

## 1 Fully automatic segmentation of brain lacunas resulting from resective surgery using a 3D 2 deep learning model

3 Raphael Fernandes Casseb<sup>1</sup>, Brunno Machado de Campos<sup>1</sup>, Wallace Souza Loos<sup>3</sup>, Marcelo Eduardo  
4 Ramos Barbosa<sup>2</sup>, Marina Koutsodontis Machado Alvim<sup>1</sup>, Gabriel Chagas Lutfala Paulino<sup>1</sup>, Francesco  
5 Pucci<sup>4</sup>, Samuel Worrell<sup>4</sup>, Roberto Medeiros de Souza<sup>1</sup>, Lara Jehi<sup>4</sup>, Fernando Cendes<sup>1</sup>

6 <sup>1</sup>Universidade Estadual de Campinas (UNICAMP), Neuroimaging Laboratory, Campinas, SP, Brazil

7 <sup>2</sup>School of Medical Sciences, Pontifical Catholic University of Campinas, Campinas, SP, Brazil

8 <sup>3</sup>Advanced Imaging and Artificial Intelligence Lab, University of Calgary, Calgary, AB, Canada

9 <sup>4</sup>Cleveland Clinic Foundation, Cleveland, OH, United States of America

10

## 11 ABSTRACT

12 The rapid and constant development of deep learning (DL) strategies is pushing forward the quality of  
13 object segmentation in images from diverse fields of interest. In particular, these algorithms can be very  
14 helpful in delineating brain abnormalities (lesions, tumors, lacunas, etc), enabling the extraction of  
15 information such as volume and location, that can inform doctors or feed predictive models. Here, we  
16 describe ResectVol DL, a fully automatic tool developed to segment resective lacunas in brain images of  
17 patients with epilepsy. ResectVol DL relies on the nnU-Net framework that leverages the 3D U-Net deep  
18 learning architecture. T1-weighted MRI datasets from 120 patients (57 women;  $31.5 \pm 15.9$  years old at  
19 surgery) were used to train (n=78) and test (n=48) our tool. Manual segmentations were carried out by  
20 five different raters and were considered as ground truth for performance assessment. We compared  
21 ResectVol DL with two other fully automatic methods: ResectVol 1.1.2 and DeepResection, using the  
22 Dice similarity coefficient (DSC), Pearson's correlation coefficient, and relative difference to manual  
23 segmentation. ResectVol DL presented the highest median DSC (0.92 vs. 0.78 and 0.90), the highest  
24 correlation coefficient (0.99 vs. 0.63 and 0.94), and the lowest median relative difference (9 vs. 44 and  
25 12 %). Overall, we demonstrate that ResectVol DL accurately segments brain lacunas, which has the  
26 potential to assist in the development of predictive models for postoperative cognitive and seizure  
27 outcomes.

28

29 **KEYWORDS**

30 Brain lacuna segmentation; Epilepsy; postoperative MRI; ResectVol DL; nnU-Net

31

32

## 33 1. Introduction

34 Epilepsy is a neurological disease with profound impacts on quality of life, morbidity, and mortality.

35 About 30-40 % of patients do not respond to treatment with medication (1,2) and can be referred to

36 surgical intervention, especially if structural alterations are detected (3,4). Many factors impact surgical

37 outcome and long-term seizure freedom is still limited, ranging from 40 to 80 % (5,6)

38 Studies have tried to determine the factors leading to surgical success (7–9) and nomograms were

39 created showing promising results for individualizing outcome prediction (5,10). However, most surgical

40 outcome studies still capture the surgical procedure under broad categorical classifications with

41 resolution limited to the lobe of resection (temporal/external temporal, hippocampal sparing versus

42 resecting).

43 Characterization of the lacuna can be informative to those predictive models by providing volume and

44 location of extracted brain tissue. Due to the laborious and time-consuming nature of the manual

45 annotation required to characterize resections, initiatives to automate this task increased in the last

46 years (11–13). Hence, the purpose of this study is twofold: (i) to introduce ResectVol DL, a fully

47 automatic method to segment brain lacunas, based on the 3D U-Net architecture (14); and (ii) to

48 compare ResectVol DL against two other methods: ResectVol (version 1.1.2), a rule-based processing

49 pipeline (12), and DeepResection, which is based on a 2D U-Net architecture (13).

## 50 2. Material and Methods

51 This study was conducted with approval from the Cleveland Clinical Foundation Institutional Review

52 Board. Informed consent was waived due to the retrospective nature of the data collection.

### 53 2.1. Subjects and image datasets

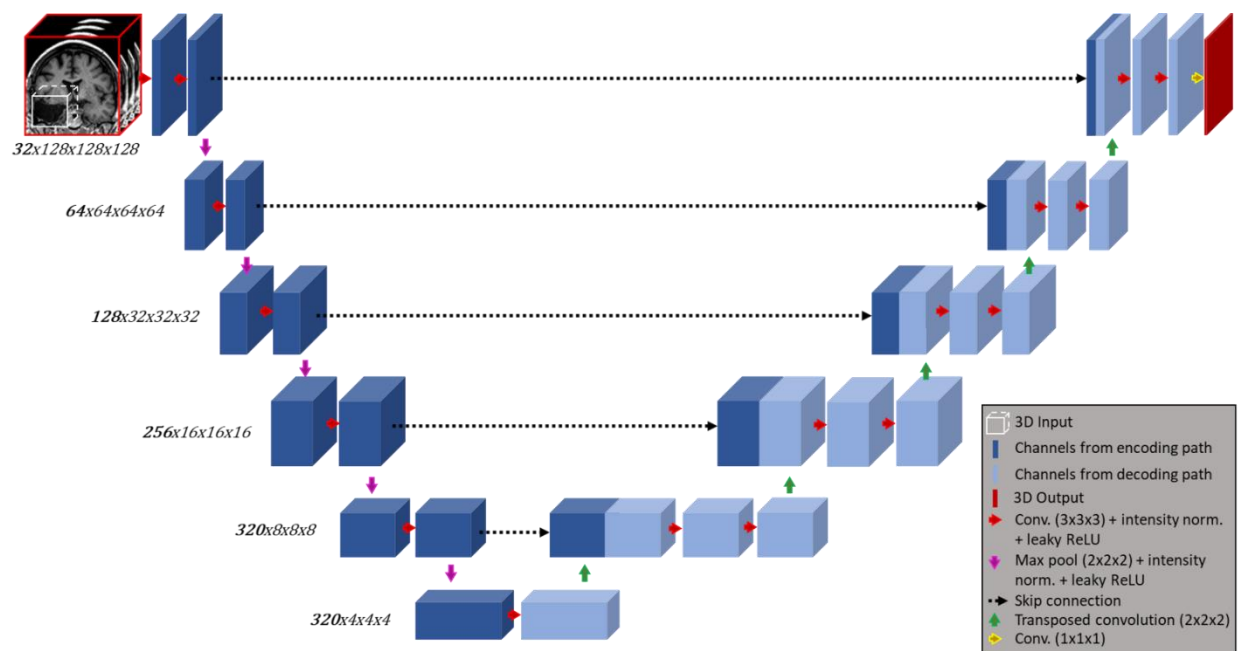
54 Image datasets from 125 patients were retrospectively selected from the Cleveland Clinic Epilepsy  
55 Center and the University of Campinas. We included subjects with temporal (n=45) and extratemporal  
56 lobe epilepsy (n=75), and volunteers with no brain lesions or resective intervention (n=5) to serve as  
57 controls for false discoveries (no-surgery data). Only resective surgeries were included.

