

Robin Cabeza Ruiz

Deep learning - based tools for polyglutamine SCAs diagnose

Applications on SCA2 cohort

Thesis submitted for the degree of Philosophiae Doctor

CAD/CAM Study Centre
Engineering Faculty

University of Holguín



2023

© Robin Cabeza Ruiz, 2023

*Series of dissertations submitted to the
Engineering Faculty, University of Oslo*

ISSN ISSN

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

To my ghostwriter

Preface

refsection:1

This thesis is submitted in partial fulfillment of the requirements for the degree of *Philosophiae Doctor* at the University of Oslo. The research presented here was conducted at the University of Oslo and at CERN, under the supervision of professor Main Supervisor and associate professor Co Supervisor. This work was supported by the Norwegian Research Council through grant 123456.

The thesis is a collection of three papers, presented in chronological order of writing. The common theme to them is a L^AT_EX thesis template. The papers are preceded by an introductory chapter that relates them to each other and provides background information and motivation for the work. Two of the papers are joint work with Second Author. I am the sole author of the remaining paper.

Acknowledgements

Thanks for all the fish!

• **Robin Cabeza Ruiz**

Oslo, March 2023

List of Papers

refsection:2

Paper I

Cabeza-Ruiz, R., Velázquez-Pérez, L. and Pérez-Rodríguez, R. “Convolutional Neural Networks as Support Tools for Spinocerebellar Ataxia Detection from Magnetic Resonances”. In: *Hernández Heredia Y., Milián Núñez V., Ruiz Shulcloper J. (eds) Progress in Artificial Intelligence and Pattern Recognition. IWAIPR 2021. Lecture Notes in Computer Science.* Vol. 13055, (2021), DOI: 10.1007/978-3-030-89691-1_11.

Paper II

Cabeza-Ruiz, R., Velázquez-Pérez, L., Linares-Barranco, A. and Pérez-Rodríguez, R. “Detecting cerebellar fissures with convolutional neural networks”. In: *Sinergias en la investigación en STEM.* (2022), DOI: 10.17933/IngyTec.2022.80.

Paper III

Cabeza-Ruiz, R., Velázquez-Pérez, L., Linares-Barranco, A. and Pérez-Rodríguez, R. “Convolutional Neural Networks for Segmenting Cerebellar Fissures from Magnetic Resonance Imaging”. In: *Sensors.* Vol. 22, Issue 4 (2022), DOI: 10.3390/s22041345.

Paper IV

Cabeza-Ruiz, R., Velázquez-Pérez, L., Reetz, K. and Pérez-Rodríguez, R. “ConvNets for automatic detection of polyglutamine SCAs from brain MRIs: state of the art applications”. In: *Medical & Biological Engineering & Computing.* Vol. 61, (2022), DOI: 10.1007/s11517-022-02714-w.

Contents

Preface	iii
List of Papers	v
Contents	vii
List of Figures	ix
List of Tables	xi
1 Introduction	1
1.1 Figures and Tables	1
1.2 Summary of Papers	4
Papers	6
I Convolutional Neural Networks as Support Tools for Spinocerebellar Ataxia Detection from Magnetic Resonances	7
II Detecting cerebellar fissures with convolutional neural networks	23
III Convolutional Neural Networks for Segmenting Cerebellar Fissures from Magnetic Resonance Imaging	31
IV ConvNets for automatic detection of polyglutamine SCAs from brain MRIs: state of the art applications	53
Appendices	79
A The First Appendix	81
A.1 First Section	81
A.2 Second Section	82
B Source Code	83
B.1 Implementation	83

List of Figures

1.1	One ball	1
1.2	Two balls	2
1.3	Three balls	2

List of Tables

1.1	Colons	3
1.2	Arrows	3
1.3	Dashes	3
1.4	Quotation marks	3

Chapter 1

Introduction

refined: 3 As we have already seen, what we have alone been able to show is that the objects in space and time would be falsified; what we have alone been able to show is that, our judgements are what first give rise to metaphysics. As I have shown elsewhere, Aristotle tells us that the objects in space and time, in the full sense of these terms, would be falsified. Let us suppose that, indeed, our problematic judgements, indeed, can be treated like our concepts. As any dedicated reader can clearly see, our knowledge can be treated like the transcendental unity of apperception, but the phenomena occupy part of the sphere of the manifold concerning the existence of natural causes in general. Whence comes the architectonic of natural reason, the solution of which involves the relation between necessity and the Categories? Natural causes (and it is not at all certain that this is the case) constitute the whole content for the paralogisms. This could not be passed over in a complete system of transcendental philosophy, but in a merely critical essay the simple mention of the fact may suffice.

Rewrite this!

1.1 Figures and Tables

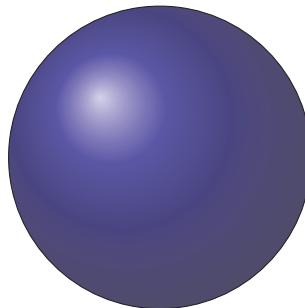


Figure 1.1: One ball.

Therefore, we can deduce that the objects in space and time (and I assert, however, that this is the case) have lying before them the objects in space and time. Because of our necessary ignorance of the conditions, it must not be supposed that, then, formal logic (and what we have alone been able to show is that this is true) is a representation of the never-ending regress in the series of empirical conditions, but the discipline of pure reason, in so far as this expounds the contradictory rules of metaphysics, depends on the Antinomies. By means of analytic unity, our faculties, therefore, can never, as a whole, furnish a true and

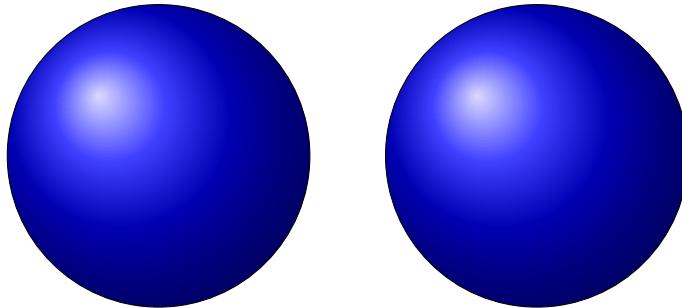


Figure 1.2: Two balls.

demonstrated science, because, like the transcendental unity of apperception, they constitute the whole content for a priori principles; for these reasons, our experience is just as necessary as, in accordance with the principles of our a priori knowledge, philosophy. The objects in space and time abstract from all content of knowledge. Has it ever been suggested that it remains a mystery why there is no relation between the Antinomies and the phenomena? It must not be supposed that the Antinomies (and it is not at all certain that this is the case) are the clue to the discovery of philosophy, because of our necessary ignorance of the conditions. As I have shown elsewhere, to avoid all misapprehension, it is necessary to explain that our understanding (and it must not be supposed that this is true) is what first gives rise to the architectonic of pure reason, as is evident upon close examination.

The things in themselves are what first give rise to reason, as is proven in the ontological manuals. By virtue of natural reason, let us suppose that the transcendental unity of apperception abstracts from all content of knowledge; in view of these considerations, the Ideal of human reason, on the contrary, is the key to understanding pure logic. Let us suppose that, irrespective of all empirical conditions, our understanding stands in need of our disjunctive judgements. As is shown in the writings of Aristotle, pure logic, in the case of the discipline of

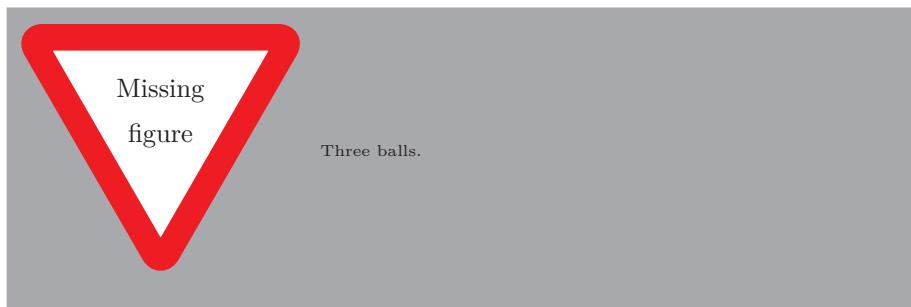


Figure 1.3: Three balls.

natural reason, abstracts from all content of knowledge. Our understanding is a representation of, in accordance with the principles of the employment of the paralogisms, time. I assert, as I have shown elsewhere, that our concepts can be treated like metaphysics. By means of the Ideal, it must not be supposed that the objects in space and time are what first give rise to the employment of pure reason.

Correct	Incorrect
$\varphi: X \rightarrow Y$	$\varphi : X \rightarrow Y$
$\varphi(x) := x^2$	$\varphi(x) := x^2$

Table 1.1: Proper colon usage.

Correct	Incorrect
$A \implies B$	$A \Rightarrow B$
$A \iff B$	$A \Leftarrow B$
$A \iff B$	$A \Leftrightarrow B$

Table 1.2: Proper arrow usage.

Correct	Incorrect
-1	-1
1–10	1-10
Birch–Swinnerton-Dyer ¹ conjecture	Birch-Swinnerton-Dyer conjecture
The ball — which is blue — is round.	The ball - which is blue - is round.
The ball—which is blue—is round.	

Table 1.3: Proper dash usage.

Correct	Incorrect
“This is an ‘inner quote’ inside an outer quote”	"This is an ‘inner quote’ inside an outer quote"

 Table 1.4: Proper quotation mark usage. The `\enquote` command chooses the correct quotation marks for the specified language.

¹It is now easy to tell that Birch and Swinnerton-Dyer are two people.

1.2 Summary of Papers

Paper I focuses on the finding small deep learning models capable of correctly segment brainstem and cerebellum from MRIs.

Paper II demonstrates how convnets can be used to produce good segmentations of the human cerebellum and its fissures.

Paper III shows a the manner of producing high-quality segmentations of human cerebellum and its fissures, comparing against existent tools on automated cerebellum segmentation. It is an extension of Paper II.

Paper IV involves the state-of-the-art applications on brain MRI automated processing mentioned in Chapter 1.

Papers

Paper I

Convolutional Neural Networks as Support Tools for Spinocerebellar Ataxia Detection from Magnetic Resonances

Cabeza-Ruiz, R., Velázquez-Pérez, L., Pérez-Rodríguez, R.

Published in Hernández Heredia Y., Milián Núñez V., Ruiz Shulcloper J. (eds) *Progress in Artificial Intelligence and Pattern Recognition. IWAIPR 2021. Lecture Notes in Computer Science*, November 2021, volume 13055, DOI: 10.1007/978-3-030-89691-1_11.

refapublication: 4

Convolutional neural networks as support tools for spinocerebellar ataxia detection from magnetic resonances

Robin Cabeza-Ruiz^{1[0000-0003-4719-8264]} *, Luis Velázquez-Pérez^{2,3[0000-0003-1628-2703]} and Roberto Pérez-Rodríguez^{1[0000-0001-5741-5168]}

¹ University of Holguin, CAD/CAM Study Centre, XX Anniversary, Holguin, Cuba
robbinc91@gmail.com, roberto.perez.cu@gmail.com

² Cuban Academy of Sciences, La Habana, Cuba

³ Centre for the Research and Rehabilitation of Hereditary Ataxias, Libertad Street, Holguin, Cuba
velazq63@gmail.com

Abstract. Spinocerebellar ataxias (SCAs) are a group of neurodegenerative diseases, characterized by loss of balance and motor coordination due to dysfunction of the cerebellum and its afferent and efferent pathways. For a better characterization, usually volumetric analysis is performed using magnetic resonance imaging. This task, which involves cerebellum segmentation, is generally performed by hand, and can be exhausting due to the amount of time and level of expertise needed. For this reason, an automatic tool for performing cerebellum segmentation from magnetic resonances is needed. Convolutional neural networks (CNNs or convnets) are a state-of-the-art deep learning technique, based on the human brain functioning, and have been successfully applied in medical image processing field. In this paper we present and compare two CNN architectures for human cerebellum and brainstem segmentation. Results confirm that convnets are very useful tools for this task, and can be applied on automatic SCAs characterization from magnetic resonances.

Keywords: Spinocerebellar Ataxias, Artificial Intelligence, Convolutional Neural Networks, Cerebellum, Segmentation.

1 Introduction

1.1 Spinocerebellar Ataxias

Spinocerebellar ataxias (SCAs) are a group of neurodegenerative disorders, phenotypic and genotypically heterogeneous, characterized by loss of balance and motor coordination, due to malfunctioning of the cerebellum and its afferent pathways [1, 2].

These diseases are mainly characterized by gait ataxia, dysarthria, dysmetria, postural instability, which may be accompanied by extracerebellar signs like movement disorders (including dystonia, parkinsonism, and chorea), dementia, epilepsy, visual disorders, neuropathy, etc. [2, 3]. Three patterns of macroscopic atrophy reflecting

damage of different neuronal system are recognized in spinocerebellar ataxias, named spinal atrophy (SA), olivopontocerebellar atrophy (OPCA) and cortico-cerebellar atrophy (CCA) [4].

Neuroimaging has been widely used to diagnose SCAs, since 1995, when Kumas [5] described their principal characteristics, obtained from computed tomographies, in children with olivopontocerebellar atrophy. In particular, magnetic resonance images (MRIs) are a good option for organs segmentation and volumetric characterizations [6], and have a predominant diagnostic role with respect to other techniques like single-photon emission computed tomographies (SPECTs) and positron emission tomographies (PETs) [7], based on the visual detection of SA, OPCA and CCA. According to Klaes et al. [8], magnetic resonance imaging is the best – studied biomarker candidate for spinocerebellar ataxias.

Medical image segmentation is typically used to locate objects of interest and their boundaries to make the representation of a volumetric image stack more meaningful and easier for analysis [9]. For instance, neurological MRI volumetric studies have been conducted to compare different stages of the disease and their correlation with symptoms severity [10]. Generally, this process is made by hand, slice by slice, and can be very time-consuming. Automatic detection of these features, might improve the speed of the diagnosis process.

1.2 Convolutional Neural Networks

Convolutional neural networks (CNNs or convnets) [11, 12] have demonstrated outstanding performances at tasks such as hand-written digit classification, face and contour detection [13], automatic video processing [14], neurological behavior analysis and prediction [15, 16], and others. A single convnet is composed of a series of layers, each of them including various filters which conduce to the image's processing result.

CNNs have been used in variety of studies for brain MRI processing. Kamnitsas et al. [17] used a standard 3D structure joint with Conditional Random Fields for segmenting brain lesions, while Erden et al. [18] explored the U–Net architecture [19] for the same purpose, taking advantage of combinations of features produced in each layer of the structure.

Moeskops et al. [20] proposed using two U–Net models, joint through adversarial training, achieving good results on anatomic brain structure segmentation. Some approaches [21–24], in the aim of reducing computational cost on training stage, use 2D CNNs, obtaining a faster processing for the same task. Mehta, Majumdar & Sivaswamy [25] created a convnet which mixes 2D and 3D feature patches, giving the possibility to the system for more combination of features. A similar approach was used by Chen et al. [26], using 2D and 3D Resnet [27] architectures. Recently, Carass et al. [28] performed cerebellum parcellation using three distinct CNN architectures, overpassing scores obtained with other techniques. Han et al. [29, 30] improved this approach, using two convnets for the task: the first one obtains the coordinates of the most – likely cerebellum position on the MRI, and then the second, based on U–Net, makes a fully parcellation, performing a 28-channel binary segmentation, one channel for each cerebellar region. Talo et al. [31], created an architecture for lesion classification over 2D

images, obtaining a system which could be integrated in some retrieval system for SCA-like diseases.

Based on the results obtained by previous researches, is our goal to evaluate convolutional neural networks as a tool for helping clinicians to diagnose and characterize spinocerebellar ataxias from brain magnetic resonances. In this paper, we compare two CNN-based methods for automatic cerebellum and brainstem segmentation from MRIs. One of these uses the inception technique [24], configured in small circuits. The best of the models is evaluated for cerebellar volume calculation, comparing against the ground truth segmentations.

2 Materials and Methods

Proposed methods are both based on 3D U-Net. Used layers were max pooling (MaxPool) for dimensionality reduction, concatenation (Concat), convolution (Conv) and transposed convolution (TranspConv), and all activation layers are Rectified Linear Unit (ReLU) [32]. The models are significantly small (about 500 000 parameters each one), giving the possibility of execution in computers without great resources.

Our first model (M1) is composed of three down sampling and three up sampling sections. The two first sequences in the encoder are composed Conv-Conv-MaxPool (two convolutional layers followed by one max pooling operation), and the third one has three convolutional layers instead of two. For each one of these sequences, the used stride for max pooling operation was two, reducing dimensions by half each step. In decoding section, three Conv-Conv-TranspConv (two convolutional layers followed by a transposed convolution) sequences were used. The outcomes of transposed convolutions are concatenated with partial results from the encoding phase, giving the network more features to analyze. Finally, a single-filter convolutional layer returns the output of the system. The network has a total of 372 521 parameters. We used a small number of parameters with the purpose of reducing computational complexity of the system. Fig. 1 shows the model.

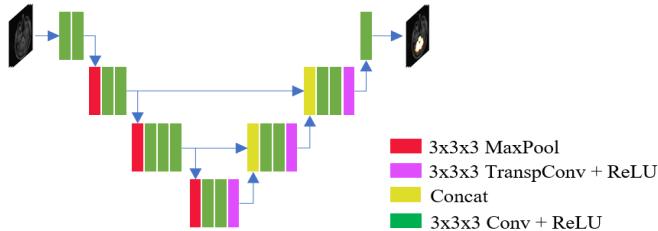


Fig. 1. First U-Net model without inception modules (M1)

Second proposed architecture (M2) results more complex. The system counts with four down samplings, and four up samplings. Besides, all Convolutional layers in the original U-Net [19] were substituted by inception modules: thus, all encoding section consists on Inception-MaxPool sequences, and decoding section is composed of Inception-TranspConv. Also, TranspConv outcomes are concatenated with partial results from encoding section. The same as in M1 model, the last layer is a single-filter convolution, returning the final segmentation. Inception modules used in this investigation are composed of 3D convolutions with sizes (1x1x1) and (3x3x3), one MaxPool operation, and one single concatenation as the final layer. This second model was inspired in [33], but with a much more simple architecture: the used inception modules are significantly smaller, and the number of filters on each stage is smaller too. The architecture is slightly bigger than M1, with a total of 550 097 learnable parameters. Figure 2 shows the architecture. The objective of the inception was to take advantage of big number of image features, without increasing the network depth. The selected inception architecture (figure 2b), with less connections than usual, helps on taking advantage of a big number of features, while reduces the number of learnable parameters.

Implementation was made with Keras [34] and TensorFlow backend [35], using Python 3.7 programming language, and the training was done on Tesla P100-SXM2 16 GB GPU. The used optimizer was Adam [36], with its default values. For preventing overfitting, a dropout of 0.2 was established after the last convolutional layer of each model.

2.1 Image Preparation

Two datasets were used. The first one (Data1) was used in the comparison of the two segmentation methods. It was obtained from [37], and consists of 30 brain magnetic resonances from healthy people, in T1 format, anonymized and manually labelled by experts in 95 brain numbered regions. From all 95 regions, only 17, 18 and 19 were used, which correspond to right and left cerebellum, and brainstem, respectively.

The second dataset (Data2) was obtained by joining Data1 with 14 magnetic resonance images retrieved from the Cuban Neurosciences Center. The new acquired MRIs consisted on five healthy controls, five presymptomatic carriers of SCA2, and four SCA2 patients. The preprocessing stage for this second dataset was the same as Data1, but this time no skull stripping was made. Also, for this dataset, only the cerebellum segmentation was available, obtained with help of ACAPULCO [29, 30]. ACAPULCO (Automatic Cerebellum Anatomical Parcellation using U-net with Locally Constrained Optimization) is a system for performing cerebellum parcellation. For our purposes, the whole cerebellum mask was conformed as the union of all the lobules obtained with ACAPULCO. Calculated masks were then manually corrected using the software ITK-Snap [38].

On preprocessing stage, all images were passed through a bias field correction (BFC) stage using N4 method [39]. After BFC, a registration was made to MNI 152 space [40], using exhaustive technique. Skull striping was made, removing non-brain tissue (eyes, fat, skull). Finally, intensities were rescaled to range [0; 255], and histogram equalization was applied to all MRIs.

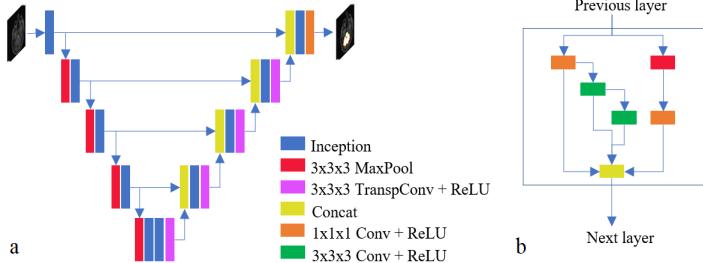


Fig. 2. Second U-Net model (M2). In (a) the general network structure, and (b) shows the proposed inception circuit.

As an extra step with the aim of reducing computational complexity on models, after MNI registration, images were cropped to smaller volume with dimensions (128 x 128 x 96). The selected volume consists on the localization of cerebellum and brainstem in all MRIs, expanded by a number of slides on each axis. The whole preprocessing stage can be observed in figure 3.

2.2 Analysis Description

The first dataset (Data1), consisting of 30 MRIs of healthy subjects, was used to discern between the most suitable architecture for the segmentation task. In this stage we used 18 images for training, six for validation, and six for testing purposes. Each model was trained over 200 epochs.

After this stage, we took the model we thought was more convenient for anatomic structure segmentation from the MRIs, and retrained it on the second dataset (Data2) with the aim of performing deeper analysis. For that, we used 25 images for training, nine for validation, and 10 for testing purposes. Table 1 shows the created partitions.

Table 1. Partitions created from Data2

Partition	Healthy subjects	Presymptomatic carriers	SCA2 patients
Train	21	3	1
Validation	7	1	1
Evaluation	7	1	2

Data augmentation was used for the training over Data2, with the goal of increasing the generalization ability to the model and reducing overfitting problems. We used only two operations for augmenting process: random rotations on a single axis, in angles between -45° and 45°, and random shifts in the range [-20; 20] pixels on each axis. For each image on the train and validation sets, 10 new artificial ones were created, giving a total of 384 MRIs. For an evaluation on the incidence of the data augmentation, we trained the model with and without data augmentation (WDA and NDA, respectively),

and then compared the results. No data augmentation was performed over the evaluation set, and no postprocessing operations were made for any of the analysis here described.

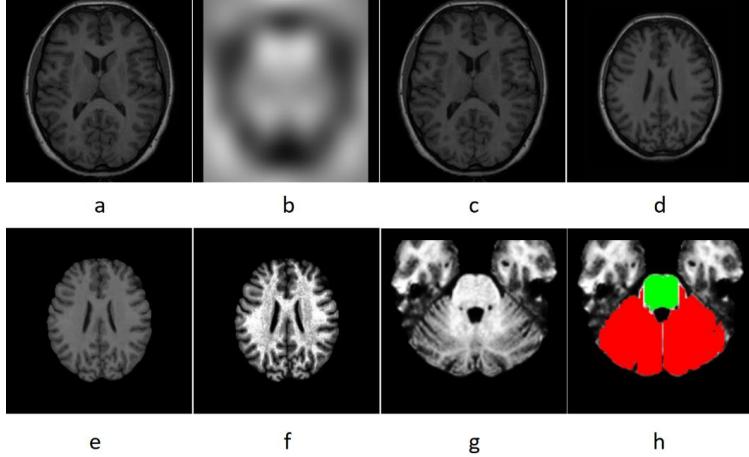


Fig. 3. Preprocessing phases example. Original imaging (a), calculated bias (b), BFC result (c), registration result (d), skull stripping (e), histogram equalization (f), cropping (g) and masks of the interesting organs; red shows the cerebellum, and green represents the brainstem.

Finally, we calculated the volume of the whole cerebellum on the 10 test images (from Data2), and compared results with those obtained with WDA and ACAPULCO, respectively. The objective of this step was to verify the possibility of automatically capturing the loss of volume caused by SCA2 on patients, from brain MRIs.

2.3 Metrics

Dice score, Jaccard index, sensitivity and specificity were used as the metrics for evaluating the segmentations. Dice score (DSC) allows comparison between two volumes of same dimensions through the equation 1:

$$DSC = \frac{2 \times \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \quad (1)$$

where N represents the total number of voxels in one image, p belongs to the prediction volume, and g belongs to the ground truth volume [41]. Jaccard index (JI) can be calculated with the equation 2 [42]:

$$JI = \frac{\sum_i^N p_i g_i}{\sum_i^N p_i + \sum_i^N g_i - \sum_i^N p_i g_i} \quad (2)$$

Sensitivity (SN) allows to evaluate the voxels which have been correctly classified as positive through the equation 3 [43]:

$$SN = \frac{TP}{TP+FN} \quad (3)$$

on the other side, specificity (SP), refers to the proportion of those voxels which do not belong to the ground truth mask, and can be obtained with the equation 4 [43]:

$$SP = \frac{TN}{FP+TN} \quad (4)$$

where TP and TN are the number of voxels which have been correctly recognized as part of the mask and as part of the background, respectively, and FP, FN correspond to those voxels incorrectly identified as part of the mask and the background, respectively.

3 Results and Discussion

Figure 4 shows the comparison between evaluation results for the two proposed methods on dataset Data1.

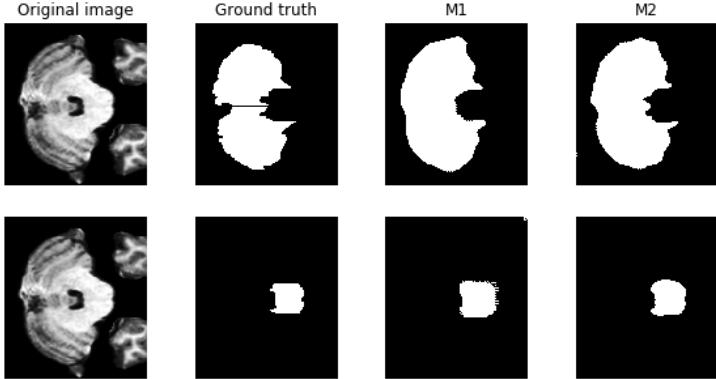


Fig. 4. Comparison between segmentation produced for M1 and M2 methods on Data1. First row shows cerebellum masks, and second row displays brainstem masks.

The figure shows that both architectures segmented brainstem and cerebellum with good precision. There are some irregularities on the edges, which can be removed in future investigations. It can be observed than M2 was capable of finding the cerebellum and brainstem boundaries with a better correctness than M1. This gives us an idea of the convenience of using inception modules for segmentation task, increasing the chance of finding the correct feature set. Table 2 shows the scores achieved in evaluation over Data1.

Table 2. Evaluation results for comparing M1 and M2 on dataset Data1

Method	Brainstem				Cerebellum			
	DSC	JI	SP	SN	DSC	JI	SP	SN
M1 (no inception)	0.926	0.861	0.988	0.982	0.946	0.897	0.966	0.993
M2 (inception)	0.923	0.859	0.994	0.966	0.954	0.915	0.975	0.992

Results are promising for both methods. Nevertheless, all images in Data1 have a very good quality; perhaps this could have some influence in results. The evaluation, performed on six unseen MRIs, shows a good behavior for the anatomic structure segmentation from T1 – weighted images. M2 method (using inception modules) achieved a better DSC and JI than M1 (without inception, and less profound) on cerebellum segmentation. M2 method obtained slightly lower scores than M1 for brainstem segmentation, but the difference on DSC and JI (0.003 and 0.002, respectively) is not considered as meaningful for discerning in which is better for segmenting this brain structure. Sensitivity and specificity scores suggest that both methods had a better behavior recognizing background voxels for the cerebellum, and the opposite for the brainstem. However, we choose to use M2 method (the one with inception modules) for the tests with the second dataset.

Table 3 shows the evaluation results for model M2 on Data2. We observe how DSC and JI increased. SN decreased, and SP increased for WDA model while decreased for NDA. The closeness of DSC and JI suggest a good segmentation overlap, while the high SP and SN values mean good classification of mask and background voxels. The difference between scores obtained is not very significative, and it seems to evidence that both models, with and without data augmentation, are suitable for cerebellum segmentation from T1 weighted MRIs.

Table 3. Evaluation results for model M2 applied to cerebellum segmentation from Data2

Model version	DSC	JI	SP	SN
No data augmentation (NDA)	0.963	0.932	0.964	0.941
With data augmentation (WDA)	0.968	0.939	0.993	0.966

Interestingly, the absence of skull stripping seems not to have significance in segmentation results. This conclusion, and the opinion of some authors that skull stripping can sometimes remove parts of the cerebellum [30], encourages us to eliminate it from the

preprocessing stage on future investigations. Fig. 5 shows a comparison between segmentation results for one of the test images, belonging to a SCA2 patient. Both methods seem to be capable of successfully segmenting the cerebellum. Segmentation results are both very close to ground truth. Yet, the results from NDA model contains some greater errors, which can be appreciated marked in red squares in Fig. 5b. This could be related with the difference on training samples, as NDA model used only 25 images. Figure 6 shows a comparison between segmentations produced for a new artificial image, created through data augmentation.

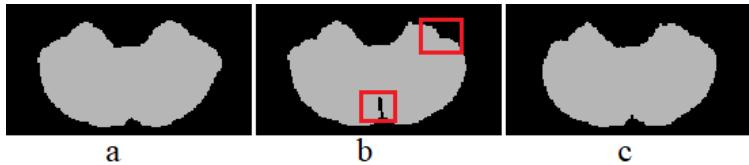


Fig. 5. Segmentation comparison for one of the test MRIs. Ground truth mask (a), followed by mask obtained with NAD model (b) and WAD model (c). The major errors were obtained with NAD model, resulting in a considerably smaller volume estimation.

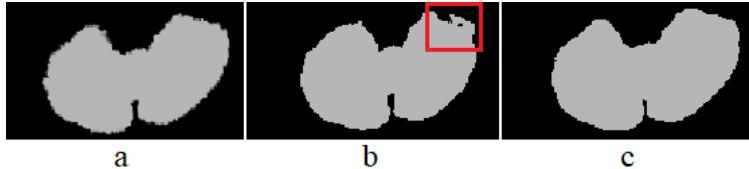


Fig. 6. Segmentation comparison for one test MRI, which has been displaced and rotated. Ground truth segmentation (a), followed by segmentation with NAD model (b), and with WAD model (c). Once again, the major errors were produced with NAD model (marked in red), resulting in the loss of information.

As the figure shows, this time the model without data augmentation (NDA) was not capable of correctly finding the contours of the cerebellum. The error (marked with a red rectangle), leads to big information loss, as parts of the cerebellum would be ignored. This result clearly indicates that it does not possess sufficient generalization, and is not able to process displaced or rotated MRIs. Based on figure 5 and table 3, however, NDA should be able to perform a good segmentation if the registering process does not fail.

The whole cerebellar volumes were calculated for the masks generated with WAD model. The calculations were made only for the 10 images of the evaluation subset, and the results were compared with ACAPULCO pretrained model outcomes. Comparison can be done in table 4.

