Medical & Biological Engineering & Computing
https://doi.org/10.1007/s11517-022-02714-w

REVIEW ARTICLE

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

Robin Cabeza-Ruiz[1] · **Luis Velázquez-Pérez**[2,3] · **Roberto Pérez-Rodríguez**[1,2] · **Kathrin Reetz**[4]

Received: 16 September 2021 / Accepted: 26 October 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 diagnose 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-pallydolusian 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 olivopontocerebellar 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 olivary 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,

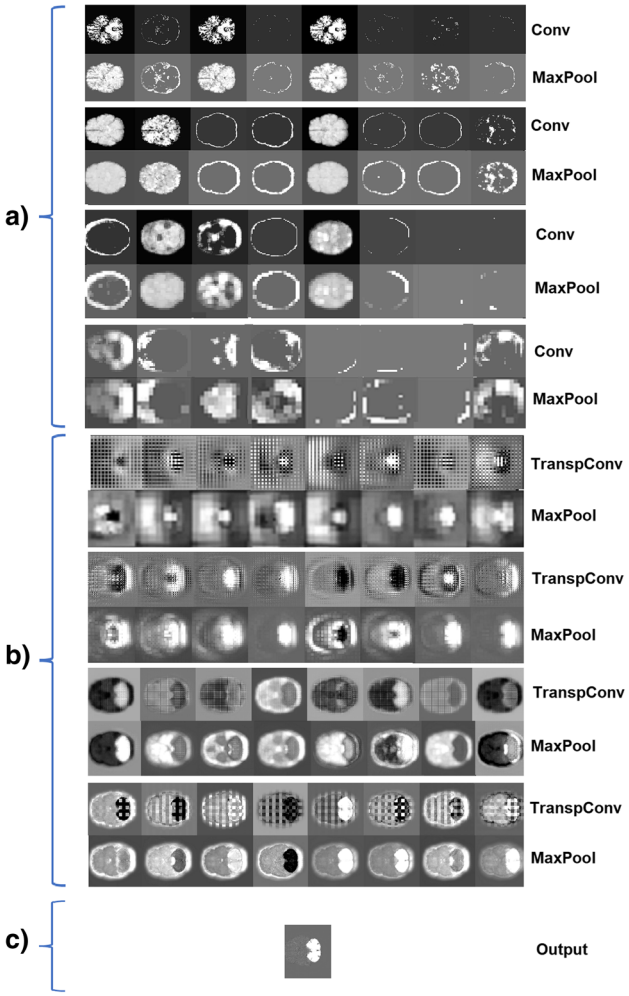

Fig. 1 From top to bottom: partial results of different layers of a CNN, trained for cerebellum segmentation from brain MRIs. Each individual row shows eight 2D sample slices, which are the partial results of one layer. The encoder section (**a**) is composed of several convolution layers (Conv) with 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].

