- 1 Development of a cloud framework for training and deployment of deep learning models in
- 2 Radiology: automatic segmentation of the human spine from CT-scans as a case-study
- 3 Rui Santos^{1*}, Nicholas Bünger^{1*}, Benedikt Herzog¹, and Sebastiano Caprara^{1#}
- 4
- 5 ¹Digital Medicine Unit, Balgrist University Hospital Zürich, University of Zürich, Forchstrasse 340,
- 6 8008, Zurich, CH, Switzerland.
- 7 *Equally contributing authors
- [#]Corresponding author: sebastiano.caprara@balgrist.ch 8
- 9
- 10

Abstract 11

12 Advancements in artificial intelligence (AI) and the digitalization of healthcare are revolutionizing

- 13 clinical practices, with the deployment of AI models playing a crucial role in enhancing
- diagnostic accuracy and treatment outcomes. Our current study aims at bridging image data 14
- collected in a clinical setting, with deployment of deep learning algorithms for the segmentation 15
- 16 of the human spine. The developed pipeline takes a decentralized approach, where selected
- 17 clinical images are sent to a trusted research environment, part of private tenant in a cloud
- 18 service provider. As a use-case scenario, we used the TotalSegmentator CT-scan dataset,
- 19 along with its annotated ground-truth spine data, to train a ResSegNet model native to the
- 20 MONAI-Label framework. Training and validation were conducted using high performance
- 21 GPUs available on demand in the Trusted Research Environment. Segmentation model performance benchmarking involved metrics such as dice score, intersection over union, 22
- accuracy, precision, sensitivity, specificity, bounding F1 score, Cohen's kappa, area under the 23
- curve, and Hausdorff distance. To further assess model robustness, we also trained a state-of-24
- 25 the-art nnU-Net model using the same dataset and compared both models with a pre-trained
- spine segmentation model available within MONAI-Label. The ResSegNet model, deployable 26
- 27 via MONAI-Label, demonstrated performance comparable to the state-of-the-art nnU-Net
- 28 framework, with both models showing strong results across multiple segmentation metrics. This
- 29 study successfully trained, evaluated and deployed a decentralized deep learning model for CT-
- 30 scan spine segmentation in a cloud environment. This new model was validated against state-
- of-the-art alternatives. This comprehensive comparison highlights the value of the MONAI-Label 31
- 32 as an effective tool for label generation, model training, and deployment, further highlighting its
- 33 user-friendly nature and ease of deployment in clinical and research settings. Further we also
- 34 demonstrate that such tools can be deployed in private and safe decentralized cloud
- environments for clinical use. 35
- 36
- 37 Key words: U-Net, ResSegNet, nnU-Net, deep-learning, framework, MONAI, cloud computing
- 38
- 39 Abbreviations:
- DL: deep learning; AI: artificial intelligence; AUC: area under the curve 40
- 41

42 Author Summary:

In the rapidly evolving field of medical imaging, the integration of artificial intelligence (AI) and 43 cloud computing is becoming increasingly critical for advancing diagnostic and treatment 44 capabilities. To address the growing demand for flexible digital frameworks in the clinical 45 environment supporting the deployment of data-driven applications, we have developed and 46 deployed a cloud-based Trusted Research Environment designed specifically for training. 47 validation, and deployment of deep learning models focused on semantic segmentation in 48 49 musculoskeletal radiology. This environment facilitates the efficient handling of large datasets and the accessibility of algorithmic output for the physicians, optimizing the interface between 50 development and clinical translation. The established framework enables significant 51 improvements in the deployment of deep learning tools for image analysis in the clinical setting. 52 53 In our current use-case, we have utilized this environment to train and evaluate two advanced 54 deep learning models for the segmentation of the human spine from CT scans. By leveraging 55 the computational power and flexibility of the cloud-based infrastructure, we were able to 56 perform rigorous training and comparison of these models, aiming to enhance the accuracy and reliability of spine segmentation in clinical practice. This approach not only streamlines the 57 process of model development but also provides valuable insights into the performance and 58 potential clinical applications of these Al-driven segmentation tools. 59

60

61 Introduction

62 With the inflection point of Artificial Intelligence (AI) algorithms, with its deep learning (DL)

- 63 models and frameworks (1), medical operations are facing the paradigm shift of both
- 64 digitalization and DL models deployment. Out of all clinical specialties, Radiology is poised to be
- at the forefront of this transformation due to its inherent reliance on visual data and the well-
- 66 established role of image analysis in diagnosis and treatment planning (2,3). As such, there is a
- 67 strong research focus on developing and deploying AI models within the Radiology field. As an
- example, in the past 5 years (2019-2023), a total of 11'392 research manuscripts (PubMed
- 69 search) were related in some form with the thematic of radiology and deep learning (DL).
- 70 Indeed, medical imaging is a key area for the development of high added-value DL models that
- streamline and accelerate the process of image segmentation and classification, which is a
- 72 major time-consuming task for physicians (4).
- 73 Several AI tools have been developed with the objective of accelerating medical image
- segmentation. Recently, the development of the TotalSegmentator DL model (5), trained to
- rs segment 117 anatomical structures (version 2.0), from CT datasets, demonstrates the immense

76 possibilities that DL-powered medical image segmentation offers. For the development of such

- a powerful DL segmentation model, the nnU-Net framework was used (6). This framework
- allows the training of state-of-the art segmentations models, based on the U-Net DL architecture
- 79 (7), with minimal to no coding needed. Still, these tools are limited in their deployment in daily
- 80 clinical scenarios. As an easier and user-oriented approach, the open-source framework MONAI
- 81 (Medical Open Network for Artificial Intelligence) (8) has been developed for the training and
- 82 easier deployment of automatic medical image segmentation models. At the forefront of MONAI,
- the sub-framework MONAI-Label (9) connects DL architectures and pretrained models with
- 84 open-source graphic user interfaces (GUI) such as 3D Slicer and OHIF (Open Health Imaging
- 85 Foundation Viewer), enabling active learning with direct visualization and correction of
- segmentation results (10,11).

87 As optimal environment for the deployment of such aforementioned frameworks, the adoption of 88 cloud computing in medical research setting, offers many benefits: it fosters improved 89 collaboration among physicians and researchers, provides a decentralized and scalable 90 computing infrastructure and greater accessibility to tools and services such as giving support for data annotation. Furthermore, cloud-based infrastructure offers simplified management of 91 resources in a cost-effective manner, allowing medical research centers to focus resources on 92 data analysis and clinical relevance of research questions rather than IT maintenance (12,13). 93 94 In this study, we have developed a cloud computing framework leveraging the open-source pipeline of MONAI-Label for the development and deployment for clinical research teams of an 95 interactive spine segmentation model out of CT data. Moreover, we benchmarked the model 96 output with an in-house trained nnU-Net model, both trained and validated with the same 97 collection of images (TotalSegmentator open-source dataset). We also compare it with a 98 pretrained model for spine segmentation present in MONAI-Label, trained on the VERSE 99 dataset (14). We demonstrate that the MONAI-Label developed model produces similar 100 101 segmentation results to the model trained with the nnU-Net framework yet providing a far

- simpler vertical integration in a clinical research setting. This *de novo* trained model also
- 103 considerably outperforms the pretrained spine segmentation model present in MONAI-Label *ab*
- *initio*. Moreover, the presented research highlights the use of MONAI-Label as a powerful tool
- for developing DL models and automatically generating anatomical segmentation of anatomical
 structures which may support surgical planning, diagnostics and future creation of digital twins
- 107 of patients (15,16).
- 108

109 Materials and Methods

110 Architecture of a trusted cloud research environment for deep learning model training and 111 deployment

The framework introduced in this work is based on a private cloud environment composed of 112 113 private endpoints, i.e., a pointer to the location of the data, which allows the hospital network to be extended in a private cloud space (figure 1A). Through an Application Programming Interface 114 (API) Manager, the DICOM (Digital Imaging and Communications in Medicine standard) or NIFTI 115 (Neuroimaging Informatics Technology Initiative) storage locations connect with cloud 116 117 computational nodes (with adequate hardware accelerators, such as graphic processing units -GPUs), and a centralized management dashboard that allows the definition of strict governance 118 policies, as the ones in place at healthcare institutions, granting secure and private access to 119 specified users. Through the combination of the cloud services mentioned above, the framework 120 composed of 3D Slicer and MONAI Label-based computational nodes was developed. In this way, 121 122 scalable computation resources are connected with a Graphical User Interface (GUI) thanks to 123 the open-source tool 3D Slicer, as shown on figure 1b.

- For segmentation model training in the trusted research environment, the TotalSegmentator open-source image dataset was stored in a project-specific NIFTI storage and connected with computational resources embedded in the computational node via the API Manager.
- For segmentation inference, an additional API allows the access to DICOM images stored in the clinical PACS (Picture Archiving and Communication System), retrieve these images to a DICOM storage location in the cloud which then directly connects to the MONAI Label computational node for automatic segmentation processing. The connection to the GUI results in a user-friendly environment where radiologists and other physicians can visualize the segmentation results, correct them, and potentially fine-tuning the models by sending the corrected results back to MONAI Label central instance to update the training.
- 134

135 Resources for deep learning model training

The TotalSegmentator dataset composed of CT images and their ground-truth vertebrae and 136 137 sacrum segmentations (part of the open-source dataset) was used for model training in either the MONAI-Label or nnU-Net frameworks (5) (https://zenodo.org/records/10047292, with an average 138 139 isotropic pixel size of 1.5 mm). The dataset was divided into training and inference dataset. 964 images were randomly chosen to be part of the training dataset, while the remaining 108 were 140 used for validation (the open-source version of the dataset contains more than 1200 images, 141 nonetheless, some cannot be used with the current version, 2.0.1). The dataset was 142 preprocessed in accordance with the requirements of both MONAI-Label and nnU-Net 143

144 (instructions found at https://github.com/Project-MONAI/MONAILabel?tab=readme-ovfile#getting-started-with-monai-label and https://github.com/MIC-DKFZ/nnUNet). The dataset 145 was uploaded into a storage location in the developed cloud-based Trusted Research 146 Environment. The default segmentation task in MONAI-Label, based on a ResSegNet 147 architecture (17), was used for spine segmentation model training. The training was performed in 148 4 A100 GPU computational nodes, for 1000 epochs, 241 iterations each, taking *circa* 25 hours of 149 150 compute time. Since for both training and results analysis a graphic user interface is necessary, the open-source 3D Slicer tool was used in a container (https://github.com/pieper/SlicerDockers, 151 v5.0.3). No validation split was used in this training as model benchmarking would be done with 152 an already assigned dataset (remaining 108 images available). Default segmentation pipeline 153 scripts were used with no modifications (segmentation.py, part of the radiology app from the 154 155 MONAI-Label framework with а tutorial found here: https://github.com/Project-156 MONAI/tutorials/tree/main/monailabel). The nnU-Net model training was performed on one A100 157 compute node for 1000 epochs, 250 iterations each, over the course of 15 hours of training. No 158 validation split was used (fold 0), as model benchmarking would be done with an already assigned dataset (same 108 datasets used to validate the MONAI-Label model). Default nnU-Net pipeline 159 scripts were used with no modifications, following the developers tutorial. 160

161

162 Deep learning model benchmarking

108 CT scans of the TotalSegmentator dataset (not part of the dataset for DL model training) 163 were used for model validation. We calculated different segmentation metrics to benchmark both 164 deep learning models' segmentation results such as dice score, intersection over union, accuracy, 165 precision, sensitivity, specificity, bounding F1 score, recall, area under the curve (AUC) and the 166 Hausdorff distance between segmentation predictions and their ground truth (18). Each metric 167 was calculated with custom scripts (available upon request). Calculation of distances between 168 169 ground truth and predicted segmentations was done in the open-source software Meshlab (https://www.meshlab.net/) v2022.02 with built-in functions. P-values were calculated based on a 170 171 non-paired t-student function.

172

173 Results

- 174 Benefits of a decentralized cloud infrastructure for DL model training and deployment
- 175 Deploying a decentralized trusted research environment on the cloud for training and deployment
- of deep learning models present two major benefits: on one hand the scalability and flexibility to

compose a tailored DL model for a specific research or clinical question, on the other hand, the
 possibility to deploy the developed models in clinical workflows using the same architecture.