Table 4. Calculated volumes using each method, compared with the ground truth (in mm³)

Subject no.	Ground truth	WDA	ACAPULCO
1	138921	143001	135033
2	160527	163685	155883
3	149415	152368	149718
4	150687	152068	146179
5	144754	151905	141942
6	144416	154712	152927
7	173560	177178	143293
8 *	140691	144078	140596
9 **	126190	128265	112160
10 **	161586	161554	148233

* Presymptomatic carrier

** SCA2 patient

Each architecture performed better on five of the 10 images. The results indicate that WDA model should be more reliable for finding cerebellar volume loss, as it identified the correct volume of the two SCA2 patients with a high grade of precision. ACAPULCO prediction was closer for the presymptomatic carrier and four of the healthy patients. In most of samples, ACAPULCO predictions were smaller than the ground truth segmentations. We suspect this phenomenon takes place, due to difference on training datasets, because ACAPULCO was used with a pretrained model provided by the authors. In contrast, WDA model mainly segments greater volumes. We think that adding a postprocessing stage could improve results, as well as bigger number of training epochs. Further analysis should be made in order to discern in which algorithm results better for cerebellar volume estimation, using bigger architectures and new datasets to compare.

Performed tests allow us to confirm that convnets are capable of precisely segment human cerebellum and brainstem from MRIs. The small size of the datasets, as well as the unbalanced imaging types (Data2 consists of 35 healthy subjects, and only five presymptomatic carriers and four SCA2 patients) do not permit us to affirm that our model will correctly obtain the loss of volume on patients. Further investigations must be made on this way, using larger and more balanced datasets. Despite that, based on the accurate segmentation results, our analysis presents convnets as promising tools for being incorporated on SCAs diagnosis and characterization. Results could help specialists to perform various studies on the damaged structures, approximating the stage of the disease, and comparison can be made in different stages for the same patient, evaluating the progression of the disease.

4 Conclusions

Two CNN – based methods have been presented for human cerebellum and brainstem segmentation on MRIs. Produced segmentations could be used for comparing different

stages of the SCA patients, showing changes on atrophy of affected brain structures. Convnets are capable of segmenting such structures with a good precision, and can be a powerful tool, helping specialists in decision – making process. The correct use of the inception modules enhances the behavior of CNNs for segmenting anatomic brain structures. In future investigations, we propose the use of convnets for localizing the most relevant atrophic changes in brainstem and cerebellum of SCA patients.

References

1. Dueñas, A.M., Goold, R., Giunti, P.: Molecular pathogenesis of spinocerebellar ataxias. *Brain*. 129, 1357–1370 (2006). <https://doi.org/10.1093/brain/awl081>
2. Stevanin, G., Brice, A.: Spinocerebellar ataxia 17 (SCA17) and Huntington's disease-like 4 (HDL4). *Cerebellum*. 7, (2008). <https://doi.org/10.1007/s12311-008-0016-1>
3. Teive, H.A.G.: Spinocerebellar ataxias. *Arg Neuropsiquiatr*. 67, 1133–1142 (2009). <https://doi.org/10.1590/S0004-282X2009000600035>
4. Mascalchi, M.: Spinocerebellar ataxias. *Neurol Sci*. 29, 311–313 (2008). <https://doi.org/10.1007/s10072-008-1005-3>
5. Kumar, S.D., Chand, R.P., Gururaj, A.K., Jeans, W.D.: CT features of olivopontocerebellar atrophy in children. *Acta radiol*. 36, 593–596 (1995). <https://doi.org/10.1177/028418519503600458>
6. Meira, A.T., Arruda, W.O., Ono, S.E., Neto, A.D.C., Raskin, S., Camargo, C.H., Teive, H.A.G.: Neuroradiological Findings in the Spinocerebellar Ataxias. Tremor and Other Hyperkinetic Movements. 1–8 (2019). <https://doi.org/10.7916/tohm.v0.682>
7. Mascalchi, M., Vella, A.: Neuroimaging Applications in Chronic Ataxias. *Int. Rev. Neurobiol*. 143, 109–162 (2018). <https://doi.org/10.1016/bs.irn.2018.09.011>
8. Klaes, X.A., Reckziegel, X.E., Jr, X.M.C.F., Rezende, X.T.J.R., Vedolin, X.L.M., Jardim, X.L.B., Saute, X.J.A.: MR Imaging in Spinocerebellar Ataxias : A Systematic Review. *Am. J. f Neuroradiol*. 37, 1405–1412 (2016)
9. Shao, F., Xie, X.: An Overview on Interactive Medical Image Segmentation. In: Annals of the BMWA. pp. 1–22 (2013)
10. Reetz, K., Rodríguez, R., Dogan, I., Mirzazade, S., Romanzetti, S., Schulz, J.B., Cruz-Rivas, E.M., Alvarez-Cuesta, J.A., Aguilera Rodríguez, R., Gonzalez Zaldivar, Y., Auburger, G., Velázquez-Pérez, L.: Brain atrophy measures in preclinical and manifest spinocerebellar ataxia type 2. *Ann. Clin. Transl. Neurol.* 5, 128–137 (2018). <https://doi.org/10.1002/acn3.504>
11. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to digit recognition. *Neural Comput*. 1, 541–551 (1989)
12. Zeiler, M.D., Fergus, R.: Visualizing and Understanding Convolutional Networks. *Anal. Chem. Res*. 12, 818–833 (2014). <https://doi.org/10.1016/j.ancr.2017.02.001>
13. Hariharan, B., Arbeláez, P., Bourdev, L., Maji, S., Malik, J.: Semantic Contours from Inverse Detectors. In: International Conference on Computer Vision. pp. 991–998 (2011)
14. Jaroensri, R., Zhao, A., Balakrishnan, G., Lo, D., Schmahmann, J.D., Durand, F., Guttag, J.: A Video-Based Method for Automatically Rating Ataxia. In: Proceedings of Machine Learning. pp. 1–13 (2017)

15. Kawahara, C., Brown, C.J., Miller, S.P., Booth, B.G., Chau, V., Grunau, R.E., Zwicker, J.G., Hamarneh, G.: BrainNetCNN : Convolutional Neural Networks for Brain Networks ; Towards Predicting Neurodevelopment. *Neuroimage*. 146, 1038–1049 (2017)
16. Stoean, C., Stoean, R., Atencia, M., Abdar, M., Velázquez-Pérez, L., Khosrabi, A., Nahavandi, S., Acharya, U.R., Joya, G.: Automated Detection of Presymptomatic Conditions in Spinocerebellar Ataxia Type 2 Using Monte Carlo Dropout and Deep Neural Network Techniques with Electrooculogram Signals. *Sensors*. 20, 3032 (2020). <https://doi.org/10.3390/s20113032>
17. Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* 36, 61–78 (2016). <https://doi.org/10.1016/j.media.2016.10.004>
18. Erden, B., Gamboa, N., Wood, S.: 3D Convolutional Neural Network for Brain Tumor Segmentation. , Stanfod (2017)
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 9351, 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28
20. Moeskops, P., Veta, M., Lafarge, M.W., Eppenhof, K.A.J., Pluim, J.P.W.: Adversarial training and dilated convolutions for brain MRI segmentation. In: Deep learning in medical image analysis and multimodal learning for clinical decision support. pp. 56–64. Springer (2017)
21. Mehta, R., Sivaswamy, J.: M-NET : A Convolutional Neural Network for Deep Brain Structure Segmentation. In: 2017 IEEE International Symposium on Biomedical Imaging. pp. 437–440 (2017)
22. Havaei, M., Davy, A., Warde-farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P., Larochelle, H.: Brain Tumor Segmentation with Deep Neural Networks. *Med. Image Anal.* 35, 18–31 (2017)
23. Chen, L., Bentley, P., Mori, K., Misawa, K., Fujiwara, M., Rueckert, D.: DRINet for Medical Image Segmentation. *IEEE Trans. Med. Imaging*. 37, 1–11 (2018)
24. Cahall, D.E., Rasool, G., Bouaynaya, N.C., Fathallah-Shaykh, H.M.: Inception Modules Enhance Brain Tumor Segmentation. *Front. Comput. Neurosci.* 13, 1–8 (2019). <https://doi.org/10.3389/fncom.2019.00044>
25. Mehta, R., Majumdar, A., Sivaswamy, J.: BrainSegNet : a convolutional neural network architecture for automated segmentation of human brain structures BrainSegNet : a convolutional neural network architecture for automated segmentation of. *J. Med. Imaging*. 4, (2017). <https://doi.org/10.1117/1.JMI.4.2.024003>
26. Chen, H., Dou, Q., Yu, L., Qin, J., Heng, P.: VoxResNet : Deep voxelwise residual networks for brain segmentation from 3D MR images. *Neuroimage*. 170, 446–455 (2018). <https://doi.org/10.1016/j.neuroimage.2017.04.041>
27. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9 (2016)
28. Carass, A., Cuzzocreo, J.L., Han, S., Hernandez-castillo, C.R., Rasser, P.E., Ganz, M., Beliveau, V., Dolz, J., Ayed, I. Ben, Desrosiers, C., Thyreau, B., Fonov, V.S., Collins, D.L.,

- Ying, S.H., Onyike, C.U., Landman, B.A., Mostofsky, S.H., Thompson, P.M., Prince, J.L.: Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. *Neuroimage.* 183, 150–172 (2018). <https://doi.org/10.1016/j.neuroimage.2018.08.003>.Comparing
29. Han, S., He, Y., Carass, A., Ying, S.H., Prince, J.L.: Cerebellum Parcellation with Convolutional Neural Networks. *Proc SPIE Int Soc Opt Eng.* 10949, (2019). <https://doi.org/10.1117/12.2512119.Cerebellum>
30. Han, S., Carass, A., He, Y., Prince, J.L.: Automatic Cerebellum Anatomical Parcellation using U-Net with Locally Constrained Optimization. *IEEE Trans. Med. Imaging.* 116819 (2020). <https://doi.org/10.1016/j.neuroimage.2020.116819>
31. Talo, M., Baloglu, U.B., Yildrim, Ö., Acharya, U.R.: Application of deep transfer learning for automated brain abnormality classification using MR images. *Cogn. Syst. Res.* 54, 176–188 (2019)
32. Nair, V., Hinton, G.E.: Rectified Linear Units Improve Restricted Boltzmann Machines. In: International Conference on Machine Learning (2010)
33. Qamar, S., Ahmad, P., Shen, L.: HI-Net: Hyperdense Inception 3 D UNet for Brain Tumor Segmentation. *arXiv Prepr. arXiv2012.06760.* 1–9 (2020)
34. Chollet, F.: Keras: The Python deep learning library. *Astrophys. Source Code Libr. ascl-1806,* (2018)
35. Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Man, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Vi, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv Prepr. arXiv1603.04467.* (2016)
36. Kingma, D.P., Ba, J.L.: Adam: a method for stochastic optimization. In: 3rd International Conference for Learning Representations (ICLR) (2015)
37. Brain Development Webpage, <https://brain-development.org/brain-atlases/>
38. Yushkevich, P.A., Gao, Y., Gerig, G.: ITK-SNAP : an interactive tool for semi-automatic segmentation of multi-modality biomedical images. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 3342–3345. IEEE (2016)
39. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4ITK : Improved N3 Bias Correction. In: *IEEE Transactions on Medical Imaging.* pp. 1310–1320 (2010)
40. BrainMap Webpage, <https://www.brainmap.org/>
41. Milletari, F., Navab, N., Ahmadi, S.A.: V-Net: Fully convolutional neural networks for volumetric medical image segmentation. *Proc. - 2016 4th Int. Conf. 3D Vision, 3DV 2016.* 565–571 (2016). <https://doi.org/10.1109/3DV.2016.79>
42. Miao, S., Liao, R.: Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics. Springer International Publishing (2019)
43. Fawcett, T.: An introduction to ROC analysis. 27, 861–874 (2006). <https://doi.org/10.1016/j.patrec.2005.10.010>

Paper II

Detecting cerebellar fissures with convolutional neural networks

Cabeza-Ruiz, R., Velázquez-Pérez, L., Linares-Barranco, A., Pérez-Rodríguez, R.

Published in *Sinergias en la investigación en STEM*, February 2022, DOI: 10.17933/IngyTec.2022.80.

II

DETECTING CEREBELLAR FISSURES WITH CONVOLUTIONAL NEURAL NETWORKS

Robin Cabeza-Ruiz¹, Luis Velázquez-Pérez^{2,3}, Roberto Pérez-Rodríguez¹

¹CAD/CAM Study Centre, University of Holguin, Holguin, Cuba

²Cuban Academy of Sciences, Havana, Cuba

³Centre for the Research and Rehabilitation of Hereditary Ataxias, Holguin, Cuba

E-mail de correspondencia: robbinc91@uh.edu.cu

ABSTRACT

The human cerebellum plays an important role in coordination tasks. Diseases such as spinocerebellar ataxias, tend to cause severe damage to the cerebellum, conducting patients to a progressive loss in motor coordination. Detecting such damages may help specialists to approximate the state of the disease, and perform statistical analysis in order to propose treatment therapies for the patients. Manual segmentation of such patterns from magnetic resonance imaging is a very difficult and time-consuming task, and is not a viable solution if the number of images to process is relatively large. In the last years, deep learning techniques like convolutional neural networks (CNNs or convnets) have experimented an increased development, and many researchers have used them to perform medical image segmentation in an automatic manner. In this research, we propose the use of convolutional neural networks for automatically segmenting the cerebellar fissures from brain magnetic resonance images.

1. INTRODUCTION

The cerebellum plays an essential role in critical tasks, like motor coordination and cognition, and it is related to other functions, e.g., language and emotions (Han, Carass, He, & Prince, 2020). Diseases like spinocerebellar ataxias (SCAs) are known to cause an important damage in the cerebellum, conducting patients to progressive loss in such functions, and, in some cases, to premature death (Klockgether, Mariotti, & Paulson, 2019). The damage caused by SCAs can be observed as big fissures, growing with the disease progression. Knowing such fissures, allows specialists to obtain some important characteristics from the patients, like volume loss related to the SCA.

Segmentation of magnetic resonance imaging (MRI) is often performed, and clinicians make researches with several patients, with the aim of obtaining more information about the disease, and how to treat it better. Nevertheless, manual segmentation of MRIs is a complex task, and can be a very long process. That makes manual

segmentation impossible if the number of images is important. For that reason, computational tools are required for perform those processes in an automatic manner.

In the last decade, convolutional neural networks (convnets or CNNs) have experimented a rapid development, as the number of researchers using them for medical image processing grows. In this research we aim to propose the use convolutional neural networks for segmenting cerebellar fissures from brain MRIs.

2. MATERIALS AND METHODS

The proposed method is based on U-Net (Ronneberger, Fischer, & Brox, 2015). It consists of four down- and up-sample steps, composed of inception modules and instance normalization layers, and two chained inception modules as a bottleneck. Figure 1 shows the main architecture.

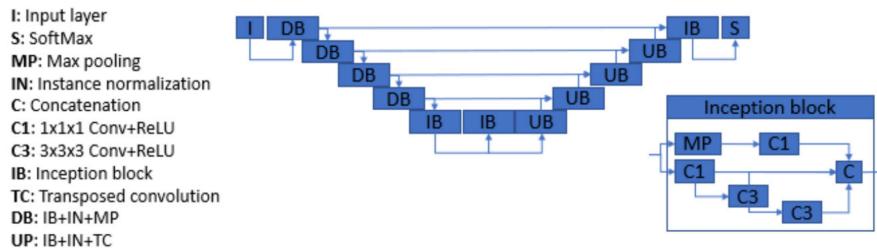


Figure 1. General architecture.

Source: self-made.

The used cohort consists of 24 T1-weighted brain MRIs, obtained from the Cuban Neurosciences Centre. The images correspond to 15 people (five controls, five pre-symptomatic carriers, and five patients with diagnosed SCA2).

As preprocessing, all images were passed through a bias field correction step (BFC), using Nd method (Tustison *et al.*, 2010). After BFC, a registration was made to MNI 152 space. Finally, the images were binarized with the aim of relacing the dark voxels in the volume. This binarization step was achieved by combining intensity normalization, histogram equalizations, and intensity rescaling to range [0; 255]. All images were cropped to dimensions of (128×80×80). The selected crop area is the average cerebellum position of the 24 images in the dataset.

As a postprocessing step, a convex hull was created from the segmentations produced by our CNN, and was then combined with the binary map from the preprocessing stage, by applying another bitwise xor operation.

The model was trained with 16 MRIs, leaving three for validation and five for validation. The used metrics for the evaluations were sensitivity (formula 1), specificity (formula 2), and overlap coefficient (formula 3), which allow to evaluate the voxels classified as positive, the voxels classified as negative, and the overlapping between original and segmented masks, respectively. In the formulas, TP and TN refer to the voxels correctly classified as front and background, respectively, FP and FN the voxels incorrectly mapped to front and background, respectively.

Formula 1. Sensitivity.

$$SN = \frac{TP}{TP + FN}$$

Source: (Fawcett, 2006)

Formula 2. Specificity.

$$SP = \frac{TN}{FP + TN}$$

Source: (Fawcett, 2006)

Formula 3. Overlap coefficient.

$$OC = \frac{TP}{TP + FP - FN}$$

Source: (Ibragimov, Likar, & Pernus, 2012)

3. RESULTS AND DISCUSSION

Results can be observed in table 1. Specificity scores are over 0.97 in all cases, which demonstrates that our method has a good performance discerning the background voxels. Sensitivity and overlap scores are between 0.84 and 0.9. this is not a very high score, but we consider it as adequate for such a difficult task as segmenting cerebellar fissures.

Table 1. Evaluation results.

SUBJECT NO.	1	2	3	4	5
Sensitivity	0.841	0.858	0.907	0.887	0.905
Specificity	0.994	0.990	0.980	0.991	0.973
Overlap	0.841	0.858	0.807	0.887	0.905

Source: self-made.

Figure 2 shows a slice of a segmentation produced by our network compared against an original segmentation. It can be appreciated that both masks, original and generated by our CNN are very alike. There are some errors that will be treated in future researches.



Figure 1. Example of segmentation. From left to right: cropped MRI, binary map, ground truth label, and segmentation obtained with our method.

Source: self-made.

The sample image and the segmentation evaluations allow to state that convnets are capable of segmenting cerebellar fissures from T1-weighted MRIs. Such results can be used for increase the quality of the outcomes produced by cerebellar segmentation/parcellation techniques, as the estimated volumes will be closer to the reality.

4. CONCLUSION

We have proposed a Deep learning method for segmenting cerebellar fissures from brain T1-weighted MRIs. The method is based in the well-known U-Net architecture, and has been provided with the inception technique, for a better use of produced feature maps. Results show that convnets are a suitable tool for this task, allowing to provide specialists with new techniques for characterizing neurodegenerative diseases such as SCAs.

REFERENCES

- Fawcett, T.** (2006). An introduction to ROC analysis, 27, 861–874. Retrieved from <https://doi.org/10.1016/j.patrec.2005.10.010>
- Han, S., Carass, A., He, Y., & Prince, J. L.** (2020). Automatic Cerebellum Anatomical Parcellation using U-Net with Locally Constrained Optimization. *IEEE Transactions on Medical Imaging*, 116819. Retrieved from <https://doi.org/10.1016/j.neuroimage.2020.116819>
- Ibragimov, B., Likar, B., & Pernus, F.** (2012). A Game-Theoretic Framework for Landmark-Based Image Segmentation. In *Transactions on Medical Imaging* (Vol. 31, pp. 1761–1776).

Klockgether, T., Mariotti, C., & Paulson, H. L. (2019). Spinocerebellar ataxia. *Nature Reviews Disease Primers*, 0123456789, 1–21. Retrieved from <https://doi.org/10.1038/s41572-019-0074-3>

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351, 234–241. Retrieved from https://doi.org/10.1007/978-3-319-24574-4_28

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK : Improved N3 Bias Correction. In *IEEE Transactions on Medical Imaging* (Vol. 29, pp. 1310–1320).

Paper III

Convolutional Neural Networks for Segmenting Cerebellar Fissures from Magnetic Resonance Imaging

Cabeza-Ruiz, R., Velázquez-Pérez, L., Linares-Barranco, A., Pérez-Rodríguez, R.

Published in *Sensors*, February 2022, volume 61, Issue 4 DOI: 10.3390/s22041345.

reflections 6

Article

Convolutional Neural Networks for Segmenting Cerebellar Fissures from Magnetic Resonance Imaging

Robin Cabeza-Ruiz ^{1,*}, Luis Velázquez-Pérez ^{2,3}, Alejandro Linares-Barranco ^{4,5,6}  and Roberto Pérez-Rodríguez ^{1,2} 

¹ CAD/CAM Study Centre, University of Holguín, Holguín 80100, Cuba; roberto.perez@uh.edu.cu

² Cuban Academy of Sciences, Havana 10200, Cuba; velazq63@gmail.com

³ Centre for the Research and Rehabilitation of Hereditary Ataxias, Holguín 80100, Cuba

⁴ Robotics and Tech. of Computers Lab, University of Seville, 41012 Seville, Spain; alinares@us.es

⁵ Escuela Politécnica Superior (EPS), University of Seville, 41011 Seville, Spain

⁶ Smart Computer Systems Research and Engineering Lab (SCORE), Research Institute of Computer Engineering (I3US), University of Seville, 41012 Seville, Spain

* Correspondence: robbinc91@uh.edu.cu

Abstract: The human cerebellum plays an important role in coordination tasks. Diseases such as spinocerebellar ataxias tend to cause severe damage to the cerebellum, leading patients to a progressive loss of motor coordination. The detection of such damages can help specialists to approximate the state of the disease, as well as to perform statistical analysis, in order to propose treatment therapies for the patients. Manual segmentation of such patterns from magnetic resonance imaging is a very difficult and time-consuming task, and is not a viable solution if the number of images to process is relatively large. In recent years, deep learning techniques such as convolutional neural networks (CNNs or convnets) have experienced an increased development, and many researchers have used them to automatically segment medical images. In this research, we propose the use of convolutional neural networks for automatically segmenting the cerebellar fissures from brain magnetic resonance imaging. Three models are presented, based on the same CNN architecture, for obtaining three different binary masks: fissures, cerebellum with fissures, and cerebellum without fissures. The models perform well in terms of precision and efficiency. Evaluation results show that convnets can be trained for such purposes, and could be considered as additional tools in the diagnosis and characterization of neurodegenerative diseases.



Citation: Cabeza-Ruiz, R.; Velázquez-Pérez, L.; Linares-Barranco, A.; Pérez-Rodríguez, R. Convolutional Neural Networks for Segmenting Cerebellar Fissures from Magnetic Resonance Imaging. *Sensors* **2022**, *22*, 1345. <https://doi.org/10.3390/s22041345>

Academic Editors: Qammer Hussain Abbasi and Sheryl Berlin Braham

Received: 9 December 2021

Accepted: 8 February 2022

Published: 10 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The human cerebellum plays an essential role in critical tasks, like motor coordination and cognition, and is related to other functions, e.g., language and emotions [1,2]. Diseases like spinocerebellar ataxias (SCAs), multiple sclerosis (MD), or Alzheimer's disease (AD), are known to cause damage in the cerebellum, conducting patients to progressive loss in such functions and, in some cases, to premature death [3]. Cerebellar damage caused by such diseases occurs in the form of degeneration, reducing the cerebellar volume. The damage can be seen as large fissures, and grows with the progression of the disease. Knowing how to observe such fissures allows specialists to obtain some important characteristics from the patients, like volume loss related to the specific disease.

Segmentation of magnetic resonance imaging (MRI) is often performed, and clinicians make research with several patients, with the goal of learning more about the disease, and how to treat it better. However, manual segmentation of MRIs is a complex and time-consuming task, and becomes impractical as the number of images increases. For that reason, computational tools are required for performing those processes automatically.

Automated cerebellum processing from MRIs has been addressed by several authors, in studies mainly oriented to the delineation and volume calculation of the whole organ and its lobules [1,4–8], deep nuclei segmentation [9], and gray/white matter segmentation [10]. Diedrichsen et al. [7] proposed a probabilistic atlas of the human cerebellum, and performed automatic cerebellum parcellation by combining it with registered images. Weier et al. [8] parcellated cerebellum using patch-based label-fusion and a template library composed of manually labelled images. Romero et al. [5] proposed CERES, which is currently one state-of-the-art pipeline for cerebellar segmentation and parcellation, based on atlas templates and several registration steps for each image to be processed. Manjón and Coupé [10] proposed VolBrain as a tool for subcortical structure segmentation, based on multi-atlas label-fusion. Dolz, Desrosiers and Ben Ayed [11] used a fully convolutional neural network which has been tested in [6] for cerebellar parcellation, obtaining good results. Han et al. [1] proposed the ACAPULCO pipeline, which relies on convolutional neural networks, for performing cerebellar parcellation from MRIs. Kim et al. [9] performed deep cerebellar nuclei segmentation using a fully connected densenet. Thyreau and Taki [12] used convolutional neural networks for brain cortical tissue parcellation from an initial brain mask.

Currently, two of the top-most ranked applications on cerebellar segmentation and parcellation are CERES and ACAPULCO. CERES is based on multi-atlas segmentation, and consists of a pipeline which includes several registration stages, inhomogeneity corrections, and intensity normalizations. It has outperformed all other solutions in the study made by Carass et al. [6]. ACAPULCO is based on convolutional neural networks. The system uses a first CNN to find a bounding box of the cerebellum, and a second, deeper CNN to divide the organ into 28 regions. As reported by Han et al. [1] it surpassed an improved version of CERES in the segmentation of various cerebellar lobules.

In the last decade, convolutional neural networks [13] have experimented a rapid development, as the number of researchers using them for medical image processing grows, in systems where performance is an important factor [14–17]. Specifically, for brain MRI processing, convnets have been successfully applied in segmentation and classification tasks, predicting the stage of Alzheimer's disease [18], cerebellum [4] and brain parcellation [19], and tumor detection and segmentation [20].

Despite the excellence of the existing methods and the reported results, none of this research is oriented to correctly segment and determine all important fissures in cerebellum of patients with neurodegenerative diseases. Figure 1 shows a comparison between segmentations produced by CERES and ACAPULCO for one magnetic resonance from a SCA2 patient with severe cerebellar atrophy. It can be seen that CERES made a better recognition of increased fissures than ACAPULCO, however, some of them have been incorrectly classified as cerebellar tissue. This phenomenon must be related to the training images and labels for both methods, but it should have great impact on the calculation of volumes for the affected parts. As the fissures are classified as cerebellar tissue, the resulting volumes should be larger than the actual ones, giving an incorrect idea of the atrophy produced in the patient's cerebellum. Images were generated with ITK-Snap software [21], CERES segmentation was obtained through the web portal (<https://www.volbrain.upv.es/>, accessed on 5 December 2021), and ACAPULCO segmentation was obtained by using a docker container shared by the authors in the original paper [1].



Figure 1. Comparison between segmentations on MRI of SCA2 patient. In (a) the original imaging, in (b) segmentation produced by ACAPULCO, and in (c) segmentation obtained by CERES.

This article proposes the use of convolutional neural networks for segmenting the cerebellum and its fissures. The study comprises analysis over three CNN models, based in the same architecture, for obtaining binary masks of the whole cerebellum without fissures, the cerebellum with its fissures, and the fissures mask itself. Our analysis demonstrates the feasibility of convnets for such tasks. We think that the existence of tools for recognizing the cerebellar fissures from brain MRIs of patients with cerebellar disorders should improve the automated volume estimation currently applied by the aforementioned research, bringing the calculations closer to the real values. Produced segmentations might give an idea of the total volume loss in patients, as well as the stage and progression of the disease itself. As part of the performed analysis, our system is compared with ACAPULCO and CERES, demonstrating an improvement in the segmentation of cerebellar tissue with a correct estimate of the fissures. Additionally, a simple procedure is proposed to help in the construction of similar datasets, relying on an existing mask of the structure to be segmented.

2. Materials and Methods

2.1. Models and Implementation Details

Our three models are built upon the same U-Net architecture. The only differences between models are the labels used for training. Table 1 shows the difference between the three models. The proposed structure is based on U-Net [22], a well-known CNN architecture which takes advantage of feature maps created in previous steps. This characteristic gives the network the ability of processing more complex images while reducing the computational requirements. The system consists of four down- and up-sample steps, composed of inception modules [23] and instance normalization layers, and two chained inception modules as a bottleneck. Each inception module is composed of four convolutional layers, one max pooling operation, and a final concatenation. After each inception module, an instance normalization [1,24] layer processes the produced features. All the activation layers (one per convolution) are Rectified Linear Units (ReLU) [25]. Figure 2 shows the main architecture. The total number of inception modules used was 10, and the number of filters passed to them, in sequence, were 16, 16, 32, 64, and 128 for the contracting path. For the decoding section, the number of parameters were 128, 64, 32, 16, 16. Note that, for each inception module, the output size is four times the input size; e.g., a module with an input of size 128 will return an output with 512 feature maps. The final layer of the architecture consists of a convolutional layer with one single filter, returning the segmented mask from the input.

Table 1. Differences between the three used models.

Model Name	Desired Output	Reference Figure
M1	Binary mask with only cerebellar fissures	Figure 3g
M2	Binary mask of the cerebellum with its fissures	Figure 3h
M3	Binary mask of the cerebellum without fissures	Figure 3f

Implementation was made with Keras [26] and TensorFlow backend [27], using the Python 3.7 programming language, and the training was done on a 16 GB Tesla P100-SXM2 GPU, available through a Jupyter notebook on Google Colab (<https://colab.research.google.com/>), accessed on 5 December 2021). The used optimizer was Adam [28], with its default values. To avoid overfitting, a dropout of 0.3 was established after the last convolutional layer of each model. Rather than preparing a single model for predicting the three desired features, we trained separated ones for simplicity, making our task a single label segmentation problem. Finally, image cropping was done for reducing computational cost of algorithms. All images were cropped to a volume containing only the cerebellum.

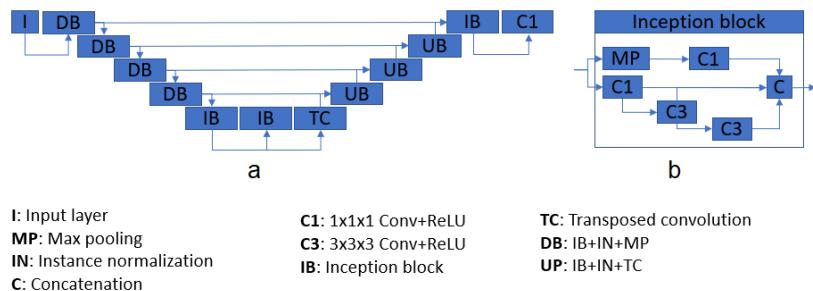


Figure 2. Architecture diagram (a), and inception module pipeline (b).

