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

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

[1] University of Holguin, CAD/CAM Study Centre, XX Anniversary, Holguin, Cuba
robbinc91@gmail.com, roberto.perez.cu@gmail.com
[2] Cuban Academy of Sciences, La Habana, Cuba
[3] Centre for the Research and Rehabilitation of Hereditary Ataxias, Libertad Street, Holguin, Cuba
velazq63@gmail.com

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

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

1 Introduction

1.1 Spinocerebellar Ataxias

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

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

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

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

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

1.2 Convolutional Neural Networks

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

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

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

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

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

2 Materials and Methods

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

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

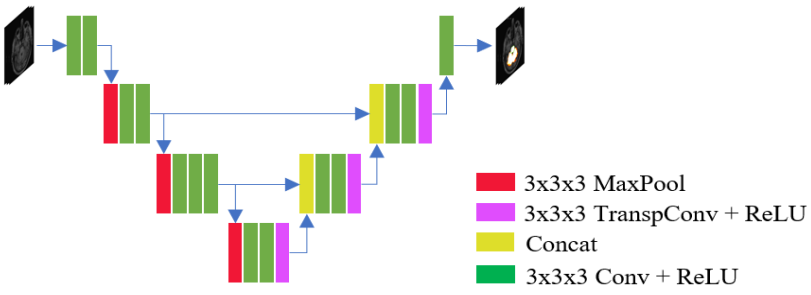

Fig. 1. First U-Net model without inception modules (M1)

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

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

2.1 Image Preparation

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

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

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

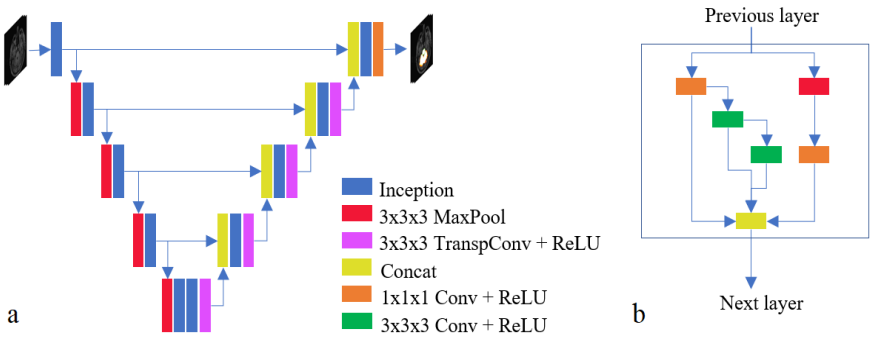

Fig. 2. Second U–Net model (M2). In (a) the general network structure, and (b) shows the proposed inception circuit.

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

2.2 Analysis Description

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

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

Table 1. Partitions created from Data2

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

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

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

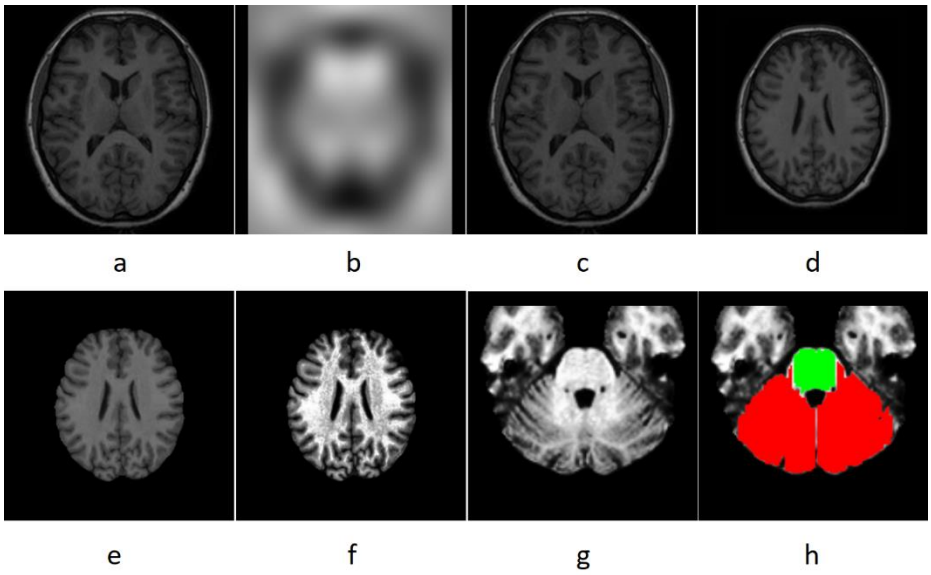

Fig. 3. Preprocessing phases example. Original imaging (a), calculated bias (b), BFC result (c), registration result (d), skull stripping (e), histogram equalization (f), cropping (g) and masks of the interesting organs; red shows the cerebellum, and green represents the brainstem.