## 58 2.2. Manual segmentation

59 Five raters trained in neuroanatomy performed the manual segmentation of the surgical lacunas using  
60 MRlcron (15). Forty-eight image datasets were segmented by either two (n=42) or three (n=6) raters and  
61 were analyzed for inter-rater agreement.

## 62 2.3. Automatic Segmentation

63 To perform the automatic segmentation, we created ResectVol DL, which is based on the nnU-Net deep  
64 learning framework (16). nnU-Net derives from the U-Net (14), a popular deep learning architecture  
65 composed of an encoder, a decoder, and skip connections that concatenate features from the encoder  
66 to the decoder part of the network. Each processing stage of the encoding and decoding streams has  
67 two blocks of operation comprising steps of convolution, intensity normalization, and linear  
68 regularization (**Fig. 1**). Data augmentation strategies included rotation, scaling, addition of Gaussian  
69 noise, Gaussian blurring, changes to brightness and contrast, simulation of low resolution, and  
70 mirroring. The model training was carried out with five cross-validation folds and 1000 epochs for each  
71 fold.



72  
 73 **Figure 1.** Representation of the 3D nnU-Net architecture. The number of channels and resolution at each  
 74 of the six stages are shown on the left (eg: there are 320 channels and a resolution of 4x4x4 in the  
 75 feature map of the sixth stage). Contiguous blocks after the green arrows represent the concatenation of  
 76 channels from the lower stage of the decoding branch (light blue) with channels at the same stage from  
 77 the encoding branch (dark blue).

78  
 79 To compare our algorithm against other freely available methods, we also performed the lacuna  
 80 prediction with two other tools: ResectVol version 1.1.2 (available at [lniunicamp.com/resectvol](http://lniunicamp.com/resectvol)), a rule-  
 81 based algorithm previously developed by our group that does not employ deep learning methods (12);  
 82 and the DeepResection tool (13) (available at [github.com/penn-cnt/DeepResection](https://github.com/penn-cnt/DeepResection)) which uses a set of  
 83 three 2D U-Nets, each in one anatomical plane, followed by voxel-wise majority voting. In ResectVol  
 84 1.1.2, users must provide the pre- and postoperative images to run the analysis, whereas in  
 85 DeepResection users are required to provide the postoperative image only.

## 86 2.4. Computational Resources

87 All automatic processing was performed on a 3.6 GHz Intel(R) Core(TM) i7-12700KF CPU with 32-GB  
88 RAM, a 48-GB NVIDIA RTX A6000, and a 20.04.1 Ubuntu operational system.

## 89 2.5. Performance Assessment

90 The Dice similarity coefficient (DSC) (17) was used as an indirect measure of overlap between manual  
91 and automatic segmentations. We also calculated the volume obtained from the manual and the  
92 automatic segmentations, and estimated Pearson's correlation and the volume difference percentage  
93 (manual volume as reference) between them.

94 Additionally, to investigate if there were differences associated with lacuna size, we split lacunas into  
95 two groups (large and small) based on their median volume and compared the DSCs. Regarding the  
96 inter-rater segmentation performance, we calculated the DSC for the 48 images that were manually  
97 segmented more than once.

## 98 3. Results

### 99 3.1. Subjects

100 One hundred and twenty patient volunteers (57 women;  $31.5 \pm 15.9$  years old at surgery) had temporal  
101 (n=45; 23 women;  $37.7 \pm 13.8$  years old) and extratemporal (n=75; 34 women;  $27.8 \pm 16.0$  years old)  
102 epilepsy, whereas five subjects (2 women;  $29.0 \pm 4.5$  years old) were included as controls who were not  
103 submitted to any brain surgery. Nearly all postoperative images were acquired 5 months after surgery  
104 (median: 6.2; Q1-Q3: 6.0 - 8.0; range: .2 - 63.9), to avoid necrotic tissue that could be misleading in the  
105 segmentation process. Seventy-three patients (~61 %) were seizure-free (Engel Ia) after surgery.

### 106 3.2. Imaging protocol

107 T1-weighted anatomical MR images were obtained in a total of eight different MRI systems (1.5 and 3 T)  
108 from two different vendors (Siemens and Philips). Acquisition parameters and image quality varied,  
109 benefiting the robustness of the performance assessment. Parameters can be summarized in the  
110 following ranges: voxel size =  $0.41 \times 0.45 \times 0.41$  to  $1 \times 2 \times 1$  mm<sup>3</sup>; TR = 7 to 2200 ms; TE = 2.3 to 46 ms;  
111 and image matrix = 180 x 96 to 512 x 512.