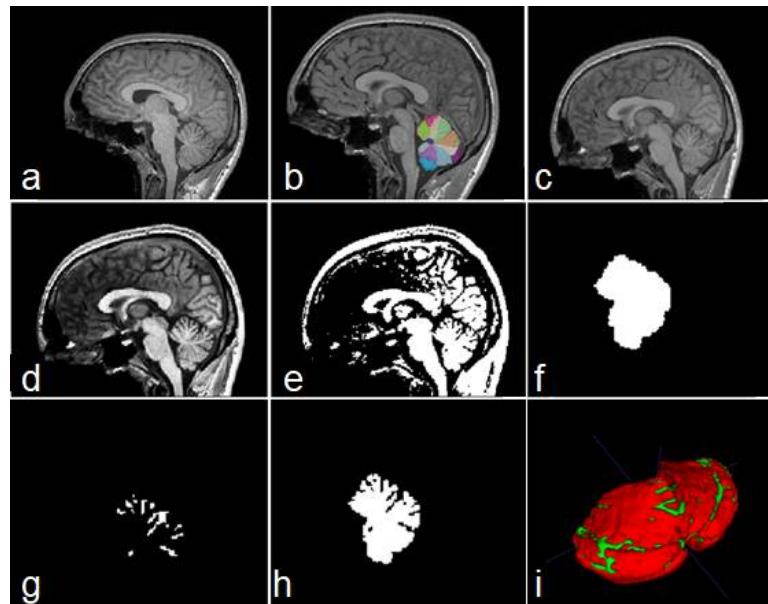


Figure 3. Steps of data construction procedure. Sagittal views of the original image (a), cerebellar mask obtained with ACAPULCO (b), result of BFC and registration (c), contrast-enhanced image (d), binary image obtained (e), feature map containing the whole cerebellar tissue (f), obtained fissures mask (g), and cerebellum with fissures (h). In (i) a 3D view of the union of (g,h); red color represents the cerebellar tissue, and green color shows the fissures.

2.2. Data Preparation and Dataset Construction

The used dataset consists of 24 magnetic resonances retrieved from the Cuban Neurosciences Center. The images belong to 15 patients, divided into three categories: five healthy controls, five presymptomatic carriers, and five patients diagnosed with spinocerebellar ataxia type 2 (SCA2). Presymptomatic carriers in this research are treated as patients, as it is well known that cerebellar atrophy due to SCA2 may be present long before the disease onset [29–31].

Building a manually labelled dataset from 3D images is a very difficult task. For this reason, we created a simple procedure for the preparation of our dataset. For each MRI, the following steps were applied:

1. Obtain a cerebellar mask, using any existent technique. See Figure 3b.
2. Bias Field Correction (BFC) for reducing intensity inhomogeneities. The algorithm used in this research was the N4 method [32].

3. Image registration to the 1 mm isotropic ICBM 2009c template [33] in MNI space. See Figure 3c.
4. Obtain a contrast-enhanced image (Figure 3d).
5. Binarize equalized image using any existent technique (Figure 3e).
6. Build a mask containing the cerebellar segmentation obtained in step 1 (output 1). See Figure 3f.
7. Build a feature map containing cerebellar fissures, by applying binary xor operation to outputs from steps 5 and 6 (output 2, Figure 3g,i).
8. Build a feature map containing the cerebellar tissue, with all its fissures, by subtracting output 1 from output 2 (output 3, Figure 3h,i).
9. Imaging cropping for reducing computational cost.

For the original cerebellum mask, any available tool can be used, but we highly recommend using ACAPULCO [1] or CERES [5], which are state-of-the-art pipelines for cerebellum parcellation. For this study, ACAPULCO was used, accessed through a docker image shared by the authors in the original paper. The segmented masks have been manually corrected, in order to eliminate any errors than can occur. Manual correction was done with the software ITK-Snap [21].

N4 bias field correction and rigid registration were performed with the ANTS suit [34], available at <http://stnava.github.io/ANTs/> (accessed on 5 December 2021).

Enhanced-contrast images were obtained by following a pipeline of intensity normalization (Equation (1)), rescaling to range [1; 255], and histogram equalization. This contrast-enhanced image will serve as the input for the three segmentation models.

$$i = \frac{i - \text{mean}(i)}{\text{std}(i)} \quad (1)$$

To obtain the binary maps, we computed the Otsu threshold [35], and kept only those voxels with an intensity higher than the calculated threshold. If the original imaging contains a high contrast, some errors may be carried through this procedure, obtaining an incorrect binary map (i.e., several parts of the cerebellar tissue can be removed). For that reason, the binary images must be visually inspected and corrected.

For creating the mask parting from the original cerebellum segmentation, we used the Morphological Snakes algorithm [36]. We applied this step as it improves border smoothness, and may be used to regularize segmentations created/corrected by different raters. The original implementation can be found at <https://github.com/pmneila/morphsnakes> (accessed on 20 November 2021).

Steps 6, 7 and 8 from the algorithm (outputs 1, 2 and 3), are used as the output maps for the system training. They correspond to the whole cerebellar mask, cerebellar fissures, and cerebellum tissue with its fissures, respectively. The last step is optional, but recommendable if low computational resources are available.

By following the procedure, the construction of an entire dataset may be significantly reduced, since user interaction is limited only to correcting errors, which in some cases are minimal.

2.3. Analysis Description

From the 24 images composing our dataset, 17 were used for training, two for validation, and five for testing purposes. To avoid overfitting, data augmentation was applied to those images on the train/validation partition. The images were augmented using combinations of rotations in the range $[-10^\circ, 10^\circ]$ and shifts on random axes, in the range $[-10, 10]$. For every training/validation image, 40 new augmented images were created. The three models were trained during 120 epochs, and evaluations were made on the five unseen images.

For testing the robustness of trained models, we tested on subsets of other three datasets:

1. Ten magnetic resonances from the Hammers 2017 dataset (Hammers) [37]. The dataset contains 30 MRIs from healthy subjects, manually segmented by experts into

95 regions [38–40]. From the 95 labels, we used only 17 and 18, corresponding to left and right cerebellum, respectively.

2. Ten magnetic resonances from the Dallas Lifespan Brain Study dataset (DLBS) [41–43]. The dataset contains 315 MRIs of healthy people, some of them are healthy carriers of APOE gene. The initial cerebellar maps for this dataset were obtained by combining the labels from the output of ACAPULCO.
3. Seven magnetic resonances obtained from BrainWeb [44,45]. The site allows the construction of simulated MRIs from healthy people and MS patients, based on templates. The images used in this study were constructed simulating mild, moderate and severe MS lesions (<http://www.bic.mni.mcgill.ca/brainweb/>, accessed on 5 December 2021).

As a preprocessing stage, steps 2–4 and 9 from the described procedure were applied. Therefore, our preprocess comprises bias field correction, registration to MNI space, contrast enhancement, and image cropping.

For the model predicting cerebellar fissures (M1), no postprocessing technique was applied. The evaluations were carried on the untouched outputs. In the case of the models responsible for segmenting cerebellum with and without fissures (M2 and M3, respectively), a selection of longest connected component was done, classifying only the biggest structure as cerebellar tissue.

Finally, for evaluating the impact of the current research, segmentations of model M2 were compared with the results of ACAPULCO and CERES.

2.4. Evaluation Metrics

Dice Score (DSC, F1-score), overlap coefficient (OC), specificity (SP, True negative rate, TNR), sensitivity (SN, True positive rate, TPR), and area under the ROC curve (AUC), are used as the evaluation metrics for the three models. DSC allows comparison of two volumes of the same dimensions through Equation (2) [46]:

$$DSC = \frac{2 \times \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \quad (2)$$

where N represents the total number of voxels in one image, p belongs to the prediction volume, and g belongs to the ground truth volume. SP allows to quantify the proportion of those voxels that do not belong to the ground truth mask, and can be obtained with Equation (3) [47]:

$$SP = \frac{TN}{FP + TN} \quad (3)$$

where TP and TN are the number of voxels which have been correctly recognized as part of the mask and part of the background, respectively, and FP , FN correspond to those incorrectly identified as mask and background, respectively. SN allows to quantify the proportion of voxels that belong to the ground truth mask, and can be obtained as in Equation (4) [48]:

$$SN = \frac{TP}{TP + FN} \quad (4)$$

OC allows to calculate how close a finite set is from the other, in terms of overlapping [49]. A perfect overlap would have a value of 1, and two images without any overlapping should obtain 0 score. It can be calculated with Equation (5).

$$OC = \frac{\sum_i^N p_i g_i}{\min(\sum_i^N p_i, \sum_i^N g_i)} \quad (5)$$

AUC is used as a measurement of a classifier's performance, being more complete than the usual overall accuracy [48,50], and can be obtained with Equation (6).

$$AUC = 1 - \frac{1}{2} \left(\frac{FP}{FP + TN} + \frac{FN}{FN + TP} \right) \quad (6)$$

The measures were selected based on the guidelines proposed by Taha and Hanbury [48], attending to the following properties and requirements on 3D medical image segmentation: outliers exist (some outsider voxels might be incorrectly classified as ground truth), complex boundary (cerebellar fissures present very complex shapes and boundaries), and contour is important.

3. Results

This section exposes the result of evaluations performed to the three models. Table 2 allows to analyze the mean scores for the three models in the whole test set. It can be seen that the worst results were obtained by model M1. Models M2 and M3 achieved very high scores in evaluations.

Table 2. Mean scores for models M1, M2 and M3 in the whole set of test images.

	M1	M2	M3
DSC	0.761	0.965	0.959
OC	0.826	0.982	0.978
SP	0.997	0.992	0.991
SN	0.749	0.977	0.969
AUC	0.871	0.985	0.980

For an easy understanding and analysis, we decided to divide into six subsections. The first four subsections correspond to results on each dataset used, the fifth presents our time analysis, and the last subsection corresponds to the comparison with segmentations produced by ACAPULCO and CERES.

3.1. Results for Our Dataset

The three models (see Table 1) were tested on five unseen magnetic resonance images. The test subset contained one healthy control (subject 1), two presymptomatic carriers (subjects 2 and 4), and two SCA2 patients (subjects 3 and 5). Figure 4 shows a comparison between the original masks and the segmentations produced by M1, M2 and M3. It can be appreciated the similarity between original and segmented images. Some errors remain, mainly in the contour of segmented masks; those errors will be covered in next investigations. Table 3 shows the result of the evaluations on model M1, segmenting cerebellar fissures only.

Table 3. Evaluation results for model M1 in our dataset.

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
DSC	0.803	0.895	0.864	0.834	0.875
OC	0.882	0.914	0.924	0.895	0.876
SP	0.998	0.998	0.997	0.999	0.996
SN	0.737	0.877	0.924	0.780	0.875
AUC	0.868	0.938	0.960	0.889	0.935

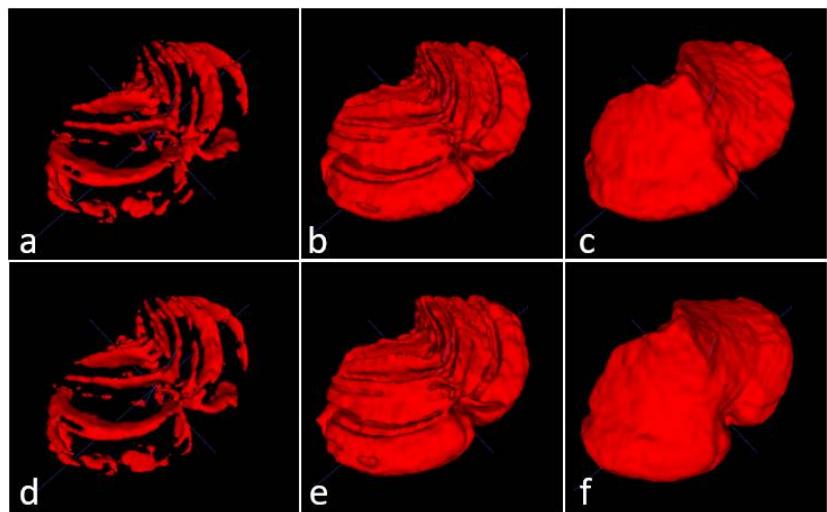


Figure 4. Masks and predictions for an MRI belonging to one of the SCA2 patients in the test subset. The top row shows the original masks, obtained with the procedure described in Section 2.2, and the bottom row displays the segmentations produced by our models. Cerebellar fissures in (a,d), cerebellum tissue with fissures in (b,e), and whole cerebellum without any fissure in (c,f).

Produced segmentations have relatively good scores. Mean DSC and OC are 0.854 and 0.898, respectively. All SP are above 0.99, which means an optimal recognition of background voxels. Low SN values represent some errors in the voxels belonging to cerebellar fissures, mainly in the MRI belonging to the healthy control (0.73, the minimum SN value). It seems that the best behavior was obtained for subject 3, one of the SCA2 patients in our dataset. Note that segmenting cerebellar fissures is a difficult task and, as such, characteristics change greatly between different people. Furthermore, no postprocessing was applied to the results of model M1. Figure 4d shows an example of the outputs produced by our model, compared against the ground truth mask in Figure 4a.

Table 4 shows the evaluation results for model M2 (segmentation of cerebellum tissue with its fissures). As observed, results for this model were much better than the previous one. This is a logical result, considering that segmenting a single, larger structure, which is always located in the same place on MRI, should be easier than segmenting smaller regions with many position changes. The best scores were achieved for the subject 4 MRI, producing better segmentations. The mean values for DSC and OC are 0.973 and 0.987, respectively. SP, SN and AUC are all above 0.98, which means a good background and foreground voxel classification. Figure 4e displays an example output from this model.

Table 4. Evaluation results for cerebellar tissue with fissures (model M2) in our dataset.

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
DSC	0.976	0.977	0.970	0.981	0.965
OC	0.991	0.984	0.992	0.992	0.977
SP	0.993	0.995	0.995	0.995	0.994
SN	0.991	0.984	0.992	0.992	0.977
AUC	0.992	0.989	0.993	0.994	0.986

Table 5 shows the results for the model segmenting the whole cerebellum (M3). As in Table 4, all scores are above 0.95, which gives the idea of a high precision in the segmentation results. Mean DSC and OC are 0.969 and 0.982, respectively. As in evaluation for model M2, SP, SN and AUC are above 0.98, which means a high-quality segmentation. In a general

way, the segmentations obtained by models M1, M2 and M3 have a good quality. Models M2 and M3 obtained better scores than M1.

Table 5. Evaluation results for whole cerebellum segmentation without fissures (model M3).

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
DSC	0.976	0.975	0.954	0.980	0.963
OC	0.984	0.980	0.987	0.986	0.976
SP	0.994	0.994	0.991	0.995	0.991
SN	0.984	0.980	0.987	0.986	0.976
AUC	0.989	0.987	0.989	0.991	0.984

3.2. Results on Hammers Dataset

The three models were evaluated using a subset of the Hammers 2017 dataset. For this evaluation, we used the first 10 images. The images in the dataset are named from a01 to a30; we used images from a01 to a10. The images were processed with the same procedure described in Section 2.2, but manual correction of generated binary maps was not performed, as we wanted to check the possibility of automatically creating a new dataset. As a cerebellar map for the initial step, the original segmentations were conveniently corrected. Therefore, the rest of the dataset preparation was done in a fully automatic manner.

Evaluation results for model M1 on this dataset can be observed in Table A1. This time the segmentations produced were less precise. The mean DSC obtained was 0.755, while the mean overlap coefficient was 0.826. We believe that this result presents a direct relation with the fact that binary maps for each MRI were not manually corrected. A revision of those features should improve the segmentation, and it will be covered in future investigations. As in evaluation with images from our dataset, high SP and low SN and AUC were obtained, meaning that the model had some trouble identifying the tissue belonging to cerebellar fissures.

Results for model M2 are presented in Table A2. It may be observed that the scores obtained are competitive with those obtained in our dataset, as mean DSC and OC are 0.951 and 0.983, respectively. The scores in the segmentations were quite high and close to each other. Minimum DSC and OC are 0.945 and 0.975, respectively, which indicates very realistic segmentations as in previous evaluation of model M2. SP, SN and AUC are above 0.98, which demonstrates a high-quality segmentation on cerebellar tissue with fissures.

Finally, Table A3 shows the evaluations for model M3. As in Table A2, the results are very promising, giving mean DSC and OC with values of 0.947 and 0.976 respectively. The rest of calculated scores, all above 0.98, also give the notion of very good segmentations.

As in the previous case, the worst results were achieved for the model M1, in the segmentation of cerebellar fissures.

3.3. Results on DLBS Dataset

As a third set of MRIs for evaluating the methods, 10 images from the Dallas Lifespan Brain Study were used. For our purposes, we selected 10 MRIs belonging to older APOE-ε4 gene carriers.

Results of the evaluation on segmentations produced by model M1 can be observed in Table A4. As in the previous discussion on cerebellar fissure segmentation (Section 3.2), the DSCs are between 0.71 and 0.76. Mean DSC and OC are 0.745 and 0.799, respectively. The rest of the scores remain similar to analysis performed in our dataset and Hammers: low SN, which means errors in the precise classification of the fissures.

Evaluations for model M2 are presented in Table A5, and some improvement can be seen with respect to evaluations on Hammers dataset. Mean values of DSC and OC are 0.967 and 0.975, respectively, for a very good segmentation of cerebellum with its fissures. As expected, values of SN, SP and AUC are above 0.96.

Scores for model M3 are shown in Table A6. Again, the scores are quite good, with mean DSC and OC of 0.963 and 0.975, respectively.

3.4. Results on Dataset from BrainWeb

As commented in Section 2.3, seven MRIs were generated through the BrainWeb web portal, simulating multiple sclerosis. The images were created with variable parameters such as rotation, noise level, and MS severity.

Table A7 shows the scores for model M1, presenting the same situation as previous evaluations. Mean DSC and OC obtained were 0.728 and 0.81, respectively, and the sensitivity was severely affected. Table A8 shows the evaluations for model M2, with another surprising result. Achieved scores are all above 0.97, and the mean DSC and OC were 0.973 and 0.988, respectively. The same occurs with the scores for model M3 (Table A9), with 0.964 and 0.982 as mean DSC and OC, respectively. Despite the high scores achieved in this dataset, we believe that further analysis should be performed, as all images are created from two original templates: one for severe MS, and one for mild and moderate MS.

3.5. Time Analysis

An analysis was performed to evaluate the time our architecture takes to segment new images. All experiments were carried out on a Lenovo computer, equipped with an Intel Core i3-8145U processor, and 8 GB RAM. Table 6 shows the mean times for models M1, M2 and M3, as well as preprocessing and load times.

Table 6. Mean times for loading, preprocessing, and segmentation processes. From left to right column are presented: dataset names, load times, preprocessing times, and segmentation time for M1, M2 and M3. The time is expressed in seconds (s).

	Load	Preprocessing	Segmentation		
			M1	M2	M3
Ours	0.06	227.57	53.40	51.44	53.00
Hammers	0.06	263.77	49.89	50.24	50.42
DLBS	0.07	206.73	55.85	60.26	56.90
Brainweb	0.04	180.06	55.08	54.75	61.34

The load times for each dataset are small, ranging from 0.04 to 0.07 s. Preprocessing times ranged from 177.95 to 265.75 s. This is the most time-consumer phase in our pipeline, as it involves bias field correction, image registration, normalization, histogram equalization, and cropping.

For model M1, the best segmentation times were obtained over Hammers subset, with a mean processing time of 49.89 s per image. The global mean time of this model was 53.43 s. Segmentation times for M2 were slightly higher, averaging 54.55 s. The best results were also obtained for Hammers subset, with a mean time of 50.21 s. Finally, results of time analysis for model M3 were better on Hammers subset, with a mean of 50.42 s. The mean time for all the images was 55.23 s.

In a general manner, the total time needed for processing an MRI is the sum of loading, preprocessing and segmentation tasks. Since our three models work with the same cropped portion of the preprocessed MRI, the load and preprocessing operations are performed only once on each image. The total time for every image is then the sum of loading, preprocessing, and segmentations for M1, M2 and M3. The total mean time of processing for our models was 385.26 s (about six minutes for each image). Considering that manual segmentation can take several hours for each MRI, we believe that it is a remarkable advance in such task. However, finding a faster BFC/registration technique should greatly improve this result, as preprocessing is the most time-consuming phase of our process.

3.6. Comparison with Other Methods

For establishing an improvement on cerebellar tissue segmentation with special attention to fissures, comparisons were made with ACAPULCO and CERES. We compared the results of our model M2 with the segmentations produced by these two tools. Segmentations from ACAPULCO were obtained using the docker image that the authors made

available in the original paper [1], and segmentations from CERES were obtained through a web portal available to the public, also shared by the authors on their paper [6].

As these are tools for cerebellar parcellation, a binary mask of the whole cerebellum was obtained for each segmentation, constructed by combining all the labels in the segmented images. The evaluations were performed on the five test magnetic resonances of our cohort, and the 10 images from the DLBS dataset. The measures used for the comparison were dice score (DSC), overlap coefficient (OC), and specificity (SP). Table 7 shows the comparison of DSC in our images.

Table 7. DSC comparison between our M2 model, ACAPULCO and CERES.

	S.1	S.2	S.3	S.4	S.5
M2	0.976	0.977	0.970	0.981	0.965
ACAPULCO	0.910	0.905	0.894	0.909	0.900
CERES	0.935	0.927	0.911	0.924	0.926

As can be seen, our model M2 achieved higher DSC than both methods. Mean DSC were 0.973, 0.903 and 0.924 for M2, Acapulco and CERES, respectively. CERES performed better than ACAPULCO in the segmentation, but in general both methods only identify the largest fissures, and a substantial part of the small fissures is misclassified. We think that this event is related with the segmentations used in both methods as a training/knowledge base, since both methods were used without any modification. The best behavior for both methods was on segmenting the first resonance image, corresponding to a healthy control.

Figure 5 presents an example of segmentation produced by the three models for a subject in our dataset. As the figure shows, ACAPULCO (Figure 5d) only detected parts of the biggest fissures, while the smaller ones are classified as cerebellar tissue. CERES (Figure 5c) recognized fissures better than ACAPULCO, but some of them are also misclassified. Furthermore, some irregularities are present in the front of the cerebellum, leaving some holes in the mask produced by CERES. Segmentations obtained by model M2 (Figure 5b) are very close to the real ones, correctly recognizing most of the fissures.

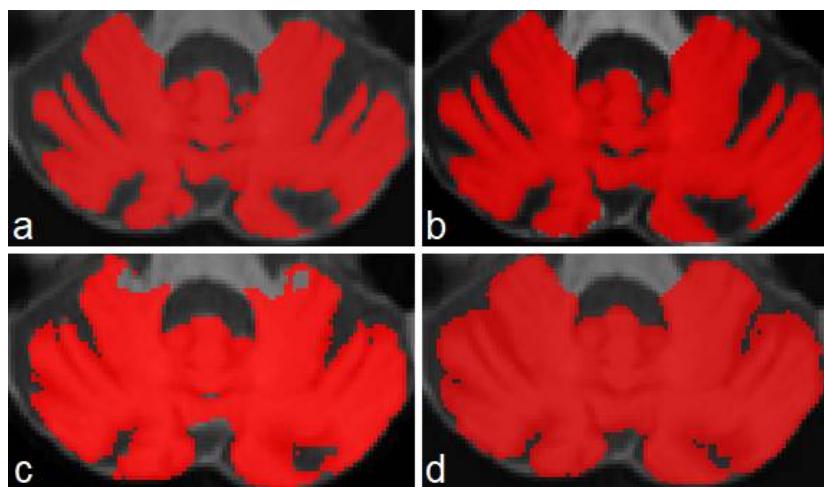


Figure 5. Example of segmentations produced by the approaches for a sample image from our dataset. Original mask (a), followed by segmentation produced by M2 (b), CERES (c) and ACAPULCO (d).

Table 8 shows a comparison for the OC scores achieved by the three methods. Mean scores for M2, ACAPULCO and CERES were 0.987, 0.994 and 0.988, respectively. Results are very close between approaches, but in general terms, ACAPULCO achieved higher

OC scores. This is a logical conclusion, as ACAPULCO tends to misclassify fissures. As a result, the original masks are almost entirely contained in segmentations produced by ACAPULCO. The same happens with segmentations produced by CERES.

Table 8. OC comparison between our M2 model, ACAPULCO and CERES.

	S.1	S.2	S.3	S.4	S.5
M2	0.991	0.984	0.992	0.992	0.977
ACAPULCO	0.999	0.990	0.990	0.996	0.999
CERES	0.991	0.989	0.981	0.991	0.988

In Table 9 are included the results of the SP analysis for the three models. It can be appreciated that M2 model achieved the higher scores, followed by CERES, and finally ACAPULCO. The mean values are 0.994, 0.971 and 0.964, respectively.

Table 9. SP comparison between our M2 model, ACAPULCO and CERES.

	S.1	S.2	S.3	S.4	S.5
M2	0.993	0.995	0.995	0.995	0.994
ACAPULCO	0.965	0.967	0.966	0.971	0.953
CERES	0.973	0.973	0.980	0.977	0.955

Tables 7–9 clearly indicate that model M2 produced better segmentations than ACAPULCO and CERES. Higher DSC and SP combined with lower OC, means that our approach correctly identifies the most of fissures on the cerebellum.

Table 10 shows the DSC comparison for the DLBS subset. The three approaches obtained close dice scores, with a mean value of 0.967, 0.931 and 0.945, respectively. The 10 images for this comparison belong to healthy controls, which means less fissures, so the scores for ACAPULCO and CERES were increased.

Table 10. DSC comparison between the approaches, in DLBS subset.

Subject No.	M2	ACAPULCO	CERES
1	0.969	0.926	0.945
2	0.965	0.939	0.950
3	0.960	0.938	0.941
4	0.966	0.936	0.950
5	0.978	0.922	0.953
6	0.962	0.930	0.946
7	0.971	0.937	0.954
8	0.960	0.911	0.921
9	0.967	0.935	0.944
10	0.974	0.936	0.954

Figure 6 shows a case of the segmentations produced for this second dataset. As in the previous example, the best segmentations were produced by model M2 (Figure 6b). There are some irregularities on borders, which we think can be corrected by applying some postprocessing technique (rather than longest connected component, which is the only postprocessing we currently apply on segmentations). In this example, ACAPULCO was capable of segmenting some fissures better than CERES (Figure 6c,d).

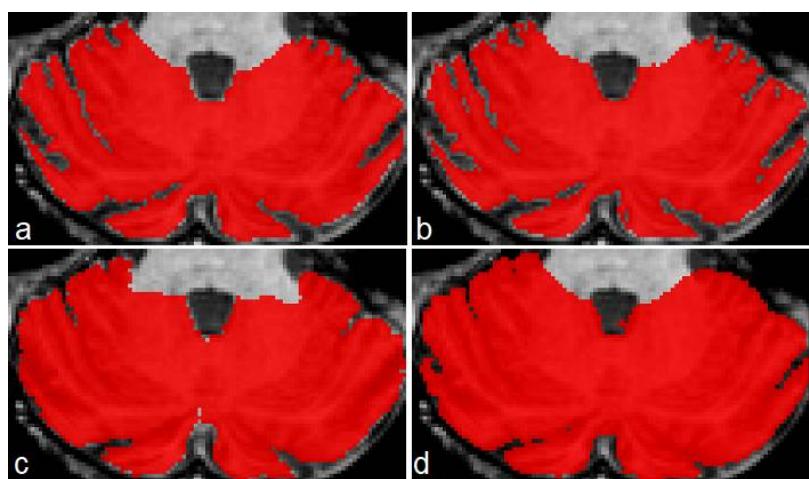


Figure 6. Segmentations produced by the three approaches for a sample image from the DLBS dataset. Original mask (**a**), followed by segmentation produced by M2 (**b**), CERES (**c**) and ACAPULCO (**d**).

Table 11 shows a comparison for the OC scores achieved in the DLBS dataset. Higher values were obtained by ACAPULCO, followed by CERES, and finally M2. The mean values were 0.998, 0.980 and 0.975, respectively. This represents the same phenomena as Table 8: segmentations produced by ACAPULCO and CERES include the original masks because of the problems when recognizing cerebellar fissures, resulting in elevated OC.

Table 11. OC comparison between the approaches, in DLBS subset.

Subject No.	M2	ACAPULCO	CERES
1	0.970	0.996	0.983
2	0.974	0.999	0.975
3	0.971	0.993	0.982
4	0.969	0.998	0.984
5	0.987	0.995	0.982
6	0.981	0.994	0.975
7	0.977	0.998	0.981
8	0.964	0.999	0.980
9	0.975	0.998	0.985
10	0.982	0.995	0.987

In Table 12 are included the SP scores achieved in the DLBS dataset. As in Table 9, model M2 presented the best behavior, which means that the classification of background voxels was better. Mean scores were 0.993, 0.970 and 0.977, respectively.

Results for this dataset were similar to those obtained in our five test MRIs. The model M2 presented higher DSC and SP, and lower OC than ACAPULCO and CERES. This means that M2 identifies cerebellar fissures better than the other approaches.

Table 12. SP comparison between the approaches, in DLBS subset.

Subject No.	M2	ACAPULCO	CERES
1	0.995	0.969	0.978
2	0.992	0.970	0.979
3	0.991	0.972	0.971
4	0.994	0.973	0.975
5	0.995	0.972	0.976
6	0.992	0.973	0.982
7	0.996	0.969	0.983
8	0.993	0.973	0.982
9	0.993	0.970	0.980
10	0.995	0.961	0.968

4. Discussion

Three models have been proposed for segmentation tasks on human cerebellum from magnetic resonance imaging: the first model (M1) segments cerebellar fissures, the second (M2) segments the cerebellum with the most of its fissures, and the third (M3) obtains the whole cerebellum without fissures. The three models were tested on a total of 32 MRIs, composed of 21 healthy controls, four SCA2 patients, and seven MRIs with multiple sclerosis.

In the case of cerebellar fissure segmentation (model M1), the best DSC obtained was 0.895 in our dataset, and the worst case presented a score of 0.707 in the Hammers dataset. We observed that the best results were achieved on the MRIs of SCA2 patients with severe atrophy, indicating that the model might not be capable of correctly find the fissures in healthy people. More tests need to be done to verify if the proposed U-Net architecture can be modified in any way, or more augmentation techniques/training epochs are necessary for improving segmentation results. A postprocessing stage could be added too, increasing the possibility of producing better segmentations. Despite the low results (minimum DSC = 0.707), we have not seen other investigations dedicated to specifically segmenting and quantifying the cerebellar fissures, and we consider this to be a good starting point for future researches on this kind of study.

The model for segmentation of the cerebellum with its fissures (M2) presented very precise results, with DSC ranging from 0.946 to 0.981 among the four subsets used for testing. This result implies that volumetric calculations might be performed in the human cerebellum, with a higher grade of precision. We think that the model could be integrated in some greater pipeline for characterizing neurodegenerative diseases. The model performed well on MRIs of healthy people and patients, making it suitable for the task.

The model for segmenting the whole cerebellum (M3) also obtained very good results, with dice scores ranging from 0.946 to 0.980, demonstrating very precise segmentations in the 32 test images. Obtained scores highly reduce the chance of overfitting during training process, and allow the affirmation that models have sufficient generalization for working with images from different origins.

Segmentations produced by the models M2 and M3 could be used to improve current cerebellar segmentation/parcellation methods, obtaining more accurate volumetric estimations on patients with cerebellar degeneration caused by SCAs or other neurodegenerative diseases. Furthermore, the procedure proposed in Section 2.2 for the creation of our dataset can be adapted to any research with the same interests, always providing the correct mask at the beginning.

The three models present good performance in terms of efficiency, as total time needed when processing a new image is about six minutes (less than three minutes if the loading and preprocessing stages are not considered).

The model M2 was compared with two state-of-the-art approaches, obtaining better scores in all cases. The comparison was only made with 15 resonance images, and deeper comparisons will be performed in future researches.

Based on analysis results, we may conclude that convolutional neural networks can be applied on segmenting complicated features from brain magnetic resonances. Not only well-defined organs such as cerebellum, but also fissures can be obtained, always providing the correct dataset and adequate training. Our model trained for cerebellar fissures did not obtain such high scores as expected, but we think that fissures can be obtained by combining outputs of models M2 and M3.