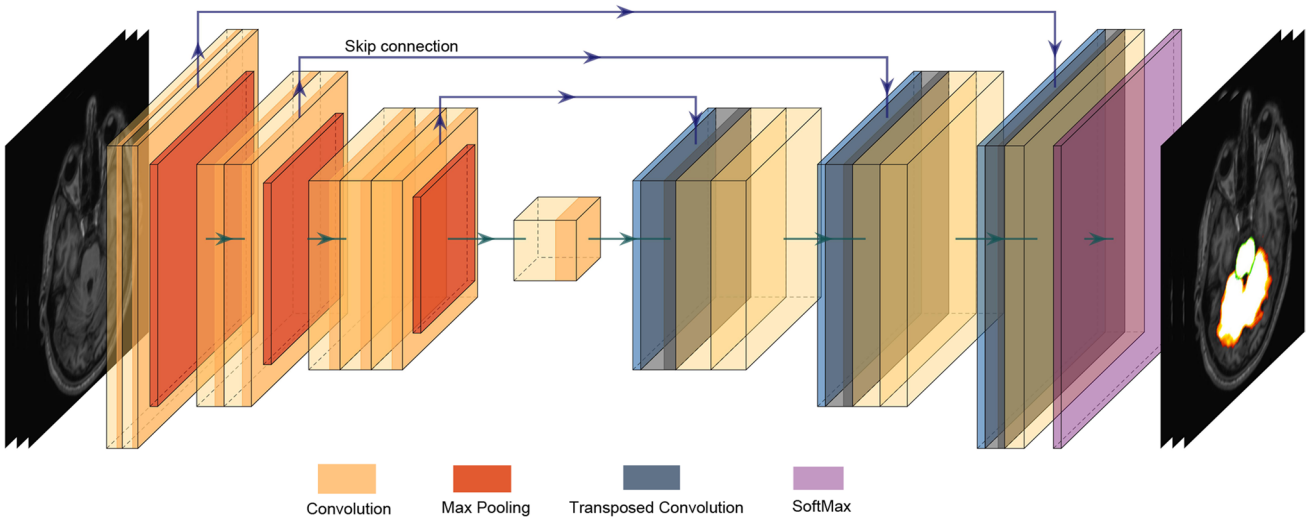

Fig. 2 Representation of a CNN applied to semantic segmentation. The network is composed of three down-sample operations, consisting of joining convolutional layers (yellow color) and max pooling operations (orange color). In the middle, a single convolutional layer processes the more local features. 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
				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
2018	[86]	ResNet	3D	MrBrains2013 [87]	0.91	-	-	-	-	20
	[74]	DenseNet	Pixel	IBSR	-	0.92	-	-	-	18
2019	[88]	SegNet	Pixel	MICCAI12	0.79	-	-	-	-	35
	[89]	U-Net	3D	iSeg-2017 [90]	0.95	-	-	-	-	23
				iSeg-2019 [91]	0.87	-	-	-	-	39
	[92]	CNN	3D	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 investigations[1].

[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 parcellating 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 Tyreau 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 fed-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 o 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 Yİğİt & Işik

Yİğİt & Işik [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 Yİğİt & Işik [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 lobule 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, bout 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