### 112 3.3. Manual segmentation

113 All the images containing brain lacunas (n=120) were manually segmented at least once. Three of the  
114 five raters timed the duration of manual delineation in a subset of images yielding the approximated  
115 median times of: 114 min (MERB, n=31); 70 min (RFC, n=14); 54 min (GCLP, n=19). Forty-eight images  
116 were segmented by two different raters at least, and the median DSC was .88 (Q1-Q3: .78 - .91; range:  
117 .00 - .97). We visually selected the best manual mask in these cases to compose the final set of 120  
118 manual masks used as ground truth in the subsequent analyses.

### 119 3.4. ResectVol DL

120 Seventy-two images were used to train the nnU-Net architecture. We found a median DSC of .92 (Q1-  
121 Q3: .88 - .94) and .91 (Q1-Q3: .86 - .94) for the test and the cross-validation set, respectively (**Table 1**).  
122 For 47 of the 48 test images, ResectVol DL was able to correctly identify the lacuna (DSC  $\geq$  .612) and it  
123 failed for only one extratemporal case – the smallest lacuna in the sample. The correlation between the  
124 manual and the automatic volumes was  $r(46) = .99$  ( $p < .001$ ), with a median relative difference of 9 %  
125 (Q1-Q3: 4 – 15 %) (**Table 1** and **Fig. 2**). ResectVol DL correctly identified the absence of lacunas in the  
126 five control datasets. The prediction for each image took, on average, 41.7 seconds.

127

**Table 1.** Dice Similarity Coefficient (DSC), correlation, and relative difference obtained with each method for the test set

		<b>ResectVol DL</b>	<b>ResectVol 1.1.2</b>	<b>DeepResection</b>
<b>All data</b>	median DSC (Q1-Q3)	<b>.92 (.88 - .94)</b>	.78 (.65 - .84)	.07 (0 - .89)*
	range	0 - .96	0 - .90	0 - .95*
	correlation (p-val)	.99 (< .001)	.63 (< .001)	----
	median rel. diff. (Q1-Q3)	9 (4 - 15) %	44 (29 - 75) %	----
<b>Temporal</b>	median DSC (Q1-Q3)	<b>.93 (.89 - .95)</b>	.79 (.75 - .86)	.90 (.86 - .92)
	range	.85 - .96	.21 - .90	.20 - .95
	correlation (p-val)	.97 (< .001)	.72 (<.001)	.94 (< .001)
	median rel. diff. (Q1-Q3)	8 (4 - 11) %	45 (30 - 65) %	12 (6 - 20) %
<b>Extra-temporal</b>	median DSC (Q1-Q3)	<b>.92 (.88 - .94)</b>	.73 (.61 - .79)	0 (0 - 0)*
	range	0 - .96	0 - .89	0 - .08*
	correlation (p-val)	.99 (< .001)	.62 (.001)	----
	median rel. diff. (Q1-Q3)	10 (4 - 20) %	37 (15 - 84) %	----
<b>No-surgery data</b>	<b>5 cases</b>	<b>0</b>	<b>5</b>	<b>0</b>

\* *DeepResection* authors do not recommend applying it to extra-temporal cases. These values are just reported for completeness. The associated correlation values, however, are not displayed due to the impossibility related to data distribution (see **Fig. 2A**, bottom plot).

