Automatic neonatal cranial ultrasound segmentation using deep learning: A review

Roa'a Khaled[1,2,*], Arantxa Ortega-Leon[3], Joaquín Pizarro[2], Isabel Benavente Fernández[4], Simón P. Lubián López[4] and Lionel C. Gontard[1,5]

[1]Applied Optics and Magnetism Research Group, University of Cádiz, 11510 Puerto Real, Spain
[2]Department of Computer Engineering, University of Cádiz, 11519 Puerto Real, Spain
[3]Intelligent Modelling of Systems Research Group, University of Cádiz, 11202 Algeciras, Spain
[4]Biomedical Research and Innovation Institute of Cádiz (INiBICA) Research Unit, Puerta del Mar University Hospital, 11009 Cádiz, Spain
[5]IMEYMAT, University of Cádiz, 11510 Puerto Real, Spain

Abstract

Ultrasound is widely used as a clinical routine tool for neonates' brain assessment, especially for preterm neonates. This population is at high risk of developing serious complications leading to neurocognitive and motor impairments. However, the analysis of Cranial Ultrasound requires experienced personnel to perform a time-consuming visual assessment, which is nontrivial due to the low quality and artifacts in the images. For this analysis to be more objective, fast, and accurate, many automatic methods have been proposed. Such methods usually rely on segmenting brain structures or regions of interest for the extraction of subsequent clinically useful measurements. Deep Learning methods are being more adopted recently as they proved to have a huge potential in many medical image analysis tasks.

In this review article, we present and discuss the Deep Learning-based methods developed for the automatic segmentation of preterm neonatal ultrasound images, more specifically the methods developed for segmenting the Cerebral Ventricle System. The performance and evaluation results of these methods are compared, and their major contributions are outlined. Furthermore, we discuss the main challenges of neonatal ultrasound automatic segmentation and possible ways to address these challenges. Finally, we discuss the future directions in this very specific context.

Keywords

Cranial Ultrasound analysis, Deep Learning, Medical image segmentation, Preterm neonates, Cerebral Ventricle System segmentation,

1. Introduction

Ultrasound (US) imaging has been widely used in clinical practice as the first screening and diagnostic tool in many medical domains, including fetal and neonatal care. In neonatal care,

DETERMINED 2022: Neurodevelopmental Impairments in Preterm Children — Computational Advancements, August 26, 2022, Ljubljana, Slovenia
*Corresponding author.
roaa.khaled@uca.es (R. Khaled); arantxa.ortega@uca.es (A. Ortega-Leon); joaquin.pizarro@uca.es (J. Pizarro); isabel.benavente@uca.es (I. B. Fernández); simonplubian@gmail.com (S. P. L. López); lionel.cervera@uca.es (L. C. Gontard)
0000-0002-5231-8462 (R. Khaled); 0000-0002-0793-3677 (A. Ortega-Leon); 0000-0002-4295-6743 (J. Pizarro); 0000-0001-9276-1912 (I. B. Fernández); 0000-0001-8603-7119 (L. C. Gontard)

Cranial US is extensively used for routine brain assessment of newborn infants and more specifically preterm infants. This widespread use of US is due to its several advantages over other imaging modalities, such as the low cost, non-invasive nature, non-ionizing radiation, real-time display, operator comfort, portability, and accessibility [1, 2, 3, 4].

Cranial US can only be used in the newborn period during which the anterior fontanelle is still open (usually until 18 months of age), but it is mostly used during the first 5-6 months of age when the best US images can be obtained. After that age brain structures will start to be less visible due to the processes of brain membranes thickening and fontanelle closure [5, 6].

Cranial US allows the detection of most neonatal hemorrhagic and ischemic lesions in addition to the main congenital and maturational anomalies [7]. However, the use of US entitles some challenges. For instance, US has low imaging quality and suffers from noise and artifacts. Moreover, it requires trained and experienced operators to acquire good images and perform a tedious visual assessment, which leads to high inter- and intra-observer variability across different institutes and US systems manufacturers [1, 7].

Since the standard clinical practice is based on visual assessment and some 2D linear measurements, much research has been conducted to propose the use of better quantitative analysis over the visual assessment and to prove the usefulness of other measurements than the 2D linear ones, such as volumetric measurements. This has the potential to improve the diagnosis and prognosis of neurodevelopmental disorders in preterm neonates. However, this could not be adopted in clinics yet because it requires manually segmenting anatomical structures of interest in the brain, which is time-consuming and prone to inter- and intra-observer variability [7].

Developing automatic methods for the analysis of Cranial US images can alleviate these challenges by making such analysis more objective, accurate, and fast. Automatic methods include segmentation as an important preliminary step for the extraction of clinical parameters that neonatologists need in order to perform an assessment and diagnosis based on quantitative measurements [8].

One of the important structures to be segmented in Cranial US images of preterm neonates is the Cerebral Ventricle System (CVS). CVS can be affected by some serious complications such as germinal matrix-intraventricular hemorrhage leading to posthemorrhagic ventricular dilatation (PHVD). This happens because of preterm birth and causes neurocognitive and motor impairments.

Currently, the clinical standard is to perform 2D measurements manually on 2D US images to estimate the CVS volume. This practice, apart from being time-consuming and subjective, is imprecise due to the unavailability of 3D information [2, 7]. Therefore, developing automatic segmentation methods and quantitative analysis methods based on 3D US can help clinicians to perform timely medical interventions and improve the outcome of those infants [9, 10]. However, the task of automatically segmenting anatomical structures from Cranial US is very challenging due to several reasons, such as the variable image quality, presence of noise and shading artifacts, unclear and incomplete boundaries, similar intensities among different structures, variable size and anatomical shape of the ventricles for neonates with abnormalities. Moreover, the differences in shape, size, and texture characteristics caused by the change in blood pressure [11].