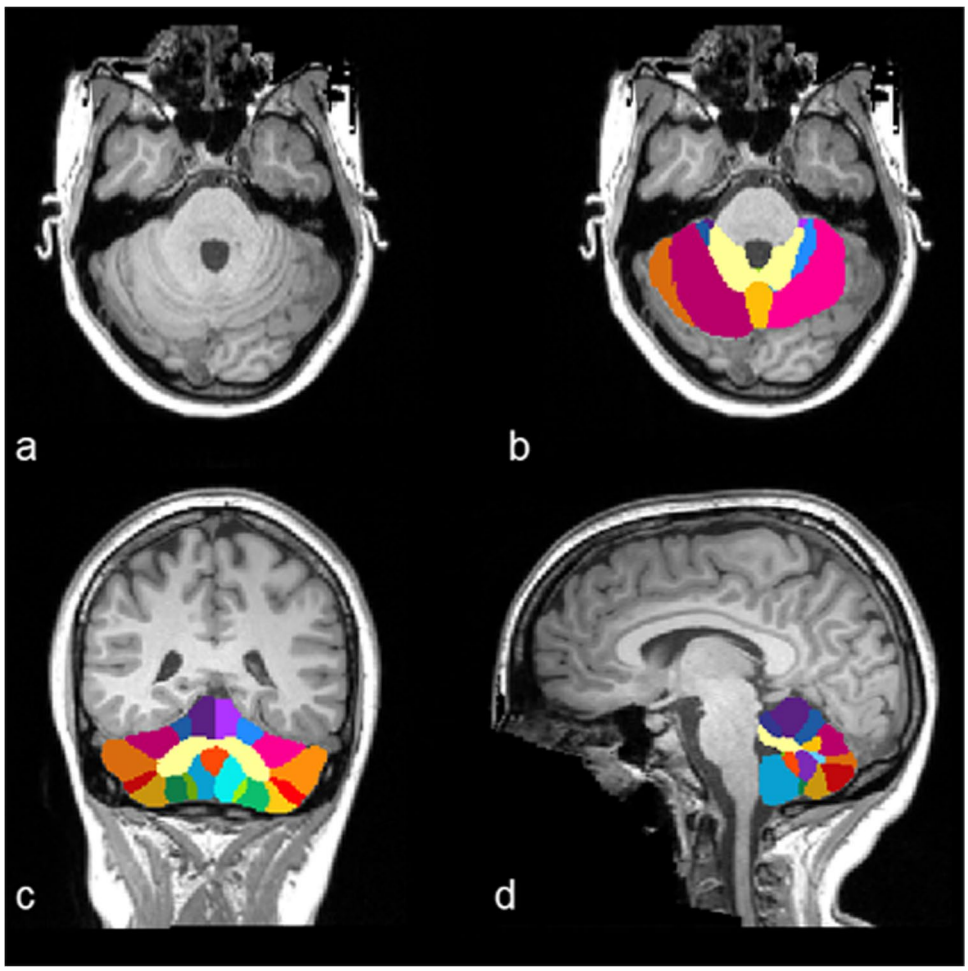

Fig. 3 Images from a human cerebellum parcellated from MRI. Original axial slice (a), followed by axial (b), coronal (b) and sagittal (c) slices of the parcellation. Image generated with free distribution of ACAPULCO [64]