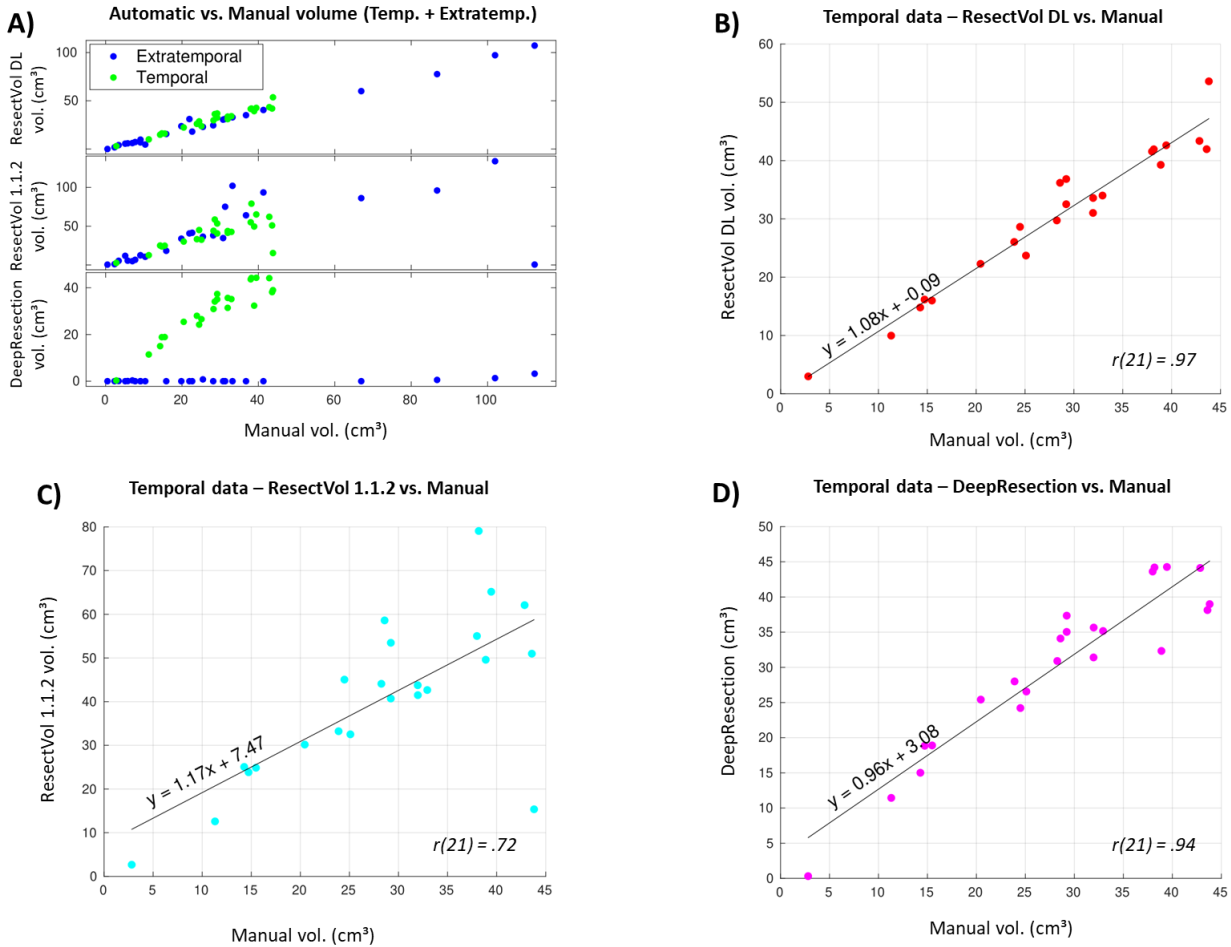
128

### 129 3.5. Performance of other approaches

130 The test set of 53 images (48 lacunas and five no-surgery controls) was also processed by ResectVol  
 131 1.1.2 and DeepResection.

132 ResectVol 1.1.2 requires no pre-processing of the images, although the origin of image space (x=0, y=0,  
 133 z=0) must not be too far off the anterior commissure, in which case reorienting is necessary. It took on  
 134 average nine minutes and 17 seconds to process each dataset composed of the pre- and the  
 135 postoperative images. We found a median DSC = .78 (Q1-Q3: .65 - .84), a correlation between the  
 136 manual and automatic volumes of  $r(46) = .63$  ( $p < .001$ ), and a median relative difference of 44 % (Q1-  
 137 Q3: 29 – 75 %). For the five control images, it incorrectly found a lacuna in all five cases (volume range:  
 138 .3 - 4.2 cm<sup>3</sup>) (**Table 1** and **Fig. 2**).





139

140 **Figure 2.** Scatter plots for (A) all data (temporal + extratemporal cases) and for (B-D) temporal data only

141 between each segmentation method and manual delineation. Plots B-D include the linear fit, the

142 equation that best adjusts to the data, and the correlation value  $r$ .

143

144 Like the other tools, DeepResection does not require any image preprocessing. The authors do

145 recommend, however, that images are in Left-Anterior-Superior orientation because predictions can be

146 inaccurate otherwise. We used *fslorient* function from the FMRIB Software Library (FSL) (18) to perform

147 the reorientation. The available version of DeepResection is only intended for temporal resections, thus

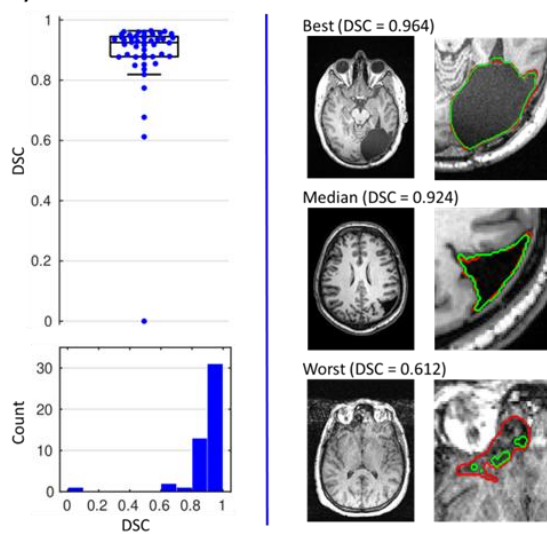
148 results for this tool are mostly restricted to this brain region. Temporal lacuna predictions overlapped

149 with the manual segmentation in all cases, yielding a DSC = .90 (.86 - .92). Median correlation and  
 150 relative difference were  $r(46) = .94$  ( $p < .001$ ) and 12 % (Q1-Q3: 6 – 20 %), respectively (Table 1 and Fig.  
 151 2).

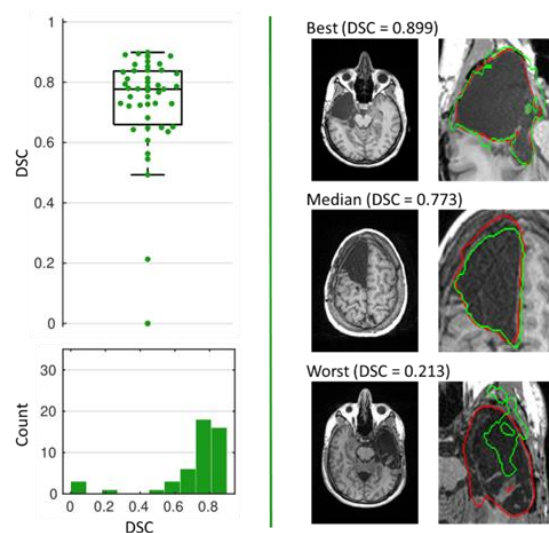
152 Fig. 3 exhibits the best, median, and worst DSC-associated cases for all tools.