Depending on the size and architecture of the model, the compute resources of the computational nodes can be quickly scaled-up or scaled-down, providing the creation of an optimized setup for each use-case requiring a training, validation and/or deployment of a research or clinical DL model. Further, this decentralized approach makes it far simpler to securely and privately connect different medical devices (CT scanners, MRI machines, ultrasound machines, etc.) as endpoints to the developed trusted research environment and get outputs in real-time.

Within the MONAI Label framework, multiple models can easily be retrieved and fine-tuned by 185 labeling additional data via a user-friendly GUI, in this way a local version can be trained. The 186 versatility of this process enables various healthcare operators, e.g. radiologists, to retrain models 187 on top of open-source available models (transfer learning). Furthermore, they can easily test 188 189 models that have been previously created in the framework and validate them thanks to 190 commonly available software for 3D visualization. This way, the interaction between model's 191 development and radiologists is considerably improved and physicians are a central component in the development of new ML models. 192

193

194 Training a ResSegNet based CT-scan spine segmentation model within the MONAI-Label 195 framework: spineCT-ResSegNet model

Out of 1072 images present in the TotalSegmentator, 964 images (90% of the dataset), were 196 randomly chosen to compose the training dataset, while 108 (remaining dataset) served as the 197 validation dataset. Default 3D ResSegNet parameters of the segmentation (ResSegNet) task of 198 199 MONAI-Label radiology app, were used. Qualitative analysis of the segmentation results (figure 2A) demonstrates the capacity of robust prediction of the entire human spine and sacrum (25 200 segmented classes). Vertebral bodies identification is correctly done, with their shape being in 201 202 accordance with anatomical expectancy. Quantitative segmentation quality metrics over all the 203 25 classes reveal a robust model for spine segmentation from CT-scans with an average dice 204 score of 0.856±0.066, an average intersection over union score of 0.801±0.074, average accuracy of 1.000±0.000, a mean precision of 0.862±0.066, a mean sensitivity (or recall) of 205 206 0.861±0.070, a mean specificity of 1.000±0.000, an average bounding F1 score of 0.837±0.057, a mean area under the curve (AUC) of 0.932±0.034 and a Kappa of 0.853±0.068 (figure 2B, all 207 numbers rounded to 3 decimal places). Detailed results, with 95% confidence intervals, are found 208 209 in supplementary figure 1.

210

211 Training a nnU-Net based CT-scan spine segmentation model: spineCT-nnUNet model

212 Serving as a benchmark for DL segmentation model comparison, we trained the nnU-Net framework on the same dataset with the 3D-UNet full resolution model (1.5 mm), using the 213 default configuration and trained the model over 1000 epochs. Validation was performed with the 214 same 108 CT scans as for the MONAI-Label ResSegNet model, native to MONAI-Label. 215 Segmentation output is qualitatively robust with anatomical structures, i.e. vertebrae, are localized 216 217 to their correct position and have a correct tridimensional conformational (figure 3A). the segmentation prediction and their ground-truth, reveal a robust prediction model, with an 218 219 average dice score over the 25 classes (anatomical structures) of 0.914±0.036, an average intersection over union of 0.879±0.045, an accuracy of 1.000±0.000, a precision of 0.911±0.046, 220 a sensitivity (or recall) of 0.910±0.050, a specificity of 1.000±0.000, a boundary F1 score of 221 222 0.908±0.029, an AUC score of 0.961±0.018 and a Kappa score of 0.904±0.049 (figure 3B, 5 all 223 numbers rounded to 3 decimal places). Detailed results, with 95% confidence intervals, are found 224 in supplementary figure 2.

225

226 Comparison of spineCT-ResSegNet, spineCT-nnUNet and pretrained spine segmentation 227 models

To further compare the segmentation predictions between the ResSegNet model and the nnU-228 229 Net model to the respective ground truth, we overlaid selected examples as exemplified in figure 4A (spineCT-ResSegNet) and 4C (spineCT-nnUNet). Hausdorff distances between each 230 segmentation model predictions and their ground truths were calculated (spineCT-ResSegNet 231 and spineCT-nnUNet in figure 4B and figure 4D respectively). The spineCT-ResSegNet model 232 presents an average Hausdorff distance of 2.604±0.450 mm while the spineCT-nnUNet model 233 predictions present a Hausdorff distance of 2.214±0.398. While the nnU-Net based model 234 outperforms the ResSegNet model in regard to the quality metric of Hausdorff distance, the same 235 236 is not observed for the remaining metrics calculated. Thus, we find sufficient evidence to say both 237 models produce on-par segmentation results (supplementary figure 5A, all p-valued rounded to 5 238 decimal plates).

We also benchmarked the pretrained vertebrae segmentation model which is part of the MONAI-Label core. The model spine segmentation predictions average a dice score over the 24 classes (as this model only predicts from C1 to L5) of 0.601±0.153, an average intersection over union of 0.533±0.154, an accuracy of 1.000±0.000, a precision of 0.697±0.150, a sensitivity (or recall) of 0.565±0.157, a specificity of 1.000±0.000, a boundary F1 score of 0.588±0.144, an AUC score of 0.785±0.080 and a Kappa score of 0.597±0.151 (supplementary figure 3, all numbers rounded

to 3 decimal places). The output segmentations present a Hausdorff distance of 3.129±0.322 mm
on average. Thus, we can demonstrate that the *de novo* trained model spineCT-ResSegNet
outperforms the pretrained model in MONAI-Label in most quality metrics calculated
(supplementary figure 5B, all p-valued rounded to 5 decimal plates). Furthermore, it also trained
in one more class, the sacrum.

250

251 Discussion

252 A decentralized trusted research environment as a core of novel digital ecosystems in medical 253 research has the potential to simplify communication between physicians and researchers, enable effective testing and validation of novel tools, provide secure and private data exchange channels 254 - and ultimately reduce costs by sharing computational resources, with non-complicated 255 256 scalability. Further benefits are the ability to access common datasets, shared code, and 257 validation methods at any time reduces errors and improves usability - and thus accelerates the 258 clinical translation of digital technologies. Altogether, such platforms facilitate testing and 259 validation of digital and DL pipelines, making data-driven medical research more reproducible and better suited for the deployment in clinical settings. With these principles in mind, we developed 260 a radiology-specific cloud infrastructure, based on a combination of DICOM storage services and 261 MONAI-Label-compatible computational nodes for the training and deployment of deep learning 262 segmentation models. 263

As a case-study, we trained and deployed a ResSegNet model for human spine segmentation 264 out the open-source dataset of CT-scans from TotalSegmentator. This model was trained as part 265 of the MONAI-Label (8,19) framework. As a benchmark, we used the nnU-Net framework (6) and 266 267 also trained a segmentation model with the same training and inference dataset. Both models yielded similar quality metric results with only average Hausdorff distance being noticeable better 268 for the nnU-Net model, than for the ResSegNet model (figure 4 and supplementary figure 5). 269 270 Nonetheless, given the native deployment of the ResSegNet model via MONAI-Label, we can 271 directly deploy it via a graphic user interface such as 3D Slicer or OHIF (10,11), which is not possible with the nnU-Net, at least not in a straightforward way. This in-place visualization 272 provides direct access to the segmentation results to the health professionals who can use it to 273 274 make informed decisions, e.g. for surgical planning.

We also compared our trained model, spineCT-ResSegNet, with the pre-trained and pre-built spine segmentation model which MONAI-Label already offers (14). Our newly trailed model greatly outperforms the prebuilt model in all metrics, providing substantially more accurate segmentations (supplemental figure 4 and 5).

279 While our developed models underperformed compared with the dice scores reported for the 280 TotalSegmentator model (average 0.95 for vertebrae) (5), this nnU-Net based model was trained for 8000 epochs compared to ours which was trained for only 1000 epochs. This meant a far lower 281 282 training time for our model, fewer computational resources were needed, while still performing adequately for the task at hand. 283

Naturally, given the dice coefficient of around 0.85 and a hausdorff coefficient of approximately 284 285 2.6 mm, the ResSegNet model incurs in mispredictions and proper segmentation correction may 286 be necessary. Nonetheless, the iterative nature and easy of use of MONAI-Label and a chosen GUI such as 3d Slicer, makes any necessary corrections easy to perform. Further, this corrections 287 may be included as part of a subsequent finetuned model, as part of the labelling correction 288 process, which will then, increase the predictive power of the model. 289 290

291 Conclusion

292 The present research demonstrates the potential of cloud-based deep learning for training and 293 deployment of medical image segmentation models, in a trusted research environment. By leveraging the efficiency and scalability of cloud computing, we have created a framework that 294 empowers medical professionals, such as radiologists, with faster, more accessible segmentation 295 tools, based on the MONAI-Label open-source framework. This paves the way for improved 296

- diagnostics, treatment planning, and ultimately, better patient outcomes. 297
- 298

Author Contribution 299

NB developed and deployed the trusted cloud environment. RS worked on the development and 300 301 testing of the deep learning segmentation models. BH assisted NB in the infrastructure setting of the trusted cloud environment. NB, RS and SC wrote the manuscript and designed the figures. 302 SC supervised the design of the infrastructure and pipeline and contributed to the design of the 303 research and interpretation of the results. MF contributed to the final version of the manuscript. 304

306 Acknowledgements

307 The authors acknowledge the Digital Society Initiative, University of Zurich. This work is part of the DSI-Infrastructure / Lab-Program. 308

309

305

310 References

- Sevilla J, Heim L, Ho A, Besiroglu T, Hobbhahn M, Villalobos P. Compute Trends Across Three Eras 311 1. 312 of Machine Learning. Proceedings of the International Joint Conference on Neural Networks. 313 Institute of Electrical and Electronics Engineers Inc.; 2022. doi: 10.1109/IJCNN55064.2022.9891914. 314
- 315 2. Saba L, Biswas M, Kuppili V, et al. The present and future of deep learning in radiology. Eur J Radiol. Elsevier Ireland Ltd; 2019. p. 14–24. doi: 10.1016/j.ejrad.2019.02.038. 316

- Chea P, Mandell JC. Current applications and future directions of deep learning in 317 3. musculoskeletal radiology. Skeletal Radiol. Springer; 2020. p. 183–197. doi: 10.1007/s00256-019-318 319 03284-z.
- 320 4. Langlotz CP. The Future of AI and Informatics in Radiology: 10 Predictions. Radiology. Radiological Society of North America Inc.; 2023. doi: 10.1148/radiol.231114. 321
- Wasserthal J, Breit HC, Meyer MT, et al. TotalSegmentator: Robust Segmentation of 104 322 5. 323 Anatomic Structures in CT Images. Radiol Artif Intell. Radiological Society of North America Inc.; 324 2023;5(5). doi: 10.1148/RYAI.230024/ASSET/IMAGES/LARGE/RYAI.230024.VA.JPEG.
- 325 6. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for 326 deep learning-based biomedical image segmentation. Nature Methods 2020 18:2. Nature 327 Publishing Group; 2020;18(2):203–211. doi: 10.1038/s41592-020-01008-z.
- 328 7. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image 329 segmentation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag; 2015. p. 234–241. doi: 330 10.1007/978-3-319-24574-4 28. 331
- 332 8. Jorge Cardoso M, Li W, Brown R, et al. MONAI: An open-source framework for deep learning in 333 healthcare. 2022;
- 334 9. Diaz-Pinto A, Alle S, Nath V, et al. MONAI Label: A framework for AI-assisted Interactive Labeling 335 of 3D Medical Images. 2022; doi: 10.1016/j.media.2024.103207.
- 336 10. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging. 2012;30(9):1323–1341. doi: 337 338 10.1016/j.mri.2012.05.001.
- 339 11. Ziegler E, Urban T, Brown D, et al. Open Health Imaging Foundation Viewer: An Extensible Open-340 Source Framework for Building Web-Based Imaging Applications to Support Cancer Research. JCO 341 Clin Cancer Inform. 2020. https://doi.org/10.
- 342 12. Griebel L, Prokosch HU, Köpcke F, et al. A scoping review of cloud computing in healthcare. BMC 343 Med Inform Decis Mak. BioMed Central Ltd; 2015. doi: 10.1186/s12911-015-0145-7.
- 13. 344 Navale V, Bourne PE. Cloud computing applications for biomedical science: A perspective. PLoS Comput Biol. Public Library of Science; 2018;14(6). doi: 10.1371/journal.pcbi.1006144. 345
- 346 14. Sekuboyina A, Husseini ME, Bayat A, et al. VERSE: A Vertebrae labelling and segmentation 347 benchmark for multi-detector CT images. Med Image Anal. Elsevier B.V.; 2021. doi: 348 10.1016/j.media.2021.102166.
- 349 15. Caprara S, Carrillo F, Snedeker JG, Farshad M, Senteler M. Automated Pipeline to Generate 350 Anatomically Accurate Patient-Specific Biomechanical Models of Healthy and Pathological FSUs. 351 Front Bioeng Biotechnol. Frontiers Media S.A.; 2021;9. doi: 10.3389/fbioe.2021.636953.