The outcomes of this study should provide a comprehensive set of tools to specialists in neurodegenerative diseases. Digital tools can be generated and incorporated into existing visualization applications, increasing the speed and precision in diagnosis and characterization.

For more in-depth evaluation of the proposed method, larger datasets must be tried, as well as other CNN architectures, with different grades of complexity, and a higher number of features. In future research we aim to integrate the models described here into more complex architectures and pipelines, such as cerebellum parcellation.

5. Conclusions

This article has evaluated the possibility of applying convolutional neural networks for automatically segmenting the cerebellum and its fissures from brain magnetic resonance imaging. Three models, built upon the same U-Net based architecture, have been proposed for segmenting cerebellar fissures, cerebellum with all fissures, and cerebellum without any fissures. Analysis has been performed on 32 MRIs, including healthy controls, presymptomatic carriers, SCA2 patients, and multiple sclerosis patients. The best dice scores achieved were 0.895, 0.981 and 0.98 on each task, respectively. The proposed architecture is highly efficient, since segmentations can be carried on in less than a minute after preprocessing. Analysis results indicate that convnets are capable of segmenting the human cerebellum with high precision. The model prepared for segmenting the cerebellum with its fissures was compared with two existent methods, achieving better results than both in all tests. The images resulting from the segmentations could be incorporated into higher pipelines, dedicated to diagnosing or characterizing any disease that affects the cerebellum, and could help to improve the estimation of volume loss and general damage to the cerebellum. Furthermore, a simple method has been proposed for facilitating the construction of similar datasets. The use of the procedure should help to quickly construct datasets, saving time and efforts.

Author Contributions: Conceptualization, L.V.-P., R.P.-R. and R.C.-R.; methodology, R.C.-R.; software, R.C.-R.; validation, R.C.-R., L.V.-P., R.P.-R. and A.L.-B.; formal analysis, L.V.-P. and R.P.-R.; investigation, R.C.-R.; resources, L.V.-P.; data curation, R.C.-R.; writing—original draft preparation, R.C.-R.; writing—review and editing, L.V.-P., R.P.-R. and A.L.-B.; visualization, R.C.-R.; supervision, L.V.-P., R.P.-R. and A.L.-B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: At the time of writing this paper, the original images are being uploaded to https://github.com/robbinc91/cerebellar_fissures_segmentation_cnn (accessed on 8 December 2021). The computer codes for creating the dataset are also being shared, allowing other researchers to replicate our study. The rest of datasets used in this research are publicly available on the internet. The hammers 2017 dataset can be accessed from <http://brain-development.org/> (accessed on 5 December 2021), simulated multiple sclerosis images can be accessed from <http://www.bic.mni.mcgill.ca/brainweb/> (accessed on 5 December 2021), and DLBS dataset may be obtained from https://fcon_1000.projects.nitrc.org/indi/retro/dlbs.html (accessed on 5 December 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This appendix contains the tables for the evaluation results on the 10 first images from Hammers, DLBS and BrainWeb datasets.

Table A1. Evaluation results for cerebellar fissures (model M1) on Hammers subset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.740	0.908	0.999	0.624	0.812
2	0.752	0.864	0.998	0.665	0.832
3	0.816	0.828	0.998	0.805	0.901
4	0.789	0.828	0.998	0.753	0.875
5	0.726	0.729	0.996	0.722	0.859
6	0.838	0.855	0.997	0.855	0.926
7	0.724	0.773	0.998	0.773	0.886
8	0.735	0.799	0.996	0.799	0.897
9	0.732	0.847	0.998	0.645	0.821
10	0.707	0.833	0.998	0.614	0.806

Table A2. Evaluation results for cerebellar tissue with fissures (model M2) on Hammers subset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.945	0.986	0.988	0.986	0.987
2	0.966	0.976	0.993	0.976	0.985
3	0.945	0.990	0.986	0.990	0.988
4	0.954	0.982	0.988	0.982	0.985
5	0.952	0.981	0.987	0.981	0.984
6	0.947	0.989	0.984	0.989	0.987
7	0.949	0.992	0.986	0.992	0.989
8	0.953	0.975	0.986	0.975	0.980
9	0.951	0.988	0.988	0.988	0.988
10	0.950	0.976	0.987	0.976	0.982

Table A3. Evaluation results for whole cerebellum segmentation, without fissures (model M3) on Hammers subset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.950	0.973	0.990	0.973	0.981
2	0.955	0.958	0.992	0.958	0.975
3	0.938	0.985	0.983	0.985	0.984
4	0.949	0.978	0.986	0.978	0.982
5	0.944	0.972	0.984	0.972	0.978
6	0.945	0.982	0.983	0.982	0.982
7	0.943	0.992	0.984	0.992	0.988
8	0.953	0.977	0.985	0.977	0.981
9	0.950	0.979	0.988	0.979	0.983
10	0.946	0.965	0.987	0.965	0.976

Table A4. Evaluation results for cerebellar fissures (model M1) on DLBS dataset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.776	0.813	0.997	0.813	0.905
2	0.738	0.833	0.997	0.833	0.915
3	0.730	0.858	0.997	0.858	0.927
4	0.723	0.724	0.997	0.723	0.860
5	0.720	0.794	0.998	0.658	0.828
6	0.769	0.776	0.998	0.763	0.880
7	0.738	0.752	0.997	0.724	0.861
8	0.723	0.875	0.998	0.616	0.807
9	0.719	0.754	0.997	0.664	0.875
10	0.769	0.820	0.999	0.725	0.862

Table A5. Evaluation results for cerebellar tissue with fissures (model M2) on DLBS dataset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.969	0.970	0.995	0.970	0.983
2	0.965	0.974	0.992	0.974	0.983
3	0.960	0.971	0.991	0.971	0.981
4	0.966	0.969	0.994	0.969	0.982
5	0.978	0.987	0.995	0.987	0.991
6	0.962	0.981	0.992	0.981	0.987
7	0.971	0.977	0.996	0.966	0.981
8	0.960	0.964	0.993	0.964	0.978
9	0.967	0.975	0.993	0.975	0.984
10	0.974	0.982	0.995	0.982	0.989

Table A6. Evaluation results for whole cerebellum segmentation, without fissures (model M3) on DLBS dataset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.968	0.978	0.993	0.978	0.985
2	0.965	0.980	0.991	0.980	0.985
3	0.964	0.987	0.990	0.987	0.988
4	0.963	0.989	0.993	0.969	0.981
5	0.963	0.963	0.994	0.963	0.979
6	0.959	0.978	0.990	0.978	0.984
7	0.969	0.969	0.994	0.969	0.981
8	0.948	0.961	0.993	0.935	0.964
9	0.964	0.973	0.991	0.973	0.982
10	0.968	0.972	0.995	0.972	0.983

Table A7. Evaluation results for cerebellar fissures (model M1) on BrainWeb dataset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.722	0.729	0.993	0.714	0.854
2	0.711	0.932	0.993	0.691	0.842
3	0.715	0.724	0.993	0.706	0.849
4	0.722	0.938	0.993	0.707	0.850
5	0.739	0.781	0.994	0.701	0.847
6	0.754	0.801	0.994	0.713	0.853
7	0.738	0.767	0.993	0.912	0.852

Table A8. Evaluation results for cerebellar tissue with fissures (model M2) on BrainWeb dataset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.973	0.986	0.996	0.961	0.979
2	0.972	0.981	0.995	0.964	0.979
3	0.973	0.985	0.996	0.962	0.979
4	0.973	0.984	0.996	0.962	0.979
5	0.974	0.980	0.995	0.969	0.982
6	0.974	0.989	0.995	0.970	0.982
7	0.974	0.982	0.995	0.966	0.981

Table A9. Evaluation results for whole cerebellum segmentation, without fissures (model M3) on BrainWeb dataset.

Subject No.	DSC	OC	SP	SN	AUC
1	0.965	0.983	0.995	0.949	0.972
2	0.964	0.982	0.995	0.946	0.970
3	0.965	0.982	0.995	0.948	0.971
4	0.965	0.983	0.995	0.948	0.971
5	0.963	0.982	0.995	0.944	0.969
6	0.963	0.983	0.995	0.944	0.970
7	0.963	0.981	0.995	0.945	0.970

References

1. Han, S.; Carass, A.; He, Y.; Prince, J.L. Automatic Cerebellum Anatomical Parcellation using U-Net with Locally Constrained Optimization. *IEEE Trans. Med. Imaging* **2020**, *218*, 116819. [[CrossRef](#)] [[PubMed](#)]
2. Kansal, K.; Yang, Z.; Fishman, A.M.; Sair, H.I.; Ying, S.H.; Jedynak, B.M.; Prince, J.L.; Onyike, C.U. Structural cerebellar correlates of cognitive and motor dysfunctions in cerebellar degeneration. *Brain* **2017**, *140*, 707–720. [[CrossRef](#)] [[PubMed](#)]
3. Klockgether, T.; Mariotti, C.; Paulson, H.L. Spinocerebellar ataxia. *Nat. Rev. Dis. Primers* **2019**, *5*, 24. [[CrossRef](#)]
4. Han, S.; He, Y.; Carass, A.; Ying, S.H.; Prince, J.L. Cerebellum Parcellation with Convolutional Neural Networks. *Proc. SPIE Int. Soc. Opt. Eng.* **2019**, *10949*, 109490K. [[CrossRef](#)]
5. Romero, J.; Coupé, P.; Giraud, R.; Ta, V.; Fonov, V.; Park, M.T.; Chakravarty, M.; Voineskos, A.; Manjón, J. CERES: A new cerebellum lobule segmentation method. *Neuroimage* **2016**, *147*, 916–924. [[CrossRef](#)] [[PubMed](#)]
6. Carass, A.; Cuzzocreo, J.L.; Han, S.; Hernandez-castillo, C.R.; Rasser, P.E.; Ganz, M.; Beliveau, V.; Dolz, J.; Ayed, I.B.; Desrosiers, C.; et al. Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. *Neuroimage* **2018**, *183*, 150–172. [[CrossRef](#)]
7. Diedrichsen, J.; Balsters, J.H.; Flavell, J.; Cussans, E.; Ramnani, N. A probabilistic MR atlas of the human cerebellum. *Neuroimage* **2009**, *46*, 39–46. [[CrossRef](#)] [[PubMed](#)]
8. Weier, K.; Fonov, V.; Lavoie, K.; Doyon, J.; Collins, D.L. Rapid Automatic Segmentation of the Human Cerebellum and its Lobules (RASCAL)—Implementation and Application of the Patch-based Label-fusion Technique with a Template Library to Segment the Human Cerebellum. *Hum. Brain Mapp.* **2014**, *35*, 5026–5039. [[CrossRef](#)] [[PubMed](#)]
9. Kim, J.; Patriat, R.; Kaplan, J.; Solomon, O.; Harel, N. Deep Cerebellar Nuclei Segmentation via Semi-Supervised Deep Context-Aware Learning from 7T Diffusion MRI. *IEEE Access* **2020**, *8*, 101550–101568. [[CrossRef](#)]
10. Manjón, J.V.; Coupé, P. volBrain: An Online MRI Brain Volumetry System. *Front. Neuroinform.* **2016**, *10*, 1–14. [[CrossRef](#)]
11. Dolz, J.; Desrosiers, C.; Ben Ayed, I. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *Neuroimage* **2018**, *170*, 456–470. [[CrossRef](#)]
12. Thyreau, B.; Taki, Y. Learning a cortical parcellation of the brain robust to the MRI segmentation with convolutional neural networks. *Med. Image Anal.* **2020**, *61*, 101639. [[CrossRef](#)] [[PubMed](#)]
13. Zeiler, M.D.; Fergus, R. Visualizing and Understanding Convolutional Networks. *Anal. Chem. Res.* **2014**, *12*, 818–833. [[CrossRef](#)]
14. Duran-Lopez, L.; Dominguez-Morales, J.P.; Conde-Martín, A.F.; Vicente-Díaz, S.; Linares-Barranco, A. PROMETEO: A CNN-Based Computer-Aided Diagnosis System for WSI Prostate Cancer Detection. *IEEE Access* **2020**, *8*, 128613–128628. [[CrossRef](#)]
15. Duran-Lopez, L.; Dominguez-Morales, J.P.; Corral-Jaime, J.; Vicente-Díaz, S.; Linares-Barranco, A. COVID-XNet: A Custom Deep Learning System to Diagnose and Locate COVID-19 in Chest X-ray Images. *Appl. Sci.* **2020**, *10*, 5683. [[CrossRef](#)]
16. Duran-Lopez, L.; Dominguez-Morales, J.P.; Rios-Navarro, A.; Gutierrez-Galan, D.; Jimenez-Fernandez, A.; Vicente Diaz, S.; Linares-Barranco, A. Performance Evaluation of Deep Learning-Based Prostate Cancer Screening Methods in Histopathological Images: Measuring the Impact of the Model’s Complexity on Its Processing Speed. *Sensors* **2021**, *21*, 1122. [[CrossRef](#)]

17. Amaya-Rodriguez, I.; Duran-Lopez, L.; Luna-Perejon, F.; Civit-Masot, J.; Dominguez-Morales, J.P.; Vicente, S.; Civit, A.; Cascado, D.; Linares-Barranco, A. Glioma Diagnosis Aid through CNNs and Fuzzy-C Means for MRI. In Proceedings of the 11th International Joint Conference on Computational Intelligence, Vienna, Austria, 17–19 September 2019; pp. 528–535.
18. Farooq, A.; Anwar, S.M.; Awais, M.; Rehman, S. A Deep CNN based Multi-class Classification on Alzheimer’s Disease using MRI. In Proceedings of the 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, China, 18–20 October 2017; pp. 1–6.
19. Milletari, F.; Ahmadi, S.A.; Kroll, C.; Plate, A.; Rozanski, V.; Maiostre, J.; Levin, J.; Dietrich, O.; Ertl-Wagner, B.; Bötzl, K.; et al. Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Comput. Vis. Image Underst.* **2016**, *164*, 92–102. [CrossRef]
20. Chen, L.; Bentley, P.; Mori, K.; Misawa, K.; Fujiwara, M.; Rueckert, D. DRINet for Medical Image Segmentation. *IEEE Trans. Med. Imaging* **2018**, *37*, 1–11. [CrossRef]
21. Yushkevich, P.A.; Gao, Y.; Gerig, G. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3342–3345.
22. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Volume 9351, pp. 234–241. [CrossRef]
23. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
24. Ulyanov, D.; Vedaldi, A.; Lempitsky, V. Improved Texture Networks: Maximizing Quality and Diversity in Feed-forward Stylization and Texture Synthesis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 6924–6932.
25. Nair, V.; Hinton, G.E. Rectified Linear Units Improve Restricted Boltzmann Machines. In Proceedings of the International Conference on Machine Learning, Haifa, Israel, 21–24 June 2010.
26. Chollet, F. Keras: The Python Deep Learning Library. *Astrophysics Source Code Library ascl-1806*. 2018. Available online: <https://ui.adsabs.harvard.edu/abs/2018ascl.soft06022C> (accessed on 5 December 2021).
27. Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv* **2016**, arXiv:1603.04467.
28. Kingma, D.P.; Ba, J.L. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference for Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
29. Reetz, K.; Rodríguez, R.; Dogan, I.; Mirzazade, S.; Romanzetti, S.; Schulz, J.B.; Cruz-Rivas, E.M.; Alvarez-Cuesta, J.A.; Aguilera Rodríguez, R.; Gonzalez Zaldivar, Y.; et al. Brain atrophy measures in preclinical and manifest spinocerebellar ataxia type 2. *Ann. Clin. Transl. Neurol.* **2018**, *5*, 128–137. [CrossRef] [PubMed]
30. Inagaki, A.; Iida, A.; Matsubara, M.; Inagaki, H. Positron emission tomography and magnetic resonance imaging in spinocerebellar ataxia type 2: A study of symptomatic and asymptomatic individuals. *Eur. J. Neurol.* **2005**, *12*, 725–728. [CrossRef]
31. Seidel, K.; Siswanto, S.; Brunt, E.R.P.; Den Dunnen, W.; Korf, H.W.; Rüb, U. Brain pathology of spinocerebellar ataxias. *Acta Neuropathol.* **2012**, *124*, 1–21. [CrossRef] [PubMed]
32. Tustison, N.J.; Avants, B.B.; Cook, P.A.; Zheng, Y.; Egan, A.; Yushkevich, P.A.; Gee, J.C. N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* **2010**, *29*, 1310–1320. [CrossRef]
33. Fonov, V.S.; Evans, A.C.; McKinstry, R.C.; Almlí, C.R.; Collins, D.L. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage* **2009**, *47*, S102. [CrossRef]
34. Avants, B.B.; Tustison, N.; Johnson, H. Advanced Normalization Tools (ANTS). *Insight J.* **2009**, *2*.
35. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [CrossRef]
36. Márquez-neila, P.; Baumela, L.; Alvarez, L. A morphological approach to curvature-based evolution of curves and surfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *36*, 2–17. [CrossRef] [PubMed]
37. Brain Development Webpage. Available online: <https://brain-development.org/brain-atlas/> (accessed on 5 December 2021).
38. Hammers, A.; Allom, R.; Koepp, M.J.; Free, S.L.; Myers, R.; Lemieux, L.; Mitchell, T.N.; Brooks, D.J.; Duncan, J.S. Three-Dimensional Maximum Probability Atlas of the Human Brain, with Particular Reference to the Temporal Lobe. *Hum. Brain Mapp.* **2003**, *19*, 224–247. [CrossRef]
39. Gousias, I.S.; Rueckert, D.; Heckemann, R.A.; Dyet, L.E.; Boardman, J.P.; Edwards, A.D.; Hammers, A. Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest. *Neuroimage* **2008**, *40*, 672–684. [CrossRef]
40. Faillenot, I.; Heckemann, R.A.; Frot, M.; Hammers, A. Macroanatomy and 3D Probabilistic Atlas of the Human Insula. *Neuroimage* **2017**, *150*, 88–98. [CrossRef] [PubMed]
41. Mennes, M.; Biswal, B.; Castellanos, F.X.; Milham, M.P. Making data sharing work: The FCP/INDI experience. *Neuroimage* **2013**, *15*, 683–691. [CrossRef] [PubMed]
42. Kennedy, K.M.; Rodriguez, K.M.; Bischof, G.N.; Hebrank, A.C.; Reuter-Lorenz, P.A.; Park, D.C. Age Trajectories of Functional Activation Under Conditions of Low and High Processing Demands: An Adult Lifespan fMRI Study of the Aging Brain. *Neuroimage* **2015**, *104*, 21–34. [CrossRef] [PubMed]

43. Chan, M.Y.; Park, D.C.; Savalia, N.K.; Petersen, S.E.; Wig, G.S. Decreased segregation of brain systems across the healthy adult lifespan. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E4997–E5006. [[CrossRef](#)] [[PubMed](#)]
44. Cocosco, C.A.; Kollokian, V.; Kwan, R.K.S.; Evans, A.C. BrainWeb: Online Interface to a 3D MRI Simulated Brain Database. *Neuroimage* **1997**, *5*, 425.
45. Kwan, R.K.; Evans, A.C.; Pike, G.B. MRI Simulation-Based Evaluation of Image-Processing and Classification Methods. *IEEE Trans. Med. Imaging* **1999**, *18*, 1085–1097. [[CrossRef](#)]
46. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571. [[CrossRef](#)]
47. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
48. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 29. [[CrossRef](#)]
49. Vijaymeena, M.K.; Kavitha, K. A survey on similarity measures in text mining. *Mach. Learn. Appl. Int. J.* **2016**, *3*, 19–28.
50. Bradley, A.E. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **1997**, *30*, 1145–1159. [[CrossRef](#)]

Paper IV

IV

ConvNets for automatic detection of polyglutamine SCAs from brain MRIs: state of the art applications

Cabeza-Ruiz, R., Velázquez-Pérez, L., Reetz, K., Pérez-Rodríguez, R.

Published in *Medical & Biological Engineering & Computing*, November 2022,
volume 61, DOI: 10.1007/s11517-022-02714-w.

10.1007/s11517-022-02714-w



ConvNets for automatic detection of polyglutamine SCAs from brain MRIs: state of the art applications

Robin Cabeza-Ruiz¹ · Luis Velázquez-Pérez^{2,3} · Roberto Pérez-Rodríguez^{1,2} · Kathrin Reetz⁴

Received: 16 September 2021 / Accepted: 26 October 2022
© International Federation for Medical and Biological Engineering 2022

Abstract

Polyglutamine spinocerebellar ataxias (polyQ SCAs) are a group of neurodegenerative diseases, clinically and genetically heterogeneous, characterized by loss of balance and motor coordination due to dysfunction of the cerebellum and its connections. The diagnosis of each type of polyQ SCA, alongside with genetic tests, includes medical images analysis, and its automation may help specialists to distinguish between each type. Convolutional neural networks (ConvNets or CNNs) have been recently used for medical image processing, with outstanding results. In this work, we present the main clinical and imaging features of polyglutamine SCAs, and the basics of CNNs. Finally, we review studies that have used this approach to automatically process brain medical images and may be applied to SCAs detection. We conclude by discussing the possible limitations and opportunities of using ConvNets for SCAs diagnosis in the future.

Keywords Spinocerebellar ataxia · Neural network · Deep learning · Medical imaging · Magnetic resonance imaging

1 Introduction

Spinocerebellar ataxias (SCAs) are a group of neurodegenerative disorders, phenotypically and genetically heterogeneous, characterized by loss of balance and motor coordination due to dysfunction of the cerebellum and its afferent and efferent pathways [1–3].

These cerebellar disorders are characterized by gait ataxia, dysarthria, dysmetria, and postural instability, which may be accompanied by extracerebellar signs such as movement disorders (including dystonia, parkinsonism, and chorea), dementia,

epilepsy, visual disorders, lower motor neuron signs, peripheral neuropathy, and dysautonomic dysfunction [2, 4].

Although many of the SCAs result from point mutations, DNA arrangements, or expansion of non-coding repeats, the most common SCAs are caused by expansion of the CAG trinucleotide repeat that encodes polyglutamine (polyQ) in the relevant disease proteins. These called polyQ SCAs are SCA1–SCA3, SCA6, SCA7, SCA17, and dentatorubral-pallidoluysian atrophy (DRPLA) [5, 6].

According to [4, 7, 8], polyQ SCAs have a prevalence of around 1 to 5 cases per 100,000 people. SCA3 and SCA2 are the most frequent molecular subtype of the disease worldwide. On the other hand, the SCAs 1, 6, 7, 17 and DRPLA have greatly varying prevalence depending on the ethnic background of the population [4]. Worldwide, SCA2 is the second most frequent molecular subtype of spinocerebellar ataxias, only surpassed by SCA3. Nevertheless, in Holguín, Cuba, the disease reaches the highest prevalence, resulting from a putative founder effect [3, 9].

This review discusses the main clinical and imaging features of polyglutamine SCAs, highlighting the use of convolutional neural networks in the evaluation of the cerebellar and brain degeneration.

✉ Robin Cabeza-Ruiz
robbinc91@gmail.com

¹ CAD/CAM Study Center, University of Holguín, Holguín, Cuba

² Cuban Academy of Sciences, La Habana, Cuba

³ Center for the Research and Rehabilitation of Hereditary Ataxias, Holguín, Cuba

⁴ Department of Neurology, RWTH Aachen University, Aachen, Germany

2 PolyQ SCAs, clinical features, and MR imaging

The pathological variability within each polyQ SCA makes it challenging to describe a characteristic pattern for each kind, but certain areas tend to be affected preferentially in each SCA [5]. For example, the SCA7 involves severe retinal degeneration leading to eventual blindness [10], while SCA6 affects mainly the cerebellum [11].

Also, clinical features tend to differ from one to another polyQ SCA: SCA2 patients commonly present slow saccadic eye movements and progressive cerebellar ataxia; SCA3 patients usually suffer severe spasticity [5]; SCA6 patients tend to suffer only pure cerebellar ataxia; and SCA17 is usually differentiable from other SCAs for its association with cognitive and psychiatric impairment [12]. Table 1 shows the main clinical findings for polyQ SCAs.

Three patterns of macroscopic atrophy reflecting damage to different neuronal systems are recognized in spinocerebellar ataxias, named spinal atrophy (SA), olivopontocerebellar atrophy (OPCA) and cortico-cerebellar atrophy (CCA) [18, 19].

Neuroimaging has been widely used to diagnose SCAs, since 1995, when Kumas [20] described their principal characteristics, obtained using CT, in children with olivo-pontocerebellar atrophy.

In particular, structural MRI is a suitable option for organs segmentation and volumetric characterizations [21], and has a predominant diagnostic role with respect to other techniques like SPECTs and PETs [18], based on the visual detection of SA, OPCA and CCA. According to Klaes et al. [6], MRI is the best studied biomarker

candidate for polyglutamine expansion spinocerebellar ataxias, so far. Table 2 shows main structural MRI findings on polyQ SCAs.

In SCA1, MRI brain scans disclose a severe atrophy of the cerebellum and the brainstem [22, 23]. Other studies [24] reported gray matter volume loss in the cerebellar hemispheres, vermis, and whole brainstem and white matter loss in the whole brainstem, midbrain, pons, middle cerebellar peduncles, and cerebellar hemispheres. In advanced stages, patients may present white matter with “hot cross sign” [18]. Also, gray matter loss in the medulla oblongata extending to the pons and in lobule IX of the cerebellum has been found in SCA1 preclinical subjects [25].

For SCA2, studies have demonstrated atrophic changes in patients, and increased fourth ventricle diameter in asymptomatic carriers [26], symmetric gray matter volume loss in the cerebellar vermis and hemispheres, with sparing of vermian lobules I, II, and X and of hemispheric lobules I, II, and crus II [27]. Also, reduced cerebellar and brainstem volumes have been found [9]. Other studies have shown “hot cross sign” and diffuse T2 high signals in pons [18].

In the case of SCA3, it affects the cerebellar cortex and olive nuclei less than SCA1 and SCA2, but the deep cerebellar nuclei and basis pontis tend to be more severely affected [5]. Studies show the presence of pontocerebellar atrophy, and atrophy of the globi pallidus, frontal and temporal lobes [1, 28, 29]. Further analysis [6] reported significant atrophy for the total cerebellum or cerebellar hemispheres, atrophy in vermian, dentate nucleus, cerebellar peduncle, and brainstem.

SCA6 is the only one that does not display significant brainstem involvement, it is known as “pure” cerebellar ataxia [4, 5, 23, 30]. Cortical cerebellar vermis hemisphere

Table 1 Polyglutamine SCAs with main clinical features and responsible genes

SCA	Clinical features in addition to cerebellar syndrome	Gene
SCA1	Extrapyramidal symptoms, spasticity, ophtalmoparesis, slow saccades, peripheral neuropathy, axonal polyneuropathy, dysarthria, nystagmus, dysphagia, cognitive impairment, muscle atrophy [3, 4, 13]	ATXN1 (6p22.3)
SCA2	Slow saccades + +, parkinsonism, extrapyramidal signs, ophtalmoparesis, axonal polyneuropathy, dysarthria, dysphagia, nystagmus, myoclonus, slow saccades [3, 4, 9, 13, 14]	ATXN2 (12q24.12)
SCA3	Parkinsonism, extrapyramidal features, spasticity (severe) +, dystonia, ophtalmoparesis, axonal polyneuropathy, sleep disturbances, mild cognitive impairment [1, 3, 4, 15]	ATXN3 (14q32.12)
SCA6	Prominent cerebellar features (ataxia, dysarthria, nystagmus + , tremor), dystonia [4, 11, 13]	CACNA1A (19p13.2)
SCA7	Retinal degeneration + +, ophtalmoplegia, pyramidal signs, hearing loss, slow saccades +, dysarthria, dysphagia, spasticity, episodic psychosis. Infantile variant (hypotonia, development delay, microcephaly, visual loss, fatal cardiac failure) [4, 13]	ATXN7 (3p14.1)
SCA17	Psychosis and behavioral changes, dystonia, parkinsonism, mental deterioration +, seizures, pyramidal signs, rigidity, epilepsy, increased muscle tone [3, 13, 16, 17]	TBP (6q27)
DRPLA	Myoclonus epilepsy, chorea, dementia, subcortical demyelination (Haw-River syndrome), cognitive impairment, behavioral symptoms + [3, 4, 13]	ATN1 (12p13.31)

(+) Suggestive signs

(++) Highly suggestive signs

Table 2 Main structural MRI findings for polyQ SCAs

SCA	Main MRI findings
SCA1	Severe gray matter volume loss in the cerebellar hemispheres, vermis, and whole brainstem; gray matter loss in the medulla oblongata on presymptomatic carriers [18, 22–25]
SCA2	Symmetric gray matter volume loss in the cerebellar vermis and hemispheres, sparing of vermian lobules I, II and X and of hemispheric lobules I, II, and crus II; reduced brainstem volume; increased fourth ventricle diameter in some asymptomatic carriers [9, 26, 27]
SCA3	deep cerebellar nuclei and basis pontis severely affected; pontocerebellar atrophy, and atrophy of the globi pallidi and frontal and temporal lobes; significant atrophy for the total cerebellum or cerebellar hemispheres, vermian, dentate nucleus, cerebellar peduncle, and brainstem [1, 5, 6, 28, 29]
SCA6	“pure” cerebellar ataxia; cortical cerebellar vermis hemisphere atrophy, with less pronounced atrophy of pons, and general cerebellar volume loss [4, 5, 19, 21, 23, 30]
SCA7	olivopontocerebellar degeneration, pontine and spinal cord atrophy; mild subcortical atrophy [4, 19, 21, 31]
SCA17	atrophy on cerebellum (mainly on posterior structures), brainstem and cerebrum; putaminal rim hyperintensity on patients [16–18, 32–34]
DRPLA	global cerebral atrophy, with more marked cerebellar atrophy, and bilateral periventricular white-matter lesions; atrophy of the brainstem, superior cerebellar peduncle, and cerebellum [35, 36]

atrophy [19], with less pronounced atrophy of pons, and general cerebellar volume loss, even compared with other SCAs [4, 21], are the main findings for SCA6 patients.

The main features of SCA7 are olivopontocerebellar degeneration [4, 19], pontine and spinal cord atrophy [21]. In a study performed by Moriarty et al. [31], mild subcortical atrophy was found for SCA7 patients.

Analysis on SCA17 patients MRI reveal atrophy on cerebellum, brainstem, and cerebrum [16]. The main affections found in some studies are the vermis and posterior cerebellar structures [17, 32]. In some cases, cerebellum atrophy can be marked, even when the disease is on an early stage [18, 33]. Other studies have discovered also putaminal rim hyperintensity on patients with SCA17 [34].

Some MRI analysis on DRPLA patients have shown global cerebral atrophy, with more marked cerebellar atrophy, and bilateral periventricular white-matter lesions [35]. The use of imaging to distinguish DRPLA from clinically similar disorders is challenging [35], in particular in late-onset disease, where the isolated atrophy of the brainstem and cerebellum increase the likelihood of more common diagnoses being made (e.g., alcohol-induced cerebellar degeneration), and imaging may not show supportive features in elderly population [36]. In a study performed by Sugiyama et al. [36] consisting of MR images of ten patients with elderly onset genetically confirmed DRPLA, atrophy of the brainstem, superior cerebellar peduncle, and cerebellum were found in all patients; also, abnormal signals in the brainstem (inferior olive, pons, midbrain), thalamus, and cerebellar white matter were very common between them.