153

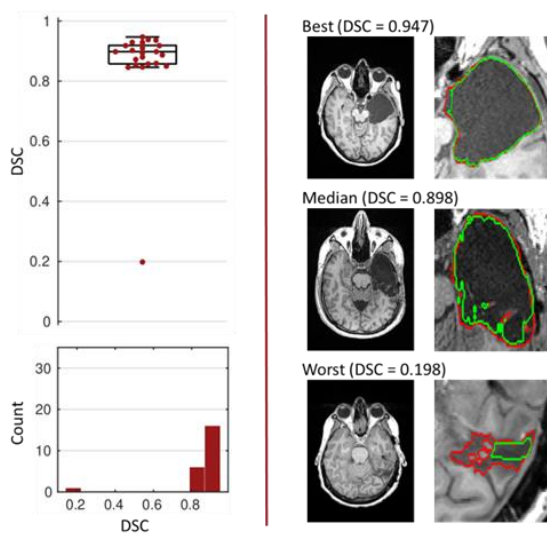
### A) ResectVol DL – all cases



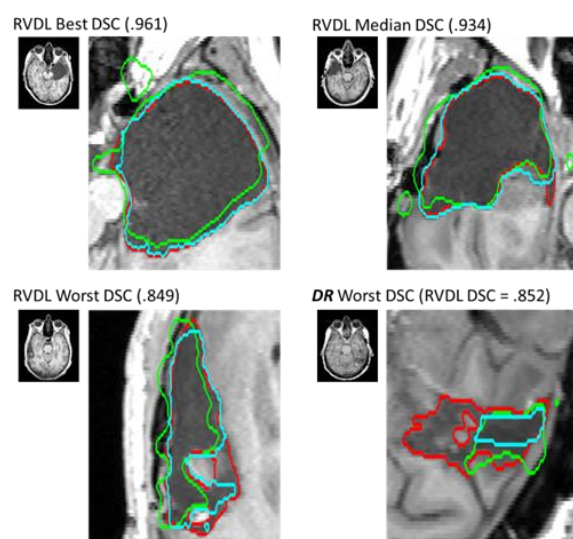
### B) ResectVol 1.1.2 – all cases



### C) DeepResection – temporal cases



### D) All tools – temporal cases

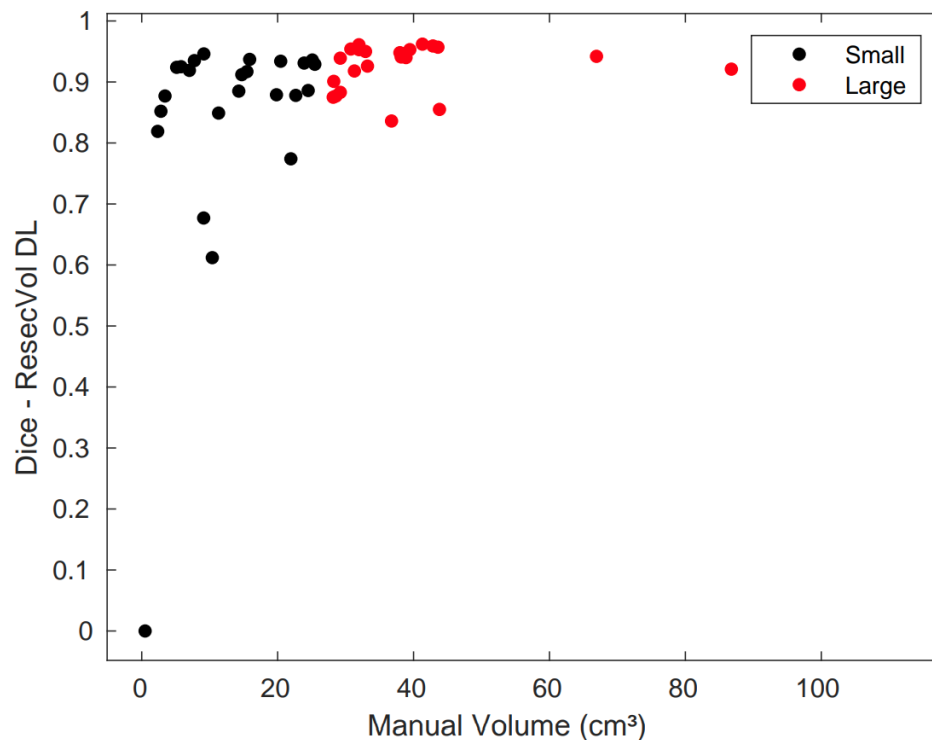


154

155 **Fig. 3.** Boxplots, histograms, and noteworthy cases for (A) ResectVol DL (RV2), (B) ResectVol 1.1.2  
156 (RV112), and (C) DeepResection (DR), along with Dice similarity coefficient (DSC). Red and green contours  
157 in A-C represent manual and automatic segmentation. Both versions of ResectVol (A and B) were applied  
158 to temporal and extratemporal data ( $n = 48$ ), whereas DeepResection (C) is only applied to temporal  
159 cases ( $n = 23$ ) due to its limitation. In D we display contours obtained with RVDL (red), RV112 (green),  
160 and DR (cyan) in temporal cases (to allow visual comparison across the three tools). Although included in  
161 boxplots and histograms, cases with  $DSC = 0$  are not shown in the figures.

### 162 3.6. Large vs. Small lacunas

163 The Spearman's correlation between volume (in  $\text{cm}^3$ ) and DSCs was  $r_s(46) = .60$  ( $p < .001$ ). To further  
164 assess this significant relationship, we ranked lacunas by their size obtained with the manual delineation  
165 and split our dataset into two subgroups based on the median lacuna volume ( $25.5 \text{ cm}^3$ ). We  
166 investigated whether there was any difference in DSC across the small ( $< 25.5 \text{ cm}^3$ ) and the large lacunas  
167 ( $> 25.5 \text{ cm}^3$ ) (Fig. 4). Since the data in each group was not normally distributed, we performed a non-  
168 parametric Mann-Whitney U test that revealed a significant difference between groups ( $U = 450.5$ ,  $p =$   
169  $.001$ ).



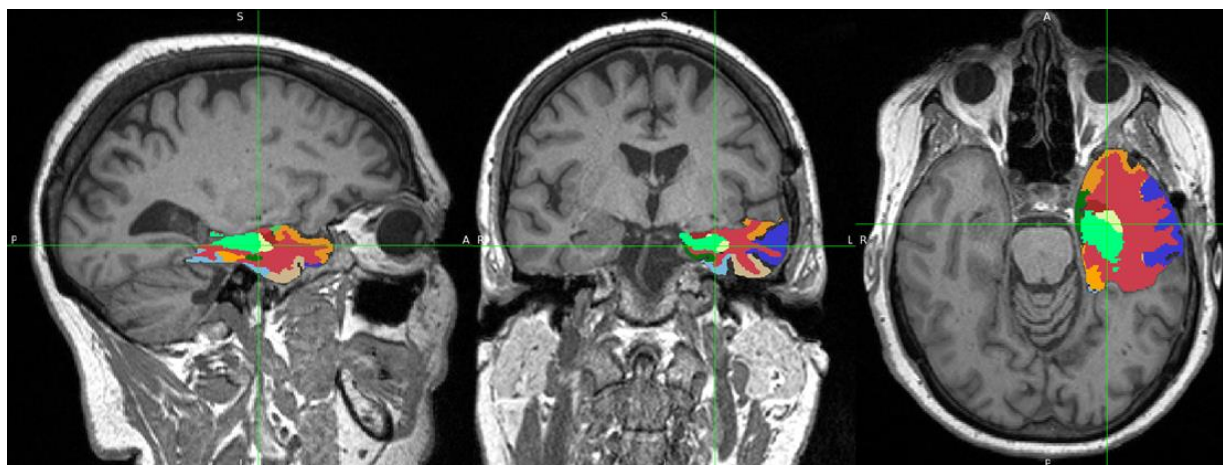
170  
171 **Fig. 4.** Distribution of Dice similarity coefficients (DSC) as a function of lacuna volume. Small lacunas  
172 (black dots) show more diverse values of DSC than large ones (red dots).