- Iqbal JD, Krauthammer M, Biller-Andorno N. The Use and Ethics of Digital Twins in Medicine.
 Journal of Law, Medicine and Ethics. Cambridge University Press; 2022;50(3):583–596. doi:
 10.1017/jme.2022.97.
- Saxena N, Babu NK, Raman B. Semantic segmentation of multispectral images using res-seg-net model. Proceedings - 14th IEEE International Conference on Semantic Computing, ICSC 2020.
 Institute of Electrical and Electronics Engineers Inc.; 2020. p. 154–157. doi: 10.1109/ICSC.2020.00030.
- 359 18. Müller D, Soto-Rey I, Kramer F. Towards a guideline for evaluation metrics in medical image
 360 segmentation. BMC Res Notes. BioMed Central Ltd; 2022. doi: 10.1186/s13104-022-06096-y.
- 19. Diaz-Pinto A, Alle S, Nath V, et al. MONAI Label: A framework for AI-assisted Interactive Labeling
 of 3D Medical Images. 2022; http://arxiv.org/abs/2203.12362.
- 363

364

365 Figure Legends

366

367 Figure 1: Architecture of the cloud infrastructure behind the trusted research environment and deployed framework with 3D Slicer and MONAI-Label. (A) Overview of the components 368 of the architecture of the trusted research environment: storage and compute. The first one has 369 a DICOM blob storage which is exposed with a private endpoint: while the second one has more 370 371 components. The compute workspace is exposed as well with a private endpoint, while the other component such as storage account, key vault Jupyter notebook and container registry are 372 connected to the workspace. Both systems have group permissions that can be managed 373 individually. To access the trusted research environment, users can use a Virtual Desktop 374 application (B) Using two compute nodes, 3D Slicer and MONAI Label are deployed separately. 375 376 With an SSH connection users can access an online environment to use the 3D Slicer interface 377 to interact with the model's output that runs on the first compute node (from the left side), which 378 process the DICOM images retrieved by the second compute node on which MONAI Label is 379 running from the DICOM blob storage.

380

Figure 2 – SpineCT-ResSegNet spine segmentation model predictions and segmentation

382 **metrics.** (A) Example of segmentation results with the spineCT-ResSegNet, a trained MONAI-

Label native model, in both 2D and 3D visualizations. Vertebrae are color coded. The entire human spine is represented in the image: cervical spine (C1-C7), thoracic spine (T1-T12),

Lumbar spine (L1-L5) and sacrum. (B) Validation metrics for spineCT-ResSegNet spine

segmentation model, based on the prediction of 108 cases. Color encodes the accuracy of
 segmentation class predictions. Blue indicates high similarity to the ground truth, while red

- 388 represents increasing deviation from the actual values.
- 389

Figure 3 – SpineCT-nnUNet spine segmentation model predictions and segmentation

391 **metrics.** (A) Example of segmentation results with the nnU-Net framework derived model (3D-

full resolution) in both 2D and 3D visualizations. Vertebrae are color coded. The entire human

- spine is represented in the image: cervical spine (C1-C7), thoracic spine (T1-T12), Lumbar
- spine (L1-L5) and sacrum. (B) Validation metrics for spineCT-nnUNet spine segmentation
- 395 model. Color encodes the accuracy of segmentation class predictions. Blue indicates high
- similarity to the ground truth, while red represents increasing deviation from the actual values.
- 397

398 Figure 4 – Comparison of spineCT-ResSegNet and spineCT-nnUNet semantic segmentation models. (A) Example of the Hausdorff distance calculation between the spineCT-399 ResSegNet model (native to MONAI-Label) predictions and ground truth. (B) Average Hausdorff 400 distance calculation for spineCT-ResSegNet predictions versus their ground truth. Predictions 401 with the spineCT-ResSegNet model average a Hausdorff Distance of 2.604±0.450 mm to the 402 ground truth. (C) Example of the Hausdorff distance calculation between the spineCT-nnUNet 403 404 model predictions and ground truth. (D) Average Hausdorff distance calculation for spineCTnnUNet predictions versus their ground truth. Predictions with the spineCT-nnUNet model 405 average a Hausdorff Distance of 2.214±0.398 mm to the ground truth. Color encoding in (B) and 406 (D) correlates with the accuracy of segmentation class predictions. Blue indicates high similarity 407 to the ground truth, while red represents increasing deviation from the actual values. 408

409

410 Supplemental Figure 1 – Comprehensive metrics evaluation of the spineCT-ResSegNet. 411 Average metrics for the spineCT-ResSegNet model based on comparison between the 412 predictions of the model on 108 datasets and their ground truth.

413

Supplemental Figure 2 – Comprehensive metrics evaluation of the spineCT-nnUNet. 414 Average metrics for the spineCT-ResSegNet model based on comparison between the 415 predictions of the model on 108 datasets and their ground truth. 416

417

Supplemental Figure 3 – Comprehensive metrics evaluation of the open-source vertebrae 418 segmentation model of MONAI-Label. Average metrics for the spineCT-ResSegNet model 419 based on comparison between the predictions of the model on 108 datasets and their ground 420 421 truth.

422

423 Supplementary Figure 4 – Hausdorff distance between prediction and ground truth by the open-source vertebrae segmentation model of MONAI-Label. (A) Example of the Hausdorff 424 425 distance calculation between the spineCT-ResSegNet model (native to MONAI-Label) predictions and ground truth. (B) The pre-trained model present in MONAI-Label, trained on the VERSE 426 427 dataset averages a Hausdorff Distance of 3.129±0.322 mm between model segmentation prediction and ground truth. 428

429

Supplementary Figure 5 – Statistical significance analysis between the quality 430 segmentation metrics for each method used. (A) On average, both models, spineCT-431

ResSegNet and spineCT-nnUNet outputs are similar with the later one outperforming the MONAI-Label native model, only in regard to Hausdorff distance. Thus, both models appear to have a relatively equal prediction quality. (B) The newly trained spineCT-ResSegNet model mostly outperforms the pre-built spine segmentation model present in the MONAI-Label framework. Further, the model was trained in one extra class, the sacrum, capable of performing robust predictions for the anatomical class.

Figure 1





Figure 2

А

 Raw CT
 Overlay
 ResSegNet 3D Semantic Segmentation

 Image: A state of the state o

В

Metric	DICE So	core	lol	loU		асу	Precis	sion	Sensi	ivity	Specif	icity	Boundary	F1 Score	AUC	;	Кар	ра
Class	Average	SD	Average	SD	Average	SD	Average	SD										
C1	0.807	0.330	0.758	0.314	1.000	0.000	0.796	0.327	0.820	0.336	1.000	0.000	0.792	0.325	0.910	0.168	0.806	0.330
C2	0.891	0.219	0.846	0.226	1.000	0.000	0.877	0.233	0.915	0.212	1.000	0.000	0.880	0.212	0.957	0.106	0.891	0.219
C3	0.891	0.211	0.840	0.207	1.000	0.000	0.907	0.204	0.877	0.213	1.000	0.000	0.894	0.134	0.939	0.107	0.890	0.211
C4	0.896	0.110	0.826	0.155	1.000	0.000	0.895	0.145	0.911	0.090	1.000	0.000	0.890	0.107	0.955	0.045	0.896	0.110
C5	0.647	0.405	0.589	0.382	1.000	0.000	0.668	0.423	0.655	0.397	1.000	0.000	0.654	0.399	0.827	0.198	0.647	0.405
C6	0.737	0.311	0.656	0.306	1.000	0.000	0.762	0.324	0.725	0.343	1.000	0.000	0.734	0.296	0.872	0.165	0.723	0.324
C7	0.812	0.222	0.724	0.225	1.000	0.000	0.874	0.204	0.792	0.219	1.000	0.000	0.783	0.196	0.896	0.110	0.812	0.222
T1	0.883	0.159	0.814	0.172	1.000	0.000	0.839	0.166	0.945	0.146	1.000	0.000	0.846	0.134	0.973	0.073	0.883	0.159
T2	0.892	0.215	0.846	0.221	1.000	0.001	0.905	0.212	0.882	0.218	1.000	0.000	0.883	0.183	0.941	0.109	0.892	0.215
T3	0.811	0.277	0.748	0.289	1.000	0.000	0.840	0.268	0.783	0.309	1.000	0.000	0.815	0.237	0.899	0.147	0.799	0.292
T4	0.804	0.282	0.739	0.295	1.000	0.001	0.775	0.289	0.850	0.268	1.000	0.000	0.787	0.251	0.925	0.134	0.803	0.282
T5	0.830	0.267	0.771	0.279	1.000	0.001	0.838	0.267	0.826	0.270	1.000	0.001	0.823	0.230	0.913	0.135	0.830	0.267
T6	0.782	0.311	0.722	0.319	1.000	0.001	0.783	0.310	0.786	0.313	1.000	0.000	0.780	0.269	0.893	0.157	0.782	0.312
T7	0.816	0.259	0.748	0.280	1.000	0.001	0.798	0.293	0.813	0.298	1.000	0.000	0.805	0.224	0.920	0.133	0.793	0.289
T8	0.879	0.195	0.819	0.214	1.000	0.001	0.858	0.234	0.888	0.207	1.000	0.000	0.850	0.173	0.951	0.089	0.866	0.219
Т9	0.878	0.229	0.830	0.241	1.000	0.001	0.883	0.237	0.877	0.221	1.000	0.001	0.853	0.202	0.938	0.111	0.878	0.229
T10	0.894	0.223	0.855	0.237	1.000	0.001	0.911	0.217	0.882	0.231	1.000	0.001	0.877	0.195	0.941	0.116	0.894	0.223
T11	0.917	0.171	0.877	0.194	1.000	0.001	0.915	0.169	0.923	0.175	1.000	0.000	0.895	0.147	0.961	0.088	0.917	0.172
T12	0.915	0.166	0.872	0.190	1.000	0.001	0.909	0.195	0.909	0.194	1.000	0.000	0.885	0.152	0.961	0.081	0.903	0.195
L1	0.900	0.198	0.856	0.220	1.000	0.001	0.917	0.197	0.906	0.175	1.000	0.000	0.870	0.191	0.953	0.087	0.900	0.198
L2	0.870	0.211	0.815	0.247	1.000	0.001	0.920	0.189	0.864	0.189	1.000	0.000	0.851	0.181	0.932	0.095	0.870	0.211
L3	0.904	0.162	0.853	0.197	1.000	0.002	0.912	0.172	0.912	0.154	1.000	0.001	0.869	0.150	0.956	0.077	0.904	0.163
L4	0.883	0.211	0.833	0.230	1.000	0.001	0.908	0.218	0.889	0.177	1.000	0.001	0.846	0.188	0.944	0.088	0.883	0.211
L5	0.914	0.134	0.859	0.154	1.000	0.001	0.906	0.103	0.936	0.137	1.000	0.000	0.867	0.130	0.968	0.068	0.914	0.134
Sacrum	0.956	0.033	0.917	0.055	1.000	0.000	0.965	0.044	0.949	0.039	1.000	0.000	0.905	0.044	0.975	0.020	0.956	0.033
Average	0.856		0.801		1.000		0.862		0.861		1.000		0.837		0.932		0.853	
SD	0.066		0.074		0.000		0.066		0.070		0.000		0.057		0.034		0.068	