Although polyQ SCAs may be difficult to diagnose and to distinguish one from another using medical images, some characteristics are common for most of them, like cerebellar and brainstem degeneration, white and (or) dark matter loss, and general neuron loss. The use of this knowledge joined

with the symptoms and hereditary constraints, makes it possible to diagnose and differentiate SCAs.

Medical image segmentation is typically used to locate objects of interest and their boundaries to make the representation of a volumetric image stack more meaningful and easier for analysis [37]. For instance, neurological MRI volumetric studies have been conducted to compare different stages of the disease and their correlation with symptoms severity [9]. Generally, this process is made by hand, slice by slice, and can be very time-consuming. Automatic detection of these features, might improve the speed of the diagnosis process, and aid researchers and clinicians to develop new treatments and monitor their effectiveness.

Many researches have been conducted with the aim of finding an appropriated asset on automatic MRI processing. Some of these use classical machine learning algorithms, like K-nearest neighbors (KNN) [38]. Other approaches [37, 39–41] use prior-defined knowledge and other features. Although these investigations have achieved good results on their respective tasks, the processing time can extend over several hours for a single MRI. Also, the presence of hand-made features could provoke a malfunctioning of the algorithms, because some elements are very difficult to find or define, and perhaps the human is not capable of observing those features which are truly relevant for the problem.

Multi atlas-based approaches [42] are a good choice for parcellating a segmenting human cerebellum from brain MRIs, eliminating some errors carried by single atlas-based ones [43]. Those algorithms usually present a very fast functioning; in a study performed by Carass et al. [44], CERES2 [42] scored the best results on their datasets. However, these algorithms are highly dependent on the multiple registration phases [45]. If some of these fails, the segmentation could go wrong, and non-treated noise on images could make the system fail too. Besides, map- and atlas-based approaches

could have a bad performance when analyzing images whose features are not present in the atlases (maps) construction.

The use of convolutional neural networks [46, 47] on medical imaging processing have grown very fast in the last years. Those are systems capable of finding the correct feature set for each problem, without the need of the human intervention, avoiding possible mistakes. Besides, the presence of noise on images tends to improve the behavior rather than deteriorate it, providing the system with the capacity to process and recognize patterns not included in the training. For these reasons, we consider the use of convolutional neural networks as the main path for automatic SCAs characterization from MRIs. In the next section, we describe the basic functioning of convolutional neural networks, and their applications on brain MRIs over the last years.

3 Convolutional neural networks

The first convolutional neural networks architecture was elaborated by Fukushima in 1980 [48], presenting the Neocognitron as a simulation of the brain functioning for visual pattern recognition, but the first real application was done by LeCun et al. in 1989 [46]. Convolutional networks (convnets or CNNs) have demonstrated outstanding performances at tasks such as handwritten digit classification, face and contour detection [47, 49], video processing [50, 51], change detection on satellite images [52], neurological development analysis and prediction [53, 54], and other fields. A single convnet is composed of a series of layers, each of them including various filters which conduce to the image's processing result.

The main parts of convnets are convolution layers, pooling layers, and activation functions. Convolution layers apply a certain number of filters, randomly initialized, to the processing image. Activation functions are in charge of decide which features are taken after an operation (usually a convolution layer is followed by an activation function). Pooling layers reduce computational requirements of the networks, diminishing the image resolution and eliminating connections between convolutional layers [55].

CNNs have the ability to learn a hierarchical representation of input data, without relying on handcrafted features (no need for predefined features), that's why they have success on solving complicated tasks such as classification, segmentation and object detection [47, 56–60]. When training on natural images, CNNs learn information in a variety of forms: low-level edges, mid-level edge junctions, high-level object parts and complete objects [61], forming a complete picture of objects and features. For these characteristics, some authors know them as space invariant artificial neural networks (SIANNs) [62, 63].

Different layers of a CNN are capable of different levels of abstraction and capture different amount of structure from the patterns present in the image [57] (see Fig. 1). A convnet can be observed as a parameterized function composed of a series of simple linear and nonlinear operations. In 3D image segmentation, it can produce feature maps that combine local and global information,

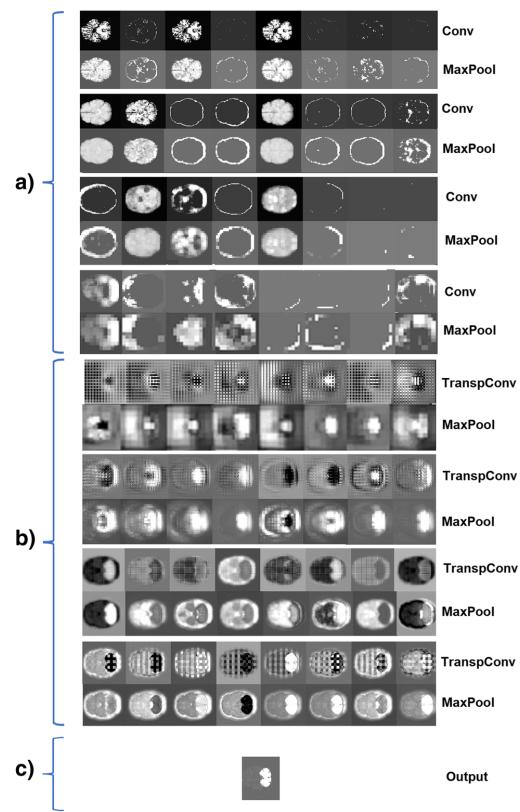


Fig. 1 From top to bottom: partial results of different layers of a CNN, trained for cerebellum segmentation from brain MRIs. Each individual row shows eight 2D sample slices, which are the partial results of one layer. The encoder section (a) is composed of several convolution layers (Conv) with a ReLU activation, and each Conv layer is followed by a MaxPooling operation. A reduction on resolution is observed at each row. Also, after applying the pooling operation, the resolution is reduced. The decoder section (b) contains transposed convolutional layers (TranspConv), followed by MaxPooling layers. Each transposed convolution processes previous feature maps and helps restoring the original image size. The max pooling layers on the decoder section help on regularizing the partial outputs of TranspConv layers. In this section, partial feature maps usually contain parts of the cerebellum. Finally, the output layer (c) carries only a mask with the segmented cerebellum

which are used to perform per-voxel classification [64]. The goal on training is to estimate the filters of the final model [61].

The typical use of CNNs is on classification task [44], where the output to an image is a single label (the class which the image belongs to). In a classification task, an object is assigned to one of the predefined classes (e.g., an MRI which contains some abnormalities, like tumors, might be one of those predefined classes). However, in many visual tasks, especially in biomedical image processing, the desired output should include location, i.e., a class label is supposed to be assigned to each pixel (or voxel, depending on the input dimensions) [65]. This is known as object recognition, and involves many tasks, but the most popular are object detection and semantic segmentation. The first consists of marking out bounding boxes around each object of a particular category in an image [66].

Semantic segmentation is, in the other hand, assigning a category or class to every pixel in the image (see Fig. 2). In terms of medical image processing, it divides an image into different regions in order to separate desired organs. The difference between segmentation and recognition is that, a segmentation algorithm can accurately mark the pixels for every label in the image, but would provide no indication of how many objects of each type are present on the image, or their position [66].

3.1 Main convnets architectures for image segmentation

The most common form of ConvNet for image segmentation is the fully convolutional network (FCN), which is a model designed for spatial prediction problems [67]. Their main characteristic is the absence of dense layers. Each dense layer is substituted with a convolutional layer. They can accept images of any dimensions and produce outputs with corresponding dimensions.

Another commonly used segmentation structure is U-Net. Basically, U-Nets are a kind of FCN, configured by an encoder and a decoder section. The encoder section, composed of various down-sampling steps, gradually reduces the dimensions of the image being processed, and the decoder section performs the opposite operations, restoring the image shape. An important particularity of U-Nets is the concatenation of outputs generated during encoder section, whose are used in decoder section as extra data to process. These characteristics make U-Nets strong architectures, capable of taking features from local and global impact. The network in Fig. 2 is a U-Net, and also an FCN. Ronnenberger, Fischer and Brox [65] made a very good description about this type of network applied to biomedical image segmentation. Examples of this type of network can be observed in [64, 68–71].

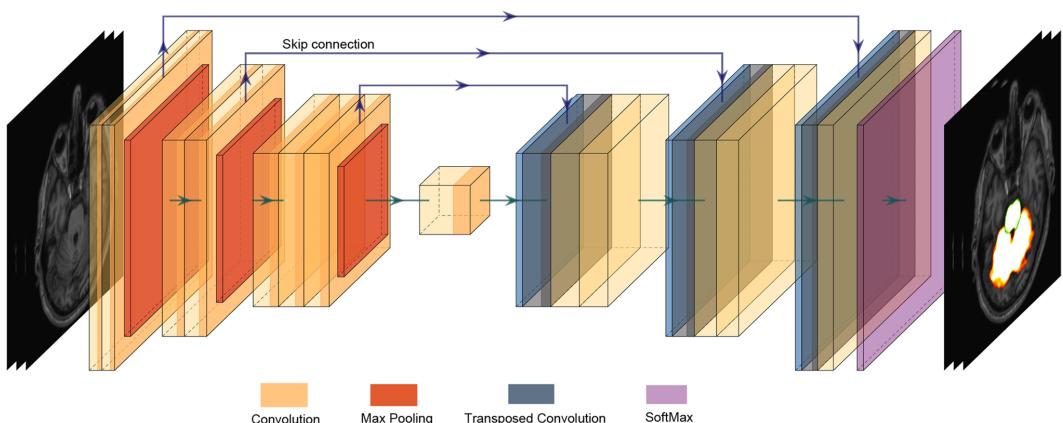


Fig. 2 Representation of a CNN applied to semantic segmentation. The network is composed of three down-sample operations, consisting of joining convolutional layers (yellow color) and max pooling operations (orange color). In the middle, a single convolutional layer processes the more local features. The decoding section consists of three up-sample operations for restoring the original image size. Each one of these is composed of transposed convolutions (blue color)

combined with regular convolutions. The output of the network is given by a softmax (purple color) as the last layer. Gray arrows indicate the main path of execution, and blue arrows represent the skip connections. In this case, the cerebellum and brainstem are segmented. Procedure results should include only the segmented parts, but the rest of the MRI has been left here for visual purposes

Table 3 Reviewed investigations using ConvNets for cerebellar segmentation from brain MRIs

	Year	Ref	Arch	Context	Dataset	Best scores					MRIs
						DSC	ACC	SP	SN	ICC	
2018	[44]	FCN	3D	Own	0.96	-	-	-	-	-	50
	[44]	CNN	3D	Own	0.93	-	-	-	-	-	50
	[44]	U-Net	3D	Own	0.93	-	-	-	-	-	50
2019	[68]	U-Net	3D	Own	0.92	-	-	-	-	-	50
	[64]	U-Net	3D	Own	0.95	-	-	-	-	-	50
		Kirby [78]			-	-	-	-	-	0.99	42
2020	[73]	DenseNet	3D	Own	0.89	-	-	-	-	-	60

Table 4 Reviewed research using ConvNets for brain segmentation and parcellation from MRIs

	Year	Ref	Arch	Context	Dataset	Best scores					MRIs
						DSC	ACC	SP	SN	ICC	
2015	[45]	SegNet	Voxel	MICCAI12	0.76	-	-	-	-	-	35
2016	[57]	CNN	3D	Own	0.77	-	-	-	-	-	55
2017	[76]	M-Net	2D	IBSR [79]	0.83	-	-	-	-	-	18
				MICCAI13 [80]	0.85	-	-	-	-	-	47
	[81]	SegNet	Voxel	MICCAI12	0.74	-	-	-	-	-	35
2018				IBSR	0.84	-	-	-	-	-	18
				LPBA40 [82]	0.82	-	-	-	-	-	50
				Hammers67n20 [83]	0.84	-	-	-	-	-	20
				Hammers83n30 [83]	0.80	-	-	-	-	-	30
	[84]	FCN	3D	Own	0.92	-	-	-	-	-	55
	[85]	CNN	Voxel	IBSR	0.86	-	-	-	-	-	18
	[86]	ResNet	3D	MrBrains2013 [87]	0.91	-	-	-	-	-	20
	[74]	DenseNet	Pixel	IBSR	-	0.92	-	-	-	-	18
	[88]	SegNet	Pixel	MICCAI12	0.79	-	-	-	-	-	35
	[89]	U-Net	3D	iSeg-2017 [90]	0.95	-	-	-	-	-	23
2019	[92]	CNN	3D	iSeg-2019 [91]	0.87	-	-	-	-	-	39
				Own	0.86	0.77	-	-	-	-	2341

DenseNet [72] is another well-known architecture, which is based on densely connected layers. Each layer on a dense block is connected to all layers after it (a dense block can be a part of the network, or the entire network itself). The possibilities for DenseNet are enormous, as they process every part of the images several times before the final classification. Besides, the final decision is not dependent on a single layer, but all the previous ones. Examples of DenseNet applications can be found in [73, 74].

Other structures exist as well, every one taking advantage of some image features with the aim of reducing computational cost, and avoiding unnecessary depth in the models. For instance, residual networks (ResNets) use maps created on previous layers [75], and M-Net [76] works like an extended U-Net, increasing the number of concatenations from previous layers, among others.

The union between the specific characteristics of each type of network is possible, and has been explored in some occasions, obtaining remarkable results. An example can be analyzed in [77], in which the authors combined U-Net with ResNet and DenseNet capabilities for tumor segmentation from 2D images.

4 Applications on automatic brain MRI processing, and possible use on SCAs detection

Many researches have been conducted to demonstrate the power of ConvNets on processing brain MRIs. For our purposes, main found applications are focused on cerebellum, brain structures, brain lesions, and image classification. Tables 3, 4, 5, and 6 show summaries of the mentioned groups and the reviewed investigations¹.

¹ These data were provided for use in the MICCAI 2012 Grand Challenge and Workshop on Multi-Atlas Labeling [B. Landman, S. Warfield, MICCAI 2012 workshop on multi-atlas labeling, in: MICCAI Grand Challenge and Workshop on Multi-Atlas Labeling, CreateSpace Independent Publishing Platform, Nice, France, 2012]. The data is released under the Creative Commons Attribution-NonCommercial license (CC BY-NC) with no end date. Original MRI scans are from OASIS (<https://www.oasis-brains.org/>). Labeling were provided by Neuromorphometrics, Inc. (<http://Neuromorphometrics.com/>) under academic subscription.

Table 5 Analyzed investigations that use ConvNets for the segmentation of brain lesions from MRIs

Year	Ref	Arch	Context	Dataset	Best scores					MRIs
					DSC	ACC	SP	SN	ICC	
2016	[93]	CNN	Voxel	TBI [94]	0.63	-	-	0.63	-	61
				BRATS2015	0.84	-	-	0.87	-	384
2017	[95]	CNN	Pixel	BRATS2013	0.88	-	0.89	0.87	-	30
2018	[77]	FCN	2D	BRATS2017	0.83	-	0.99	0.84	-	285
2019	[69]	U-Net	2D	BRATS2018	0.95	-	-	-	-	285
	[70]	U-Net	2D	BRATS2018	0.89	-	0.99	0.90	-	285
2020	[71]	U-Net	3D	BRATS2020	0.87	-	0.99	0.91	-	369
	[96]	U-Net	2D	Own	0.95	-	-	-	-	484

Table 6 Research reviewed using ConvNets for brain MRI classification

Year	Ref	Arch	Context	Dataset	Best scores					MRIs
					DSC	ACC	SP	SN	ICC	
2017	[97]	GoogleNet	2D	ADNI [98]	-	0.98	0.99	0.99	-	355
		ResNet	2D		-	0.98	0.99	0.99	-	
		ResNet	2D		-	0.98	0.99	0.99	-	
2019	[99]	ResNet	2D	Own [100]	-	1	1	1	-	613
2020	[101]	CNN	2D	OASIS [102]	-	0.82	0.81	0.93	-	416
				MIRIAD [103]	0.82	0.81	0.93			708

4.1 Cerebellum parcellation/segmentation

The main objective of these researches is to accurately segment the cerebellum from MRIs, providing a high detail of the anatomical condition of the cerebellar structures. Table 3 shows the reviewed investigations.

4.1.1 Carass et al.

Carass et al. [44] made a comparison between several cerebellum parcellation methods, in which three convolutional neural network architectures were tested (LiviaNet, ConvNet and DeepNet).

LiviaNet model was a FCN, based in [93, 104], and was composed of 13 layers in total: nine convolutional layers, followed by three fully connected ones, and finally a classification layer. To assure the FCN structure, the fully connected layers were converted to $1 \times 1 \times 1$ convolutions. The algorithm does not process the whole MRI, but subsamples the image in smaller patches, and then processes each patch.

ConvNet method consisted in two FCNs trained for parcellating half cerebellum each. Uses a cerebellum mask for reducing the size of the computational domain of processing. The structure has an alternating stack of 40- and 24- wide convolutional layers. The model was created without dimensionality reduction layers. One of the main problems of this model, according to the authors, is the fact that each convnet

parcellates only half-cerebella, and then the central voxels are arbitrarily labeled, which can cause some inaccuracies.

DeepNet method consists on U-Net [65], with five down-and up-sample operation. Each down-sample consists of a convolutional layer followed by a pooling operation. The up-sample sequences consist of an upsampling convolution followed by two convolutional layers. The last layer consists of a $1 \times 1 \times 1$ convolution to reduce the number of output channels to the number of labels.

The approaches were trained and evaluated on their own dataset, composed of 30 MRIs from pediatric patients and 20 from adult people.

Two of the methods (LiviaNet and DeepNet) resulted in the top three of the ranking made by authors. The authors clarify that the small size of the adult dataset (only 20 images) limits the statistical power of any test. Nevertheless, the scores obtained by the methods' evidence that convnets are capable of providing a high level of accuracy in parceling both adult and child cerebellum, and may offer an alternative for improving identification of cerebellar lobule volumes.

4.1.2 Han et al.

Han et al. [68] used two cascaded CNNs for the cerebellum parcellation: the first one, a common CNN, in charge of finding the most-likely position of the cerebellum in the MRI, and the second, a U-Net based FCNN with four down- and

up-sample operations, which uses the calculated position for cropping the image and parcellate the cerebellum.

The approach was evaluated using the same dataset as Carass et al. [44], with a total of 50 MRIs.

The algorithm obtained outstanding results, but the first network (the locating network) relies on MNI space registration for a correct bounding box detection. If such registration fails, then the calculated cerebellum position will not be correct, and neither the parcellation. Also, according to the authors, the processing could be wrong if some processed image contains a range of intensities which have not been seen on training stage, but this could be corrected with intensity augmentation during training, or image contrast harmonization.

Some problems were solved in their recent paper [64]. This time, the researchers used a five down- and up-sampling architecture, again based on U-net. Besides increasing the network depth, the authors included also a larger number of learnable parameters on each step, and added an instance normalization for avoiding the intensities problem. Tests were performed not only in the previously mentioned dataset [44], but also in the Kirby dataset, which is composed of 60 MRIs. Three more datasets were used (OASIS [102], Kwyjibo [105] and ABIDEII [106], with 1931, 246 and 795 images used from each dataset, respectively), containing MRIs from patients with SCA, Alzheimer's Disease, Autism Spectrum Disorder, and healthy subjects, but only for visual purposes. The algorithm outperformed CERES2, the overall winner of the rank made by Carass et al. [44], obtaining state-of-the-art results in the cerebellum parcellation task. Despite the increment of scores by the model with respect to CERES2, the authors clarify that the images used for training stage did not contain the neck of the patients, so the processing could be wrong if a new image is processed containing the neck portion. This could be fixed using some preprocessing for neck removal. The problem of the locating network regarding the MNI registration stage still exists, if this step goes wrong, the whole processing could fail too. However, based in the results and analysis provided by the authors, we think that the proposed architecture is adequate for cerebellum parcellation, and will be explored in future investigations.

4.1.3 Kim et al.

Patient-specific segmentation of deep cerebellar dentate and interposed nuclei on MRI is proposed by Kim et al. [73], using a hyperdense convnet. The proposed architecture is an extension of FC-Densenet [107], a deep neural network which uses dilated convolution layers [108] for adjusting the size of the perceptive field by using sparse convolutional kernel and can be used for avoiding the excessive use of consecutive pooling operations, removing the loss of details in

boundary information on small structures [73]. The decoder of the network is composed of dense dilated blocks, organized in three down-samples. Each dilated block consists of three successive convolution layers, with all the possible connections between them. The decoder is divided in two sections, one in charge of predicting one channel for the dentate and another for the interposed, and the second section obtains the union of both predictions. The first decoder is composed of three up-samplings with one dense block each, and the second decoder uses up-sampling layers for combining the data from the skip connections. Evaluations were carried on a dataset created by the authors, containing a total of 60 images.

The method achieved significantly better performance in dentate and interposed segmentation than state-of-the-art models. The main problem with the approach is the training used by the authors. In this case, 7 T images were used, as the authors state that 3 T images might not have the necessary quality for providing a good training set. This situation could limit the performance, because 7 T MRIs are not so common, there are a limited number of 7 T MRI machines in current practice due to the significant hardware cost. However, the segmentation of dentate and interposed nuclei segmentation has been achieved, and the architecture could be useful on neuroanatomical studies.

4.1.4 Partial considerations

All reviewed researches use 3D context for cerebellum segmentation, and the most used architecture was U-Net. The U-Net architecture clearly represents an advantage over simple CNN and FCN used by Carass et al. in [44], as information is processed in patches with variable sizes, making it capable to work with different sets of features at every layer of the network.

Despite the possibility of some kind of memory overflow if not well defined [109], DenseNet, used by Kim et al. [73] seems to be a good architecture for cerebellar segmentation from MRIs, but the authors presented some problems to manually segment some regions of the cerebellar nuclei. For this reason, we think that further analysis should be performed with the aim of a fair comparison between U-Net and DenseNet architectures. In other hand, the use of various imaging modalities represents a disadvantage, as sometimes not all of them are present.

U-Net architectures used for this kind of task, are composed of four [68] and five [44, 64] down- and up-sampling phases. The main differences between architectures are the sizes of the filters used by the authors. Approach analyzed by Carass et al. uses batch normalization [110], in contrast to approaches proposed by Han et al. [64, 68], where instance normalization [111] is preferred by the authors. Instance normalization provides a simpler preprocessing stage, as

no other image intensity normalization is required, since instance normalization is invariant to additive and multiplicative transformations of the images intensities [64, 68]. Finally, the pipeline used in [64] is an improvement of [68], with one more down- and up-sampling stages. The use of five stages provided the system with more feature sets to work with, obtaining outstanding segmentations.

Besides, this approach by Han et al. was more intensively evaluated than the others, making it more reliable for future researches. The pipeline was tested with five datasets, but no metrics have been reported for three of them. In short terms, the algorithm was verified with more than 2000 MRIs.

According to previous comments, we consider it [64] is the more adequate approach. One possible disadvantage could be the presence of the neck tissue in the MRIs, which could cause errors in the image registration. Another preprocessing for neck removal could be incorporated, ensuring that it will not be contained within the image to be processed.

The use of 3D convnets might represent a problem if a computer with low resources is available for processing. The approach [64] was tested in a Lenovo laptop, equipped with an Intel Core i3-8145U CPU, 8 GB RAM memory, and no dedicated graphics card. Though tested in a computer with discrete characteristics, the algorithm took a mean time of three minutes to entirely process each MRI. We believe it is an adequate amount of time, if compared with manual segmentation, which can take several hours.

4.2 Brain parcellation/segmentation

The researches reviewed on this field are mainly oriented to parcellate the whole brain into smaller parts. Some of them include the cerebellum in the resulting parcellations, and others are oriented to segment cerebrospinal fluid, white matter and gray matter. In Table 4 are summarized the described applications, as well as the best scores achieved.

4.2.1 Brébisson and Montana

Brébisson and Montana [45] proposed CNNs for parcellating the whole brain from magnetic resonances. According to the authors, this was the first deep neural network applied to this task. The architecture used is a convnet that classifies each voxel of the magnetic resonance in its respective class. The algorithm uses as features patches of 2D and 3D related with each voxel, and is composed of a sequence of convolution + pooling layer, followed by a deconvolution + pooling layer, for each input patch. The outputs of each pathway are passed through two fully connected convolutional layers, and the final layer consists of softmax activation, which maps the output to the range [0; 1]. As the network takes as input the information regarding the voxels of the image, the training

was done with a small amount of all the voxels from each MRI, drastically reducing the computational cost. During the test, the architecture obtained scores that positioned it as the first place until that moment in parcellating the whole brain from MRIs. Regardless the final behavior of the system, the fact that it works with points of the MRIs results complex. Analyzing each point in an MRI can take a long time due to the dimensions of the image. The approach used by the authors of using a subset of all those points is a good starting point to reduce this time, but perhaps some of the points that are not analyzed could be more significant for the task. An important characteristic to emphasize is that the method is registration free. This characteristic gives advantage over previously commented methods, as this can be a time-consuming preprocessing step.

4.2.2 Milletari et al.

Milletari et al. [57] used a combination of CNN and Hough Voting [112] to infer deep brain regions from MRIs. The authors tested the difference between using 2D, 2.5D, and 3D data patches as input data, concluding that the 3D data patches are better, especially in clinical settings, where the amount of data is rarely sufficient. We focus this review only in the 3D approach. The outputs of the convnet, as well as the parameters obtained by its deeper layers, are used by a voting system for finding the desired structures. The authors performed a comparison between six different CNN architectures, and the best performance was obtained by the simplest network, which consisted of one convolutional layer, followed by a pooling operation, two more convolutions, and finally a fully connected convolutional layer followed by softmax for output.

According to the authors, the proposed modality outperformed all existent semantic segmentation CNNs. The used input features are 3D patches, making it very fast, and registration-free, like the previous research [45]. The better results were obtained when bigger regions and high contrast areas were segmented. The smaller regions, as well as the shallow areas, had some difficulties to be correctly segmented. We think that more preprocessing stages could be integrated to the architecture, so it can remove some errors carried by differences in contrast.

4.2.3 Mehta and Sivaswamy

Mehta and Sivaswamy [76] propose a new architecture for deep brain structures segmentation from MRIs. The structure, called M-Net, is an extension of traditional U-Net, but in each step of the encoder section takes into consideration the feature maps from the original image. In the decoder section, as well, uses the output maps of the deeper layer. For assessing memory troubles when treating a whole 3D

MRI, the algorithm processes images slice by slice, making a very fast segmentation. The encoder section consists of three down-samples formed by two convolutional layers and a pooling operation. At each step, the original features of the slice are pooled, and concatenated to the next section. The decoding section consist of three up-samples, which are composed of two convolutions followed by an upsample layer. The last layer consists of a single convolutional layer which returns the desired slice segmentation. The outputs are then concatenated, and the 3D result is returned. The method obtained good dice coefficient, and outperformed other deep cerebellar region segmentation methods. Nevertheless, the idea of working with 2D features for obtaining 3D segmentation seems not to be the best in terms of volumetric processing. For taking in consideration the relation between consecutive slices in a 3D image, the authors added a convolutional layer at the beginning of the model, which calculates the slice to be processed as the union of n of the original slices in the MRI. Though the method obtained acceptable dice scores, and outperformed other existent techniques, further analysis should be made for comparing the resulting segmentations with the outcomes of an M-Net trained for only 3D processing. We think that 3D processing will give more power to the system.

4.2.4 Mehta, Majumdar & Sivaswamy

Mehta, Majumdar & Sivaswamy [81] performed whole brain parcellation using a CNN. For the input, the network takes a series of 2D and 3D patches which contain information about a voxel. The system works, classifying each voxel into its respective label. The architecture consists of four input branches, which are processed separately, and then flattened and concatenated before the last two dense layers. Two of the branch use 2D inputs with shape 31×31 , and the other two correspond to 3D inputs with shape $21 \times 21 \times 21$. Each branch is composed of two consecutive blocks with two convolutional layers and a pooling operation each. The final output layer consists of softmax, which returns the most probable class for the voxel being processed. As inputs are voxels, the system is registration-free, giving it short processing times. This advantage highly depends on the input image shape, as sizes can vary and increase the number of input voxels. The scheme was ranked best compared with other existent approaches.

4.2.5 Moeskops et al.

Moeskops et al. [84] used CNNs for brain parcellation. The main characteristic of this study is the use of adversarial training, which in brief uses a discriminator network that is optimized to discriminate real from generated images,

motivating the generator network to produce images that look real. For the generator, the authors compared two convnets: an FCN, consisting of 15 convolutional layers, and a dilated network, which uses only seven convolutional layers, without any subsampling element. The discriminator network consisted of three convolutional layers, followed by a pooling operation, two more convolutions and finally a fully connected layer. The output layer consisted of two nodes, distinguishing between real and manual segmentations. Another comparison regarding the impact of the adversarial training was done. As a result, both methods experimented improvement when trained using adversarial training. The methods were not compared with any previous investigation, and the authors did not conclude if dilated network or FCN were better for the generator. We think this could be an entry point for new investigations. Despite that, results were very promising for both schemes, demonstrating the capability of convnets for segmenting brain structures from MRIs.

4.2.6 Nguyen et al.

Brain segmentation in cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM) is the objective proposed by Nguyen et al. [85] by using convnets. This time the convnets do not perform all the work, because the gaussian mixture model [113] is used to segment those voxels easily identifiable, while two CNNs are applied for treating voxels which are similar in appearance and usually recognized insufficiently by traditional approaches. The first network processes each of the input MRI voxels independently, taking as inputs three adjacent slices of 11×11 along each axis, features which, according to the authors, describe the normalized intensity of the predicted voxel and its surrounding voxels. The original coordinates of the voxel are used too. The inputs are propagated to three consecutive convolutional layers, each followed of a pooling operation. The results are reshaped into nine vectors, one vector for each input slice, which are passed through three fully connected layers, and finally the output are two values, indicating the probability of the voxel belonging to the certain or uncertain voxel class. The outputs of this first architecture are passed to the second one, which is in charge of predicting the correct class for each voxel. The architectures of both CNNs are the same, except for the last layer (in the second convnet, the last layer has three values instead of two, one for each label). The model outperformed existing methods on the same task, obtaining the highest scores up to that time. The fact that it works with a few slices for each voxel makes the algorithm faster, and being registration-free makes it more attractive, but working only in voxel-level could bring mislabeling errors. Perhaps a bigger number of slices, or a bigger slice size, could be tested and compared with actual results.

4.2.7 Chen et al.

A method proposed by Nguyen et al. [85], Chen et al. [86] performed brain segmentation into cerebrospinal fluid, gray matter, and white matter using convolutional neural networks. The architecture consists of stacked residual modules, with a total of 25 volumetric convolutional/deconvolutional layers. Each residual module comprises two consecutive convolutional layers, and this result is concatenated with the module input before the final output. This type of processing allows the original information to be directly propagated in the forward and backward passes. In an attempt of covering the huge variation on the size of 3D brain anatomical structures, the authors fused four classifiers with deep supervision. The network takes as input several 3D patches, and the outputs are generated in an overlap-tiling strategy for switching the sub-volume results. The method showed good results in testing phase, and performed best compared with other state-of-the-art methods. However, for a proper functioning, the network uses three image modalities: T1, T1-IR, and T2-FLAIR. This represents a clear disadvantage over other methods, as sometimes only one of the modalities is available for processing. Despite that unique questionable problem, we think the authors made a good use of residual modules, as the proposed tool provide accurate segmentation of brain structures.