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

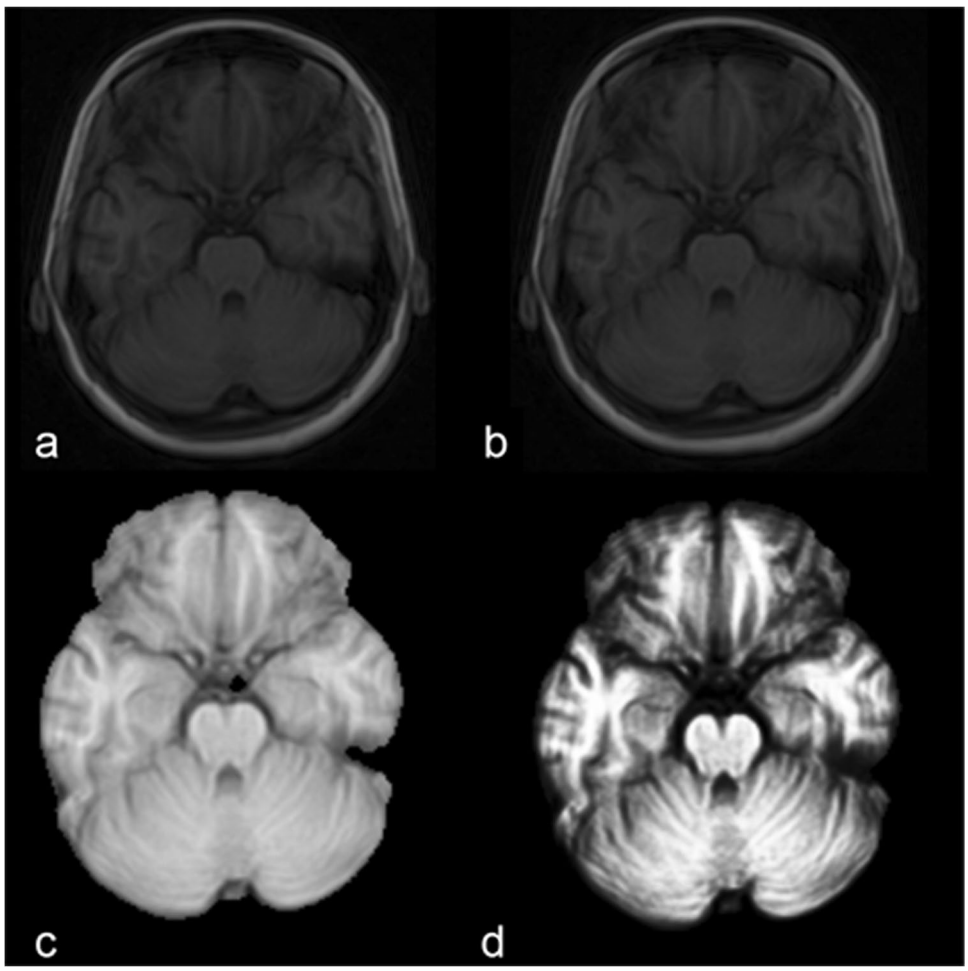

Fig. 4 Example of preprocessing over MRI. Original MRI slice (a), bias field corrected image (b), skull stripping (c), and histogram equalization (d)

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

1. Dueñas AM, Goold R, Giunti P (2006) Molecular pathogenesis of spinocerebellar ataxias. Brain 129:1357–1370. https://doi.org/10.1093/brain/awl081
2. 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
3. Manto M (2005) The wide spectrum of spinocerebellar ataxias (SCAs). The Cerebellum 4:2–6. https://doi.org/10.1080/14734220510007914
4. Teive HAG (2009) Spinocerebellar ataxias. Arq Neuropsiquiatr 67:1133–1142. https://doi.org/10.1590/S0004-282X2009000600035
5. 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
6. 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
7. 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
8. Erichsen AK, Koht ÃJ, Stray-pedersen ÃA, Abdelnoor M, Tallaksen CME (2009) Prevalence of hereditary ataxia and spastic paraplegia in southeast Norway : a population-based study. Brain 132:1577–1588. https://doi.org/10.1093/brain/awp056
9. 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/s13311-018-00696-y
12. Nethisinghe S, Lim WN, Ging H, Zeitlberger 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.2018.00429
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, Paneque M, Medina E, Hechavarría 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.v0.682
22. Seidel K, Siswanto S, Brunt ERP, Den Dunnen W, Korf HW, Rüb U (2012) Brain pathology of spinocerebellar ataxias. Acta Neuropathol 124:1–21. https://doi.org/10.1007/s00401-012-1000-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, quanti fi cation 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, Linnemann 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. Decubitis 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) Spincocerebellar ataxias: An update. Curr Opin Neurol 33:150–160. https://doi.org/10.1097/WCO.0000000000000774
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, Fancellu 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, Haselgrove 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, Chakravarty 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.ancr.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, Guttag 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, Bötzel 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. Commun ACM 60:84–90. https://doi.org/10.1201/9781420010749
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-atlases/. 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 Bruijne 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 Datasets 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, Basnet 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, Borowski B, Britson PJ, Whitwell JL, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DLG, Killiany R, Schuff N, Fox-bosetti 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, Yildrim 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. Yİğİt A, Işik 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 , Nondemented , 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, Byrge L, Castellanos FX, Dapretto M, Delorme R, Fair DA, Fishman I, 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 0. 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, Geremia 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, Rozycki M, Prastawa M, Alberts E, Lipkova J, Freymann J, Kirby J, Bilello M, Fathallah-Shaykh H, Wiest R, Kirschke J, Wiestler B, Colen 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, Amrutkar 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, Bisdas 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, Grosges 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, Kao P-Y, Karnawat A, Kellermeier 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 S, 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 AFI, Ou Y-N, Ourselin 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, Prkovska V, Pridmore TP, Puch S, Puybareau É, Qian B, Qiao X, Rajchl M, Rane S, Rebsamen M, Ren H, Ren X, Revanuru 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, Schellingerhout 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, Vasiliev E, Ventura J, Vera L, Vercauteren T, Verrastro CA, Vidyaratne 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, Menze 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. Juntu 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

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.

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.

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.