Figure 3

А

 Raw CT
 Overlay
 nnU-Net 3D Semantic Segmentation

 Image: A state of the state of

В

Metric	DICE So	core	lol	J	Accur	acy	Precis	sion	Sensitivity	(recall)	Specif	icity	Boundary	F1 Score	AU	:	Kap	ра
Class	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD	Average	SD
C1	0.843	0.303	0.805	0.308	1.000	0.000	0.847	0.337	0.790	0.353	1.000	0.000	0.838	0.296	0.917	0.155	0.806	0.345
C2	0.911	0.220	0.880	0.227	1.000	0.000	0.901	0.230	0.928	0.213	1.000	0.000	0.908	0.213	0.964	0.107	0.911	0.220
C3	0.936	0.086	0.890	0.123	1.000	0.000	0.970	0.025	0.914	0.122	1.000	0.000	0.943	0.033	0.957	0.061	0.936	0.086
C4	0.901	0.170	0.848	0.203	1.000	0.000	0.940	0.105	0.900	0.192	1.000	0.000	0.915	0.103	0.950	0.096	0.901	0.170
C5	0.794	0.271	0.720	0.287	1.000	0.000	0.795	0.336	0.750	0.332	1.000	0.000	0.822	0.250	0.904	0.138	0.747	0.324
C6	0.909	0.086	0.843	0.128	1.000	0.000	0.791	0.296	0.852	0.299	1.000	0.000	0.911	0.064	0.978	0.033	0.815	0.291
C7	0.923	0.148	0.879	0.164	1.000	0.000	0.907	0.170	0.959	0.080	1.000	0.000	0.907	0.132	0.979	0.040	0.923	0.148
T1	0.933	0.164	0.901	0.181	1.000	0.000	0.930	0.180	0.942	0.142	1.000	0.000	0.924	0.147	0.971	0.071	0.933	0.164
T2	0.916	0.208	0.888	0.225	1.000	0.001	0.922	0.203	0.914	0.214	1.000	0.000	0.914	0.185	0.957	0.107	0.916	0.209
T3	0.889	0.264	0.863	0.272	1.000	0.000	0.881	0.284	0.881	0.264	1.000	0.000	0.886	0.244	0.948	0.121	0.876	0.284
T4	0.878	0.274	0.850	0.279	1.000	0.000	0.881	0.281	0.893	0.243	1.000	0.000	0.872	0.255	0.946	0.121	0.878	0.274
T5	0.921	0.194	0.889	0.204	1.000	0.001	0.926	0.192	0.917	0.198	1.000	0.000	0.914	0.164	0.958	0.099	0.921	0.194
T6	0.894	0.196	0.848	0.227	1.000	0.001	0.887	0.211	0.912	0.189	1.000	0.000	0.890	0.174	0.956	0.094	0.894	0.196
T7	0.885	0.227	0.843	0.255	1.000	0.000	0.864	0.268	0.867	0.268	1.000	0.000	0.884	0.205	0.947	0.112	0.860	0.267
T8	0.934	0.154	0.900	0.177	1.000	0.000	0.935	0.154	0.936	0.159	1.000	0.000	0.920	0.141	0.968	0.079	0.933	0.154
Т9	0.946	0.140	0.919	0.164	1.000	0.000	0.949	0.138	0.947	0.140	1.000	0.000	0.933	0.128	0.974	0.070	0.946	0.140
T10	0.939	0.177	0.917	0.195	1.000	0.000	0.946	0.170	0.935	0.178	1.000	0.000	0.930	0.159	0.968	0.089	0.939	0.177
T11	0.945	0.166	0.923	0.180	1.000	0.000	0.952	0.153	0.941	0.171	1.000	0.000	0.936	0.140	0.971	0.085	0.945	0.166
T12	0.935	0.181	0.910	0.200	1.000	0.000	0.921	0.211	0.933	0.201	1.000	0.000	0.926	0.152	0.973	0.085	0.922	0.209
L1	0.919	0.201	0.890	0.224	1.000	0.000	0.925	0.216	0.913	0.215	1.000	0.000	0.911	0.182	0.963	0.093	0.906	0.227
L2	0.917	0.185	0.883	0.217	1.000	0.001	0.920	0.190	0.939	0.125	1.000	0.001	0.912	0.161	0.970	0.063	0.917	0.185
L3	0.939	0.161	0.912	0.186	1.000	0.001	0.943	0.165	0.943	0.155	1.000	0.001	0.925	0.154	0.971	0.078	0.939	0.161
L4	0.926	0.199	0.902	0.218	1.000	0.001	0.936	0.180	0.932	0.193	1.000	0.000	0.918	0.163	0.966	0.097	0.926	0.199
L5	0.953	0.090	0.921	0.123	1.000	0.000	0.934	0.156	0.944	0.154	1.000	0.000	0.933	0.082	0.981	0.044	0.937	0.154
Sacrum	0.968	0.035	0.940	0.054	1.000	0.000	0.961	0.055	0.977	0.011	1.000	0.000	0.931	0.043	0.989	0.006	0.968	0.035
Average	0.914		0.879		1.000		0.911		0.910		1.000		0.908		0.961		0.904	
SD	0.036		0.045		0.000		0.046		0.050		0.000		0.029		0.018		0.049	

Figure 4



Metric Class C1 C2 C3 C4 C5 C5 C6 C7 T1 T2 T3 T4 T6 T7 T7 T8 T6 T7 T8 T6 T7 T8 T9 T10 T11 L1 L2 L2 L3 L4 L5 Sacrum	Average SD 0.807 0.530 0.807 0.530 0.801 0.210 0.802 0.211 0.804 0.211 0.807 0.031 0.807 0.031 0.802 0.215 0.883 0.159 0.811 0.277 0.844 0.227 0.844 0.227 0.870 0.267 0.879 0.195 0.879 0.196 0.870 0.223 0.917 0.716 0.915 0.768 0.870 0.211 0.864 0.221 0.917 0.711 0.945 0.721 0.946 0.221 0.970 0.211 0.864 0.221 0.946 0.133	Dice Score 95%, Confidence Interval Lower 0.669 0.768 0.788 0.847 0.507 0.652 0.788 0.845 0.745 0.745 0.768 0.768 0.765 0.834 0.826 0.834 0.828 0.834 0.828 0.834 0.878 0.857 0.858 0.859 0.858 0.859 0.878 0.859 0.878 0.859 0.878 0.859 0.878 0.859 0.979 0.859 0.979 0.947	95% Confidence Interval Upper 0.944 0.945 0.985 0.944 0.787 0.821 0.944 0.877 0.876 0.986 0.921 0.944 0.877 0.876 0.924 0.924 0.925 0.926 0.944 0.953 0.944 0.953 0.946 0.922 0.944 0.953 0.946 0.942 0.946 0.946 0.946 0.946 0.946 0.946 0.946	Metric Class Average C1 0.756 0.756 C3 0.846 0.876 C5 0.826 0.866 C5 0.826 0.866 C7 0.724 0.814 T1 0.814 12 T4 0.739 0.746 T4 0.739 0.746 T6 0.774 0.746 T9 0.830 0.712 T6 0.872 1.1 U2 0.814 1.2 16 0.825 7.11 0.819 0.872 1.1 13 0.855 7.11 0.815 0.815 0.3 1.3 0.825 0.815 1.4 0.826 0.2315 1.3 0.826 0.2415 1.3 0.826 0.826	SD 95%. Confidence Inte 0.314 0.627 0.207 0.746 0.155 0.758 0.324 0.456 0.336 0.575 0.227 0.670 0.172 0.773 0.289 0.668 0.279 0.714 0.289 0.668 0.279 0.714 0.280 0.682 0.279 0.714 0.280 0.682 0.279 0.714 0.280 0.682 0.241 0.775 0.237 0.801 0.190 0.822 0.247 0.755 0.197 0.803 0.247 0.755 0.805 0.803 0.190 0.829 0.264 0.270 0.270 0.805 0.197 0.803 0.264 0.270 0.274 0.755 0.805 0.903	loU rval Lower 085% Confidence Interval Upper 0.889 0.933 0.933 0.933 0.934 0.721 0.739 0.779 0.856 0.900 0.817 0.810 0.840 0.840 0.841 0.869 0.869 0.885 0.908 0.908 0.920 0.915 0.908 0.920 0.915 0.908 0.920 0.921 0.932	Metric Class G12 G2 G2 G3 G3 G3 G4 G5 G6 C7 T1 T2 T3 T4 T6 T6 T7 T7 T9 T10 T11 T12 L1 L2 L3 L4 L5 Sacrum	Average SD I 1,000 0,000 1,000 1,000 0,000 1,000 1,000 0,000 1,000 1,000 0,000 1,000 1,000 0,000 1,000 1,000 0,000 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,000 1,000 0,001 1,001 1,000 0,001 1,001 1,00	Accuracy 95%. Confidence interval Lower 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000 0.969 0.969 1.000 1.	95% Confidence Interval Upper 1.000