173

### 174 3.7. Structure Identification

175 After segmenting the lacuna, ResectVol DL combines the lacuna mask with the brain-extracted image to  
176 estimate the original 3D brain shape. This recreated whole brain undergoes a series of registration steps  
177 that register a labeled template in the Montreal Neurological Institute (MNI) space onto the space of the  
178 original image (native). We use the Desikan-Killiany-Tourville (DKT) atlas (19) as a reference for naming  
179 brain structures. Our pipeline uses functions from the Advanced Normalization Tools (ANTs) package  
180 (20), the Statistical Parametric Mapping (SPM) toolbox (21), and the FMRIB Software Library (FSL) (18).

181 **Fig. 5** illustrates an example of the labeled regions of a temporal patient along with the table containing  
182 the volumetric information.



Cluster 1:				
Volume: 253376 voxels (42624.98 mm <sup>3</sup> )				
ROI Index	ROI Name*	Resected vol (voxel)	Resected vol (mm <sup>3</sup> )	Resected %
1006	ctx-lh-entorhinal	6205	1043.86	80
17	Left-Hippocampus	14909	2508.11	71
18	Left-Amygdala	4255	715.81	69
1016	ctx-lh-parahippocampal	3061	514.95	44
1009	ctx-lh-inferiortemporal	27193	4574.63	43
1015	ctx-lh-middletemporal	29366	4940.19	38
1030	ctx-lh-superiortemporal	32002	5383.64	34
1007	ctx-lh-fusiform	7822	1315.88	21
2	Left-Cerebral-White-Matter	102551	17251.96	7
999	Undetermined**	25213	4241.54	---

183

184 **Fig. 5.** Region labeling in the lacuna segmentation (top) and the corresponding volumetric information

185 table (bottom).

## 186 4. Discussion

187 In this study, we compared three different segmenting methods against the manual delineation of brain

188 lacunas in T1-weighted MR images. We found that ResectVol DL, which relies on the nnU-Net

189 architecture, could outperform its previous rule-based implementation (version 1.1.2) and

190 DeepResection, a deep learning method based on the 2D U-Net architecture. To the best of our

191 knowledge, DeepResection, ResectVol 1.1.2, and ResectVol DL are the only freely available, fully

192 automatic tools dedicated to this task.

193 Lacuna segmentation can offer a means of estimating characteristics of the removed brain region and  
194 informing predictive models of surgery outcome. Manual segmentation is the obvious first approach  
195 considered, and arguably the one with the best quality, but it is susceptible to human error and  
196 consistency factors, such as inter and intra-rater variability. A subset of images from our study ( $n = 48$ )  
197 were segmented by two (or three) raters to estimate the inter-rater DSC. When more than two  
198 segmentations were available, we chose the highest associated DSC, yielding a more human-favorable  
199 estimate (DSC = .88) to be compared against the automatic methods.

200 Semi-automated methods have been implemented in which users need to click on a point inside the  
201 lacuna to carry out the segmentation. Gau et al. (22), for instance, employed itk-SNAP segmenting  
202 software (23) to perform the lacuna delineation by manually setting a seed inside the lacuna, followed  
203 by the itk-SNAP's region-growing algorithm, and found a median DSC = .78 (range: .53 - .94). Billardello  
204 and colleagues (11) created a tool that also requires the user to click on (or set the coordinates of) a  
205 voxel to be used as the seed for the region-growing algorithm. They found a median DSC = .83 (Q1-Q3:  
206 .72 - .85). The major advantage of these approaches is that the seed is always positioned in the right  
207 location where the lacuna is expected to be identified. Nevertheless, this approach still demands human  
208 intervention, is prone to bias—since manual positioning of the seed may lead to different results—, and,  
209 although we did not make a direct comparison on the same dataset, the DSC found for ResectVol DL  
210 (.93) is better than in these studies.

211 Among the fully automatic tools we tested, DeepResection was the fastest (~35 s per image), ResectVol  
212 DL was second (~42 s), whereas ResectVol 1.1.2 took much longer (~9 min) due to the nature of its  
213 processing algorithm which is based on a series of processing steps. When comparing median DSCs  
214 for temporal cases, ResectVol DL was the best one (DSC = .93) followed by DeepResection (.90) and  
215 ResectVol 1.1.2 (.79). Although DeepResection is capable of being fine-tuned to work with

216 extratemporal resections—as tested in the original publication (exploratory DSC=.75)—authors do not  
217 recommend using this preliminary fine-tuned model. Hence, for extratemporal cases, we only  
218 considered the results from ResectVol 1.1.2 and ResectVol DL, and, again, version DL was the winner  
219 (.92 vs .73).

220 As previously reported in other studies (13,22), we found a relationship between DSC and lacuna size  
221 indicating that there is a trend towards better performance for larger resection. Since DSC is dependent  
222 upon a volume ratio, large values in the numerator (intersection) and denominator (total volume) will  
223 be less sensitive to intersection mismatches when comparing two volumetric shapes. Hence, small  
224 lacunas are more prone to be affected by discrepancies and present more variability in DSCs (**Fig. 4**).