Recently, there has been an increasing trend in the use of Deep Learning (DL) algorithms to segment CVS from neonatal Cranial US images. This is due to the success that DL methods have been achieving in the field of medical image analysis in the last years.

There are several reviews focused on neonatal neuroimage segmentation, but most of them focused on MRI [12, 13]. Although few reviews have been conducted on US segmentation methods based on DL [1, 14], they were generic and included studies on different medical domains (i.e. not focused on neonatal US segmentation). To this date, and to the best of our knowledge, no reviews have been written on segmentation methods of neonatal US images specifically.

Therefore, we conducted a literature search for all studies published in this field from 2018 until 2022 July 1st, by specifying keywords such as (preterm neonatal AND cerebral ventricles AND ultrasound AND segmentation AND deep learning) in Google Scholar database. Abstracts of papers resulting from this search were screened and only (8) relevant papers were chosen.

In this review we present a systematic overview of DL methods in segmenting Cranial US images of preterm neonates, more specifically, segmenting the CVS. In Section 2 we briefly mention the evolution of segmentation methods in this specific context and review several DL methods developed for preterm neonatal ventricles segmentation from US images. In Section 3 we discuss the challenges of US image segmentation and the possible ways to address these challenges in the future. Finally, in Section 4 we present our conclusions.

2. Cranial Ultrasound Image Segmentation

2.1. Non-DL Based Image Segmentation

Many studies have been conducted for automating US image analysis of different organs but very few studies have focused on US neuroimaging [15, 16]. Most segmentation methods were initially based on well-established image processing techniques. In those methods, images are first pre-processed for denoising using image filtering. Then segmentation is carried out using intensity thresholding or edge detection filters. Finally, image analysis of binary images is carried out using morphological operations. For instance, Gontard et al. [17] used median filtering followed by a global intensity threshold calculated automatically from the 3D volume for segmenting cerebrospinal fluid (CSF).

Nevertheless, boundary incompleteness in US images raises great challenges to automatic segmentation. Therefore, most methods were semi-automatic where some input from the user is required. Additionally, shape prior can provide strong guidance in estimating the missing boundary. Qiu et al. [18] proposed a semi-automatic convex segmentation algorithm for ventricle segmentation in 3D US images. In [19] a geometric-based method using a 3D ellipsoid estimation technique was proposed for ventricle segmentation. However, traditional shape models often suffer from being reliable on hand-crafted descriptors and losing local information in the fitting procedure, hence, such methods had poor generalization.

A semi-automatic approach was proposed in [20] for ventricle segmentation. In this study, the image is denoised using complex wavelets and then 3 seed points are required to be manually selected in order to perform an active contour segmentation where the contour is parametrized implicitly using a level-set function. The most advanced non-DL-based segmentation method for segmenting ventricles was developed by Qiu et al [21]. This method made use of a phase congruency map, multi- atlas initialization technique, atlas selection strategy, and a multiphase geodesic level-sets evolution combined with a spatial shape prior derived from multiple pre-segmented atlases. Nevertheless, the method proposed required 54 min to segment one volume, which is too long to be used in clinical routine.

In addition to the previously mentioned methods, Machine Learning based methods have also been proposed for US image segmentation. Tabrizi et al. [22] proposed an automatic method for ventricle segmentation in 2D US images based on a hybrid approach consisting of fuzzy c-means, adaptive thresholding, template matching, phase congruency, and active contour algorithms.

2.2. Deep Learning for Image Segmentation

Nowadays, DL methods represent the state-of-the-art methods for image analysis and have outperformed any other conventional methods in both performance and speed in terms of the specific task.

Two main methodologies are currently used to address boundary detection-segmentation in US:
1) A top-down manner that takes advantage of prior shape information to guide segmentation. For example, Yang et al. [23] formulated boundary completeness as a sequential problem and a model of the shape in a dynamic manner using Recurrent Neural Networks. Authors in [24] modified Convolutional Neural Network (CNN) architectures like the Hough-CNN which include explicitly transforms for edge detection.
2) A bottom-up manner that classifies each pixel into foreground (object) or background in a supervised manner. Most studies apply this approach by classifying each pixel in an image in an end-to-end and fully supervised learning manner employing CNNs with encoder-decoder architectures.

The first widely recognized encoder-decoder network was Seg-Net [25]. Later, UNet [26] brought a major breakthrough in medical image segmentation, and became the backbone of almost all the leading methods recently, such as UNet++, UNet3+, 3D UNet, V-Net, Res-UNet, and Dense-UNet. In these extensions of UNet, the contribution was either in skip connections, using better convolutional layer connections, or in applications. For instance, UNet++ [27], [28] utilizes nested and dense skip connections for further reduction of the semantic gap between the encoder and decoder feature maps. In UNet3+ [29], skip connections between different scales are used. 3D UNet [30] and V-Net [31] are extensions of UNet for volumetric segmentation of 3D medical images. In Res-UNet [32] the encoder and decoder convolutional blocks consist of residual connections [33], while in Dense-UNet [34], they consisted of dense blocks [35].

Due to the difficulty of 3D DL, the DL methods that are currently applied in medical US analysis mostly use 2D images as inputs, although these 2D images might be taken from available 3D volumes. In fact, 3D DL is still a challenging task, due to the following limitations:
1) Training a deep network on a large volume might be too computationally expensive for real clinical application (i.e. with a significantly increased memory and computational requirement).
2) A deep network with a 3D volume as input requires more training samples since a 3D network contains parameters that are orders of magnitude higher than a 2D network. This may dramatically increase the risk of overfitting, given the limited training samples. Alternatively, there are authors that formulate the problem of optimizing 3D image segmentation as a patch-

level classification task, as was proposed in [36].