Average SD Metric Class C1 C2 C3 C3 C3 C4 C5 C6 C6 C7 T1 T3 T3 T4 T3 T4 T5 T6 T7 T7 T8 T9 T10 T11 T12 L1 L2 L2 L3 L4 L5 Sacrum	0.856 0.856 0.066 0.066 0.066 0.075 0.0752 0.0752 0.023 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.233 0.907 0.234 0.855 0.455 0.752 0.232 0.907 0.234 0.859 0.752 0.232 0.907 0.234 0.839 0.752 0.232 0.966 0.968 0.425 0.752 0.234 0.839 0.752 0.234 0.839 0.752 0.234 0.839 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.245 0.752 0.254 0.752 0.254 0.752 0.254 0.752 0.257 0.258 0.275 0.257 0.258 0.234 0.752 0.257 0.258 0.234 0.752 0.257 0.258 0.234 0.752 0.257 0.258 0.234 0.257 0.258 0.234 0.257 0.258 0.234 0.257 0.258 0.234 0.257 0.258 0.234 0.257 0.555 0.257 0.555 0.254 0.257 0.555 0.257 0.555 0.257 0.555 0.257 0.555 0.257 0.555 0.257 0.555 0.257 0.555 0.257 0.555 0.257 0.155	Precision 95%. Confidence Interval Lower 0.057 0.808 0.808 0.827 0.516 0.672 0.824 0.798 0.853 0.775 0.704 0.775 0.706 0.833 0.833 0.828 0.828 0.828 0.828 0.877 0.874 0.871 0.874 0.875 0.875 0.875 0.861 0.851 0.953	95% Confidence Interval Upper 0.941 0.941 0.941 0.952 0.962 0.962 0.820 0.851 0.925 0.880 0.957 0.995 0.946 0.996 0.905 0.946 0.905 0.946 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.996 0.997 0.964 0.977 0.977	Average 0.801 SD 0.804 SD 0.074 Matric 0.803 GB3 Average GB3 0.803 G2 0.805 G2 0.805 G3 0.816 G5 0.855 G6 0.725 G7 0.726 G13 0.875 T4 0.845 T4 0.856 T6 0.826 T7 0.813 T9 0.877 T10 0.882 T11 0.926 T2 0.844 U3 0.912 U3 0.936 L3 0.936 Sacrum	SD Ser 0.336 0.01dence http 0.212 0.810 0.212 0.811 0.213 0.811 0.214 0.811 0.212 0.811 0.213 0.812 0.397 0.812 0.343 0.630 0.219 0.738 0.146 0.909 0.288 0.744 0.270 0.840 0.288 0.743 0.297 0.840 0.208 0.743 0.209 0.842 0.75 0.884 0.75 0.845 0.75 0.845 0.75 0.865 0.75 0.845 0.77 0.841 0.77 0.841 0.77 0.841 0.76 0.871 0.177 0.843 0.177 0.843 0.177 0.841 0.177 0.843 0.177	stilivity (recall) val Lower 05% Confidence Interval Upper 05% 0786 0788 0,788 0,819 0,846 0,881 0,885 0,885 0,885 0,885 0,885 0,885 0,885 0,885 0,885 0,986 0,985 0,995 0,	Average B Class Clas	1.000 0.000 Attation 0.001 1.000 0.000 <tr< th=""><th>Specificity 95%. Confidence Interval Lower 1000 100</th><th>95% Confidence Interval Upper 1000</th></tr<>	Specificity 95%. Confidence Interval Lower 1000 100	95% Confidence Interval Upper 1000
Average SD Metric Class C1 C2 C3 C4 C5 C6 C7 T1 T3 T4 T6 T7 T6 Sacrum Average SD	0.862 0.066 0.762 0.252 0.762 0.252 0.864 0.134 0.850 0.134 0.850 0.134 0.850 0.134 0.850 0.134 0.850 0.134 0.850 0.134 0.850 0.134 0.846 0.134 0.846 0.237 0.754 0.226 0.753 0.253 0.777 0.251 0.846 0.134 0.850 0.152 0.850 0.152 0.850 0.152 0.850 0.151 0.855 0.152 0.855 0.151 0.855	Boundary F1 Score 95%. Confidence Interval Lower 0.659 0.633 0.633 0.736 0.736 0.736 0.736 0.755 0.755 0.754 0.754 0.754 0.752 0.754 0.752 0.810 0.807 0.833 0.825 0.826 0.851 0.826 0.833 0.825 0.833 0.833	95% Confidence Interval Upper 0.201 0.21 0.24 0.37 0.37 0.792 0.814 0.830 0.878 0.227 0.871 0.886 0.886 0.889 0.821 0.927 0.919 0.914 0.895 0.908 0.894 0.901 0.916	Average 0.861 SD 0.070 Metric Class Average Class Average 0.81 Class Average 0.81 Class 0.910 0.910 Class 0.930 0.935 C6 0.927 0.973 C7 0.986 0.972 C7 0.986 0.973 T2 0.941 0.918 T6 0.918 177 T7 0.925 179 T9 0.931 171 T11 0.951 179 T12 0.941 1 T13 0.951 1 T9 0.924 1 T11 0.951 1 U2 0.941 0.941 T13 0.942 1 U2 0.942 0.944 U3 0.942 0.944 U3 0.942 0.944 U3 0.942 </td <td>SD 95%. Confidence Inte 0.160 0.835 0.107 0.837 0.107 0.837 0.108 0.837 0.109 0.934 0.148 0.756 0.109 0.914 0.147 0.863 0.147 0.863 0.135 0.877 0.135 0.879 0.135 0.879 0.135 0.879 0.135 0.899 0.136 0.931 0.147 0.863 0.028 0.942 0.088 0.942 0.081 0.942 0.082 0.942 0.084 0.942 0.085 0.935 0.020 0.969</td> <td>AUC vval Lower 95% Confidence Interval Upper 0.604 1 990 0.907 0.899 0.917 0.923 0.990 0.968 0.395 0.968 0.395 0.968 0.395 0.964 0.972 0.972 0.974 0.967 0.961 0.379 0.974 0.965 0.396 0.39 0.39 0.39 0.39 0.39 0.39 0.39 0.39</td> <td>Average SD White Class Total Tal <t< td=""><td>1.000 Averation CD 0.806 0.230 0.806 0.240 0.806 0.241 0.806 0.211 0.886 0.211 0.886 0.212 0.828 0.123 0.823 0.252 0.833 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.833 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.846 0.249 0.864 0.249 0.864 0.211 0.964 0.163 0.970 0.211 0.944 0.163 0.955 0.333 0.853 0.253</td><td>Kappa 95%. Confidence Interval Lower 0.660 0.780 0.633 0.757 0.844 0.633 0.757 0.844 0.839 0.729 0.724 0.725 0.845 0.825 0.842 0.875 0.842 0.859 0.859 0.853 0.818 0.859 0.853 0.818 0.828 0.878 0.878</td><td>95% Confidence Interval Upper 0.933 0.931 0.922 0.948 0.793 0.812 0.867 0.922 0.945 0.873 0.873 0.874 0.945 0.873 0.930 0.945 0.945 0.930 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945</td></t<></td>	SD 95%. Confidence Inte 0.160 0.835 0.107 0.837 0.107 0.837 0.108 0.837 0.109 0.934 0.148 0.756 0.109 0.914 0.147 0.863 0.147 0.863 0.135 0.877 0.135 0.879 0.135 0.879 0.135 0.879 0.135 0.899 0.136 0.931 0.147 0.863 0.028 0.942 0.088 0.942 0.081 0.942 0.082 0.942 0.084 0.942 0.085 0.935 0.020 0.969	AUC vval Lower 95% Confidence Interval Upper 0.604 1 990 0.907 0.899 0.917 0.923 0.990 0.968 0.395 0.968 0.395 0.968 0.395 0.964 0.972 0.972 0.974 0.967 0.961 0.379 0.974 0.965 0.396 0.39 0.39 0.39 0.39 0.39 0.39 0.39 0.39	Average SD White Class Total Tal Tal <t< td=""><td>1.000 Averation CD 0.806 0.230 0.806 0.240 0.806 0.241 0.806 0.211 0.886 0.211 0.886 0.212 0.828 0.123 0.823 0.252 0.833 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.833 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.846 0.249 0.864 0.249 0.864 0.211 0.964 0.163 0.970 0.211 0.944 0.163 0.955 0.333 0.853 0.253</td><td>Kappa 95%. Confidence Interval Lower 0.660 0.780 0.633 0.757 0.844 0.633 0.757 0.844 0.839 0.729 0.724 0.725 0.845 0.825 0.842 0.875 0.842 0.859 0.859 0.853 0.818 0.859 0.853 0.818 0.828 0.878 0.878</td><td>95% Confidence Interval Upper 0.933 0.931 0.922 0.948 0.793 0.812 0.867 0.922 0.945 0.873 0.873 0.874 0.945 0.873 0.930 0.945 0.945 0.930 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945</td></t<>	1.000 Averation CD 0.806 0.230 0.806 0.240 0.806 0.241 0.806 0.211 0.886 0.211 0.886 0.212 0.828 0.123 0.823 0.252 0.833 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.833 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.832 0.247 0.846 0.249 0.864 0.249 0.864 0.211 0.964 0.163 0.970 0.211 0.944 0.163 0.955 0.333 0.853 0.253	Kappa 95%. Confidence Interval Lower 0.660 0.780 0.633 0.757 0.844 0.633 0.757 0.844 0.839 0.729 0.724 0.725 0.845 0.825 0.842 0.875 0.842 0.859 0.859 0.853 0.818 0.859 0.853 0.818 0.828 0.878 0.878	95% Confidence Interval Upper 0.933 0.931 0.922 0.948 0.793 0.812 0.867 0.922 0.945 0.873 0.873 0.874 0.945 0.873 0.930 0.945 0.945 0.930 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945 0.945