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

2.3 Metrics

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

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

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

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

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

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

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

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

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

3 Results and Discussion

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

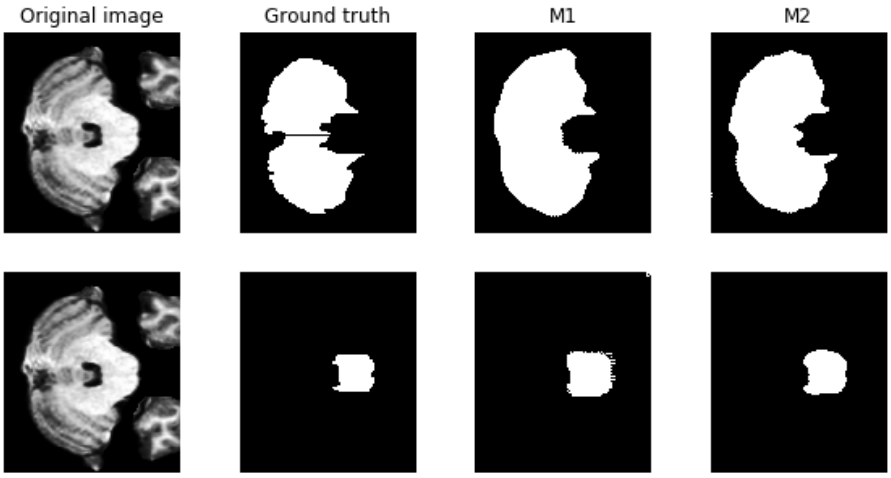

Fig. 4. Comparison between segmentation produced for M1 and M2 methods on Data1. First row shows cerebellum masks, and second row displays brainstem masks.

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

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

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

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

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

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

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

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

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

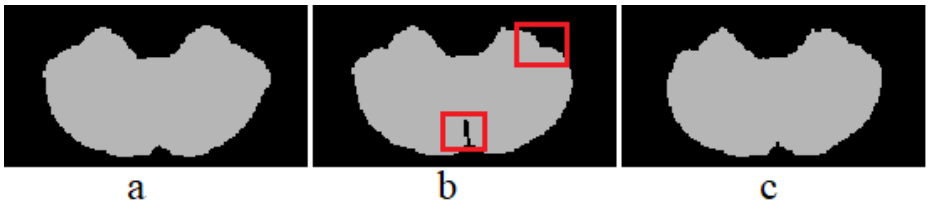

Fig. 5. Segmentation comparison for one of the test MRIs. Ground truth mask (a), followed by mask obtained with NAD model (b) and WAD model (c). The major errors were obtained with NAD model, resulting in a considerably smaller volume estimation.

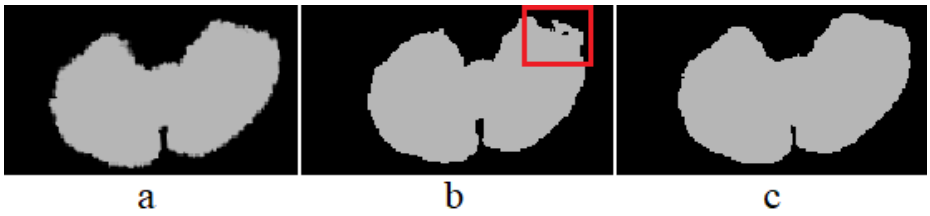

Fig. 6. Segmentation comparison for one test MRI, which has been displaced and rotated. Ground truth segmentation (a), followed by segmentation with NAD model (b), and with WAD model (c). Once again, the major errors were produced with NAD model (marked in red), resulting in the loss of information.

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

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

Table 4. Calculated volumes using each method, compared with the ground truth (in mm^3)

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

* Presymptomatic carrier ** SCA2 patient

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

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

4 Conclusions

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

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

References

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

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

Ying, S.H., Onyike, C.U., Landman, B.A., Mostofsky, S.H., Thompson, P.M., Prince, J.L.: Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. Neuroimage. 183, 150–172 (2018). https://doi.org/10.1016/j.neuroimage.2018.08.003.Comparing

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