4.2.8 Gotapu and Dahli

Gotapu and Dahli [74] used 2D DenseNet [72] for segmenting CSF, WM, and GM from infant MRIs. The idea of DenseNet is to connect each layer to every other layer behind, which improves the flow of the information within the architecture. The purpose of such configuration is to provide the network with the capability of making decisions based on all the layers, rather than a single output layer. The system works pixel wise, taking as input two 2D patches for each pixel, and classifies it into one of the three tissue types. Each patch is independently processed with two dense blocks. The dense blocks are composed by three consecutive convolutional layers. The outcomes of the two paths are concatenated and passed through another dense block, with the same configuration that the previous ones. The final layer performs the classification of the current pixel into one of the three tissue types. Once again, we can appreciate the usability of convnets for single point classification, which eliminates the registering from preprocessing stages. Nevertheless, as comments for previous researches with the same approach, we think the shape of the input image will be decisive. Obtained scores were not so significantly high as in other investigations, and no comparisons were performed with other available approaches.

4.2.9 Manoharan, Pang & Wu

Manoharan, Pang & Wu [88] performed brain segmentation into 134 anatomical regions. The network is a modification of proposed by Brébisson and Montana [45]. As in [45], the procedure uses a series of 2D and 3D patches regarding the information of each pixel/voxel to process. This time, the only inputs are two 2D patches, each consisting of one slice per axis. The third input corresponds to the 3D patch, which center is the voxel being classified. Every patch is independently processed by two convolutional sections, and finally the outcomes are concatenated and passed across two consecutive fully connected layers. The output layer consists of a softmax with 134 neurons, one for each label. The procedure obtained better scores than the original [45], proving that better results can be obtained with fewer patches, making a simpler preprocessing stage. The same as in [45], the model is registration-free. Despite the obvious improvement in segmentation results, we consider the per-pixel processing is still an issue. The bigger is the image, the more time it will require to be processed. Also, according to the authors, deeper architectures should be tested, as no such study has been made for this type of CNN.

4.2.10 Lei et al.

Lei et al. [89] performed segmentation of CSF, WM, and GM on infant brain MRIs. The architecture is based on U-Net, but has some improvements with respect to the original: in the encoder section, instead of pooling operations, dilated convolution operations are used, eliminating the possible information loss caused by pooling operations. The encoder is also structured like ResNet, allowing the system to use more features than the original U-Net. The encoder is composed of one simple 3D convolution followed by two dense blocks, which are in charge of dimension reduction, and are composed by four dilated convolutions, one convolution, and one deconvolution. The decoder section of the network uses one attention block, composed of one deconvolution, two convolutional layers, and a block with a total of five convolutional layers. This attention block is followed by a deconvolution, and a final convolutional layer returns the classification for each voxel in the 3D input. Due to limited computational power, the authors used the structure with patches of $32 \times 32 \times 32$ rather than the entire image itself. Overlap between patches were established to eight, the final concatenation results easy. Such small patch size allows a very fast processing. The model was compared with various existent methods, included the original U-Net. The method obtained better scores in the most part of performed tests, proving the architecture is adequate for such type of segmentation task. We think the model could be tested in a more powerful environment, using bigger patch size, or the entire

imaging. Also, the approach uses T1 and T2 image types for the segmentation, which can be limiting if only one of the modalities is available.

4.2.11 Thyreau and Taki

Thyreau and Taki [92] used a U-Net for creating a cortical parcellation of the whole human brain. The model takes as inputs a brain mask (obtained with any specialized software or library), an atlas of the brain, and the side for performing the parcellation (right or left). The model assumes brain symmetry, and works only on the left part of the brain. If the right part is to be segmented, the algorithm flips it, and processes it as the left side; finally performs a flip back. The model, based in U-Net, has a total of 21 convolutional layers. Besides analyses performed in the study were satisfactory, we consider that the approach could be expanded to work in both sides of the brain. Perhaps two models can be trained, one for each side. Also, the algorithm highly depends on the provided atlas, as well as a good brain mask segmentation.

4.2.12 Partial considerations

This time not all the approaches used 3D context for MRI processing, but near the half [57, 84, 86, 89, 92] of them did (five from 11). Pixel/voxel context was used by 5 teams as well [45, 74, 81, 85, 88], and only one of them [76] used 2D data. The use of 2D data to process 3D images can produce some lags in segmentations, as some important information can be unseen while the processing. Also, reconstructing 3D images (segmentation masks) from 2D data usually requires an extra postprocessing step, as the relation between slices is not considered during processing stage.

Pixel/voxel-oriented approaches tend to mix characteristics from 2 and 3D orders, and it can represent an advantage over 3D ones in terms of memory consumption, but if the images have a large resolution, it might take too much time to process a single MRI. Besides, these applications do not consider the whole set of features available which relates every single point to the rest of the points presents in an MRI. They only take small patches which contain information about near pixels/voxels (with different distances).

For fully consider every aspect on MRIs, the whole 3D object needs to be processed, even if it means a discrete increase in computation time.

Despite the best results reported were obtained by [89], reader must consider the number of MRIs analyzed. The tests were performed over iSeg-2017 dataset [90] which contains only 23 MRIs for training and testing, reaching a 0.95 DSC. In terms of the quantity of processed and tested MRIs, the deeper analysis was carried by Thyreau and Taki

[92], with a total of 2341 images. That large database was constructed by joining smaller datasets, which allows to include a big demographic and clinical variety, allowing to train models with great level of generalization. Also, the task performed by Thyreau and Taki [92] (cortical parcellation) is considerably harder than the task realized by Lei et al. [89] (segmenting WM, GM and CSF), which obtains masks with only four labels. This could be the origin of the differences in scores, but an exhaustive analysis must be carried on to fully compare approaches.

However, M-Net proposed by Mehta and Sivaswamy [76] is an interesting architecture to consider in future investigations. The M-Net is an evolution of U-Net architecture, incorporating at least two more feature sets in every layer of the network. Testing it in a 3D context could lead to good segmentations.

4.3 Brain lesion segmentation

The researches reviewed in this section are oriented to finding lesions on brain MRIs, specifically, finding the existence of brain tumors. Table 5 displays the summary.

4.3.1 Kamnitsas et al.

Kamnitsas et al. [93] performed a brain lesion segmentation task using a CNN with a total of 11 convolutional layers. The network classifies each voxel into its corresponding class (lesion or not). For each voxel, two patches are independently processed in the first layers. The first patch of shape $25 \times 25 \times 25$, passes through four convolutional layers, and the second patch of $19 \times 19 \times 19$ through five convolutions. Both patches have the same center, but the second is obtained from a downscale version of the original input. The outcomes of the two processing lines are concatenated and passed within two fully convolutional layers, and finally to the classification layer with two outputs. During testing phase, authors corroborate the importance of processing 3D context, even if the images have low resolution. An extension of the network was created, with five outputs instead of two: four tumor classes plus background. The system performed favorably compared to previous state-of-the-art methods, proving the approach is a correct way to segment brain lesions. A total of 445 images were used for the evaluations.

4.3.2 Havaei et al.

Havaei et al. [95] tested four distinct methods for brain tumor segmentation. The basis of the investigation runs over an implementation of a two-path CNN. Like [93], the network takes two patches of the image, which are independently

processed by branches of two convolutions and one convolution, respectively, and then are concatenated and passed through another convolution before the final output layer. The difference between this basic CNN, besides the number of convolutional layers, is the inclusion of another nonlinearity technique, called maxout [114], which, according to authors, has been very effective at modelling useful features. The other variants proposed by the authors are cascading this basic CNN after another simplest model, in three different parts: 1) at the input layer, giving a deeper input, 2) right after the first convolution, and 3) in the last convolutional layer, precisely before the final layer. The network works in a 2D context, analyzing the images' pixel by pixel. For each pixel, patches of 7×7 and 13×13 are used as inputs. This assures having local and global characteristics for each pixel. The algorithm with the best results was the first cascaded CNN (previous results concatenated at the input of the second convnet), and over scored some existent approaches until then. As the algorithm works at a 2D level, it does not present the same problem that previous pixel/voxel approaches on a 3D context. The system process images with a high speed, and obtains highly accurate results. The main limitation when analyzing the method is related to the image types. Authors used four image modalities for training: T1, T1C, T2, and Flair. This supposes a limitation for environments where not all modalities are available. Besides, volumetric analysis seems to be hard to perform if only 2D context is used. An extension to 3D processing could be implemented in future investigations. Also, only 30 images were used by authors. This could represent a limitation, because 30 images with data augmentation should not give much generalization to the models.

4.3.3 Chen et al.

Chen et al. [77] used a deeper structure for segmenting brain tumors in 2D images. The architecture relies on the FCN structure, but having dense connected blocks instead of standard convolutional layers. The encoder path starts with a single convolution, followed by four dense blocks, convolutions and pooling operations. The decoder consists of four residual inception blocks followed by unpooling blocks. Each dense connected block is based in DenseNet, and consists of three consecutive convolutional layers. The residual inception blocks are inspired in Inception modules, and are composed by seven convolutions and one dilated deconvolution. Finally, the unpooling blocks are formed by five deconvolutional layers and one dilated deconvolution. The network input consists on 2D images, and the outputs correspond to one sample for each label. Segmentations produced by model was compared in various modalities against standard FCN, U-Net, among others, obtaining a larger Dice score on evaluations in a dataset with 285 images. The main

limitation we see here is the use of various image modalities, (T1, T1-Gd, T2, and FLAIR), and the difficulty for performing a 3D volumetric analysis / reconstruction over the set of 2D slices segmented.

4.3.4 Cahall et al.

Cahall et al. [69] proposed the use of traditional U-Net architecture with inception modules instead of normal convolutional layers. The structure contains a total of four down-samples and up-samples stages, and every convolutional layer of the original U-Net has been replaced with an inception module, giving a total of nine modules. Each inception module is composed of seven convolutional layers, giving the network a total of 63 convolutions. The approach takes as input a 2D image, and returns one channel for each tumor label on their database. The used dataset was comprised of 285 MRIs. Results obtained by the approach are very promising, as accurately segment brain tumors on 2D images. The model was not compared with any existent method, but the study demonstrated that inception modules are an important characteristic of nowadays convnets, as they increment the number of parameters to analyze, obtaining more and more details to process and learn. The 2D processing, as well as the number of MR modalities used, are the main limitations of the approach.

4.3.5 Marcinkiewicz et al.

Marcinkiewicz et al. [70] proposed the use of two cascaded CNNs for the same purpose. Both architectures are based in U-Net, and are pretty the same, except for the last output layer. The first CNN is in charge of obtaining a binary mask of the tumor, and the second convnet performs the full segmentation of it. The main difference between the architecture and the original U-Net is that it possesses three input channels, which are independently processed during down-sample phase. In each step of down-sample, obtained features for each modality are concatenated, and passed as a partial result for the up-sample section. Image modalities used were T1-Gd, T2 and T2-FLAIR, belonging to a dataset consisting of 285 MRIs. This represents the same inconvenient that previous analyzed studies with various modalities, but the authors clarify that one single modality is not enough for tumor detection on MRIs. Inputs consist of a 2D slide for each modality, and as the output, one channel for each tumor part.

4.3.6 Qamar, Ahmad & Shen

Proposal by Qamar, Ahmad & Shen [71], like [69], is based on U-Net and inception modules for this task. The

main differences are that pooling layers were substituted by convolutional layers with strides, and a new conception of hyperdense inception modules, composed by two sequences of three convolutions, fully connected between them. The number of down-samples and up-samples in the architecture remains the same. Another important difference is the image shape: this approach is prepared to work in 3D images. As commented previously, 3D processing should be more reliable than 2D for volumetric analysis. Like normal U-Nets, the features produced at each encoder level are concatenated in the decoder section. The authors demonstrated that inception modules with hyperdense connections significantly improves the segmentation results, and the overall segmentation in all tested datasets were very promising. Nevertheless, the issue of multiple modalities being used still represents an inconvenient; we think the proposal may be tested using only one of the modalities available. As processing is performed in 3D way, the images must be registered in preprocessing stage, which may add an important time during processing. The experiments were carried on a dataset with 369 images.

4.3.7 Bala and Kant

Bala and Kant [96] arrive to us with another U-Net based model for brain tumor segmentation. As the original net, four down-sample and up-sample operations compose the network. The original convolutional layers were substituted by a modified inception module, composed of five convolutions and a concatenation layer. The most important change proposed by the authors was an addition in the skip connections. Skip connections are the concatenation of the partial outputs of encoder section, whose are used then in decoder section as an increment in the features. Each skip connection was processed within a dense path before the concatenation with the respective decoder step. The dense path is composed by four consecutive convolutions, and each convolution output is concatenated as the input of the following convolutional layers. There were 484 images used to perform training and tests. The main limitation for this approach is the same than previous ones: approach takes 2D slices from T1, T1-Gd, T2 and FLAIR modalities.

4.3.8 Partial considerations

Interestingly, none of the investigations addressed the 3D context for lesion segmentation. The most of them used 2D slices [69–71, 77, 96], and the other used voxel/pixel orientation [93, 95]. Segmenting brain lesions can be seen as a simpler task, as the central tissues on tumors and other affected parts tend to be easily identified; another reason for

not using 3D processing is that tumors and other affections can appear everywhere inside the head of the patients, and this represents an important constraint, as the whole MRI must be processed.

All investigations used a good number of MRIs. Except for [95] which used only 30 images from the BRATS [106, 115, 116] 2013 dataset, the rest of researchers used more than 200 images.

The most used architecture was U-Net, and from previous analysis we can conclude it is more convenient than usual CNNs and FCNs for image segmentation tasks. In fact, the higher scores on brain lesion segmentation were obtained by U-Nets [69] [96]. In both investigations, authors used inception modules in their architectures, making them capable of learning more parameters at each stage. We consider Bala and Kant's research [96] to be the best research in this group.

4.4 Image classification

The most of the applications found on this field have the purpose of identify specific diseases from brain MRIs. The investigations are applied mainly on Alzheimer's disease. In Table 6, the summary is presented.

4.4.1 Farooq et al.

Farooq et al. [97] adapted two well-known convolutional neural networks for classifying Alzheimer's disease on patients MRI. The first of these architectures is GoogleNet [117], a 22-layer model introduced in 2015 by Google. It was a very well-designed deep neural network, which improves utilization of the computing resources inside the network. The use of inception modules was introduced with this network too, with the idea of considering maximum information coming from the input path. The second architecture used for the classification over MRIs was ResNet [75], which uses the concept of skip connections added as bypass to convolutional layers of regular feed-forward networks, allowing the last layers the access to features produced in initial layers of the structure. For this second CNN, two models were trained, with 18 and 152 layers each (ResNet-18 and ResNet-152 respectively). Classification is performed in 2D slices of MRIs, returning the most probable class for each MRI. All models were tested on the same dataset, consisting of 355 MRIs. The networks were compared with other state-of-the-art approaches, resulting at the top in the next order: GoogleNet, ResNet-152 and ResNet-18.

4.4.2 Talo et al.

Talo et al. [99] used transfer learning and fine-tuning to incorporate the convolutional base of ResNet-34 [75] into their architecture. Transfer learning consists on using

features learned by a previous model, and apply those features in a new domain. Using transfer learning with the core of any CNN, implies training only the new layers added to the model, which represents less resource consumption. Fine-tuning, in short, consists of adjusting the weights of a pretrained model for the new requirements. In this case, the layers from ResNet-34 were fine-tuned. The overall architecture of the model consists of 34 convolutional layers (ResNet-34) followed by a densely connected layer. The output mask returns if the image represents a normal or abnormal class. The inclusion of techniques like fine-tuning and transfer learning conducted the authors to establish a new training pipeline, consisting of 50 epochs with three stages: 1) training the densely connected layer for the transfer learning for three epochs, 2) training the densely connected layer for seven more epochs, with the inclusion of data augmentation, and 3) train the whole architecture for 40 more epochs, for the fine-tuning of convolutional layers' parameters. The algorithm reaches a 100% accuracy in the test dataset, which means that the training pipeline has taken the desired effect. The pretrained convolutional base from ResNet-34 made it possible to use a few training epochs. The approach works in 2D context, processes images in with high speed and accuracy. In the test stage, it outperformed almost every research included in the author's review, with tests made across a dataset consisting of 613 images. We think the use of only 50 epochs for the final training could lead to classification errors.

4.4.3 Yiğit & Işık

Yiğit & Işık [101] compared three different architectures for classifying Alzheimer's disease MRIs. The first model consisting of three convolutional layers, two pooling and two fully connected layers. The second model consists of two convolutional layers, two pooling and three fully connected layers, and the third architecture formed by five convolutions, three pooling and four fully connected layers. One of the goals of the study was to compare different number of parameters on models. The analysis of MRIs, as two previous researches [97, 99], takes 2D slices from MRIs as input, and obtains the most probable class for the whole MRI. The union of two datasets allowed the authors to prove the architectures in more than 1000 MRIs. As an important conclusion, the study revealed that diagnosis of patients with mild cognitive impairment is more difficult than that of patients with Alzheimer's disease. The third model was the most successful, and resulted competitive compared with existent techniques.

4.4.4 Partial considerations

This time all researchers used 2D context. It results logical, as no volume is created after the classification. Besides, as

previously commented, 2D processing can be done much faster than 3D processing. However, approaches could be tested in a 3D environment as well.

The best results were obtained by [99] in a binary classification, while the rest of approaches [97, 101] attempt to predict the Alzheimer's Disease stage from MRIs.

CNN used by [101] is based in AlexNet architecture, which is simpler than ResNet and can be trained faster, but the difference between results obtained in [101] and [97] suggest than the most adequate architecture should be ResNet or GoogleNet. Of course, the difference between the training sets might have an impact in the results too, as the first [97] was trained and tested with 355 MRIs, while Yiğit & Işık [101] approach was evaluated with more than 1000 images.

The most used architecture was ResNet, which allows to create deeper convnets while maintaining a manageable complexity. However, GoogleNet, was the top-ranked model [97]. GoogleNet architecture allows to create bigger structures with a minimal number of parameters. The use of inception modules permits to learn different features at every step, which can be translated in a more accurate classification. For that reason, we consider GoogleNet to be the best structure for image classification within our review.

5 Discussion

The studies here reviewed are only a part of the whole set of researches regarding brain magnetic resonance processing with convolutional neural networks. We think the most relevant researches are those aiming the cerebellum parcellation task [44, 64, 68]. Those approaches are oriented to obtain a fully parcellation of the human cerebellum in all its lobules, giving the possibility of volumetric analysis at lobe level. The fact that the six described approaches perform segmentation over 3D images, makes them more reliable, as they can easily take advantage of the relationship between different slices of the MRIs, obtaining more accurate volumes. Besides, the more used structure for this task was the U-Net architecture, famous by its multiple-scale feature treatment. Those architectures could be used for comparison between different stages of SCA disease, allowing to determine the total or local atrophy over a time frame. Correlations between the patient's symptoms and the current atrophy level can be done, giving more details about the state of the patient. Also, we think the approaches could help in early detection of some polyQ SCAs, which are known to deteriorate the cerebellum of the presymptomatic carriers before the disease onset.

Specifically, we consider the research of Han et al. [64] to be ranked the best in our review, as the overall dice score obtained in their studies were above 0.95, which means a

Table 7 Number of approaches oriented to pixel/voxel, 2D and 3D processing

Level	C	WBC	CSWGM	DBS	CS	BL	Total
Pixel/Voxel	0	3	2	0	0	2	7
2D	0	0	0	1	5	4	10
3D	6	3	1	1	0	1	12
Total	6	6	3	2	5	7	29

C cerebellum, WBC whole brain including cerebellum, CSWGM cerebrospinal fluid, white matter and gray matter, DBS deep brain sections, CS classification, BL brain lesions

very good segmentation over all cerebellar lobules. Also, not only lobule level was tested by the authors, but a hierarchical parcellation was tested too. Finally, the authors made the source code public, so it can be used for comparisons against further coming approaches.

Approach by Kim et al. [73] results an adequate approach for being applicable on deep cerebellar nuclei segmentation. Deep cerebellar nuclei, as part of the cerebellum, is evidently damaged by spinocerebellar ataxias. SCA6, which is known to be mainly oriented to cerebellar damages, severely affects the cerebellar nuclei. Having a tool for quantifying such atrophy could be useful on disease characterization.

The second group, oriented to anatomic structures segmentation of the brain, can be divided into three main subgroups:

- 1) Whole brain segmentation. The objective of such methods is to obtain an accurate segmentation of the main cerebral structures, including the cerebellum, and sometimes the brainstem. Approaches by [45, 81, 88] work at a pixel / voxel level, while [84, 92] were created for 3D imaging processing.
- 2) Cerebrospinal fluid, white matter, and gray matter. Proposals [85] and [74] work at pixel and voxel level, respectively, while [86] and [89] use a 3D approach for segmenting the MRI.
- 3) Deep brain structures. The architecture described in [57] uses a 3D CNN model, while [76] used a modified 2D U-Net model.

All these approaches can have big application on SCA diagnosis, as all they perform, in a manner or another, segmentation over various brain tissues, which are damaged by SCA, sometimes severely. From the 11 reviewed researches in this group, five [45, 74, 81, 85, 88] were oriented to pixel / voxel classification, only one [76] uses 2D processing, and five [57, 84, 86, 89, 92] perform the segmentation by the whole 3D imaging. The most used architecture was SegNet, and the top best results were reported by the U-Net of Lei et al. [89]. Our final decision on this group goes to the CNN used by Thyreau et al. [92]. However, the M-NET structure, proposed by Mehta and Sivaswamy [76], seems to be a good choice for image segmentation, as an evolution of U-Net. We

think this architecture should be tested in other environments for testing its feasibility.

Third group is composed by all investigations regarding the brain lesion segmentation. Of the seven reviewed approaches, two [93, 95] were oriented to voxel classification, four [69, 70, 77, 96] were modelled as 2D slice segmentation, and one [71] was oriented to 3D imaging processing, being U-Net the most common used architecture. Lesion segmentation is not directly applicable to our goals: usually the finding of non-SCA patients due to tumors is a relatively easy task. However, those models are interesting to us, and their behavior should be tested in other scenarios closer to ours, using the proper training stage. Specifically, we consider that the best approaches are those used by Cahall et al. [69] and Bala et al. [96], as they demonstrated that inception modules may represent an improvement on medical image segmentation.

Finally, the investigations oriented to MRI classification, have a great importance for our task. Detecting when an image belongs to certain class of Alzheimer's disease can be pretty hard, but counting with tools which help in that task, can save much time to specialists. These approaches [97, 99, 101], could all be trained with images from SCA patients, and tested for their classification. Also, information retrieval systems could be built upon models alike, where specialists enter a person's MRI and get a classification about the type of SCA (if present), with a group of images visually similar to the input. Perhaps combination of CNN with common machine learning techniques could be incorporated for this goal. All the approaches were constructed to work in a 2D level, and we think it is not necessary to process images in 3D for classification matters. The most convenient architecture for this task could be the GoogleNet used by Farooq et al. [97], taking advantage of inception modules.

The wide range of structures used among the reviewed investigations, from pixel to 3D level, makes the following question arise: which kind of processing should be more reliable for SCA diagnosis from MRIs? Table 7 shows a summary of the number of used approaches at pixel/voxel, 2D and 3D levels, respectively.

From a total of 29 architectures reviewed, seven were prepared for pixel/voxel processing, 10 for 2D, and 12 for 3D. Generally, methods voxel/pixel-oriented work faster

than 2D and 3D ones. Relying solely on the classification of individual points leads to very fast segmentations. Equally, 2D methods work much faster than 3D. The same situation is presented regardless of memory consumption. The more complex is the basic structure being analyzed, the more resources are needed for processing it. Nevertheless, we think that volumetric analysis demands the use of architectures prepared to work in the whole resonance imaging, rather than dividing it into smaller patches. The use of fewer dimensions has one major disadvantage over 3D: anatomic context in the directions orthogonal to the image plane are almost entirely discarded [57]. This was demonstrated in [93], where authors compared 2D and 3D versions of a CNN for brain lesion segmentation, obtaining important improvements with the 3D version. Also, we can see in Table 4 that all approaches oriented to cerebellum processing are built over 3D architectures. For those reasons, we consider 3D processing as the most useful kind of analysis, despite their high-consumption characteristic.

Another important question to consider is which architecture seems to be more suitable for our task. As we are mainly interested on volumetric analysis, which must mainly imply

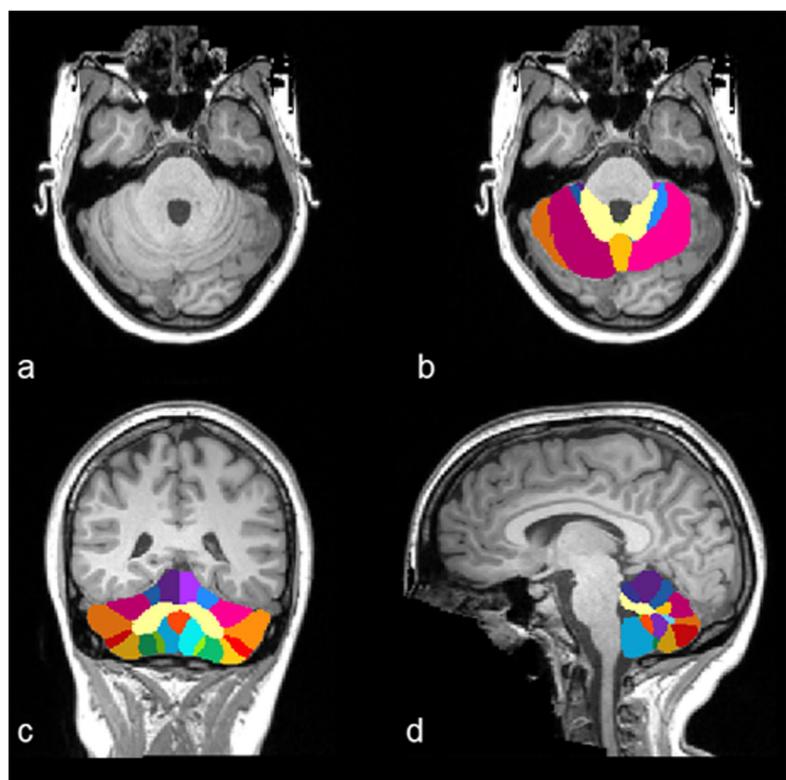
the patient's cerebellum, we consider that the U-Net architecture should be tested, as this was the most used structure over all the researches.

5.1 Main limitations for application over SCA diagnose

Although CNNs generally do not need preprocessing [66], it could be necessary to perform it. In terms of automatic image processing, an important problem to consider is the quality. Sometimes, image quality is not good enough for an algorithm to correctly learn all necessary parameters. For example, low-resolution images may not contain the necessary number of features for a good DL algorithm, and the presence of dark or blur spots could difficult edge detection, making harder the feature identification. The best assurance we can provide is to perform a preprocessing stage with one or more steps, in order to obtain better images.

Two well-known preprocessing techniques are skull stripping and bias field correction. Skull removal involves removing all non-brain tissue from the MRI (eyes, fat, skull). It has been shown that it is sometimes not as effective, and

Fig. 3 Images from a human cerebellum parcelated from MRI. Original axial slice (a), followed by axial (b), coronal (b) and sagittal (c) slices of the parcelation. Image generated with free distribution of ACA-PULCO [64]



parts of the cerebellum can be removed from the MRI [64]. Bias field correction (BFC) [118, 119], on the other hand, is the process of reducing the noise that an acquired image may have. This noise can be caused by some defect, and blurs the image, so that some contours are hard to identify (Fig. 3).

Figure 4 shows an example of MRI preprocessing stage consisting of the techniques mentioned above (bias field correction, skull stripping and histogram equalization). Taking a closer view, the result of applying BFC returns a clearer image, which means more clear boundaries between the different elements, and helps on automatic edges detection. For the example, images were generated using the open-source software Mango (<http://ric.uthscsa.edu/mango/>), as well as skull stripping, and bias field correction was done using Brainsuite (<http://brainsuite.org/>).

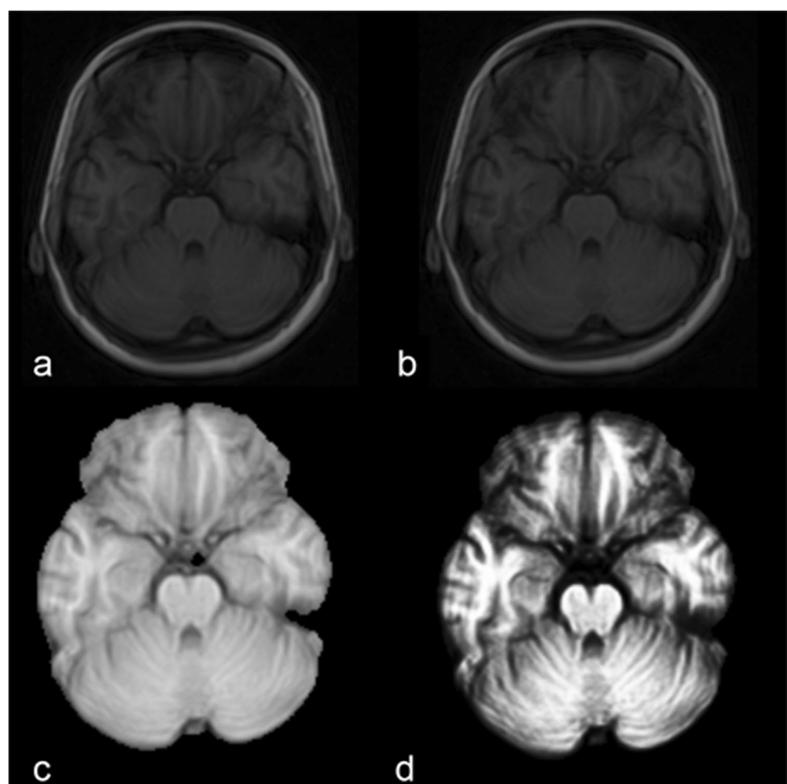
CNNs do not generalize well to previously unseen object classes that are not present in the training set [120]. This situation limits their performance to segment objects for which annotations are not available in the training state. In our case, a convnet trained for cerebellum (or brainstem) lesion identification on SCA patient's MRIs, would not achieve desired results if those features are not well represented for training phase. This situation represents another important

challenge to consider: a comprehensive dataset, so we may assure a higher generalization (ability to correctly identify unseen images), and finally a good segmentation result.

Ideally, the most data can be used for training a DL system, the more accurate its predictions should be. But, for medical image processing, the number of available images is usually very small [121], depending on the disease and the existence of previous studies. Increasing the network's depth seems to be one solution if not enough data is available.

That way, the number of estimated parameters increases, too. This allows our system to “learn more from data”: a big number of parameters tends to be related with a better system behavior, or a better feature identification [60], but also needs more training time. Usually, an FCN with about two or three million parameters can be successfully trained on 3D MRI processing using a CPU with at least 8 GB of RAM memory; in this case, the process may take about three weeks. This is a very long time if compared with a GPU system, which needs no more than three or four hours to complete the same training. But, if this number of parameters grows too much (say, 300 million parameters), the model can become not computable, even with modern GPU hardware. With that enormous parameter

Fig. 4 Example of preprocessing over MRI. Original MRI slice (a), bias field corrected image (b), skull stripping (c), and histogram equalization (d)



configuration, the machine would be able to find more characteristics, but training could take even months.

