015. doi:10.1371/journal.pone.0125477.

[21] H. Hwang, H. Z. U. Rehman, and S. Lee. 3D U-Net for skull stripping in brain MRI. *Applied Sciences*, 9(3):569, 2019. doi:10.3390/app9030569.

[22] F. Isensee, M. Schell, I. Pflueger, et al. Automated brain extraction of multisequence MRI using artificial neural networks. *Human Brain Mapping*, 40(17):4952–4964, 2019. doi:10.1002/hbm.24750.

[23] T. Kim, K. Lee, S. Ham, et al. Active learning for accuracy enhancement of semantic segmentation with CNN-corrected label curations: Evaluation on kidney segmentation in abdominal CT. *Scientific Reports*, 10:366, 2020. doi:10.1038/s41598-019-57242-9.

[24] J. Kleesiek, G. Urban, A. Hubert, et al. Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. *NeuroImage*, 129:460–469, 2016. doi:10.1016/j.neuroimage.2016.01.024.

[25] Z. Kuang, X. Deng, L. Yu, et al. Skull R-CNN: A CNN-based network for the skull fracture detection. In T. Arbel et al., editors, *Proc. 3rd Conf. Medical Imaging with Deep Learning MIDL*, volume 121 of *Proceedings of Machine Learning Research*, pages 382–392, Montreal, Canada, 06-08 Jul 2020. `http://proceedings.mlr.press/v121/kuang20a.html`.

[26] K. Kushibar, S. Valverde, S. González-Villà, et al. Automated sub-cortical brain structure segmentation combining spatial and deep convolutional features. *Medical Image Analysis*, 48:177–186, 2018. doi:10.1016/j.media.2018.06.006.

[27] P. J. LaMontagne, T. L. S. Benzinger, J. C. Morris, et al. Oasis-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. *medRxiv*, 2019. doi:10.1101/2019.12.13.19014902.

[28] P. J. LaMontagne, T. L. S. Benzinger, J. C. Morris, et al. OASIS Open Access Series of Imaging Studies, 2019. `https://www.oasis-brains.org`. [Accessed 10 Oct 2020].

[29] K. Landheer, R. F. Schulte, M. S. Treacy, et al. Theoretical description of modern ^{1}H in Vivo magnetic resonance spectroscopic pulse sequences. *Journal of Magnetic Resonance Imaging*, 51(4):1008–1029, 2020. doi:10.1002/jmri.26846.

[30] Y. Li, H. Li, and Y. Fan. ACEnet: Anatomical context-encoding network for neuroanatomy segmentation. *arXiv*, 2020. arXiv:2002.05773 [eess.IV]. `https://arxiv.org/abs/2002.05773`.

[31] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):318–327, 2020. doi:10.1109/TPAMI.2018.2858826.

[32] T.-Y. Lin, P. Dollár, R. Girshick, et al. Feature pyramid networks for object detection. In *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition CVPR*, pages 936–944, Honolulu, Hawaii, 22-25 Jul 2017. doi:10.1109/CVPR.2017.106.

[33] T.-Y. Lin, G. Patterson, M. R. Ronchi, et al. COCO. Common Objects in Context, 2020. `https://cocodataset.org`. [Accessed 10 Oct 2020].

[34] X.-B. Lin, X.-X. Li, and D.-M. Guo. Registration error and intensity similarity based label fusion for segmentation. *IRBM*, 40(2):78–85, 2019. doi:10.1016/j.irbm.2019.02.001.

[35] Y. Liu, Y. Wei, and C. Wang. Subcortical brain segmentation based on atlas registration and linearized kernel sparse representative classifier. *IEEE Access*, 7:31547–31557, 2019. doi:10.1109/ACCESS.2019.2902463.

[36] O. Lucena, R. Souza, L. Rittner, et al. Silver standard masks for data augmentation applied to deep-learning-based skull-stripping. In *Proc. 2018 IEEE 15th International Symposium on Biomedical Imaging ISBI*, pages 1114–1117, Washington, USA, 4-7 Apr 2018. doi:10.1109/ISBI.2018.8363766.

[37] O. Lucena, R. Souza, L. Rittner, et al. Convolutional neural networks for skull-stripping in brain MR imaging using silver standard masks. *Artificial Intelligence in Medicine*, 98:48–58, 2019. doi:10.1016/j.artmed.2019.06.008.

[38] J. V. Manjón, J. E. Romero, R. Vivo-Hernando, et al. Deep ICE: A deep learning approach for MRI intracranial cavity extraction. *arXiv*, 2020. arXiv:2001.05720 [q-bio.QM]. `https://arxiv.org/abs/2001.05720`.

[39] R. Mehta, A. Majumdar, and J. Sivaswamy. BrainSegNet: a convolutional neural network architecture for automated segmentation of human brain structures. *Journal of Medical Imaging*, 4(2):1–11, 2017. doi:10.1117/1.JMI.4.2.024003.

[40] S. Moldovanu, L. Moraru, and A. Biswas. Robust skull-stripping segmentation based on irrational mask for magnetic resonance brain images. *Journal of Digital Imaging*, 28(6):738–747, 2015. doi:10.1007/s10278-015-9776-6.