In fact, there are not so many DL methods proposed for neonatal US segmentation, and in this review (Section 2.3) we are reviewing the DL methods for CVS segmentation specifically.

2.3. Deep Learning for CVS segmentation from Cranial US images

In this subsection, we review 8 papers that were selected after a literature search for studies published on the use of DL for segmenting lateral ventricles or the whole CVS from Cranial US. The search included papers that were published from 2018 until 2022 July 1st. Those 8 papers were the only studies found that utilize DL for this task and they are summarized in Table 1.

	Architecture (2D/3D)	Dataset	2D/3D segmentation	Augmentation	Loss function	Evaluation	Inference time (s)
Martin et al. (2018) [10]	UNet (2D)	15 volumes (private)	3D	-	Soft Dice	DSC = 0.816 HD = 13.6 MAD = 0.62	5 s
Wang et al. (2018) [37]	UNet and SegNet combination (2D)	687 slices (private)	2D	horizontal flip, random crop	MAE	DSC = 0.908 IoU = 84.84% Pix. Acc. = 92.14%	0.022 s
Valanarasu et al. (2020) [38]	CBAS (2D)	1629 (private)	2D	horizontal and vertical flips, random crop	confidence guided	DSC = 0.8901 IoU = 81.03%	0.01 s
Tabrizi et al. (2020) [39]	UNet like (2D)	1253 slices (private)	2D	vertical flip, affine transformation	probabilistic atlas-based	DSC = 0.86 HD = 0.3 mm	17.4 s
Gontard et al. (2021) [40]	pretrained SegNet (2D)	152 volumes (private)	3D	translation, rotation, scale, shear	weighted BCE	DSC = 0.8	< 60 s
Martin et al. (2021) [41]	V-Net/ UNet with CPPN (2D and 3D)	25 volumes (private)	2D and 3D	-	BCE then soft Dice	(for V-Net) DSC = 0.822 MAD = 0.5 mm δ Va = 0.35 cm^3 δ Vr = 11.1%	3.5 s (for 2D)
Valanarasu et al. (2022) [42]	KiUNet (2D)	1629 slices (private)	2D	-	BCE	DSC = 0.8943	-
Szentimrey et al. 2022 [43]	UNet ensemble (3D)	190 volumes (private)	3D	translation	combined BCE and Dice loss (with MSE for the 3rd model)	DSC = 0.72 VD = 3.7 cm^3 MAD = 1.14 mm	5 s

Table 1: Comparison of DL based methods for automatic CVS segmentation from Cranial US images. Loss Functions: **MAE** is Mean Absolute Error loss, **BCE** is Binary Cross Entropy loss, and **MSE** is Mean Squared Error loss. Evaluation Metrics: **DSC** is Dice Similarity Coefficient, **HD** is Hausdorff Distance, **MAD** is Mean Absolute Distance, **IoU** is Intersection over Union, **Pix. Acc.** is Pixel Accuracy, δ **Va** is Absolute volume difference, δ **Vr** is Relative volume difference, and **VD** is Absolute Volumetric Difference

It is worth mentioning that to get an estimation of the CVS volume, clinicians usually obtain various linear measurements manually from 2D images. However, this practice is imprecise (since 3D information is missing), time-consuming, and operator dependent. Therefore, the studies reviewed here mainly aimined to improve the accuracy and reduce the time required to perform manual segmentation by automating this task and therefore paving the way for obtaining clinical measurements automatically. Some of the reviewed studies were also aiming to improve the performance of automatic segmentation by utilizing 3D information, which may result in more accurate and representative volumetric clinical measurements.

In 2018 Martin et al. [10] extended CVS volume estimation to 3D. They used a 2D UNet to first segment 2D angular image sequence. Then they propose an algorithm for 3D reconstruction to reconstruct 3D segmentation. This method can significantly reduce the extensive computation cost and memory requirement of 3D processing. A limitation of this study is the small dataset, which affects the ability of the model to generalize.

Wang et al. [37] proposed a CNN that combines the advantages of both UNet and SegNet architectures to segment lateral ventricles from 2D US. The proposed network consists of two components: a pre-trained DenseNet as the encoder to extract deep features, and a multi-scale decoder that first applys pooling of the feature maps (resulted from the encoder) into four different sizes and then applies a series of transposed convolutions to transform lower dimensional feature maps into higher ones in steps. Moreover, the output of each transposed convolution is concatenated with existing feature maps of the same size and then fed into the next transposed convolution.

Since the resolution of small features is gradually lost along the deeper layers of a CNN, the resulting coarse features can miss the details of small structures. This leads to poor performance of traditional CNN architectures in segmenting small anatomical structures (as in the case of normal ventricles for example). To address that, Valanarasu et al. [38] propose a network (Confidence-guided Brain Anatomy Segmentation-CBAS), where segmentation and corresponding confidence maps are estimated at different scales. Aleatoric uncertainty is computed as the confidence scores to indicate how confident the CBAS network is about the segmentation output. This allows CBAS to learn how to differentiate regions with higher error (low confidence score) and therefore focus more on those regions in subsequent layers and block the propagation of error while computing the segmentation output.

Tabrizi et al. [39] proposed a method to segment lateral ventricles from 2D US images. The proposed method integrates anatomical information into a CNN by defining a new weighted loss function and an image-specific adaption. First, a deep CNN was used to detect the cranium and brain interhemispheric fissure to estimate the anatomical position of ventricles and correct the cranium rotation. Then, lateral ventricles were segmented using a CNN with a similar structure to that of a 2D UNet. The CNN learning was integrated with a prior model of the lateral ventricles through a probabilistic atlas-based weighted loss function and an image-specific adaption. Moreover, the authors performed posthemorrhagic hydrocephalus (PHH) outcome prediction (necessity of intervention) using a support vector machine classifier that was trained on ventricular morphology and clinical parameters. The segmentation performance was affected by the unclear boundaries caused by the build-up of hemorrhage pressure, but this is a challenge that experts also experience when doing manual segmentation. Regarding PHH output prediction, although the prediction performance was good, the features used were hand-crafted and based on 2D measurements. We believe that 3D features learned by the DL model may improve the PHH output prediction accuracy.