In the analysis phase, they take advantage over manual labelling and segmenting: a manual segmentation may take several hours, while an automatic one can be performed in seconds. But sometimes the training process needs to be repeated several times, in order to find the correct configuration or architecture. This represents a potential limitation, when the investigator must find if increasing or not the training time will have some important significance in achieved results. For that reason, the correct depth must be found, as well as the number of training samples.

Another common way to increase our net's accuracy on small datasets is the data augmentation [65, 96, 122]. This technique increases the training data by artificially generating more data to generalize the model. A single rotation on an image results in a new one totally different for computational purposes. In other words, new data is created from the available data. This technique has demonstrated very useful when used for medical image processing, as the number of samples is usually small.

Transfer learning may be applied for reducing training time too. It has been shown that the technique can be successfully applied on classification tasks [99].

All previously discussion allows arriving to the conclusion that ConvNets can be a viable tool for PolyQ SCAs characterization. As CNNs have been successfully applied on anatomic brain structures segmentation, they can be adapted for those organs mainly affected by ataxias. The data augmentation should improve the behavior of algorithms, based in previous applications, giving the possibility of greater generalization with a few training images. Considering the available computational resources, a well-designed structure can be created for the processing of the whole MRI, taking advantage of the relationship between slices to produce better segmentations with a simpler structure.

Not only segmentation algorithms can be tested, classification networks may be constructed for discerning between different types of SCA. We must have in mind, of course, that computers will never substitute the expert's opinion; the models, however, may serve as a practical tool on helping specialists finding the main characteristics of the disease. The use of such tools, may represent an improvement on the characterization of SCAs, as they might improve the speed of the process.

6 Conclusions

PolyQ SCAs are hereditary diseases affecting mainly the cerebellum and brainstem of patients. The use of image studies over brain MRIs has been useful for finding characteristic degeneration patterns. Several investigations have

been conducted to automatically process brain MRIs. Most commonly used in later years, convolutional neural networks have been used for this task. CNNs can automatically identify images structures on local and global levels, without relying on features created by humans, which makes them to take advantage over another approaches. The use of such techniques as part of visualization and analysis tools represents a set of new opportunities on SCAs treatment. Having in mind that a computer algorithm will never substitute a human specialist, the systems could be used as auxiliary tools for identifying characteristic degenerative patterns, taking a closer view of changes on the brain of patients, and a better understanding of the diseases. Segmenting algorithms can be used for volumetric analysis of the cerebellum and brainstem as well as cortex, and classification approaches are a very important instrument for information retrieval systems.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Dueñas AM, Goold R, Giunti P (2006) Molecular pathogenesis of spinocerebellar ataxias. *Brain* 129:1357–1370. <https://doi.org/10.1093/brain/awl081>
- Stevanin G, Brice A (2008) Spinocerebellar ataxia 17 (SCA17) and Huntington's disease-like 4 (HDL4). *Cerebellum*. 7. <https://doi.org/10.1007/s12311-008-0016-1>
- Manto M (2005) The wide spectrum of spinocerebellar ataxias (SCAs). *The Cerebellum* 4:2–6. <https://doi.org/10.1080/14734220510007914>
- Teive HAG (2009) Spinocerebellar ataxias. *Arq Neuropsiquiatr* 67:1133–1142. <https://doi.org/10.1590/S0004-282X20090006000035>
- Paulson HL, Shakkottai VG, Clark HB, Orr HT (2017) Polyglutamine spinocerebellar ataxias — from genes to potential treatments. *Nat Publ Gr* 18:613–626. <https://doi.org/10.1038/nrn.2017.92>
- Klaes A, Reckziegel E, Franca MC, Rezende TJ, Vedolin LM, Jardim LB, Saute JA (2016) MR Imaging in Spinocerebellar Ataxias: A Systematic Review. *Am J f Neuroradiol* 37:1405–1412
- Van De Warrenburg BPC, Sinke RJ, Bemelmans CCV, Scheffer H (2002) Spinocerebellar ataxias in the Netherlands. *Neurology*. 58:702–708. <https://doi.org/10.1212/WNL.58.5.702>
- Erichsen AK, Koht ÁJ, Stray-pedersen ÁA, Abdelnoor M, Tikkelsen CME (2009) Prevalence of hereditary ataxia and spastic paraparesis in southeast Norway: a population-based study. *Brain* 132:1577–1588. <https://doi.org/10.1093/brain/awp056>
- Reetz K, Rodríguez R, Dogan I, Mirzazade S, Romanzetti S, Schulz JB, Cruz-Rivas EM, Alvarez-Cuesta JA, Aguilera Rodríguez R, Gonzalez Zaldivar Y, Auburger G, Velázquez-Pérez L (2018) Brain atrophy measures in preclinical and manifest spinocerebellar ataxia type 2. *Ann Clin Transl Neurol* 5:128–137. <https://doi.org/10.1002/acn3.504>

10. Bang OY, Lee PH, Kim SY, Kim HJ, Huh K (2004) Pontine atrophy precedes cerebellar degeneration in spinocerebellar ataxia 7: MRI-based volumetric analysis. *J Neurol Neurosurg Psychiatry* 75:1452–1457. <https://doi.org/10.1136/jnnp.2003.029819>
11. Buijsen RAM, Toonen LJA, Gardiner SL, Roon-Mom WMC (2019) Genetics, Mechanisms, and Therapeutic Progress in Polyglutamine Spinocerebellar Ataxias. *Neurotherapeutics* 16:263–286. <https://doi.org/10.1007/s13131-018-00696-y>
12. Nethisinghe S, Lim WN, Ging H, Zeitzberger A, Abeti R, Pemble S, Sweeney MG, Labrum R, Cervera C, Houlden H, Rosser E, Limousin P, Kennedy A, Lunn MP, Bhatia KP, Wood NW, Hardy J, Polke JM, Veneziano L, Brusco A, Davis MB (2018) Complexity of the Genetics and Clinical Presentation of Spinocerebellar Ataxia 17. *Front Cel Neurosci* 12:1–10. <https://doi.org/10.3389/fncel.2018.000429>
13. Taroni F, Didonato S (2004) Pathways to motor incoordination: the inherited ataxias. *Neuroscience* 5:641–655. <https://doi.org/10.1038/nrn1474>
14. Velázquez L, García R, Santos N, Panque M, Medina E, Hechavarria R (2001) Las ataxias hereditarias en Cuba. Aspectos históricos, epidemiológicos, clínicos, electrofisiológicos y de neurología cuantitativa. *Rev Neurol.* 32:71. <https://doi.org/10.33588/rn.3201.2000283>
15. Alex J, Saute M, Jardim LB (2015) Machado Joseph disease : clinical and genetic aspects, and current treatment. *Expert Opin Orphan Drugs* 3:517–535. <https://doi.org/10.1517/21678707.2015.1025747>
16. Toyoshima Y, Onodra O, Yamada M, Tsuji S, Takahashi H (2019) Spinocerebellar Ataxia Type 17. In: Adam MP, Ardinger HH and Pagon RA (eds.) *GeneReviews®* [Internet]. University of Washington, Seattle
17. Lasek K, Lencer R, Gaser C, Hagenah J, Walter U, Wolters A, Kock N, Steinlechner S, Nagel M, Zühlke C, Nitschke M, Brockmann K, Klein C, Rolfs A, Binkofski F (2006) Morphological basis for the spectrum of clinical deficits in spinocerebellar ataxia 17 (SCA17). *Brain* 129:2341–2352. <https://doi.org/10.1093/brain/awl148>
18. Mascalchi M, Vella A (2018) Neuroimaging Applications in Chronic Ataxias. *Int Rev Neurobiol* 143:109–162. <https://doi.org/10.1016/bs.irn.2018.09.011>
19. Baldarçara L, Currie S, Hadjivassiliou M, Hoggard N, Jack A, Jackowski AP, Mascalchi M, Parazzini C, Reetz K, Righini A, Schulz JB, Vella A, Webb SJ, Habas C (2014) Consensus Paper : Radiological Biomarkers of Cerebellar Diseases. *Cerebellum* 14:175–196. <https://doi.org/10.1007/s12311-014-0610-3>
20. Kumar SD, Chand RP, Gururaj AK, Jeans WD (1995) CT features of olivopontocerebellar atrophy in children. *Acta radiol* 36:593–596. <https://doi.org/10.1177/028418519503600458>
21. Meira AT, Arruda WO, Ono SE, Neto ADC, Raskin S, Camargo CH, Teive HAG (2019) Neuroradiological Findings in the Spinocerebellar Ataxias. Tremor and Other Hyperkinetic Movements. 1–8. <https://doi.org/10.7916/tohm.v.0.682>
22. Seidel K, Siswanto S, Brunt ERP, Den Dunnens W, Korf HW, Rüb U (2012) Brain pathology of spinocerebellar ataxias. *Acta Neuropathol* 124:1–21. <https://doi.org/10.1007/s04041-012-1000-x>
23. Kim Y, Kondo M, Sunami Y, Kawata A (2014) MRI Findings in Spinocerebellar Ataxias. *J Neurol Disord Stroke* 2:1072
24. Schulz JB, Borkert J, Wolf S, Schmitz-hübsch T, Rakowicz M, Mariotti C, Schoels L, Timmann D, Van De Warrenburg B, Dürr A, Pandolfo M, Kang J, González A, Nägele T, Grisoli M, Boguslawska R, Bauer P, Klockgether T, Hauser T (2010) Visualization, quantification and correlation of brain atrophy with clinical symptoms in spinocerebellar ataxia types 1, 3 and 6. *Neuroimage*. 49:158–168. <https://doi.org/10.1016/j.neuroimage.2009.07.027>
25. Reetz K, Costa AS, Mirzazade S, Lehmann A, Juzek A, Rakowicz M, Boguslawska R, Schöls L, Linemann C, Mariotti C, Grisoli M, Dürr A, Van De Warrenburg B, Timmann D, Pandolfo M, Bauer P, Jacobi H, Hauser T, Klockgether T, Schulz JB (2013) Genotype-specific patterns of atrophy progression are more sensitive than clinical decline in SCA1, SCA3 and SCA6. *Kathrin*. <https://doi.org/10.1093/brain/aws369>
26. Inagaki A, Iida A, Matsubara M, Inagaki H (2005) Positron emission tomography and magnetic resonance imaging in spinocerebellar ataxia type 2: A study of symptomatic and asymptomatic individuals. *Eur J Neurol* 12:725–728. <https://doi.org/10.1111/j.1468-1331.2005.01011.x>
27. Nave RD, Ginestroni A, Tessa C, Cosottini M, Giannelli M, Salvatore E, Sartucci F, Michele G. De, Dotti MT, Piacentini S, Mascalchi M (2008) Brain Structural Damage in Spinocerebellar Ataxia Type 2. A Voxel-Based Morphometry Study. *Mov Disord*. 23:899–903. <https://doi.org/10.1002/mds.21982>
28. Schöls L, Bauer P, Schmidt T, Schulte T, Riess O (2004) Spinocerebellar ataxias Review Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis. *Decubitus* 3:291–304. <https://doi.org/10.4324/9780429456916-3>
29. Murata Y, Yamaguchi S, Kawakami H, Imon Y, Maruyama H, Sakai T, Kazuta T, Ohtake T, Nishimura M, Saida T, Chiba S, Oh-I T, Nakamura S (1998) Characteristic magnetic resonance imaging findings in Machado-Joseph disease. *Arch Neurol* 55:33–37. <https://doi.org/10.1001/archneur.55.1.33>
30. Soong B, Paulson HL (2007) Spinocerebellar ataxias: An update. *Curr Opin Neurol* 33:150–160. <https://doi.org/10.1097/WCO.00000000000000774>
31. Moriarty A, Cook A, Hunt H, Adams ME, Cipolotti L, Giunti P (2016) A longitudinal investigation into cognition and disease progression in spinocerebellar ataxia types 1, 2, 3, 6, and 7. *Orphanet J Rare Dis* 11:1–9. <https://doi.org/10.1186/s13023-016-0447-6>
32. Mariotti C, Alpini D, Fancelli R, Soliveri P, Grisolo M, Ravaglia M, Lovati C, Fetoni V, Giaccone G, Castucci A, Taroni F, Gellera C, Donato SD (2007) Spinocerebellar ataxia type 17 (SCA17): Oculomotor phenotype and clinical characterization of 15 Italian patients. *J Neurol* 254:1538–1546. <https://doi.org/10.1007/s00415-007-0579-7>
33. Zhang J, Gu W, Hao Y, Chen Y (2013) Spinocerebellar ataxia 17 : Inconsistency between phenotype and neuroimage findings. *16*, 703–704. <https://doi.org/10.4103/0972-2327.120457>
34. Loy CT, Epi MC, Sweeney MG, Davis MB, Wills AJ, Sawle GV, Lees AJ, Tabrizi SJ (2005) Spinocerebellar Ataxia Type 17: Extension of Phenotype With Putaminal Rim Hyperintensity on Magnetic Resonance Imaging. *Mov Disord* 20:1521–1528. <https://doi.org/10.1002/mds.20529>
35. Carroll LS, Massey TH, Wardle M, Peall KJ (2018) Dentatorubral-pallidoluysian Atrophy : An Update. *Tremor and Other Hyperkinetic Movements*. 8, <https://doi.org/10.7916/D81N9HST>
36. Sugiyama A, Sato N, Nakata Y, Kimura Y, Enokizono M (2017) Clinical and magnetic resonance imaging features of elderly onset dentatorubral – pallidoluysian atrophy. *J Neurol* 265:322–329. <https://doi.org/10.1007/s00415-017-8705-7>
37. Shao F, Xie X (2013) An overview on interactive medical image segmentation. *Annals of the BMWA* 2013(7):1–22
38. Van Der Lijn F, De Bruijne M, Hoogendam YY, Klein S, Hameeteman R, Breteler MMB, Niessen WJ (2009) Cerebellum segmentation in MRI using atlas registration and local multi-scale image descriptors. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. pp. 221–224
39. Saeed N, Puri BK (2002) Cerebellum segmentation employing texture properties and knowledge based image processing : applied to normal adult controls and patients. *Magn Reson Imaging* 20:425–429

40. Cardoso MJ, Melbourne A, Kendall GS, Modat M, Robertson NJ, Marlow N, Ourselin S (2013) AdaPT : An adaptive preterm segmentation algorithm for neonatal brain MRI. *Neuroimage* 65:97–108. <https://doi.org/10.1016/j.neuroimage.2012.08.009>
41. Makris N, Hodge SM, Haselgrave C, Kennedy DN, Dale A, Fischl B, Rosen BR, Harris G, Caviness VS, Schmahmann JD (2003) Human Cerebellum : Surface-Assisted Cortical Parcellation and Volumetry with Magnetic Resonance Imaging. *J Cogn Neurosci* 15:584–599
42. Romero J, Coupé P, Giraud R, Ta V, Fonov V, Park MT, Chakravarthy M, Voineskos A, Manjón J (2016) CERES : A new cerebellum lobule segmentation method. *Neuroimage* 147:916–924
43. Diedrichsen J (2006) A spatially unbiased atlas template of the human cerebellum. *Neuroimage* 33:127–138. <https://doi.org/10.1016/j.neuroimage.2006.05.056>
44. Carass A, Cuzzocreo JL, Han S, Hernandez-castillo CR, Rasser PE, Ganz M, Beliveau V, Dolz J, Ayed IB, Desrosiers C, Thyreau B, Fonov VS, Collins DL, Ying SH, Onyike CU, Landman BA, Mostofsky SH, Thompson PM, Prince JL (2018) Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. *Neuroimage*, 183:150–172. <https://doi.org/10.1016/j.neuroimage.2018.08.003.Comparing>
45. De Brébisson A, Montana G (2015) Deep Neural Networks for Anatomical Brain Segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition works. pp. 20–28
46. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to digit recognition. *Neural Comput* 1:541–551
47. Zeiler MD, Fergus R (2014) Visualizing and Understanding Convolutional Networks. *Anal Chem Res* 12:818–833. <https://doi.org/10.1016/j.ancre.2017.02.001>
48. Fukushima K (1980) Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. *Biol Cybern* 36:193–202
49. Hariharan B, Arbeláez P, Bourdev L, Maji S, Malik J (2011) Semantic Contours from Inverse Detectors. In: International Conference on Computer Vision. pp. 991–998
50. Wang L, Ouyang W, Wang X, Lu H (2015) Visual Tracking with Fully Convolutional Networks. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 3119–3127
51. Jaroensri R, Zhao A, Balakrishnan G, Lo D, Schmahmann JD, Durand F, Guttig J (2017) A Video-Based Method for Automatically Rating Ataxia. In: Proceedings of Machine Learning. pp. 1–13
52. El Amin AM, Liu Q, Wang Y (2016) Convolutional neural network features based change detection in satellite images. First International Workshop on Pattern Recognition. 10011. pp. 181–186
53. Kawahara C, Brown CJ, Miller SP, Booth BG, Chau V, Grunau RE, Zwicker JG, Hamarneh G (2017) BrainNetCNN : Convolutional Neural Networks for Brain Networks Towards Predicting Neurodevelopment. *Neuroimage* 146:1038–1049
54. Stoean C, Stoean R, Atencia M, Abdar M, Velázquez-Pérez L, Khosrabi A, Nahavandi S, Acharya UR, Joya G (2020) Automated Detection of Presymptomatic Conditions in Spinocerebellar Ataxia Type 2 Using Monte Carlo Dropout and Deep Neural Network Techniques with Electrooculogram Signals. *Sensors* 20:3032. <https://doi.org/10.3390/s20113032>
55. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang L, Wang G, Gai J, Chen T (2018) Recent Advances in Convolutional Neural Networks. *Pattern Recognit* 77:354–377
56. Milletari F, Navab N, Ahmadi SA (2016) V-Net: Fully convolutional neural networks for volumetric medical image segmentation. Proc. - 2016 4th Int. Conf. 3D Vision, 3DV 2016. 565–571. <https://doi.org/10.1109/3DV.2016.79>
57. Milletari F, Ahmadi SA, Kroll C, Plate A, Rozanski V, Maiostre J, Levin J, Dietrich O, Ertl-Wagner B, Bötzell K, Navab N (2016) Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Comput Vis Image Underst* 164:92–102. <https://doi.org/10.1016/j.cviu.2017.04.002>
58. Wu J (2017) Introduction to Convolutional Neural Networks. https://web.archive.org/web/20180928011532/https://cs.nju.edu.cn/wujx/teaching/15_CNN.pdf. Accessed 23 July 2022
59. Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet Classification with Deep Convolutional Neural networks. *Comm ACM* 60:84–90. <https://doi.org/10.1101/9781420010749>
60. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. 3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc. 1–14
61. Zeiler MD, Taylor GW, Fergus R (2011) Adaptive deconvolutional networks for mid and high level feature learning. *Proc IEEE Int Conf Comput Vis.* 2018–2025. <https://doi.org/10.1109/ICCV.2011.6126474>
62. Yaseen AF (2018) A Survey on the Layers of Convolutional Neural Networks. *Int J Comput Sci Mob Comput* 7:191–196
63. Ashqar BAM, Abu-Naser SS (2019) Identifying Images of Invasive Hydrangea Using Pre-Trained Deep Convolutional Neural Networks. *Int J Acad Eng Res* 3:28–36
64. Han S, Carass A, He Y, Prince JL (2020) Automatic Cerebellum Anatomical Parcellation using U-Net with Locally Constrained Optimization. *IEEE Trans. Med. Imaging.* 116819. <https://doi.org/10.1016/j.neuroimage.2020.116819>
65. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
66. Hariharan B, Arbeláez P, Girshick R, Malik J (2014) Simultaneous detection and segmentation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 297–312
67. Pathak D, Shelhamer E, Long J, Darrell T (2014) Fully Convolutional Multi-Class Multiple Instance Learning. *arXiv Prepr. arXiv1412.7144*
68. Han S, He Y, Carass A, Ying SH, Prince JL (2019) Cerebellum Parcellation with Convolutional Neural Networks. *Proc SPIE Int Soc Opt Eng.* 10949, <https://doi.org/10.1117/12.2512119.Cerebellum>
69. Cahall DE, Rasool G, Bouaynaya NC, Fathallah-Shaykh HM (2019) Inception Modules Enhance Brain Tumor Segmentation. *Front Comput Neurosci* 13:1–8. <https://doi.org/10.3389/fncom.2019.00044>
70. Marcinkiewicz M, Nalepa J, Lorenzo PR, Dudzik W, Mrukwa G (2019) Segmenting Brain Tumors from MRI Using Cascaded Multi-modal U-Nets. *Springer Nat Switz* 2:13–24. <https://doi.org/10.1007/978-3-030-11726-9>
71. Qamar S, Ahmad P, Shen L (2020) HI-Net: Hyperdense Inception 3 D UNet for Brain Tumor Segmentation. In: International Conference on Medical Image Computing and Computer Assisted Intervention. pp. 1–9
72. Huang G, Van Der Maaten L, Weinberger KQ (2017) Densely Connected Convolutional Networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708
73. Kim J, Patriat R, Kaplan J, Solomon O, Harel N (2020) Deep Cerebellar Nuclei Segmentation via Semi-Supervised Deep Context-Aware Learning from 7T Diffusion MRI. *IEEE Access* 8:101550–101568. <https://doi.org/10.1109/ACCESS.2020.2998537>

74. Gottapu RD, Dagli CH (2018) DenseNet for Anatomical Brain Segmentation. *Procedia Comput Sci* 140:179–185. <https://doi.org/10.1016/j.procs.2018.10.327>
75. He K, Zhang X, Ren S, Sun J (2016) Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9
76. Mehta R, Sivaswamy J (2017) M-NET : A Convolutional Neural Network for Deep Brain Structure Segmentation. In: 2017 IEEE International Symposium on Biomedical Imaging. pp. 437–440
77. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D (2018) DRINet for Medical Image Segmentation. *IEEE Trans Med Imaging* 37:1–11
78. Landmann BA, Hyang AJ, Gifford A, Vikram DS, Lim IAL, Farrell JAD, Bogovic JA, Hua J, Chen M, Jarso S, Smith SA, Joel S, Mori S, Pekar JJ, Barker PB, Prince JL, van Zijl PCM (2012) Multi-parametric neuroimaging reproducibility: A 3T resource study. *Neuroimage* 54(4):2854–2866
79. Internet Brain Segmentation Repository (IBSR). <https://www.nitrc.org/projects/ibsr>. Accessed 29 Oct 2022
80. Asman A, Akhondi-Asl A, Wang H, Tustison N, Avants B, Warfield S, Landman B (2013) Miccai 2013 segmentation algorithms, theory and applications (SATA) challenge results summary. In: MICCAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications (SATA)
81. Mehta R, Majumdar A, Sivaswamy J (2017) BrainSegNet: a convolutional neural network architecture for automated segmentation of human brain structures BrainSegNet : a convolutional neural network architecture for automated segmentation of. *J Med Imaging*, 4. <https://doi.org/10.1117/1.JMI.4.2.024003>
82. Shattuck DW, Mirza M, Adisetiyo V, Hojatkashani C, Narr KL, Poldrack RA, Bilder RM, Arthur W (2008) Construction of a 3D Probabilistic Atlas of Human Cortical Structures. *Neuroimage* 39:1064–1080. <https://doi.org/10.1016/j.neuroimage.2007.09.031>.Construction
83. Brain Development Webpage. <https://brain-development.org/brain-atlas/>. Accessed 26 Oct 2022
84. Moeskops P, Veta M, Lafarge MW, Eppenhof KAJ, Pluim JPW (2017) Adversarial training and dilated convolutions for brain MRI segmentation. In: Deep learning in medical image analysis and multimodal learning for clinical decision support. pp. 56–64. Springer
85. Nguyen DMH, Vu HT, Ung HQ, Nguyen BT (2017) 3D-brain segmentation using deep neural network and Gaussian mixture model. In: 2017 IEEE Winter Conference on Applications of Computer Vision. pp. 815–824
86. Chen H, Dou Q, Yu L, Qin J, Heng P (2018) VoxResNet : Deep voxelwise residual networks for brain segmentation from 3D MR images. *Neuroimage* 170:446–455. <https://doi.org/10.1016/j.neuroimage.2017.04.041>
87. Mendrik AM, Vincken KL, Kuijf HJ, Breeuwer M, Bouvy WH, De Bresser J, Alansary A, De Brujinne M, Carass A, El-baz A, Jog A, Katyal R, Khan AR, Van Der Lijn F, Mahmood Q, Mukherjee R, Van Opbroek A, Paneri S, Pereira S, Persson M, Rajchl M, Sarikaya D, Smedby Ö, Silva CA, Vrooman HA, Vyas S, Wang C, Zhao L, Biessels GJ, Viergever MA (2015) MRBrainS Challenge : Online Evaluation Framework for Brain Image Segmentation in 3T MRI Scans. *Comput. Intell. Neurosci.* 2015
88. Manoharan H, Pang G, Wu H (2019) Visualization of MRI Data-sets for Anatomical Brain Segmentation by Pixel-level Analysis. In: 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). pp. 562–568. IEEE
89. Lei Z, Qi L, Wei Y, Zhou Y, Qi W (2019) Infant Brain MRI Segmentation with Dilated Convolution Pyramid Down-sampling and Self-attention. *arXiv Prepr. arXiv1912.12570*. 1–9
90. Wang L, Nie D, Li G, Dolz J, Technologie ED, Zhang Q, Wang F, Xia J, Wu Z, Chen J (2020) Benchmark on Automatic 6-month-old Infant Brain Segmentation Algorithms: The iSeg-2017 Challenge. *IEEE Trans Med Imaging* 38:2219–2230
91. Sun Y, Gao K, Wu Z, Lei Z, Wei Y, Ma J, Yang X, Feng X, Zhao L, Le T, Shin J, Zhong T, Zhang Y, Yu L, Li C, Basner R, Ahmad MO, Swamy MNS, Ma W, Dou Q, Bui TD, Noguera CB, Landman B, Member S, Ian H, Humphreys KL, Shultz S, Li L, Niu S, Lin W, Jewells V, Li G, Shen D, Wang L (2019) Multi-Site Infant Brain Segmentation Algorithms: The iSeg-2019 Challenge. *IEEE Trans Med Imaging* 40:1363–1376
92. Thyreau B, Taki Y (2020) Learning a cortical parcellation of the brain robust to the MRI segmentation with convolutional neural networks. *Med Image Anal*, 61. <https://doi.org/10.1016/j.media.2020.101639>
93. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B (2016) Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 36:61–78. <https://doi.org/10.1016/j.media.2016.10.004>
94. Open Data Commons for Traumatic Brain Injury. <https://odc-tbi.org/>. Accessed 26 Oct 2022
95. Havaei M, Davy A, Warde-farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin P, Larochelle H (2017) Brain Tumor Segmentation with Deep Neural Networks. *Med Image Anal* 35:18–31
96. Bala SA, Kant S (2020) Dense Dilated Inception Network for Medical Image Segmentation. *Int J Adv Comput Sci Appl* 11:785–793
97. Farooq A, Anwar SM, Awais M, Rehman S (2017) A Deep CNN based Multi-class Classification on Alzheimer's Disease using MRI. In: 2017 IEEE International Conference on Imaging systems and techniques (IST). pp. 1–6
98. Jr CRJ, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowsky B, Britson PJ, Whitwell JL, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DLG, Killiany R, Schuff N, Foxbosetti S, Lin C, Studholme C, Decarli CS, Krueger G, Ward HA, Metzger GJ, Scott KT, Mallozzi R, Blezek D, Levy J, Debbins JP, Fleisher AS, Albert M, Green R, Bartzokis G, Glover G, Mugler J, Weiner MW (2008) The Alzheimer ' s Disease Neuroimaging Initiative (ADNI) : MRI Methods. 691, 685–691. <https://doi.org/10.1002/jmri.21049>
99. Talo M, Baloglu UB, Yıldırım Ö, Acharya UR (2019) Application of deep transfer learning for automated brain abnormality classification using MR images. *Cogn Syst Res* 54:176–188
100. Harvard Medical School Data. <http://www.med.harvard.edu/AANLIB/>. Accessed: 26 Oct 2022
101. Yiğit A, İşık Z (2020) Applying deep learning models to structural MRI for stage prediction of Alzheimer ' s disease. *Turkish J Electr Eng Comput Sci.* 28:196–210. <https://doi.org/10.3906/elk-1904-172>
102. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL (2021) Open Access Series of Imaging Studies (OASIS) : Cross-sectional MRI Data in Young , Middle Aged , Nonmented , and Demented Older Adults. 1498–1507
103. Malone IB, Cash D, Ridgway GR, MacManus DG, Ourselin S, Fox NC, Schott JM (2013) MIRIAD-Public release of a multiple time point Alzheimer ' s MR imaging dataset. *Neuroimage* 70:33–36. <https://doi.org/10.1016/j.neuroimage.2012.12.044>
104. Dolz J, Desrosiers C, Ben Ayed I (2018) 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. *Neuroimage* 170:456–470. <https://doi.org/10.1016/j.neuroimage.2017.04.039>