[41] W. Niblack. *An Introduction to Digital Image Processing.* Prentice Hall, 1986.

[42] M. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man and Cybernetics*, 9(1):62–66, 1979. doi:10.1109/TSMC.1979.4310076.

[43] G. Prasad, A. A. Joshi, A. Feng, et al. Skull-stripping with machine learning deformable organisms. *Journal of Neuroscience Methods*, 236:114–124, 2014. doi:10.1016/j.jneumeth.2014.07.023.

[44] H. Z. U. Rehman, H. Hwang, and S. Lee. Conventional and deep learning methods for skull stripping in brain MRI. *Applied Sciences*, 10(5):1773, 2020. doi:10.3390/app10051773.

[45] S. Ren, K. He, R. B. Girshick, and J. Sun. Faster R-CNN: towards real-time object detection with region proposal networks. *arXiv*, 2015. arXiv:1506.01497 [cs.CV]. `http://arxiv.org/abs/1506.01497`.

[46] S. Ren, K. He, R. B. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149, 2017. doi:10.1109/TPAMI.2016.2577031.

[47] S. Roy, A. Knutsen, A. Korotcov, et al. A deep learning framework for brain extraction in humans and animals with traumatic brain injury. In *Proc. 2018 IEEE 15th International Symposium on Biomedical Imaging ISBI*, pages 687–691, Washington, USA, 4-7 Apr 2018. doi:10.1109/ISBI.2018.8363667.

[48] S. Roy and P. Maji. A simple skull stripping algorithm for brain MRI. In *Proc. 2015 8th Int. Conf. Advances in Pattern Recognition ICAPR*, pages 1–6, Kolkata, India, 4-7 Jan 2015. doi:10.1109/ICAPR.2015.7050671.

[49] S. Roy and P. Maji. An accurate and robust skull stripping method for 3-D magnetic resonance brain images. *Magnetic Resonance Imaging*, 54:46–57, 2018. doi:10.1016/j.mri.2018.07.014.

[50] G. Ruffini, M. D. Fox, O. Ripolles, et al. Optimization of multifocal transcranial current stimulation for weighted cortical pattern targeting from realistic modeling of electric fields. *Neuroimage*, 89:216–225, 2014. doi:10.1016/j.neuroimage.2013.12.002.

[51] J. Sauvola and M. Pietikäinen. Adaptive document image binarization. *Pattern Recognition*, 33(2):225–236, 2000. doi:10.1016/S0031-3203(99)00055-2.

[52] D. Selvathi and T. Vanmathi. Brain region segmentation using convolutional neural network. In *2018 4th Int. Conf. Electrical Energy Systems ICEES*, pages 661–666, Chennai, India, 7-9 Feb 2018. doi:10.1109/ICEES.2018.8442394.

[53] H. Tariq, A. Muqeet, A. Burney, Akhtar H. M., and H. Azam. Otsu's segmentation: Review, visualization and analysis in context of axial brain MR slices. *Journal of Theoretical & Applied Information Technology*, 95(22), 2017. `http://www.jatit.org/volumes/Vol95No22/9Vol95No22.pdf`.

[54] H. Tariq and M. Shahbaz. MAFA: Multispectral adaptive fuzzy algorithm for edge detection on MRI of head scan. *International Journal of Computer Applications*, 182(48):49–54, 2019. doi:10.5120/IJCA2019918737.

[55] G. Valvano, N. Martini, A. Leo, et al. Training of a skull-stripping neural network with efficient data augmentation. *arXiv*, 2018. arXiv:1810.10853 [cs.CV]. `https://arxiv.org/abs/1810.10853`.

[56] A. van der Plas. MRI techniques, 2016. `https://www.startradiology.com/the-basics/mri-technique/`. [Accessed 10 Oct 2020].

[57] M. Wang and P. Li. Label fusion method combining pixel greyscale probability for brain MR segmentation. *Scientific Reports*, 9:17987, 2019. doi:10.1038/s41598-019-54527-x.

[58] X. Wang, R. Girshick, A. Gupta, and K. He. Non-local neural networks. In *Proc. 2018 IEEE Conf. Computer Vision and Pattern Recognition CVPR*, pages 7794–7803, Salt Lake City, USA, 18-23 Jun 2018. doi:10.1109/CVPR.2018.00813.

[59] A. Worth, C. Haselgrove, and D. Kennedy. IBSR. The Internet Brain Segmentation Repository, 2007. https://www.nitrc.org/projects/ibsr/. [Accessed 10 Oct 2020].

[60] L. Xu, H. Liu, E. Song, et al. Automatic labeling of MR brain images through extensible learning and atlas forests. *Medical Physics*, 44(12):6329–6340, 2017. doi:10.1002/mp.12591.

[61] B. Yilmaz, A. Durdu, and G. D. Emlik. A new method for skull stripping in brain MRI using multistable cellular neural networks. *Neural Computing and Applications*, 29(8):79–95, 2018. doi:10.1007/s00521-016-2834-2.

[62] J. Zhou, H.-Y. Heo, L. Knutsson, et al. APT-weighted MRI: Techniques, current neuro applications, and challenging issues. *Journal of Magnetic Resonance Imaging*, 50(2):347–364, 2019. doi:10.1002/jmri.26645.