Metric Dice Score	Metric Capitidance Interval Linner Class Average	IoU	95% Confidence Interval Upper	Metric	Accuracy	95% Confidence Interval Lippor
Class Average SD 95% Contidence Interval Lower B31 C1 0.044 0.30 0.716 0.716 0.716 C3 0.044 0.30 0.716 0.836 0.866 C3 0.044 0.300 0.716 0.826 0.865 C4 0.0401 0.170 0.8266 0.865 0.7 C5 0.090 0.066 0.885 0.7 0.933 0.144 0.887 C7 0.233 0.144 0.883 0.724 0.803 0.77 C7 0.232 0.144 0.872 0.74 0.812 0.74 C3 0.388 0.264 0.825 0.74 0.813 0.72 C1 0.374 0.154 0.896 0.72 0.74 0.814 T6 0.374 0.154 0.896 0.77 0.899 0.711 0.935 0.177 0.899 0.711 0.999 0.711 0.999 0.774 <	Confidences hierwall Upper Class Average 0.370 C1 0.260 0.375 C2 0.805 0.375 C4 0.805 0.375 C4 0.806 0.375 C4 0.848 0.388 C5 0.720 0.397 C7 0.828 0.359 C7 0.873 0.367 T2 0.888 0.362 T3 0.863 0.363 T6 0.883 0.363 T6 0.883 0.369 T7 0.883 0.369 T8 0.904 0.376 T9 0.931 0.377 T14 0.932 0.365 L1 0.883 0.362 L12 0.883 0.362 L13 0.917 0.365 L1 0.922 0.377 L4 0.922 0.377 Sacroma 0.430 0.240	BD BS-K confidence linkerval Lower 0.877 0.677 0.203 0.677 0.203 0.760 0.227 0.620 0.148 0.839 0.146 0.839 0.181 0.858 0.272 0.798 0.272 0.798 0.274 0.833 0.275 0.783 0.276 0.784 0.277 0.783 0.276 0.784 0.277 0.784 0.279 0.784 0.271 0.835 0.272 0.831 0.195 0.873 0.196 0.883 0.220 0.886 0.224 0.838 0.224 0.838 0.225 0.284 0.226 0.217 0.228 0.218 0.217 0.829 0.218 0.847 0.524 0.526	Box Confidence (Interval Upper 0.337 0.337 0.346 0.337 0.357 0.337 0.361 0.377 0.371 0.342 0.345 0.342 0.345 0.342 0.346 0.342 0.347 0.343 0.343 0.343 0.343 0.343 0.363 0.363 0.363 0.362 0.362 0.365 0.342 0.365 0.364 0.365	Class Average 30 C1 1.000 0.000 0.000 C3 1.000 0.000 0.000 C3 1.000 0.000 0.000 C3 1.000 0.000 0.000 C4 1.000 0.000 0.000 C5 1.000 0.000 0.000 C7 1.000 0.000 0.000 T1 1.000 0.000 100 T3 1.000 0.000 116 T3 1.000 0.000 117 T9 1.000 0.000 119 T9 1.000 0.000 111 T10 0.000 0.000 111 T10 1.000 0.000 112 T10 1.000 0.000 112 T11 1.000 0.000 112 T12 1.000 0.000 112 1.000 T11 1.000 0.000 112 1.00	35% Contractice Interval Lower 1.000	35%, Gondeance Interval Upper 1.000
Average 0.914 SD 0.036	Average 0.879 SD 0.045			Average 1.000 SD 0.000		
Motic Precision Class Average 50 95% Confidence Interval Lower 955 Class 0.967 0.337 1.000 955 Class 0.970 0.337 1.000 955 Class 0.970 0.25 0.701 0.23 0.701 Class 0.970 0.26 0.766 0.27 0.701 0.26 0.677 0.26 0.677 0.27 0.208 0.284 0.891 0.26 0.677 0.713 0.713 0.930 0.180 0.865 172 0.927 0.214 0.8712 0.865 172 0.9281 0.284 0.8712 173 0.9887 0.211 0.8777 175 0.9857 0.211 0.8177 175 0.935 0.154 0.800 179 0.946 0.170 0.917 171 0.927 0.211 0.918 1.4 0.921 0.211 0.918 1.4 0.9221 0.211 0.918 1.4 0.922 0.211	S Confidence hierval Upper Citass Average 0.000 0.993 Citass Average 0.993 Citass 0.928 0.928 0.990 Citass 0.928 0.928 0.990 Citass 0.930 Citass 0.930 0.990 Citass 0.930 Citass 0.930 0.991 Citass 0.930 Citass 0.930 0.992 Citass 0.930 Citass 0.930 0.940 Titas 0.942 0.942 0.942 0.974 Titas 0.941 0.942 0.941 0.975 Titass 0.936 0.912 0.936 0.400 T7 0.936 0.937 0.938 0.947 Titass 0.938 0.938 0.938 0.947 Titass 0.938 1.14 0.933 0.9487 L12 0.939 0.936 1.14 0.933 0.949 L13 0.942 0.93	Sensitivity (res:11) 0.353 0.633 0.633 0.213 0.835 0.132 0.212 0.855 0.122 0.353 0.634 0.233 0.122 0.851 0.322 0.233 0.534 0.244 0.260 0.939 0.774 0.042 0.807 0.241 0.842 0.907 0.244 0.862 0.264 0.813 0.744 0.862 0.264 0.740 0.815 0.264 0.759 0.868 0.044 0.719 0.804 0.171 0.717 0.903 0.215 0.715 0.862 0.204 0.715 0.802 0.125 0.125 0.862 0.125 0.125 0.862 0.125 0.125 0.862 0.125 0.190 0.881 0.154 0.193 0.881 0.154 0.193 0.881 </td <td>25% Confidence Interval Upper 0.44 1.25 0.972 0.990 0.865 0.331 0.977 0.347 0.347 0.347 0.347 0.347 0.357 0.357 0.367 0.367 0.373 0.373 0.373 0.373 0.375 0.380 0.380</td> <td>Matrix Class Average SD C1 1.000 0.000 C3 0.000 C6 0.000 C6 0.000 C6 0.000 C6 0.000 C7 1.000 0.000 C7 1.000 0.000 T1 1.000 0.000 T3 T3.000 0.000 T4</td> <td>Specificity 95% Confidence Interval Lower 95% Confidence Interval Lower 1000 1.000</td> <td>95% Confidence Interval Upper 1.000</td>	25% Confidence Interval Upper 0.44 1.25 0.972 0.990 0.865 0.331 0.977 0.347 0.347 0.347 0.347 0.347 0.357 0.357 0.367 0.367 0.373 0.373 0.373 0.373 0.375 0.380 0.380	Matrix Class Average SD C1 1.000 0.000 C3 0.000 C6 0.000 C6 0.000 C6 0.000 C6 0.000 C7 1.000 0.000 C7 1.000 0.000 T1 1.000 0.000 T3 T3.000 0.000 T4	Specificity 95% Confidence Interval Lower 95% Confidence Interval Lower 1000 1.000	95% Confidence Interval Upper 1.000
Average 0.911 SD 0.046	Average 0.910 SD 0.050			Average 1.000 SD 0.000		
Model Boundary P1 Score Ciss 0.266 0.276 0.476 0.57 Ciss 0.266 0.274 0.877 0.57 Ciss 0.266 0.714 0.817 0.57 Ciss 0.988 0.033 0.928 0.817 Ciss 0.983 0.033 0.928 0.735 Ciss 0.811 0.64 0.894 0.735 Ciss 0.922 0.147 0.894 0.875 Ciss 0.924 0.147 0.894 0.894 Ciss 0.924 0.147 0.888 0.77 Ciss 0.925 0.811 0.75 0.911 Ti 0.924 0.147 0.888 0.77 Ti 0.925 0.811 1.75 0.914 Ti 0.930 0.159 0.847 1.77 Ti 0.930 0.159 0.846 1.11 1.71 Ti 0.933 0.122 0.946	Confidence hierval Upper Cite Average 0.962 0.912 0.923 0.999 0.258 0.2 0.909 0.2 0.9 0.936 0.3 0.957 0.960 0.4 0.950 0.908 0.5 0.964 0.329 0.6 0.978 0.363 0.7 0.979 0.363 171 0.971 0.353 174 0.956 0.332 176 0.956 0.332 178 0.956 0.332 178 0.956 0.362 178 0.956 0.362 178 0.956 0.362 179 0.968 0.362 179 0.968 0.362 179 0.968 0.362 179 0.968 0.362 179 0.974 0.964 1711 0.974 0.964 13 0.971 0.963	SD 95% Confidence Interval Lower 0.155 0.056 0.107 0.915 0.066 0.905 0.33 0.959 0.033 0.969 0.041 0.915 0.050 0.933 0.071 0.951 0.171 0.951 0.121 0.916 0.069 0.933 0.064 0.932 0.112 0.921 0.079 0.957 0.088 0.951 0.063 0.954 0.079 0.954 0.079 0.954 0.079 0.954 0.079 0.954 0.065 0.954 0.077 0.944 0.054 0.955 0.054 0.969 0.060 0.987	95% Confidence Interval Upper 0.684 1012 0.965 0.965 0.967 0.969 0.969 0.969 0.969 0.969 0.969 0.969 0.976 0.969 0.976 0.980 0.990 0.990 0.990 0.992 0.992 0.990	Metric Average SD Class Average SD Class Average SD C2 0.911 0.220 C3 0.936 0.086 C4 0.901 0.170 C5 0.911 0.224 C6 0.815 0.291 C7 0.523 0.148 T1 0.323 0.144 T3 0.9176 0.274 T6 0.921 0.170 T7 0.9176 0.234 T6 0.924 0.194 T6 0.924 0.194 T6 0.924 0.194 T7 0.9860 0.165 T7 0.933 0.154 T9 0.9446 0.140 T13 0.9450 0.165 L2 0.917 0.165 L3 0.928 0.111 L4 0.9268 0.199 L5 0.937 0.154	Kappa 95%. Confidence Interval Lower 0.657 0.651 0.805 0.834 0.835 0.841 0.738 0.863 0.863 0.863 0.863 0.863 0.8645 0.871 0.845 0.797 0.914 0.896 0.972 0.874 0.896 0.874 0.896 0.874 0.875	95% Confidence Interval Upper 0955 1011 0.977 0.980 0.880 0.893 0.973 0.944 0.944 0.944 0.943 0.970 0.970 0.978 0.980 0.970 0.978 0.980 0.970 0.978 0.980 0.977 0.977
Average 0.908 SD 0.029	Average 0.961 SD 0.018			Average 0.904 SD 0.049		