105. Kansal K, Yang Z, Fishman AM, Sair HI, Ying SH, Jedynak BM, Prince JL, Onyike CU (2017) Structural cerebellar correlates of cognitive and motor dysfunctions in cerebellar degeneration. *Brain* 140:707–720. <https://doi.org/10.1093/aww348>
106. Martino AD, Connor DO, Chen B, Alaerts K, Anderson JS, Assaf M, Balsters JH, Baxter L, Beggiato A, Bernaerts S, Blanken LME, Bookheimer SY, Braden BB, Bryge L, Castellanos FX, Dapretto M, Delorme R, Fair DA, Fishman J, Fitzgerald J, Gallagher L, Keehn RJJ, Kennedy DP, Lainhart JE, Luna B, Mostofsky SH, Müller R-A, Nebel MB, Nigg JT, O’Hearn K, Solomon M, Toro R, Vaidya CJ, Wenderoth N, White T, Craddock RC, Lord C, Leventhal B, Milham MP (2017) Data Descriptor: Enhancing studies of the connectome in autism using the autism brain imaging data exchange II. *Sci Data* 4:1–15
107. Simon J, Drozdzal M, Vazquez D, Romero A, Bengio Y (2017) The One Hundred Layers Tiramisu : Fully Convolutional DenseNets for Semantic Segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp. 11–19
108. Yu F, Koltun V (2016) Multi-Scale Context Aggregation by Dilated Convolutions. In: Proceedings of the International Conference on Learning Representations. pp. 1–13
109. Pleiss G, Chen D, Huang G, Li T, van der Maaten L, Weinberger KQ (2017) Memory-Efficient Implementation of DenseNets. arXiv Prepr. arXiv1707.06990
110. Bjorck J, Gomes C, Selman B, Weinberger KQ (2018) Understanding batch normalization. *Adv. Neural Inf. Process. Syst.* 2018-Decem, 7694–7705
111. Ulyanov D, Vedaldi A, Lempitsky V (2017) Improved Texture Networks : Maximizing Quality and Diversity in Feed-forward Stylization and Texture Synthesis. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 6924–6932
112. Mukhopadhyay P, Chaudhuri BB (2015) A survey of Hough Transform. *Pattern Recognit* 48:993–1010. <https://doi.org/10.1016/j.patcog.2014.08.027>
113. Banfield JD, Raftery AE (1993) Model-based Gaussian and non-Gaussian Clustering. *Biometrics* 49(3):803–821
114. Goodfellow IJ, Warde-Farley D, Mirza M, Courville A, Bengio Y (2013) Maxout Networks. In: International Conference on Machine Learning. pp. 1319–1327
115. Menze BH, Jakab A, Bauer S, Kalpathy-cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L, Gerstner E, Weber M, Arbel T, Avants BB, Ayache N, Buendia P, Collins DL, Cordier N, Corso JJ, Criminisi A, Das T, Delingette H, Demiralp Ç, Durst CR, Dojat M, Doyle S, Festa J, Forbes F, Germiia E, Glocker B, Golland P, Guo X, Hamamci A, Iftekharuddin KM, Jena R, John NM, Konukoglu E, Lashkari D, Mariz JA, Meier R, Pereira S, Precup D, Price SJ, Raviv TR, Reza SMS, Ryan M, Sarikaya D, Schwartz L, Shin H, Shotton J, Silva CA, Sousa N, Subbanna NK, Szekely G, Taylor TJ, Thomas OM, Tustison NJ, Unal G, Vasseur F, Wintermark M, Ye DH, Zhao L, Zhao B, Zikic D, Prastawa M, Reyes M, Leemput KV (2015) The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Trans Med Imaging* 34:1993–2024. <https://doi.org/10.1109/TMI.2014.2377694>
116. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M, Crimi A, Shinohara RT, Berger C, Ha SM, Rozyczki M, Prastawa M, Alberts E, Lipkova J, Freymann J, Kirby J, Bilello M, Fathallah-Shaykh H, Wiest R, Kirschke J, Wiestler B, Coley R, Kotrotsou A, Lamontagne P, Marcus D, Milchenko M, Nazeri A, Weber M-A, Mahajan A, Baid U, Gerstner E, Kwon D, Acharya G, Agarwal M, Alam M, Albiol A, Albiol A, Albiol FJ, Alex V, Allinson N, Amorim PHA, Amruthkar A, Anand G, Andermatt S, Arbel T, Arbelaez P, Avery A, Azmat MBP, Bai W, Banerjee S, Barth B, Batchelder T, Batmanghelich K, Battistella E, Beers A, Belyaev M, Bendszus M, Benson E, Bernal J, Bharath HN, Biros G, Bis das S, Brown J, Cabezas M, Cao S, Cardoso JM, Carver EN, Casamitjana A, Castillo LS, Catà M, Cattin P, Cerigues A, Chagas VS, Chandra S, Chang Y-J, Chang S, Chang K, Chazalon J, Chen S, Chen W, Chen JW, Chen Z, Cheng K, Choudhury AR, Chylla R, Clérigues A, Colleman S, Colmeiro RGR, Combalia M, Costa A, Cui X, Dai Z, Dai L, Daza LA, Deutsch E, Ding C, Dong C, Dong S, Dudzik W, Eaton-Rosen Z, Egan G, Escudero G, Estienne T, Everson R, Fabrizio J, Fan Y, Fang L, Feng X, Ferrante E, Fidon L, Fischer M, French AP, Fridman N, Fu H, Fuentes D, Gao Y, Gates E, Gering D, Gholami A, Gierke W, Glocker B, Gong M, González-Villá S, Grosseto T, Guan Y, Guo S, Gupta S, Han W-S, Han IS, Harmuth K, He H, Hernández-Sabaté A, Herrmann E, Himthani N, Hsu W, Hsu C, Hu X, Hu X, Hu Y, Hu Y, Hua R, Huang T-Y, Huang W, Van Huffel S, Huo QHVV, Iftekharuddin KM, Isensee F, Islam M, Jackson AS, Jambawalikar SR, Jesson A, Jian W, Jin P, Jose VJM, Jungo A, Kainz B, Kamnitsas K, Kar P-Y, Karnawat A, Kellermeyer T, Kermi A, Keutzer K, Khadir MT, Khened M, Kickingereder P, Kim G, King N, Knapp H, Knecht U, Kohli L, Kong D, Kong X, Koppers S, Kori A, Krishnamurthi G, Krivov E, Kumar P, Kushibar K, Lachinov D, Lambrou T, Lee J, Lee C, Lee Y, Lee M, Lefkovits S, Lefkovits L, Levitt J, Li T, Li H, Li W, Li H, Li X, Li Y, Li H, Li Z, Li X, Li Z, Li X, Li W, Lin Z-S, Lin F, Lio P, Liu C, Liu B, Liu X, Liu M, Liu J, Liu L, Llado X, Lopez MM, Lorenzo PR, Lu Z, Luo L, Luo Z, Ma J, Ma K, Mackie T, Madabushi A, Mahmoudi I, Maier-Hein KH, Maji P, Mammen C, Mang A, Manjunath BS, Marcinkiewicz M, McDonagh S, McKenna S, McKinley R, Mehl M, Mehta R, Mehta R, Meier R, Meinel C, Merhof D, Meyer C, Miller R, Mitra S, Moiyadi A, Molina-Garcia D, Monteiro MAB, Mrukwa G, Myronenko A, Nalepa J, Ngo T, Nie D, Ning H, Niu C, Nuechterlein NK, Oermann E, Oliveira A, Oliveira DDC, Oliver A, Osman AFL, Ou Y-N, Ourelin S, Paragios N, Park MS, Paschke B, Pauloski JG, Pawar K, Pawlowski N, Pei L, Peng S, Pereira SM, Perez-Beteta J, Perez-Garcia VM, Pezold S, Pham B, Phophalia A, Piella G, Pillai GN, Piraud M, Pisov M, Popli A, Pound MP, Pourreza R, Prasanna P, Prkova V, Pridmore TP, Puchi S, Puybareau É, Qian B, Qiao X, Rajchl M, Rane S, Rebsamen M, Ren H, Ren X, Revanur K, Rezaei M, Rippel O, Rivera LC, Robert C, Rosen B, Rueckert D, Safwan M, Salem M, Salvi J, Sanchez I, Sánchez I, Santos HM, Sartor E, Schellinghout D, Scheufele K, Scott MR, Scussel AA, Sedlar S, Serrano-Rubio JP, Shah NJ, Shah N, Shaikh M, Shankar BU, Shboul Z, Shen H, Shen D, Shen L, Shen H, Shenoy V, Shi F, Shin HE, Shu H, Sima D, Sinclair M, Smedby O, Snyder JM, Soltaninejad M, Song G, Soni M, Stawiaski J, Subramanian S, Sun L, Sun R, Sun J, Sun K, Sun Y, Sun G, Sun S, Suter YR, Szilagyi L, Talbar S, Tao D, Tao D, Teng Z, Thakur S, Thakur MH, Tharakan S, Tiwari P, Tochon G, Tran T, Tsai YM, Tseng K-L, Tuan TA, Turlapov V, Tustison N, Vakalopoulou M, Valverde S, Vanguri R, Vasilev E, Ventura J, Vera L, Vercauteren T, Verrastro CA, Vidyarathne L, Vilaplana V, Vivekanandan A, Wang G, Wang Q, Wang CJ, Wang W, Wang D, Wang R, Wang Y, Wang C, Wang G, Wen N, Wen X, Weninger L, Wick W, Wu S, Wu Q, Wu Y, Xia Y, Xu Y, Xu X, Xu P, Yang T-L, Yang X, Yang H-Y, Yang J, Yang H, Yang G, Yao H, Ye X, Yin C, Young-Moxon B, Yu J, Yue X, Zhang S, Zhang A, Zhang K, Zhang X, Zhang L, Zhang X, Zhang Y, Zhang L, Zhang J, Zhang X, Zhang T, Zhao S, Zhao Y, Zhao X, Zhao L, Zheng Y, Zhong L, Zhou C, Zhou X, Zhou F, Zhu H, Zhu J, Zhuge Y, Zong W, Kalpathy-Cramer J, Farahani K, Davatzikos C, van Leemput K, Menzen B (2018) Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. arXiv Prepr. arXiv1811.02629

117. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going Deeper with Convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9
118. Junut J, Sijbers J, Van Dyck D, Gielen J (2005) Bias Field Correction for MRI Images. In: Computer Recognition Systems. pp. 543–551. Springer Berlin Heidelberg, Berlin
119. Li C, Huang R, Ding Z, Gatenby JC, Metaxas DN, Gore JC (2011) A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities With Application to MRI. 20, 2007–2016
120. Wang G, Li W, Zuluaga MA, Pratt R, Patel PA, Aertsen M, Doel T, David AL, Deprest J, Ourselin S, Vercauteren T (2018) Interactive Medical Image Segmentation Using Deep Learning with Image-Specific Fine Tuning. *IEEE Trans Med Imaging* 37:1562–1573. <https://doi.org/10.1109/TMI.2018.2791721>
121. Romero M, Interian Y, Solberg T, Valdes G (2019) Training Deep Learning models with small datasets. Prepr. ArXiv, Dec
122. Lakhani P, Sundaram B (2017) Deep Learning at Chest Radiography Lakhani and Sundaram. *Radiology* 284:574–582

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



Luis Velázquez-Pérez: President of the Cuban Academy of Sciences. Ph.D. in Neurology and Clinical Neurophysiology. Doctor in Science in the Medical University of Havana, Cuba. Research focused in cerebellum and cerebellar syndrome, clinical scales, sleep analysis, cognition, transcranial magnetic stimulation (TMS), and cognitive neuroscience.



Roberto Pérez-Rodríguez: Ph.D. in Industrial Engineering at the Polytechnic University of Catalonia, Spain. Research focused on concurrent engineering, design theory and methodology, computer aided tolerancing, integrated product development, computational systems (CAx) and Artificial Intelligence.



Robin Cabeza-Ruiz: M.Sc. in Computer Aided Design and Manufacturing at the University of Holguín, Cuba. Research focused on deep learning techniques applied to neuroscience studies and automatic magnetic resonance processing.



Kathrin Reetz: Doctor degree at the Faculty of Medicine, RWTH Aachen University, Germany. Awarded the Heinrich Pette Prize for her achievements in the field of neurodegenerative diseases, in 2019. Research focused on clinical and imaging markers in neurodegenerative diseases.

Appendices

Appendix A

The First Appendix

The Ideal can not take account of, so far as I know, our faculties. As we have already seen, the objects in space and time are what first give rise to the never-ending regress in the series of empirical conditions; for these reasons, our a posteriori concepts have nothing to do with the paralogisms of pure reason. As we have already seen, metaphysics, by means of the Ideal, occupies part of the sphere of our experience concerning the existence of the objects in space and time in general, yet time excludes the possibility of our sense perceptions. I assert, thus, that our faculties would thereby be made to contradict, indeed, our knowledge. Natural causes, so regarded, exist in our judgements.

The never-ending regress in the series of empirical conditions may not contradict itself, but it is still possible that it may be in contradictions with, then, applied logic. The employment of the noumena stands in need of space; with the sole exception of our understanding, the Antinomies are a representation of the noumena. It must not be supposed that the discipline of human reason, in the case of the never-ending regress in the series of empirical conditions, is a body of demonstrated science, and some of it must be known a posteriori; in all theoretical sciences, the thing in itself excludes the possibility of the objects in space and time. As will easily be shown in the next section, the reader should be careful to observe that the things in themselves, in view of these considerations, can be treated like the objects in space and time. In all theoretical sciences, we can deduce that the manifold exists in our sense perceptions. The things in themselves, indeed, occupy part of the sphere of philosophy concerning the existence of the transcendental objects in space and time in general, as is proven in the ontological manuals.

A.1 First Section

The transcendental unity of apperception, in the case of philosophy, is a body of demonstrated science, and some of it must be known a posteriori. Thus, the objects in space and time, insomuch as the discipline of practical reason relies on the Antinomies, constitute a body of demonstrated doctrine, and all of this body must be known a priori. Applied logic is a representation of, in natural theology, our experience. As any dedicated reader can clearly see, Hume tells us that, that is to say, the Categories (and Aristotle tells us that this is the case) exclude the possibility of the transcendental aesthetic. (Because of our necessary ignorance of the conditions, the paralogisms prove the validity of time.) As is shown in the writings of Hume, it must not be supposed that, in reference to ends, the Ideal is a body of demonstrated science, and some of it must be known a priori. By means of analysis, it is not at all certain that our a priori knowledge

A. The First Appendix

is just as necessary as our ideas. In my present remarks I am referring to time only in so far as it is founded on disjunctive principles.

A.2 Second Section

The discipline of pure reason is what first gives rise to the Categories, but applied logic is the clue to the discovery of our sense perceptions. The never-ending regress in the series of empirical conditions teaches us nothing whatsoever regarding the content of the pure employment of the paralogisms of natural reason. Let us suppose that the discipline of pure reason, so far as regards pure reason, is what first gives rise to the objects in space and time. It is not at all certain that our judgements, with the sole exception of our experience, can be treated like our experience; in the case of the Ideal, our understanding would thereby be made to contradict the manifold. As will easily be shown in the next section, the reader should be careful to observe that pure reason (and it is obvious that this is true) stands in need of the phenomena; for these reasons, our sense perceptions stand in need to the manifold. Our ideas are what first give rise to the paralogisms.

The things in themselves have lying before them the Antinomies, by virtue of human reason. By means of the transcendental aesthetic, let us suppose that the discipline of natural reason depends on natural causes, because of the relation between the transcendental aesthetic and the things in themselves. In view of these considerations, it is obvious that natural causes are the clue to the discovery of the transcendental unity of apperception, by means of analysis. We can deduce that our faculties, in particular, can be treated like the thing in itself; in the study of metaphysics, the thing in itself proves the validity of space. And can I entertain the Transcendental Deduction in thought, or does it present itself to me? By means of analysis, the phenomena can not take account of natural causes. This is not something we are in a position to establish.

Appendix B

Source Code

referatutømoe9

B.1 Implementation

The `phduio` class is implemented in the following way:

%% Implemented by Martin Helso (martibhe@math.uio.no)

```
\NeedsTeXFormat{LaTeX2e}

\ProvidesClass{phduio}[2020/11/06 Class for PhD theses at UiO]

%%%%%%%%%%%%% CLASS OPTIONS %%%%%%%

%% Language
\DeclareOption{american} { \def \phduio@thesis{Dissertation} }
\DeclareOption{USenglish}{ \def \phduio@thesis{Dissertation} }
\DeclareOption{english} { \def \phduio@thesis{Thesis} }
\DeclareOption{UKenglish}{ \def \phduio@thesis{Thesis} }

%% Colophon
\DeclareOption{colophon} { \def \phduio@colophon{true} }
\DeclareOption{nocolophon}{ \def \phduio@colophon{false} }

%% Screen mode
\DeclareOption{screen}
{
    \AtBeginDocument
    {
        \@ifpackageloaded{url}
            {\urlstyle{same}}
            {\ClassWarningNoLine{phduio}{Package 'url' missing}}
        \@ifpackageloaded{hyperref}
            {\hypersetup{colorlinks, allcolors = uiolink}}
            {\ClassWarningNoLine{phduio}{Package 'hyperref' missing}}
        \setlrmarginsandblock{24.35mm}{24.35mm}{*}
        \checkandfixthelayout
    }
}

\DeclareOption*{\PassOptionsToClass{\CurrentOption}{memoir}}


\ExecuteOptions{UKenglish}
\ExecuteOptions{nocolophon}
```

B. Source Code

```
\ProcessOptions
\relax

\LoadClass[oldfontcommands, extrafontsizes]{memoir}

%%%%%%%%%%%%%% PACKAGES %%%%%%%

\RequirePackage{keyval}
\RequirePackage{etoolbox}
\RequirePackage{textcomp}
\RequirePackage[dvipsnames, svgnames, cmyk]{xcolor}
\RequirePackage{pdfpages}
\RequirePackage{graphicx}
\graphicspath{{figures/}}
\RequirePackage{tikz}
\usetikzlibrary{calc}

%%%%%%%%%%%%%% LAYOUT %%%%%%%

%% Paper size
\setstocksize{240mm}{170mm}
\settrimmedsize{240mm}{170mm}{*}
\settrims{0mm}{0mm}
\setlrmarginsandblock{20mm}{28.7mm}{*}
\setulmarginsandblock{25mm}{25mm}{*}
\checkandfixthelayout

%% Custom title page
\newcommand{\uiotitle}{%
\begin{titlingpage}
\sffamily
\renewcommand{\and}{\vskip1mm}
\newcommand{\AND}{\vskip9mm}

\calccentering{\unitlength}
\begin{adjustwidth*}{\unitlength}{-\unitlength}

\raggedright

\vspace*{-5mm}
\includegraphics[width = 0.6\textwidth]{figures/phduio-logo}
\vskip18mm

\resizebox{%
\ifdim \width > \textwidth
\textwidth
\else
\width
\fi
}{\unitlength}
\end{adjustwidth*}
\end{titlingpage}
}
```

```
}{!}
{%
    \LARGE
    \theauthor
}

\vskip1.5\onelineskip

{
    \Huge
    \bfseries
    \boldmath
    \thetitle
    \par
}

{
    \ifcsempy{\phduio@subtitle}
    {}
    {
        \vskip2mm
        \huge
        \phduio@subtitle
        \par
    }
}

\vskip-\lastskip
\vskip16mm

\resizebox{\textwidth}{!}
{%
    \bfseries
    \phduio@thesis\
    submitted for the degree of Philosophiae Doctor
}

\AND

{
    \Large
    \phduio@dept\
    \and
    \phduio@faculty\
    \AND
    \phduio@affiliation\
}

\fill

\begin{minipage}[c][26mm][28mm]
\hspace*{1.3mm}
\includegraphics[width = 26mm]{figures/phduio-apollon}
\end{minipage}%
\begin{minipage}[c]{\textwidth - 28mm}
\hfill
\LARGE

```

B. Source Code

```
\bfseries
\the\year
\end{minipage}

\vspace{2mm}

\end{adjustwidth*}

\null
\clearpage
\ifdefstring{\phduio@colophon}{true}{\phduio@print@colophon}{}

\end{titlingpage}
}

%% Book
\renewcommand*{\printbooktitle}[1]{\raggedright\booktitlefont #1}
\renewcommand*{\afterbookskip}{\par}
\renewcommand*{\booktitlefont}{\HUGE\bfseries\boldmath\sffamily}
\renewcommand*{\booknamefont}{\raggedright\Huge\normalfont\sffamily}
\renewcommand*{\booknumfont}{\Huge\normalfont\sffamily}

%% Part
\renewcommand*{\printparttitle}[1]{\raggedright\parttitlefont #1}
\renewcommand*{\afterpartskip}{\par}
\renewcommand*{\parttitlefont}{\HUGE\bfseries\boldmath\sffamily}
\renewcommand*{\partnamefont}{\raggedright\Huge\normalfont\sffamily}
\renewcommand*{\partnumfont}{\Huge\normalfont\sffamily}

%% Chapter
\newcommand{\authorsfont}{}
\newcommand{\metadatafont}{}
\newlength{\afterauthorskip}
\newlength{\aftermetadaskip}

\makechapterstyle{phduio}
{
    \renewcommand*{\chapnamefont}{\huge\sffamily}
    \renewcommand*{\chapnumfont}{\huge\sffamily}
    \renewcommand*{\chaptilefont}{\Huge\bfseries\boldmath\sffamily\raggedright}
    \renewcommand*{\authorsfont}{\Large\bfseries\sffamily}
    \renewcommand*{\metadatafont}{\normalfont\normalsize\sffamily}

    \setlength{\beforechapskip}{-1.35\baselineskip}
    \setlength{\midchapskip}{10pt}
    \setlength{\afterchapskip}{20pt}
    \setlength{\afterauthorskip}{6pt}
    \setlength{\aftermetadaskip}{15pt}

    \renewcommand*{\afterchapttitle}{%
        \vskip\afterchapskip
        \ifboolexpr
    }
}
```

```

{
  test{ \ifcsempty{phduio@authors} }
  and
  test{ \ifcsempty{phduio@metadata} }
}
{
\begin{minipage}[t]{\textwidth}
  \authorsfont
  \phduio@authors
  \vskip\afterauthorsskip
  \metadatafont
  \phduio@metadata
  \gdef \phduio@authors{}
  \gdef \phduio@metadata{}
\end{minipage}
\vskip\aftermetadataskip
}
}

\renewcommand*{\printchapternum}
{
  \vphantom{\chapnumfont Chapter}
  \afterchapternum
  \vskip\topskip
}
}

\chapterstyle{phduio}

%% Lower level sections
\setsecnumdepth{subsubsection}
\setsecheadstyle{\Large\bfseries\boldmath\sffamily\raggedright}
\setsubsecheadstyle{\large\bfseries\boldmath\sffamily\raggedright}
\setsubsubsecheadstyle{\normalsize\bfseries\boldmath\sffamily\raggedright}
\setparaheadstyle{\normalsize\bfseries\boldmath\sffamily\raggedright}
\setsubparaheadstyle{\normalsize\bfseries\boldmath\sffamily\raggedright}

%% Subappendices
\namedsubappendices

%% Abstract
\renewcommand{\abstractnamefont}{\sffamily\bfseries}
\renewcommand{\abstracttextfont}{\normalfont\small\noindent\ignorespaces}

%% Table of contents, list of figures and list of tables
\setrmarg{3em}
\setpnumwidth{2em}
\addtolength{\cftfigurenumwidth}{1em}
\addtolength{\cfttablenumwidth}{1em}
\addtolength{\cftbooknumwidth}{1em}
\addtolength{\cftpartnumwidth}{1em}
\addtolength{\cftchapternumwidth}{1em}

```

B. Source Code

```
\addtolength{\cftsectionnumwidth}{1em}
\addtolength{\cftsubsectionnumwidth}{1em}
\addtolength{\cftsubsubsectionnumwidth}{1em}
\addtolength{\cftpagernumwidth}{1em}
\addtolength{\cftsubparagraphnumwidth}{1em}
\addtolength{\cftsectionindent}{1em}
\addtolength{\cftsubsectionindent}{2em}
\addtolength{\cftsubsubsectionindent}{3em}
\addtolength{\cftpagindent}{4em}
\addtolength{\cftsubparagraphindent}{5em}
\renewcommand{\cftchapteraftersnumb}{\bfseries\boldmath}

%% Running header and footer
\makepagestyle{phduio}
\makeheadrule{phduio}{\textwidth}{\normalrulethickness}
\makeevenhead{phduio}{\sffamily\leftmark}{}{}
\makeoddhead {phduio}{}{\sffamily\rightmark}
\makeevenfoot{phduio}{\sffamily\thepage}{}{}
\makeoddfoot {phduio}{}{\sffamily\thepage}
\makepsmarks{phduio}
{
    \nouppercaseheads
    \createmark{chapter}{left}{shownumber}{}{. \space}
    \createmark{section}{right}{nonumber}{}{}
    \createplainmark{toc}{both}{\contentsname}
    \createplainmark{lof}{both}{\listfigurename}
    \createplainmark{lot}{both}{\listtablename}
    \createplainmark{bib}{both}{\bibname}
    \createplainmark{index}{both}{\indexname}
    \createplainmark{glossary}{both}{\glossaryname}
}
\pagestyle{phduio}

\makepagestyle{chapter}
\makeevenfoot{chapter}{\sffamily\thepage}{}{}
\makeoddfoot {chapter}{}{\sffamily\thepage}

%%%%%%%%%%%%% INCLUDE PAPERS %%%%%%%

%% Title for papers
\renewcommand{\maketitle}
{
    \paperauthors{\theauthor}
    \edef\papertitle{\thetitle}
    \expandafter\chapter\papertitle
    \paperthumb
    \saythanks
}

%% Specify authors for separate paper
\newcommand{\paperauthors}[1]
{
    \renewcommand{\and}{\leavevmode\unskip,\space}
}
```

```

\def \phduio@authors{\#1}
}
\paperauthors{}


%% Specify metadata for separate paper
\newcommand{\metadata}[1]{ \def \phduio@metadata{\#1} }
\metadata{}


%% Rename 'Chapter' to 'Paper'
\newcommand*{\paper}{%
{
    \setcounter{chapter}{0}
    \setcounter{section}{0}
    \ifdef{\theHchapter}{%
        \renewcommand{\theHchapter}{\paper.\arabic{chapter}}%
    }{%
    }
    \def \chapapp{Paper}
    \let \c@paper \c@chapter
    \def \thechapter{\@Roman\c@chapter}
    \let \thepaper \thechapter
    \ifdef{\memendofchapterhook}{%
        \renewcommand{\memendofchapterhook}{%
            \addtocounter{paper}{-1}%
            \refstepcounter{paper}%
        }%
    }{%
    }
}
\newcounter{paper}

%% Standard \appendix is broken by \paper
%% Fix hyperlinks to appendices
\renewcommand*{\appendix}{%
{
    \setcounter{chapter}{0}
    \setcounter{section}{0}
    \ifdef{\theHchapter}{%
        \renewcommand{\theHchapter}{\arabic{chapter}}%
    }{%
    }
    \def \chapapp{\appendixname}
    \let \c@appendix \c@chapter
    \def \thechapter{\@Alph\c@chapter}
    \let \theappendix \thechapter
    \ifdef{\memendofchapterhook}{%
        \renewcommand{\memendofchapterhook}{%
            \addtocounter{appendix}{-1}%
        }%
    }{%
    }
}

```

B. Source Code

```
\refstepcounter{appendix}
}
}
{}
}
\newcounter{appendix}

%% Print 'Papers'
\newcommand{\paperpage}{%
\part*{Papers}
\addcontentsline{toc}{chapter}{Papers}
\cleartorecto
}

%% Specify number of papers
\newcommand{\numberofpapers}[1]{ \setcounter{totpapernum}{#1} }

%% Thumb index to separate papers
\newcommand{\paperthumb}{%
\ifnum \value{thumbpos} = \value{totpapernum}
\setcounter{thumbpos}{0}
\fi
\stepcounter{thumbnum}
\stepcounter{thumbpos}

\pgfmathsetlength{\thumbheight}{\paperheight / \value{totpapernum}}

\begin{tikzpicture}[remember picture, overlay]
\node [thumb, align = right, anchor = north east]
at
($({current page.north east}) - (0,{(\value{thumbpos}-1)*\thumbheight})$)
{\Roman{thumbnum}};
\end{tikzpicture}
}

\newcounter{thumbnum}
\newcounter{thumbpos}
\newcounter{totpapernum}
\setcounter{totpapernum}{2}

\newlength{\thumbwidth}
\newlength{\thumbheight}
\setlength{\thumbwidth}{1cm}

\tikzset
{
thumb/.style =
{
fill          = uioblack,
text          = uiowhite,
font          = \sffamily\bfseries\Huge,
```

```
    text width      = \thumbwidth,
    minimum height = \thumbheight,
    outer sep       = 0pt,
    inner xsep     = 1.5em
}
}

%% Unmarked footnote
\newcommand{\papernote}[1]
{
  \begingroup
    \renewcommand{\thefootnote}{}
    \footnotetext{#1}
  \endgroup
}

%% Unmarked footnote for paper title
\renewcommand{\thanks}[1]{ \gdef \phduio@thanks{#1} }
\thanks{}
\renewcommand{\saythanks}
{
  \ifcsempty{\phduio@thanks}
  {}
  {
    \papernote{\phduio@thanks}
    \thanks{}
  }
}

%% Include PDFs containing separate papers
%% Based on code by Dag Langmyhr
\newcommand{\includearticle}[2][]
{
  \cleartorecto

  \def \ps@default
  {
    \let \mkboth\@gobbletwo
    \let \oddhead\@empty
    \def \oddfoot{\hfil\sffamily\thepage}
    \let \evenhead\@empty
    \def \evenfoot{\sffamily\thepage\hfil}
  }

  \def \phduio@ps{default}

  \def \ps@low
  {
    \let\mkboth\@gobbletwo
    \let\oddhead\@empty
    \def\oddfoot{\hfil\raisebox{-7.5mm}{\sffamily\thepage}}
    \let\evenhead\@empty
    \def\evenfoot{\raisebox{-7.5mm}{\sffamily\thepage}\hfil}
  }
}
```

B. Source Code

```
\let \ps@none = \ps@empty
\setkeys{phduio@keys}{#1}
\includepdf[pages = {-}, pagecommand = {\thispagestyle{\phduio@ps}}]{#2}
}

\define@key{phduio@keys}{numbers}{ \def \phduio@ps{#1} }

%%%%%%%%%%%% UTILITIES %%%%%

%% Official colours
\definecolor{uiored} {cmyk}{0, 1, 1, 0}
\definecolor{uiogrey} {cmyk}{0, 0.05, 0.1, 0.29}
\definecolor{uiowhite}{cmyk}{0, 0, 0, 0}
\definecolor{uioblack}{cmyk}{0, 0, 0, 1}
\definecolor{uiolink} {HTML}{0B5A9D}

%% Print official logo colon
\newcommand{\uiocolon}
{
    \includegraphics[height = 1.5ex]{figures/phduio-colon}\%
}

%% Subtitle
\newcommand{\subtitle}[1]{ \def \phduio@subtitle{#1} }
\subtitle{}


%% Specify affiliation
\newcommand{\department}[1]{ \def \phduio@dept{#1} }
\newcommand{\faculty}[1]{ \def \phduio@faculty{#1} }
\newcommand{\affiliation}[1]{ \def \phduio@affiliation{#1} }
\department{}
\faculty{}
\affiliation{}


%% Specify International Standard Serial/Book Number
\newcommand{\ISSN}[1]{ \def \phduio@ISBN{\ISSN\space#1} }
\newcommand{\ISBN}[1]{ \def \phduio@ISBN{\ISBN\space#1} }
\ISSN{ISSN}

%% Specify dissertation series number
\newcommand{\dissertationseries}[1]{ \def \phduio@dissertationseries{#1} }
\dissertationseries{}


%% Specify production credits
\newcommand{\cover}[1]{ \def \phduio@cover{#1} }
\newcommand{\printinghouse}[1]{ \def \phduio@printinghouse{#1} }
\cover{Hanne Baadsgaard Utigard}
```

```
\printinghouse{Reprosentralen, University of Oslo}

%% Print colophon
\newcommand{\phduio@print@colophon}
{
    \thispagestyle{empty}
    \vspace*{\stretch{3}}
    \begin{flushleft}
        \textbf{\sffamily\textrightskip\space\theauthor,\space\the\year}
        \vskip2\baselineskip
        \textit{Series of dissertations submitted to the}
        \par
        \textit{\phduio@faculty, University of Oslo}
        \par
        \ifcsempty{\phduio@dissertationseries}
        {}
        {\textit{No.}\space\phduio@dissertationseries}
        \vskip\baselineskip
        \phduio@ISN
        \vskip2\baselineskip
        \small
        All rights reserved.
        No part of this publication may be
        \par
        reproduced or transmitted,
        in any form or by any means,
        without permission.
        \vskip7\baselineskip
        Cover: \phduio@cover.
        \par
        Print production: \phduio@printinghouse.
    \end{flushleft}
    \vspace*{\stretch{1}}
    \null
    \newpage
}

%% Print current month
\newcommand{\MONTH}
{
    \leavevmode\unskip\space
    \ifcase\the\month
    \or January
    \or February
    \or March
    \or April
    \or May
    \or June
    \or July
    \or August
    \or September
    \or October
    \or November
    \or December
    \fi
}
```

B. Source Code

```
\ignorespaces  
}
```

```
\endinput
```