Gontard et al. [40] utilized a pre-trained SegNet model based on VGG16 to obtain 3D ventricular segmentation from 2D thickened sagittal slices (i.e. 3 consecutive slices). After that 3D ventricular volumes were estimated using the segmented 2D slices.

Martin et al. [41] utilized both V-Net and UNet (for both 2D and 3D images) to estimate CVS volume in a dataset including both normal and dilated ventricles. Moreover, the use of a Compositional Pattern Producing Network (CPPN) was proposed to enable the CNNs to learn spatial information about the CVS location. Their results showed a comparative performance for both V-Net and UNet, with V-Net being slightly better (especially in segmenting normal ventricles). They also reported that CPPN increased the accuracy of the CNNs when having fewer layers. It would be interesting to investigate the benefits of the CPPN for multi-structural brain segmentation. Results reported in this study show that a 3D architecture is overall more accurate for this task. Nevertheless, a 2D architecture was as accurate as a 3D architecture for segmenting dilated ventricles. Moreover, it was shown that a 2D architecture enables to perform the segmentations in clinical time with hardware that requires fewer memory resources and therefore may be preferable to a 3D architecture in a clinical context.

To address the issue of poor segmentation of smaller structures and boundary regions in medical image segmentation in general, Valanarasu et al. [42] proposed an architecture (KiUNet) that consists of two branches. The first branch is an overcomplete convolutional network (Kite-Net) which learns to capture fine details and accurate edges of the input by projecting the input image into a higher dimension such that the receptive field is being constrained from increasing in the deep layers of the network. The second part is a UNet which learns high-level features. A cross-residual fusion strategy was proposed to combine the features across the two branches. Moreover, the architecture was proposed in both 2D and 3D settings, and a Res-KiUNet and a Dense-KiUNet architectures where also proposed for improving the learning of the network, where residual connections and dense blocks are utilized. Finally, the proposed method was tested on 5 different datasets of different medical image applications and modalities, including lateral ventricles from US, and was proved to generalize well to different modalities.

Nevertheless, only one metric was used for evaluation in [42], that is the Dice Similarity Coefficient (DSC), which might not be very indicative of the improvement in segmentation unless the segmented structure is small. For instance, dice values for US ventricular segmentation dataset were not significantly improved compared to other methods reported in this study, which is expected since dilated ventricles are not very small structures (compared to tumors datasets for example where improvement was reported to be clearer). Therefore, other metrics might also be more useful for showing the improvement in segmenting ventricle boundaries. Also, it would be interesting to test this method on 3D ventricles segmentation and use volumetric metrics to evaluate the performance, since volumetric measurements might be more susceptible to slight improvements in segmenting the surface or boundaries. Another contribution of this work is that the network's memory requirements are less while maintaining decent performance. However, it would be interesting to compare with a deeper KiUNet that is as deep as the UNet they compared with.

To address the limitations of 2D US, Szentimrey et al. [43] developed a method to segment lateral ventricles from 3D US images using a 3D UNet ensemble model composed of three UNet variants. Each variant highlights various aspects of the segmentation task such as the shape and boundary of the ventricles. The ensemble is made of a UNet++, attention UNet, and UNet with a DL-based shape prior combined using a mean voting strategy. The UNet++ has more skip connections compared to the basic UNet, to allow for a more flexible fusion of feature maps at the decoder pathway and make the semantic maps between the encoder and decoder more similar which is believed to make the learning task easier for the optimizer and either improve the speed and/or performance of the model. The attention UNet incorporates attention gates to improve the ventricle surface segmentation boundary (which is challenging in US images) by improving the sensitivity to foreground voxels while adding minimal complexity to the model. The UNet with a DL-based shape prior utilizes a shape prior loss function to add surface regularization by conforming the predicted ventricle shape to that of the ground truth segmentation.

Even though incorporating shape prior resulted in improving the segmentation of ventricles according to [43], it might not be the case if an unseen test image has a unique ventricle shape not captured in the training data, which is likely to happen because ventricles might have several deformations. Another limitation is that the ensemble model is computationally heavy, especially due to the UNet++ model. Therefore, GPU resources are required even during test time, which might not always be available at healthcare points. Moreover, the ventricles were manually annotated on the sagittal plane every 1mm such that slices between each manual contour required interpolation, leading to possible inaccuracies of the ground truth volume.
On the other hand, they utilized bigger data compared to previous studies, and they included scans with varying degrees of intraventricular hemorrhage and scans with only one ventricle being visible due to the limited field of view. Several metrics were used for evaluating the proposed method's performance, including metrics that are clinically useful, especially absolute volumetric difference (VD), which has been used for patients with PHVD to determine those who need intervention [44].

All methods reviewed in this section seem to have good performance according to the reported results (both in terms of accuracy and speed). Each method had its contributions and limitations. However, it is worth mentioning that the comparability of methods, in this case, is not straightforward since each method was developed using a different private dataset that varies in the number of cases, image quality...etc. Moreover, in most cases, small datasets were used, and it was not mentioned about the number of data resulting from Augmentation. This becomes more of a problem in the case of training on 3D volumes. Therefore, we believe that efforts are still needed to form large open datasets that will allow researchers to develop new methods and compare them with others.

Another area that we believe needs to be further investigated is whether segmenting 3D data would improve the performance. One would expect that incorporating 3D information using 3D architectures would increase the accuracy of segmentation. Authors in [41] reported comparative performance of both 2D and 3D architectures for segmenting dilated ventricles, however, they used a small dataset.