Image: 10 with the section of the section o	Metric		Dice Score		Metric			loU		Metric			Accuracy	
No. N	Metric Average Class Average C1 0.37 C2 0.63 C3 0.57 C4 0.56 C5 0.42 C6 0.30 C7 0.82 T1 0.50 T2 0.80 T3 0.70 T4 0.59 T5 0.57 T6 0.42 T7 0.46 T8 0.54 T9 0.22 T11 0.02 T11 0.70	10 5D 95% C 0.337 0.333 0.333 0.333 0.333 0.337 0.411 0.406 0.406 0.406 0.406 0.382 0.384 0.385 0.385 0.385 0.386 0.386 0.386 0.386 0.386 0.388 0.386 0.386 0.388 0.388 0.386 0.381 0.381 0.385	Dice Score 0.230 0.476 0.420 0.421 0.412 0.412 0.412 0.412 0.412 0.412 0.412 0.409 0.742 0.554 0.554 0.338 0.371 0.459 0.536 0.540 0.734 0	95% Confidence Interval Upper 0.511 0.708 0.708 0.566 0.414 0.383 0.594 0.786 0.696 0.696 0.550 0.550 0.550 0.550 0.550 0.550 0.573 0.708 0.837 0.708 0.837 0.708 0.837 0.708 0.837 0.708 0.842 0.852	Metric Class C1 C2 C3 C4 C5 C6 C6 C7 T1 T1 T3 T6 T6 T7 T6 T7 T7 T6 T7 T7 T6 T7 T19 T19 T10 T11 T12	Average 0.282 0.554 0.471 0.456 0.359 0.263 0.263 0.423 0.423 0.423 0.423 0.423 0.423 0.423 0.424 0.631 0.536 0.499 0.383 0.404 0.404 0.474 0.550 0.550 0.550 0.550 0.550	SD 0.281 0.335 0.298 0.308 0.358 0.358 0.362 0.357 0.363 0.364 0.387 0.361 0.387 0.361 0.381 0.364 0.366 0.360 0.366 0.360 0.320	but 95% Confidence Instruct Lower 0.165 0.165 0.321 0.321 0.235 0.165 0.165 0.165 0.165 0.464 0.464 0.469 0.469 0.265 0.315 0.316 0.347 0.467 0.586 0.586 0.585	95%, Confidence Interval Upper 0.400 0.606 0.606 0.607 0.601 0.608 0.601 0.609 0.601 0.609 0.602 0.609 0.783 0.788 0.713 0.629 0.672 0.472 0.404 0.555 0.632 0.749 0.829 0.812 0.812	Metric Class C1 C2 C3 C4 C5 C6 C7 T1 T2 T3 T4 T5 T6 T7 T6 T7 T6 T7 T6 T7 T6 T7 T10 T11 T12	Average 0.999 1.000	SD 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001 0.001	Accuracy 95% Confidence Interval Lower 0.999 1.090 0.999 1.000 1.000 1.000 1.000 1.000 1.000 0.999 0.999 0.999 0.999 0.999 0.999 0.999 0.999 0.999	95%. Confidence Interval Upper 1.0000 1.0000 1.0000 1.000 1.000 1.
Image: 10 minimized by the second	L1 0.73 L2 0.724 L3 0.746 L4 0.743 L5 0.710	0.348 0.350 0.333 0.331 0.352	0.639 0.660 0.659 0.619	0.832 0.810 0.832 0.827 0.802	L1 L2 L3 L4 L5	0.697 0.661 0.684 0.678 0.642	0.349 0.343 0.341 0.346	0.515 0.576 0.596 0.591 0.552	0.7747 0.772 0.765 0.732	L1 L2 L3 L4 L5	0.999 0.999 0.999 0.999 0.999	0.002 0.002 0.002 0.002 0.002	0.999 0.999 0.999 0.999	1.000 1.000 1.000 1.000 1.000
No No<	Average 0.601 SD 0.153	3			Average SD	0.533 0.154				Average SD	1.000 0.000			
Link of D Link of D <thlink d<="" of="" th=""> <thlink d<="" of="" th=""> <th< td=""><td>Metric Class Average Class Average Average Class Average Class C2 0.666 Class C3 0.786 Class C4 0.686 Class C5 0.422 Class C6 0.333 Class C7 0.542 Class T1 0.522 T1 C4 0.777 T1 T6 0.777 T1 T7 0.562 T1 C4 0.677 T1 C5 0.423 Class T6 0.577 T9 T7 0.562 T1 C5 0.423 Class C1 0.384 Class L1 0.384 L1 L3 0.814 L3 L4 0.778 L1</td><td>e SD 95% C 0.463 0.382 0.382 0.382 0.338 0.412 0.433 0.442 0.447 0.440 0.447 0.447 0.420 0.384 0.442 0.442 0.442 0.442 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.436 0.336 0.326 0.336 0.336 0.336 0.336 0.336 0.336 0.336 0.336 0.336 0.337 0.336 0.336 0.337 0.336</td><td>Precision Offdence interval Lower 0 0.464 0.630 0.498 0.275 0.214 0.235 0.736 0.736 0.689 0.659 0.689 0.669 0.669 0.669 0.785 0.830 0.785 0.830 0.785 0.830 0.775 0.669 0.785 0.330 0.731 0.729 0.682 0.682</td><td>95% Confidence Interval Upper 0.865 0.842 0.956 0.866 0.572 0.453 0.453 0.453 0.391 0.884 0.777 0.776 0.660 0.663 0.767 0.863 0.855 0.883 0.927 0.950 0.913 0.898 0.898 0.889 0.875</td><td>Metric Class C1 C2 C2 C3 C4 C5 C6 C6 C7 T1 T3 T4 T6 T10 T12 U1 U2 U3 U4 U5 U4 U5</td><td>Average 0.290 0.618 0.490 0.299 0.275 0.299 0.275 0.433 0.789 0.718 0.564 0.510 0.510 0.510 0.542 0.510 0.562 0.687 0.773 0.695 0.728 0.695</td><td>SD 0.291 0.368 0.309 0.312 0.427 0.409 0.403 0.368 0.368 0.352 0.403 0.375 0.370 0.375 0.370 0.376 0.375 0.370 0.376 0.332 0.332 0.332 0.348 0.348 0.348 0.348 0.348 0.348</td><td>Sensitivity (rec.ii) 95% Confidence Interval Lower 0.161 0.450 0.341 0.342 0.281 0.165 0.176 0.348 0.719 0.655 0.463 0.477 0.605 0.683 0.653 0.643 0.565</td><td>95% Conflictnce Interval Upper 0.419 0.785 0.639 0.634 0.599 0.413 0.374 0.529 0.856 0.644 0.529 0.656 0.644 0.649 0.645 0.649 0.656 0.598 0.696 0.598 0.644 0.598 0.647 0.770 0.848 0.848 0.836 0.818 0.782 0.780 0.813 0.755</td><td>Metric Class C1 C2 C3 C4 C5 C6 C6 C7 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1</td><td>Average 1,000</td><td>SD 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.001 0.001 0.001</td><td>Specificity 95% Confidence Interval Lower 95% Confidence Interval Lower 95% Confidence Interval Lower 95% Confidence Interval Lower 1000</td><td>95% Confidence Interval Upper 1.0000 1.0000 1.000 1.000 1.000 1.00</td></th<></thlink></thlink>	Metric Class Average Class Average Average Class Average Class C2 0.666 Class C3 0.786 Class C4 0.686 Class C5 0.422 Class C6 0.333 Class C7 0.542 Class T1 0.522 T1 C4 0.777 T1 T6 0.777 T1 T7 0.562 T1 C4 0.677 T1 C5 0.423 Class T6 0.577 T9 T7 0.562 T1 C5 0.423 Class C1 0.384 Class L1 0.384 L1 L3 0.814 L3 L4 0.778 L1	e SD 95% C 0.463 0.382 0.382 0.382 0.338 0.412 0.433 0.442 0.447 0.440 0.447 0.447 0.420 0.384 0.442 0.442 0.442 0.442 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.443 0.436 0.336 0.326 0.336 0.336 0.336 0.336 0.336 0.336 0.336 0.336 0.336 0.337 0.336 0.336 0.337 0.336	Precision Offdence interval Lower 0 0.464 0.630 0.498 0.275 0.214 0.235 0.736 0.736 0.689 0.659 0.689 0.669 0.669 0.669 0.785 0.830 0.785 0.830 0.785 0.830 0.775 0.669 0.785 0.330 0.731 0.729 0.682 0.682	95% Confidence Interval Upper 0.865 0.842 0.956 0.866 0.572 0.453 0.453 0.453 0.391 0.884 0.777 0.776 0.660 0.663 0.767 0.863 0.855 0.883 0.927 0.950 0.913 0.898 0.898 0.889 0.875	Metric Class C1 C2 C2 C3 C4 C5 C6 C6 C7 T1 T3 T4 T6 T10 T12 U1 U2 U3 U4 U5 U4 U5	Average 0.290 0.618 0.490 0.299 0.275 0.299 0.275 0.433 0.789 0.718 0.564 0.510 0.510 0.510 0.542 0.510 0.562 0.687 0.773 0.695 0.728 0.695	SD 0.291 0.368 0.309 0.312 0.427 0.409 0.403 0.368 0.368 0.352 0.403 0.375 0.370 0.375 0.370 0.376 0.375 0.370 0.376 0.332 0.332 0.332 0.348 0.348 0.348 0.348 0.348 0.348	Sensitivity (rec.ii) 95% Confidence Interval Lower 0.161 0.450 0.341 0.342 0.281 0.165 0.176 0.348 0.719 0.655 0.463 0.477 0.605 0.683 0.653 0.643 0.565	95% Conflictnce Interval Upper 0.419 0.785 0.639 0.634 0.599 0.413 0.374 0.529 0.856 0.644 0.529 0.656 0.644 0.649 0.645 0.649 0.656 0.598 0.696 0.598 0.644 0.598 0.647 0.770 0.848 0.848 0.836 0.818 0.782 0.780 0.813 0.755	Metric Class C1 C2 C3 C4 C5 C6 C6 C7 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1 T1	Average 1,000	SD 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.001 0.001 0.001 0.001	Specificity 95% Confidence Interval Lower 95% Confidence Interval Lower 95% Confidence Interval Lower 95% Confidence Interval Lower 1000	95% Confidence Interval Upper 1.0000 1.0000 1.000 1.000 1.000 1.00
Strand Ds 95% Onlidence Interval Uoyar 95% Conlidence Interval Uoyar	Average 0.697 SD 0.150 Metric	,)	Boundary F1 Score		Average SD Metric	0.565 0.157		AUC		Average SD Metric	1.000 0.000		Карра	
	month Figure Cita 0.073 Cita 0.073 Cita 0.073 Cita 0.073 Cita 0.073 Cita 0.073 Cita 0.058 Cita 0.586 Cita 0.566 Cita 0.666 Cita 0.666 Cita 0.777 Tita 0.777 Tita 0.678 Tita 0.678 Tita 0.678 Tita 0.678 Tita 0.678 Tita 0.778 Lita 0.071 Lita 0.071 Lita 0.071 Average 0.588	pp 5.0 9.5% C 0.331 0.3351 0.3367 0.3371 0.3361 0.337 0.3397 0.3397 0.3397 0.3397 0.3397 0.3397 0.3397 0.3387 0.3387 0.3387 0.2387 0.3384 0.3347 0.3347 0.3347 0.3347 0.3345 0.3345 0.3345 0.3345 0.3345 0.3345 0.2385 0.3346 0.2355 0.336 0.3361 0.3391 0.3361 0.3361 0.3255 0.336 0.3361 0.3255	Boomeary F = 52008 Doctory F = 52008 0.2310721 0.2452 0.452 0.452 0.452 0.452 0.452 0.452 0.452 0.452 0.452 0.452 0.279 0.188 0.721 0.629 0.511 0.491 0.382 0.583 0.618 0.708 0.658 0.658 0.593	95% Confidence Interval Upper 0 578 0 778 0 726 0 701 0 555 0 404 0 375 0 775 0 775 0 685 0 555 0 565 0 565 0 561 0 684 0 773 0 684 0 773 0 442 0 835 0 788 0 788 0 776	Class Class C1 C2 C3 C4 C5 C6 C7 T1 T2 T3 T4 T5 T6 T7 T8 T9 T10 L1 L2 L3 L4 L5	Average 0.645 0.809 0.745 0.745 0.745 0.718 0.650 0.637 0.719 0.657 0.902 0.859 0.762 0.762 0.7755 0.7755 0.7755 0.7755 0.781 0.844 0.844 0.886 0.868 0.868 0.864 0.864 0.864 0.864 0.864 0.863 0.839 0.785	SD 0.146 0.184 0.154 0.156 0.213 0.201 0.134 0.201 0.134 0.176 0.201 0.134 0.176 0.201 0.186 0.186 0.201 0.188 0.201 0.188 0.201 0.188 0.180 0.164 0.177 0.177	35% Confidence Interval Lower 0.725 0.771 0.671 0.671 0.681 0.598 0.899 0.899 0.816 0.732 0.716 0.659 0.878 0.732 0.716 0.659 0.578 1.738 0.746 0.599 0.578 0.578 0.578 0.578 0.589 0.578 0.589 0.578 0.589 0.578 0.589 0.578 0.589 0.578 0.589 0.578 0.580 0.598 0.578 0.580 0.598 0.578 0.580 0.578 0.578 0.580 0.585 0.580 0.585 0.580 0.585 0.580 0.585 0.580 0.585 0.580 0.580 0.578 0.578 0.578 0.578 0.578 0.578 0.578 0.578 0.578 0.580 0.580 0.578 0.578 0.580 0.580 0.578 0.578 0.580 0.580 0.578 0.580 0.580 0.580 0.578 0.580	95% Contidence Interval Upper 0.703 0.819 0.819 0.817 0.794 0.875 0.876 0.878 0.878 0.874 0.875 0.874 0.875 0.874 0.875 0.875 0.831 0.772 0.723 0.823 0.825 0.899 0.891 0.905 0.818	- memo- Clease C 2 C 2 C 2 C 2 C 2 C 2 C 2 C 2 C 2 C 2	Average 0.371 0.637 0.577 0.560 0.424 0.501 0.285 0.501 0.791 0.705 0.598 0.574 0.454 0.547 0.723 0.795 0.544 0.6547	SD 0.336 0.372 0.333 0.337 0.401 0.406 0.406 0.406 0.382 0.405 0.388 0.392 0.411 0.388 0.392 0.413 0.385 0.380 0.362 0.350 0.351 0.362	35% Confidence Interval Lower 0.457 0.467 0.417 0.402 0.276 0.191 0.188 0.726 0.726 0.725 0.622 0.488 0.476 0.356 0.525 0.457 0.569 0.538 0.720 0.707 0.669 0.627 0.627 0.669 0.567 0.667 0.663	95%. Confidence Interval Upper 0.500 0.806 0.738 0.738 0.718 0.572 0.416 0.365 0.855 0.865 0.697 0.673 0.697 0.697 0.673 0.552 0.500 0.700 0.808 0.865 0.833 0.812 0.812 0.812 0.812 0.783 0.783