225 Besides the volumetric information obtained with such computational segmentation methods, this line  
226 of research may offer potential avenues for surgery planning and outcome prediction. Koepp et al. (24)  
227 showed that the removal of at least 50 % of the piriform cortex increased (by a factor of 16) the odds of  
228 becoming seizure-free, and no relationship was found with other mesiotemporal structures. Leon-Rojas  
229 and colleagues (25) implemented a pipeline, based on this result and on their dataset of resective  
230 surgeries, to automatically segment the piriform cortex and calculate chances of seizure freedom during  
231 planning. They were able to plan a new patient intervention with an estimated 50 % chance of long-term  
232 seizure freedom. The automatic segmentation of resected areas can help increase knowledge and  
233 accelerate the development of such tools. Since ResectVol can list areas removed along with their  
234 volume, it can also help serve that purpose.

235 Regarding manual segmentation, raters spent more time in our study (approximately or more than 1  
236 hour) than what has been reported by other authors (approximately 30 minutes) (12,22). We found that  
237 duration depends on lesion complexity (as reported in these studies), tool used, and rater experience.  
238 Computer systems can spare humans from these repetitive tasks and accelerate their completion,

239 besides avoiding the mentioned human error and variability. In the inter-rater assessment, we found a  
240 DSC = .88, comparable to previous reports on brain tumors (.84 - .86) (26) and lacuna resection (.84)  
241 (27). We noticed idiosyncrasies among raters, that were more pronounced in temporal resections,  
242 especially in the very medial region near the hippocampus. We do recommend that researchers  
243 carefully check, or standardize—if possible—, important resection areas if they anticipate that some  
244 regions may be controversial.

245 One limitation of our tool refers to not being tested on postoperative MRIs of other surgery types, like  
246 laser interstitial thermal therapy. However, we anticipate that, due to the different profiles of lesioned  
247 brain sites, a new network should be trained on this specific dataset to accommodate these discordant  
248 characteristics. We also have not validated the labeling of resected brain structures, which is an area  
249 that does deserve attention in future studies. Overall, ResectVol DL had the best performance across the  
250 tested tools, having failed for only one case out of 53 on which it was tested. It can be utilized in  
251 temporal and extratemporal cases and is freely available at [github.com/rfcasseb/resectvol\\_dl/](https://github.com/rfcasseb/resectvol_dl/).

## 252 5. Conclusion

253 ResectVol DL successfully segmented brain lacunas in MRIs of patients with epilepsy who have  
254 undergone surgery. ResectVol DL is fast, accurate, and does not require any image pre-processing. It  
255 relies solely on postoperative images to measure lacuna volumes besides providing estimates of the  
256 volumes of anatomical structures within the lacuna. This enables more precise estimation of volumetric  
257 information, which facilitates studies such as fMRI investigations performed before surgery. Moreover,  
258 it may aid in generating predictive models of surgical results.

259

## 260 Ethics statement



261 This study was conducted with approval from the Cleveland Clinical Foundation Institutional Review  
262 Board. Informed consent was waived due to the retrospective nature of the data collection.

263

## 264 Acknowledgements

265 This work was supported by the São Paulo Research Foundation [grant numbers 2020/00019-7 and  
266 2013/07559-3] and the National Institutes of Health [grant number R01NS097719].

267

## 268 CRediT authorship contribution statement

269 **Raphael Fernandes Casseb:** Conceptualization, Methodology, Software, Data Curation, Formal Analysis,  
270 Visualization, Writing – original draft, Writing – review & editing. **Brunno Machado de Campos:**  
271 Conceptualization, Methodology, Software, Formal Analysis, Writing – original draft, Writing – review &  
272 editing, **Wallace Souza Loos:** Methodology, Software, Writing – review & editing. **Marcelo Eduardo**  
273 **Ramos Barbosa:** Data Curation, Formal Analysis, Writing – review & editing. **Marina Koutsodontis**  
274 **Machado Alvim:** Data Curation, Investigation, Writing – review & editing, **Gabriel Chagas Lutfala**  
275 **Paulino:** Data Curation, Formal Analysis, Writing – review & editing. **Francesco Pucci:** Data Curation,  
276 Formal Analysis, Writing – review & editing. **Samuel Worrell:** Data Curation, Formal Analysis, Writing –  
277 review & editing. **Roberto Medeiros de Souza:** Methodology, Software, Writing – review & editing. **Lara**  
278 **Jehi:** Conceptualization, Resources, Project administration, Supervision, Writing – review & editing.  
279 **Fernando Cendes:** Conceptualization, Supervision, Resources, Project administration, Writing – original  
280 draft, Writing – review & editing.

281

282

## 283   References

284

285

286

287

288

289

290

291

292

293   1.    Chen Z, Brodie MJ, Liew D, Kwan P. Treatment outcomes in patients with newly diagnosed  
294        epilepsy treated with established and new antiepileptic drugs a 30-year longitudinal cohort study.  
295        *JAMA Neurol.* 2018;75(3).

296   2.    Löscher W, Potschka H, Sisodiya SM, Vezzani A. Drug resistance in epilepsy: Clinical impact,  
297        potential mechanisms, and new innovative treatment options. *Pharmacol Rev.* 2020;72(3).

298   3.    Duncan JS, Winston GP, Koepp MJ, Ourselin S. Brain imaging in the assessment for epilepsy  
299        surgery. Vol. 15, *The Lancet Neurology.* 2016.

300   4.    Bernasconi A, Cendes F, Theodore WH, Gill RS, Koepp MJ, Hogan RE, et al. Recommendations for  
301        the use of structural magnetic resonance imaging in the care of patients with epilepsy: A  
302        consensus report from the International League Against Epilepsy Neuroimaging Task Force.  
303        *Epilepsia.* 2019;60(6).

304   5.    Jehi L, Yardi R, Gonzalez-Martinez J, Chagin K, Kattan MW, Bartolomei F, et al. Development and  
305        validation of nomograms to provide individualised predictions of seizure outcomes after epilepsy  
306        surgery: a retrospective analysis. *Articles Lancet Neurol [Internet].* 2015 [cited 2021 Apr  
307        26];14:283–90. Available from: <http://dx.doi.org/10.1016/>

308   6.    Jobst BC, Cascino GD. Resective epilepsy surgery for drug-resistant focal epilepsy: A review. *JAMA*  
309        - *Journal of the American Medical Association.* 2015;313(3).