Regarding the applicability of the proposed methods in clinical settings, memory, and computational requirements are also important (besides the accuracy and speed). Even though inference time was reported in most of these studies, it was not always mentioned whether the developed methods can be used in machines with lower memory and computational resources, or if they need special requirements. We believe that most of the proposed methods were computationally heavy and therefore novel methods are still needed to tackle this issue.

3. Challenges in US segmentation and Possible solutions

3.1. Limited availability of annotated data and Image Synthesis

One of the major problems in medical image analysis is the limited number of annotated data. This is due to the difficulty of sharing patient data publicly and the difficulty of obtaining clinical annotations since it is expensive and time-consuming.

However, most advanced research on automation of US analysis is based on supervised learning which is strongly dependent on the access to open and considerable amounts of data, acquired on different populations and with different operating conditions (and with different US scanners). This leads to a lack of generalization and validation of the AI models. Moreover, not having access to open large data makes it difficult to reproduce and compare the proposed methods.

In this context, federated learning or data augmentation strategies are important for developing better algorithms. Moreover, novel image synthesis methods are proposed in the literature to synthesize high-quality data that could be added to the training dataset. Generative Adversarial Networks (GANs) [45] and their variants are powerful architectures capable of generating synthetic images to be used for training other networks, for example, UNet-based networks. In addition, GANs are favored over traditional methods for handling data imbalance [46] by synthesizing realistic-looking minority class samples, thereby balancing the class distribution, and avoiding overfitting. GANs are being applied for generating 3D medical imaging data [47], however, generating realistic-looking data samples in US neuroimaging is an open research problem [48] and further research is required to improve and validate the quality of the synthezised samples. Another challenge is that while using GANs in medical imaging to synthesize new images solves the issue of limited available data, the problem of annotations still exists in this setup. Therefore, novel methods are needed to synthesize annotations as well. To tackle this issue, Valanarasu et al. [38] proposed a method for image synthesis using multi-scale self-attention generator where 2D Cranial US images are synthesized directly from manipulated segmentation masks (ventricle and septum pellucidi masks). Thus, there is no need for annotation of the synthesized data.

Alternatively, data can be generated through the simulation of US images [49]. This is a field largely unexplored in the context of neuroimaging. For example, we suggest that 3D models of neonatal brain gyrification might be generated as in [50] and then used for simulating US images using computing simulation toolboxes like MUST [51] or FIELD II [52].

3.2. Segmentation of other brain structures

MRI is used in neonatology to segment not only the lateral ventricles and external CSF but also white matter, cortical gray matter, cerebellum, or brain stem [53]. US neuroimaging might complement better MRI neuroimages if US data could provide information on other brain structures. For example, most studies with US report measurements related only to ventricular dilation but it would be more interesting to assess those measurements relative to the total brain volume [19]. With appropriate data labeling US might also be used for the detection and quantification of white matter injuries. Finally, the folding dynamics of the brain, occurring mostly before normal-term birth, are vastly unknown. US might help to better understand this process by looking into the development of cortical sulci in infants. For instance, longitudinal studies of the central brain sulcus could in principle be carried out with 3D US like it is done with MRI [54].

3.3. Inherent US image limitations and Image preprocessing

US acquisition introduces noise in the signal, which corrupts the resulting image and affects further processing steps, e.g., segmentation and quantitative analysis. US segmentation can clearly benefit from the application of preprocessing methods for improving image quality (denoising, deblurring, increasing resolution). DL is being applied to improve the resolution and contrast-to-noise ratio of the reconstruction algorithms of the signal acquired with the US sensors [55, 56]. And DL will certainly be very promising for US image enhancement and denoising using super-resolution methods [57, 58].

3.4. Novel AI architectures

The aforementioned encoder-decoder CNN architectures achieved state-of-the-art performance in medical imaging segmentation. UNet, has become the de-facto standard and achieved tremendous success. However, due to the intrinsic locality of convolution operations, UNet generally demonstrates limitations in explicitly modeling long-range dependency (i.e., they lack focus in extracting low-level features) since the networks are built to be deeper and hence more high-level features get extracted. As a result, they fail in providing a good segmentation of small structures with blurred boundaries, which is the case with US image segmentation. This implies the need for novel architectures or variants.

GANs for example are explored for image segmentation using image transfer methods [45]. And Transformers, designed for sequence-to-sequence prediction, with innate global self-attention mechanisms, have emerged strongly as alternative architectures [59] to Encoder-Decoder architectures for medical image segmentation. To name some recent examples, TransUNet [60] merits both Transformers and UNet CNNs, UNetFormer [61] increases the efficiency of conventional UNet architectures, and MedFormer can generalize to different medical domains[62].

4. Conclusions

DL has meant a change of paradigm in medical imaging analysis, and new techniques and architectures are in continuous development which will certainly impact US imaging and analysis. Synthetic data generation, transformers, and super-resolution methods can help to overcome some limitations of US image analysis with respect to MRI.

Automatic methods that yield reliable 3D measurements of the ventricles are expected to provide a more accurate assessment of preterm neonates' ventricles and other cerebral structures, which can improve the monitoring and treatment decisions of preterm born infants. Overall, the studies reviewed in this review demonstrate the possibility of achieving an accurate segmentation of preterm neonates' CVS in a clinical time in 3D US images and therefore pave the way to prove the clinical benefits of 3D US in monitoring cerebral structures of preterm neonates, not only for CVS dilation but also for brain growth, sulci formation or detection of white matter injuries.

In the future, studies that compare volumetric measurements obtained from both US and MRI are needed, to show whether the measurements obtained from 3D US can be competitive with those obtained from MRI. Moreover, models utilizing both US and MRI can be developed to study whether both modalities contain complementary information that could help improve the accuracy.

Another important future direction is automatic outcome prediction based on automatic ventricular segmentation and measurements, this can include predicting the progression of PHH which offers an opportunity for early interventions to improve outcome [39]. Developing AI tools that combine measurements of other cerebral structures, like those related to White Matter damage or Sulci malformation, can also be used to predict the long-term outcome of preterm infants and the probability of them developing neurodevelopmental impairments. To the best of our knowledge, this has not been achieved yet, but with the continuous developments of methods in this field, this can be achieved in the following few years.

References

[1] S. Liu, Y. Wang, X. Yang, B. Lei, L. Liu, S. X. Li, D. Ni, T. Wang, Deep Learning in Medical Ultrasound Analysis: A Review, Engineering 5 (2019) 261–275. URL: https://www.sciencedirect.com/science/article/pii/S2095809918301887. doi:https://doi.org/10.1016/j.eng.2018.11.020.

[2] M. J. Brouwer, L. S. De Vries, L. Pistorius, K. J. Rademaker, F. Groenendaal, M. J. Benders, Ultrasound Measurements of the Lateral Ventricles in Neonates: Why, How and When? A Systematic Review, Acta paediatrica 99 (2010) 1298–1306.

[3] M. Riccabona, Potential Role of 3D US in Infants and Children, Pediatric radiology 41 Suppl 1 (2011) S228–37. doi:10.1007/s00247-011-2051-1.

[4] I. Timor-Tritsch, A. Monteagudo, P. Mayberry, Three-dimensional Ultrasound Evaluation of the Fetal Brain: the Three Horn View, Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology 16 (2000) 302–306.

[5] L. Pogliani, G. Zuccotti, M. Furlanetto, V. Giudici, A. Erbetta, L. Chiapparini, L. Valentini, Cranial ultrasound is a reliable first step imaging in children with suspected craniosynostosis, Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery 33 (2017). doi:10.1007/s00381-017-3449-3.

[6] P. Gupta, K. Sodhi, A. Saxena, N. Khandelwal, P. Singhi, Neonatal cranial sonography: A concise review for clinicians, Journal of Pediatric Neurosciences 11 (2016) 7. doi:10.4103/1817-1745.181261.

[7] I. Benavente-Fernández, E. Ruiz-González, M. Lubian-Gutiérrez, et al., Ultrasonographic Estimation of Total Brain Volume: 3D Reliability and 2D Estimation. Enabling Routine Estimation During NICU Admission in the Preterm Infant, Frontiers in Pediatrics 9 (2021). doi:10.3389/fped.2021.708396.

[8] J. P. Kusanovic, J. K. Nien, L. F. Gonçalves, J. Espinoza, W. Lee, M. Balasubramaniam,

E. Soto, O. Erez, R. Romero, The Use of Inversion Mode and 3D Manual Segmentation in Volume Measurement of Fetal Fluid-filled Structures: Comparison with Virtual Organ Computer-aided AnaLysis (VOCAL™), Ultrasound in obstetrics & gynecology 31 (2008) 177–186.

[9] M. N. Cizmeci, N. Khalili, N. H. P. Claessens, et al., Assessment of Brain Injury and Brain Volumes after Posthemorrhagic Ventricular Dilatation: A Nested Substudy of the Randomized Controlled ELVIS Trial, The Journal of Pediatrics 208 (2019) 191–197.e2. doi:10.1016/j.jpeds.2018.12.062.

[10] M. Martin, B. Sciolla, M. Sdika, et al., Automatic Segmentation of the Cerebral Ventricle in Neonates Using Deep Learning with 3D Reconstructed Freehand Ultrasound Imaging, in: 2018 IEEE International Ultrasonics Symposium (IUS), 2018, pp. 1–4. doi:10.1109/ULTSYM.2018.8580214.

[11] K. M. Meiburger, U. R. Acharya, F. Molinari, Automated Localization and Segmentation Techniques for B-mode Ultrasound Images: A review, Computers in Biology and Medicine 92 (2018) 210–235. URL: https://www.sciencedirect.com/science/article/pii/S0010482517303888. doi:https://doi.org/10.1016/j.compbiomed.2017.11.018.

[12] C. N. Devi, A. Chandrasekharan, V. Sundararaman, Z. C. Alex, Neonatal brain MRI segmentation: A review, Computers in Biology and Medicine 64 (2015) 163–178. URL: https://www.sciencedirect.com/science/article/pii/S0010482515002346. doi:https://doi.org/10.1016/j.compbiomed.2015.06.016.

[13] A. Makropoulos, S. J. Counsell, D. Rueckert, A Review on Automatic Fetal and Neonatal Brain MRI Segmentation, NeuroImage 170 (2018) 231–248. URL: https://www.sciencedirect.com/science/article/pii/S1053811917305451. doi:https://doi.org/10.1016/j.neuroimage.2017.06.074, segmenting the Brain.

[14] Z. Wang, Deep Learning in Medical Ultrasound Image Segmentation: a Review, 2021. arXiv:2002.07703.

[15] J. A. Noble, D. Boukerroui, Ultrasound Image Segmentation: a Survey, IEEE Transactions on medical imaging 25 (2006) 987–1010.

[16] M. H. Mozaffari, W. Lee, 3D Ultrasound Image Segmentation: A Survey, arXiv preprint arXiv:1611.09811 (2016).

[17] L. C. Gontard, J. Pizarro, I. Benavente-Fernández, S. P. Lubián-López, Automatic Measurement of the Volume of Brain Ventricles in Preterm Infants from 3D Ultrasound Datasets, in: ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing, Springer, 2019, pp. 323–329.

[18] W. Qiu, J. Yuan, J. Kishimoto, J. McLeod, Y. Chen, S. de Ribaupierre, A. Fenster, User-Guided Segmentation of Preterm Neonate Ventricular System from 3-D Ultrasound Images Using Convex Optimization, Ultrasound in Medicine Biology 41 (2015) 542–556. URL: https://www.sciencedirect.com/science/article/pii/S0301562914006371. doi:https://doi.org/10.1016/j.ultrasmedbio.2014.09.019.