310   7.    Jehi L, Braun K. Does etiology really matter for epilepsy surgery outcome? *Brain Pathology.*  
311        2021;31(4).

312   8.    Kuroda N, Sonoda M, Miyakoshi M, Nariai H, Jeong JW, Motoi H, et al. Objective interictal  
313        electrophysiology biomarkers optimize prediction of epilepsy surgery outcome. *Brain Commun.*  
314        2021;3(2).

315   9.    Fitzgerald Z, Morita-Sherman M, Hogue O, Joseph B, Alvim MKM, Yasuda CL, et al. Improving the  
316        prediction of epilepsy surgery outcomes using basic scalp EEG findings. *Epilepsia.* 2021;62(10).

- 317 10. Busch RM, Hogue O, Miller M, Ferguson L, Mcandrews MP, Hamberger M, et al. Nomograms to  
318 Predict Verbal Memory Decline After Temporal Lobe Resection in Adults With Epilepsy.  
319 Neurology. 2021;97(3).
- 320 11. Billardello R, Ntolkeras G, Chericoni A, Madsen JR, Papadelis C, Pearl PL, et al. Novel User-Friendly  
321 Application for MRI Segmentation of Brain Resection following Epilepsy Surgery. Diagnostics  
322 2022, Vol 12, Page 1017 [Internet]. 2022 Apr 18 [cited 2023 Apr 18];12(4):1017. Available from:  
323 <https://www.mdpi.com/2075-4418/12/4/1017/htm>
- 324 12. Casseb RF, de Campos BM, Morita-Sherman M, Morsi A, Kondylis E, Bingaman WE, et al.  
325 ResectVol: A tool to automatically segment and characterize lacunas in brain images. Epilepsia  
326 Open [Internet]. 2021 Dec 1 [cited 2022 Apr 25];6(4):720–6. Available from:  
327 <https://onlinelibrary.wiley.com/doi/full/10.1002/epi4.12546>
- 328 13. Arnold TC, Muthukrishnan R, Pattnaik AR, Sinha N, Gibson A, Gonzalez H, et al. Deep learning-  
329 based automated segmentation of resection cavities on postsurgical epilepsy MRI. Neuroimage  
330 Clin [Internet]. 2022 Jan 1 [cited 2023 Jan 5];36:103154. Available from:  
331 [/pmc/articles/PMC9402390/](https://pubmed.ncbi.nlm.nih.gov/39402390/)
- 332 14. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image  
333 Segmentation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial  
334 Intelligence and Lecture Notes in Bioinformatics) [Internet]. 2015 [cited 2022 Apr 25];9351:234–  
335 41. Available from: [https://link.springer.com/chapter/10.1007/978-3-319-24574-4\\_28](https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28)
- 336 15. Rorden C, Brett M. Stereotaxic display of brain lesions. Behavioural Neurology. 2000;12(4):191–  
337 200.
- 338 16. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for  
339 deep learning-based biomedical image segmentation. Nature Methods 2020 18:2 [Internet]. 2020  
340 Dec 7 [cited 2022 Apr 25];18(2):203–11. Available from:  
341 <https://www.nature.com/articles/s41592-020-01008-z>
- 342 17. Dice LR. Measures of the Amount of Ecologic Association Between Species. Ecology [Internet].  
343 1945 Jul 1 [cited 2021 May 8];26(3):297–302. Available from:  
344 <https://esajournals.onlinelibrary.wiley.com/doi/full/10.2307/1932409>
- 345 18. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. Neuroimage [Internet].  
346 2012 Aug 15 [cited 2018 Nov 5];62(2):782–90. Available from:  
347 <http://www.ncbi.nlm.nih.gov/pubmed/21979382>
- 348 19. Klein A, Tourville J. 101 labeled brain images and a consistent human cortical labeling protocol.  
349 Front Neurosci. 2012;(DEC).
- 350 20. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs  
351 similarity metric performance in brain image registration. Neuroimage. 2011;54(3).
- 352 21. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ. Statistical parametric maps  
353 in functional imaging: A general linear approach. Hum Brain Mapp [Internet]. 1994 [cited 2018  
354 Mar 13];2(4):189–210. Available from: <http://doi.wiley.com/10.1002/hbm.460020402>

- 355 22. Gau K, Schmidt CSM, Urbach H, Zentner J, Schulze-Bonhage A, Kaller CP, et al. Accuracy and  
356 practical aspects of semi- and fully automatic segmentation methods for resected brain areas.  
357 *Neuroradiology* [Internet]. 2020 Dec 1 [cited 2021 May 8];62(12):1637–48. Available from:  
358 [/pmc/articles/PMC7666677/](#)
- 359 23. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour  
360 segmentation of anatomical structures: Significantly improved efficiency and reliability.  
361 *Neuroimage*. 2006;31(3).
- 362 24. Koepp MJ, Galovic M, Baudracco I, Wright-Goff E, Pillajo G, Nachev P, et al. Association of  
363 piriform cortex resection with surgical outcomes in patients with temporal lobe epilepsy. *JAMA*  
364 *Neurol*. 2019;76(6).
- 365 25. Leon-Rojas JE, Iqbal S, Vos SB, Rodionov R, Miserocchi A, McEvoy AW, et al. Resection of the  
366 piriform cortex for temporal lobe epilepsy: a Novel approach on imaging segmentation and  
367 surgical application. *Br J Neurosurg*. 2021;
- 368 26. Ermiş E, Jungo A, Poel R, Blatti-Moreno M, Meier R, Knecht U, et al. Fully automated brain  
369 resection cavity delineation for radiation target volume definition in glioblastoma patients using  
370 deep learning. *Radiation Oncology*. 2020;15(1).
- 371 27. Pérez-García F, Rodionov R, Alim-Marvasti A, Sparks R, Duncan JS, Ourselin S. Simulation of Brain  
372 Resection for Cavity Segmentation Using Self-supervised and Semi-supervised Learning. In:  
373 *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and*  
374 *Lecture Notes in Bioinformatics)*. 2020.
- 375