[19] M.-A. Boucher, S. Lippé, A. Damphousse, R. El-Jalbout, S. Kadoury, Dilatation of Lateral Ventricles with Brain Volumes in Infants with 3D Transfontanelle US, in: MICCAI, 2018.

[20] B. Sciolla, M. Martin, P. Delachartre, P. Quetin, Segmentation of the Lateral Ventricles in 3D Ultrasound Images of the Brain in Neonates, in: 2016 IEEE International Ultrasonics Symposium (IUS), IEEE, 2016, pp. 1–4.

[21] W. Qiu, Y. Chen, J. Kishimoto, S. de Ribaupierre, B. Chiu, A. Fenster, J. Yuan, Automatic Segmentation Approach to Extracting Neonatal Cerebral Ventricles from 3D Ultrasound Images, Medical image analysis 35 (2017) 181–191.
[22] P. Roshani Tabrizi, R. Obeid, J. Cerrolaza, A. Penn, A. Mansoor, M. G. Linguraru, Automatic Segmentation of Neonatal Ventricles from Cranial Ultrasound for Prediction of Intraventricular Hemorrhage Outcome, volume 2018, 2018, pp. 3136–3139. doi:10.1109/EMBC.2018.8513097.
[23] X. Yang, L. Yu, L. Wu, Y. Wang, D. Ni, J. Qin, P.-A. Heng, Fine-grained Recurrent Neural Networks for Automatic Prostate Segmentation in Ultrasound Images, 2016. URL: https://arxiv.org/abs/1612.01655. doi:10.48550/ARXIV.1612.01655.
[24] F. Milletari, S.-A. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, J. Levin, O. Dietrich, B. Ertl-Wagner, K. Bötzel, N. Navab, Hough-CNN: Deep Learning for Segmentation of Deep Brain Regions in MRI and Ultrasound, Computer Vision and Image Understanding 164 (2017) 92–102. URL: https://www.sciencedirect.com/science/article/pii/S1077314217300620. doi:https://doi.org/10.1016/j.cviu.2017.04.002, deep Learning for Computer Vision.
[25] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (2017) 2481–2495. doi:10.1109/TPAMI.2016.2644615.
[26] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, in: N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing, Cham, 2015, pp. 234–241.
[27] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, UNet++: A Nested U-Net Architecture for Medical Image Segmentation: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings, volume 11045, 2018, pp. 3–11. doi:10.1007/978-3-030-00889-5_1.
[28] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation, IEEE transactions on medical imaging (2019). doi:10.1109/TMI.2019.2959609.
[29] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, J. Wu, UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation, ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2020) 1055–1059.
[30] Çiçek, S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, 2016, pp. 424–432. doi:10.1007/978-3-319-46723-8_49.
[31] F. Milletari, N. Navab, S.-A. Ahmadi, V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, 2016, pp. 565–571. doi:10.1109/3DV.2016.79.
[32] X. Xiao, S. Lian, Z. Luo, S. Li, Weighted Res-UNet for High-Quality Retina Vessel Segmentation, in: 2018 9th International Conference on Information Technology in Medicine and Education (ITME), 2018, pp. 327–331. doi:10.1109/ITME.2018.00080.
[33] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, in: 2016

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778. doi:10.1109/CVPR.2016.90.

[34] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, P.-A. Heng, H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation From CT Volumes, IEEE Transactions on Medical Imaging 37 (2018) 2663–2674. doi:10.1109/TMI.2018.2845918.

[35] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely Connected Convolutional Networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2261–2269. doi:10.1109/CVPR.2017.243.

[36] R. Khaled, J. Vidal, J. C. Vilanova, R. Martí, A U-Net Ensemble for Breast Lesion Segmentation in DCE MRI, Computers in Biology and Medicine (2022) 105093.

[37] P. Wang, N. G. Cuccolo, R. Tyagi, et al., Automatic Real-Time CNN-based Neonatal Brain Ventricles Segmentation, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 2018, pp. 716–719. doi:10.1109/ISBI.2018.8363674.

[38] J. M. Jose Valanarasu, R. Yasarla, P. Wang, I. Hacihaliloglu, V. M. Patel, Learning to Segment Brain Anatomy From 2D Ultrasound With Less Data, IEEE Journal of Selected Topics in Signal Processing 14 (2020) 1221–1234. doi:10.1109/JSTSP.2020.3001513.

[39] P. R. Tabrizi, A. Mansoor, R. Obeid, J. J. Cerrolaza, D. A. Perez, J. Zember, A. Penn, M. G. Linguraru, Ultrasound-Based Phenotyping of Lateral Ventricles to Predict Hydrocephalus Outcome in Premature Neonates, IEEE Transactions on Biomedical Engineering 67 (2020) 3026–3034. doi:10.1109/TBME.2020.2974650.

[40] L. C. Gontard, J. Pizarro, B. Sanz-Peña, et al., Automatic Segmentation of Ventricular Volume by 3D Ultrasonography in Post Haemorrhagic Ventricular Dilatation Among Preterm Infants, Scientific Reports 11 (2021). doi:10.1038/s41598-020-80783-3.

[41] M. Martin, B. Sciolla, M. Sdika, et al., Automatic Segmentation and Location Learning of Neonatal Cerebral Ventricles in 3D Ultrasound Data Combining CNN and CPPN, Computers in Biology and Medicine 131 (2021) 104268. doi:10.1016/j.compbiomed.2021.104268.

[42] J. M. J. Valanarasu, V. A. Sindagi, I. Hacihaliloglu, V. M. Patel, KiU-Net: Overcomplete Convolutional Architectures for Biomedical Image and Volumetric Segmentation, IEEE Transactions on Medical Imaging 41 (2022) 965–976. doi:10.1109/TMI.2021.3130469.

[43] Z. Szentimrey, S. de Ribaupierre, A. Fenster, et al., Automated 3D U-Net based Segmentation of Neonatal Cerebral Ventricles from 3D Ultrasound Images, Medical Physics 49 (2022) 1034–1046. doi:10.1002/mp.15432.

[44] J. Kishimoto, A. Fenster, D. S. C. Lee, S. de Ribaupierre, Quantitative 3-D Head Ultrasound Measurements of Ventricle Volume to Determine Thresholds for Preterm Neonates Requiring Interventional Therapies Following Posthemorrhagic Ventricle Dilatation, Journal of Medical Imaging 5 (2018) 1 – 9. URL: https://doi.org/10.1117/1.JMI.5.2.026001. doi:10.1117/1.JMI.5.2.026001.

[45] S. Kazeminia, C. Baur, A. Kuijper, B. van Ginneken, N. Navab, S. Albarqouni, A. Mukhopadhyay, GANs for Medical Image Analysis, Artificial Intelligence in Medicine 109 (2020) 101938.

[46] R. J. Chen, M. Y. Lu, T. Y. Chen, D. F. Williamson, F. Mahmood, Synthetic Data in Machine Learning for Medicine and Healthcare, Nature Biomedical Engineering 5 (2021) 493–497.

[47] S. Hong, R. Marinescu, A. V. Dalca, A. K. Bonkhoff, M. Bretzner, N. S. Rost, P. Golland,

3D-stylegan: A Style-based Generative Adversarial Network for Generative Modeling of Three-dimensional Medical Images, in: Deep Generative Models, and Data Augmentation, Labelling, and Imperfections, Springer, 2021, pp. 24–34.
[48] A. Montero, E. Bonet-Carne, X. P. Burgos-Artizzu, Generative Adversarial Networks to Improve Fetal Brain Fine-grained Plane Classification, Sensors 21 (2021) 7975.
[49] H. Rivaz, D. L. Collins, Simulation of Ultrasound Images for Validation of MR to Ultrasound Registration in Neurosurgery, in: Workshop on Augmented Environments for Computer-Assisted Interventions, Springer, 2014, pp. 23–32.
[50] X. Wang, A. Bohi, M. A. Harrach, M. Dinomais, J. Lefèvre, F. Rousseau, On Early Brain Folding Patterns using Biomechanical Growth Modeling, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 146–149. doi:10.1109/EMBC.2019.8856670.
[51] D. Garcia, SIMUS: An Open-source Simulator for Medical Ultrasound Imaging. Part I: Theory & Examples, Computer Methods and Programs in Biomedicine 218 (2022) 106726.
[52] J. A. Jensen, Field: A Program for Simulating Ultrasound Systems, Medical & Biological Engineering & Computing 34 (1997) 351–353.
[53] A. Largent, J. De Asis-Cruz, K. Kapse, S. D. Barnett, J. Murnick, S. Basu, N. Andersen, S. Norman, N. Andescavage, C. Limperopoulos, Automatic Brain Segmentation in Preterm Infants with Post-hemorrhagic Hydrocephalus using 3D Bayesian U-Net, Human brain mapping 43 (2022) 1895–1916.
[54] H. de Vareilles, D. Rivière, Z.-Y. Sun, C. Fischer, F. Leroy, S. Neumane, N. Stopar, R. Eijsermans, M. Ballu, M.-L. Tataranno, et al., Shape Variability of the Central Sulcus in the Developing Brain: a Longitudinal Descriptive and Predictive Study in Preterm Infants, NeuroImage 251 (2022) 118837.
[55] A. A. Nair, T. D. Tran, A. Reiter, M. A. L. Bell, A Deep Learning based Alternative to Beamforming Ultrasound Images, in: 2018 IEEE International conference on acoustics, speech and signal processing (ICASSP), IEEE, 2018, pp. 3359–3363.
[56] S. Khan, J. Huh, J. C. Ye, Adaptive and Compressive Beamforming using Deep Learning for Medical Ultrasound, IEEE transactions on ultrasonics, ferroelectrics, and frequency control 67 (2020) 1558–1572.
[57] S. Cammarasana, P. Nicolardi, G. Patanè, Real-time Denoising of Ultrasound Images based on Deep Learning, Medical & Biological Engineering & Computing (2022) 1–16.
[58] A. Sawant, S. Kulkarni, Ultrasound Image Enhancement using Super Resolution, Biomedical Engineering Advances (2022) 100039.
[59] S. Zheng, J. Lu, H. Zhao, X. Zhu, Z. Luo, Y. Wang, Y. Fu, J. Feng, T. Xiang, P. H. Torr, et al., Rethinking Semantic Segmentation from a Sequence-to-sequence Perspective with Transformers, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 6881–6890.
[60] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, Y. Zhou, Transunet: Transformers Make Strong Encoders for Medical Image Segmentation, arXiv preprint arXiv:2102.04306 (2021).
[61] A. Hatamizadeh, Z. Xu, D. Yang, W. Li, H. Roth, D. Xu, UNetFormer: A Unified Vision Transformer Model and Pre-Training Framework for 3D Medical Image Segmentation, 2022. URL: https://arxiv.org/abs/2204.00631. doi:10.48550/ARXIV.2204.00631.

[62] Y. Gao, M. Zhou, D. Liu, Z. Yan, S. Zhang, D. N. Metaxas, A Data-scalable Transformer for Medical Image Segmentation: Architecture, Model Efficiency, and Benchmark, 2022. URL: https://arxiv.org/abs/2203.00131. doi:10.48550/ARXIV.2203.00131.