А

Class	Dice Score	loU	Accuracy	Precision	Sensitivity	Specificity	Boundary F1 Score	AUC	Kappa	Hausdorff Distance
C1	0.70459	0.62022	0.37299	0.60639	0.77369	0.23583	0.63048	0.88678	0.99876	0.10134
C2	0.76732	0.62264	0.04599	0.74846	0.84630	0.07175	0.67279	0.84609	0.76712	0.05541
C3	0.38920	0.37530	0.81518	0.19540	0.52399	0.08493	0.13669	0.52379	0.38927	0.25019
C4	0.91825	0.69562	0.70715	0.26210	0.82387	0.11134	0.45383	0.82425	0.91819	0.64195
C5	0.09377	0.12542	0.98932	0.18394	0.29629	0.24132	0.04939	0.07429	0.27292	0.38644
C6	0.00029	0.00013	0.55940	0.62289	0.03974	0.28055	0.00009	0.00002	0.11968	0.01228
C7	0.00100	0.00002	0.01311	0.32447	0.00000	0.70594	0.00004	0.00000	0.00099	0.00000
T1	0.07743	0.00520	0.29485	0.00300	0.90277	0.20333	0.00176	0.90530	0.07739	0.00000
T2	0.50221	0.28602	0.77267	0.63746	0.38511	0.82585	0.32729	0.38542	0.50240	0.01414
T3	0.09918	0.01880	0.43897	0.38910	0.05015	0.69282	0.08713	0.03676	0.12344	0.00002
T4	0.12399	0.02828	0.20584	0.03377	0.34078	0.21254	0.05540	0.34000	0.12387	0.00000
T5	0.03212	0.00840	0.33326	0.03821	0.03608	0.29437	0.01227	0.03611	0.03218	0.00273
T6	0.01649	0.01140	0.37949	0.02841	0.00710	0.48161	0.00721	0.00712	0.01653	0.00330
T7	0.09539	0.03637	0.30951	0.16098	0.25788	0.11183	0.03190	0.19187	0.14948	0.00028
T8	0.06248	0.01513	0.12454	0.01994	0.12454	0.08958	0.00857	0.23178	0.03474	0.00016
T9	0.02940	0.00959	0.13170	0.03971	0.02133	0.15602	0.00441	0.02125	0.02933	0.00145
T10	0.16943	0.07827	0.17091	0.27567	0.11463	0.24472	0.06896	0.11449	0.16902	0.00091
T11	0.30885	0.12574	0.33137	0.14648	0.50813	0.30453	0.07663	0.50723	0.30843	0.00396
T12	0.49914	0.23033	0.52564	0.69929	0.46163	0.98056	0.09686	0.38318	0.56405	0.00552
L1	0.57887	0.36898	0.78695	0.82605	0.83668	0.95313	0.18787	0.49408	0.86899	0.01320
L2	0.18699	0.10461	0.68086	0.98805	0.00890	0.76718	0.04443	0.00897	0.18701	0.00902
L3	0.24919	0.10319	0.47364	0.32719	0.27620	0.44883	0.05272	0.27548	0.24888	0.00051
L4	0.25326	0.09455	0.45636	0.43427	0.20754	0.33812	0.02724	0.20696	0.25294	0.00015
L5	0.06874	0.02030	0.44177	0.25953	0.76385	0.17086	0.00164	0.24012	0.39430	0.00006
Sacrum	0.06283	0.03119	0.15076	0.69608	0.00000	0.64117	0.00198	0.00000	0.06199	0.29829
	Class C1 C2 C3 C4 C5 C6 C7 T1 T2 T3 T4 T5 T6 T7 T6 T7 T6 T7 T6 T7 T1 T1 T12 L1 L2 L2 L3 L4 Sacrum	Class Dice Score C1 0.7459 C2 0.76732 C3 0.39920 C4 0.91825 C5 0.09377 C6 0.09027 C7 0.00100 T1 0.0743 T2 0.5021 T3 0.00248 T6 0.01649 T7 0.06548 T9 0.02344 T10 0.16943 T112 0.49914 L1 0.57887 L2 0.16899 L3 0.24914 L4 0.24532 L5 0.06238	Class Dice Score but C1 0.7049 0.62022 C2 0.76732 0.62262 C3 0.3820 0.37530 C4 0.91825 0.68562 C5 0.0937 0.12542 C6 0.00023 0.00012 C7 0.00100 0.00002 T1 0.07743 0.01820 T2 0.50221 0.28602 T3 0.09918 0.01830 T4 0.12399 0.02828 T6 0.01649 0.01140 T7 0.09539 0.03637 T10 0.16943 0.07827 T11 0.30885 0.12574 T12 0.4944 0.23033 L1 0.57887 0.36899 L2 0.18699 0.10461 L3 0.23256 0.09455 L5 0.06874 0.02330 L4 0.22363 0.03119	Class Dice Score DU Accuracy C1 0.7149 0.8224 0.04599 C2 0.76732 0.62264 0.04599 C3 0.3820 0.37530 0.81530 C4 0.91825 0.69562 0.70715 C5 0.0977 0.12542 0.98692 C7 0.00100 0.00021 0.93140 C7 0.00100 0.00021 0.93145 C8 0.00929 0.01810 0.43897 T2 0.5524 0.27465 0.77267 T3 0.09518 0.03322 0.20544 T6 0.01494 0.01740 0.37949 T6 0.01494 0.01830 0.43897 T4 0.12399 0.02324 0.00541 0.37349 T6 0.01444 0.03744 0.37349 0.3337 T7 0.05939 0.03547 0.130951 1.1370 T110 0.16944 0.02540 0.52564 1.1574 1	Class Dice Score IOU Accuracy Precision C1 0.7045 0.6224 0.3729 0.6659 0.6224 C2 0.76732 0.62264 0.04599 0.74446 0.5224 0.94599 0.74446 C3 0.38202 0.37530 0.81516 0.19540 0.62254 0.94599 0.74446 C4 0.91825 0.69562 0.70715 0.26210 0.55940 0.62289 0.77 0.70716 0.26289 0.7287 0.12845 0.00520 0.91814 0.32447 0.33811 0.32447 0.33810 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.33821 0.30851 0.13096 0.13170 0.03971 1110 0.16440 0.11400 0.33137 0.14444 0.41939 0.22441 0.43987 0.33137 0.14444 0.41934 0.32271 0.144	Class Dice Score DU Accuracy Precision Sensitivity C1 0.7749 0.6224 0.04599 0.7739 0.60539 0.77396 C2 0.76732 0.6224 0.04599 0.74446 0.64239 C3 0.3820 0.37539 0.17540 0.62239 0.7639 C3 0.3820 0.37530 0.81518 0.19540 0.62399 C4 0.91825 0.69562 0.70715 0.26210 0.82387 C5 0.00977 0.12542 0.98032 0.18394 0.03247 C7 0.00100 0.00020 0.01311 0.32447 0.03000 0.90277 T2 0.05299 0.02224 0.06302 0.72457 0.33810 0.05015 T3 0.09918 0.04880 0.43897 0.33810 0.05015 T4 0.12399 0.02224 0.06840 0.03371 0.34078 T4 0.01390 0.25783 0.03971 0.12444 0.03714 <t< th=""><th>Class Dico Score IoU Accuracy Precision Sensitivity Specificity C1 0.70459 0.62224 0.04999 0.60639 0.77369 0.22583 C2 0.76732 0.62264 0.04959 0.74846 0.84630 0.07176 C3 0.38202 0.37309 0.81518 0.18540 0.52399 0.04431 C4 0.91825 0.696562 0.70715 0.22210 0.823387 0.11134 C5 0.09077 0.12542 0.98932 0.18394 0.22629 0.23747 C7 0.00100 0.00022 0.01131 0.32447 0.03000 0.92277 0.22635 T2 0.05210 0.24845 0.03300 0.92777 0.22343 T2 0.05210 0.27845 0.03300 0.9277 0.22343 T3 0.00918 0.01880 0.43897 0.33810 0.03615 0.68285 T4 0.12399 0.02281 0.03847 0.03801 0.04716 <td< th=""><th>Class Dico Score IoU Accuracy Procision Sensitivity Specificity Boundary F1 Score C1 0.70459 0.62224 0.04599 0.60639 0.77369 0.63048 C2 0.76732 0.62264 0.04599 0.77369 0.52393 0.03745 C3 0.38202 0.37303 0.81516 0.52239 0.07175 0.67279 C3 0.38202 0.37330 0.81516 0.52239 0.04493 0.13669 C4 0.91625 0.86962 0.70715 0.22621 0.28237 0.11134 0.45383 C5 0.00977 0.12542 0.98932 0.13944 0.28055 0.00099 C7 0.0010 0.00022 0.29145 0.00300 0.9277 0.2333 0.00176 T2 0.05221 0.28062 0.7287 0.03811 0.32477 0.02171 T3 0.01800 0.43997 0.38910 0.05150 0.65282 0.06511 T4 0.12994</th><th>Class Dico Score IoU Accuracy Procision Sensitivity Specificity Boundary F1 Score AUC C1 0.7049 0.82264 0.04599 0.77369 0.02383 0.63048 0.838676 C2 0.76732 0.82264 0.04599 0.77369 0.02383 0.63048 0.838676 C3 0.38202 0.37530 0.81518 0.05239 0.04493 0.136696 0.52379 C4 0.91825 0.869562 0.70715 0.26210 0.82387 0.11134 0.45383 0.82429 C5 0.09077 0.12542 0.89894 0.029629 0.24132 0.049439 0.07429 C6 0.00021 0.05540 0.02360 0.02977 0.23851 0.02604 0.00004 0.00001 T2 0.05021 0.28467 0.03811 0.02585 0.03772 0.3842 T3 0.09186 0.43897 0.33810 0.05151 0.62285 0.03574 0.24124 0.03540 0.33400 <</th><th>Class Dice Score But Accuracy Precision Sensitivity Specificity Boundary F1 Score AUC Kappa C1 0.7749 0.23283 0.03048 0.83048 0.83048 0.89878 0.99876 C2 0.7739 0.22583 0.03048 0.84699 0.76712 0.67279 0.23638 0.03048 0.84699 0.76712 0.87719 0.23263 0.03048 0.84699 0.76712 0.87719 0.23247 0.84699 0.76712 0.82387 0.011366 0.52379 0.03643 0.82487 0.98976 0.82387 0.04493 0.82445 0.91819 C5 0.00977 0.12544 0.98926 0.11134 0.45383 0.82425 0.91819 C7 0.00001 0.00020 0.03141 0.23265 0.00004 0.00000 0.00000 0.00000 0.00000 0.00000 0.00004 0.990530 0.03746 0.12344 1 0.01390 0.02247 0.03616 0.62248 0.0377 0.23445</th></td<></th></t<>	Class Dico Score IoU Accuracy Precision Sensitivity Specificity C1 0.70459 0.62224 0.04999 0.60639 0.77369 0.22583 C2 0.76732 0.62264 0.04959 0.74846 0.84630 0.07176 C3 0.38202 0.37309 0.81518 0.18540 0.52399 0.04431 C4 0.91825 0.696562 0.70715 0.22210 0.823387 0.11134 C5 0.09077 0.12542 0.98932 0.18394 0.22629 0.23747 C7 0.00100 0.00022 0.01131 0.32447 0.03000 0.92277 0.22635 T2 0.05210 0.24845 0.03300 0.92777 0.22343 T2 0.05210 0.27845 0.03300 0.9277 0.22343 T3 0.00918 0.01880 0.43897 0.33810 0.03615 0.68285 T4 0.12399 0.02281 0.03847 0.03801 0.04716 <td< th=""><th>Class Dico Score IoU Accuracy Procision Sensitivity Specificity Boundary F1 Score C1 0.70459 0.62224 0.04599 0.60639 0.77369 0.63048 C2 0.76732 0.62264 0.04599 0.77369 0.52393 0.03745 C3 0.38202 0.37303 0.81516 0.52239 0.07175 0.67279 C3 0.38202 0.37330 0.81516 0.52239 0.04493 0.13669 C4 0.91625 0.86962 0.70715 0.22621 0.28237 0.11134 0.45383 C5 0.00977 0.12542 0.98932 0.13944 0.28055 0.00099 C7 0.0010 0.00022 0.29145 0.00300 0.9277 0.2333 0.00176 T2 0.05221 0.28062 0.7287 0.03811 0.32477 0.02171 T3 0.01800 0.43997 0.38910 0.05150 0.65282 0.06511 T4 0.12994</th><th>Class Dico Score IoU Accuracy Procision Sensitivity Specificity Boundary F1 Score AUC C1 0.7049 0.82264 0.04599 0.77369 0.02383 0.63048 0.838676 C2 0.76732 0.82264 0.04599 0.77369 0.02383 0.63048 0.838676 C3 0.38202 0.37530 0.81518 0.05239 0.04493 0.136696 0.52379 C4 0.91825 0.869562 0.70715 0.26210 0.82387 0.11134 0.45383 0.82429 C5 0.09077 0.12542 0.89894 0.029629 0.24132 0.049439 0.07429 C6 0.00021 0.05540 0.02360 0.02977 0.23851 0.02604 0.00004 0.00001 T2 0.05021 0.28467 0.03811 0.02585 0.03772 0.3842 T3 0.09186 0.43897 0.33810 0.05151 0.62285 0.03574 0.24124 0.03540 0.33400 <</th><th>Class Dice Score But Accuracy Precision Sensitivity Specificity Boundary F1 Score AUC Kappa C1 0.7749 0.23283 0.03048 0.83048 0.83048 0.89878 0.99876 C2 0.7739 0.22583 0.03048 0.84699 0.76712 0.67279 0.23638 0.03048 0.84699 0.76712 0.87719 0.23263 0.03048 0.84699 0.76712 0.87719 0.23247 0.84699 0.76712 0.82387 0.011366 0.52379 0.03643 0.82487 0.98976 0.82387 0.04493 0.82445 0.91819 C5 0.00977 0.12544 0.98926 0.11134 0.45383 0.82425 0.91819 C7 0.00001 0.00020 0.03141 0.23265 0.00004 0.00000 0.00000 0.00000 0.00000 0.00000 0.00004 0.990530 0.03746 0.12344 1 0.01390 0.02247 0.03616 0.62248 0.0377 0.23445</th></td<>	Class Dico Score IoU Accuracy Procision Sensitivity Specificity Boundary F1 Score C1 0.70459 0.62224 0.04599 0.60639 0.77369 0.63048 C2 0.76732 0.62264 0.04599 0.77369 0.52393 0.03745 C3 0.38202 0.37303 0.81516 0.52239 0.07175 0.67279 C3 0.38202 0.37330 0.81516 0.52239 0.04493 0.13669 C4 0.91625 0.86962 0.70715 0.22621 0.28237 0.11134 0.45383 C5 0.00977 0.12542 0.98932 0.13944 0.28055 0.00099 C7 0.0010 0.00022 0.29145 0.00300 0.9277 0.2333 0.00176 T2 0.05221 0.28062 0.7287 0.03811 0.32477 0.02171 T3 0.01800 0.43997 0.38910 0.05150 0.65282 0.06511 T4 0.12994	Class Dico Score IoU Accuracy Procision Sensitivity Specificity Boundary F1 Score AUC C1 0.7049 0.82264 0.04599 0.77369 0.02383 0.63048 0.838676 C2 0.76732 0.82264 0.04599 0.77369 0.02383 0.63048 0.838676 C3 0.38202 0.37530 0.81518 0.05239 0.04493 0.136696 0.52379 C4 0.91825 0.869562 0.70715 0.26210 0.82387 0.11134 0.45383 0.82429 C5 0.09077 0.12542 0.89894 0.029629 0.24132 0.049439 0.07429 C6 0.00021 0.05540 0.02360 0.02977 0.23851 0.02604 0.00004 0.00001 T2 0.05021 0.28467 0.03811 0.02585 0.03772 0.3842 T3 0.09186 0.43897 0.33810 0.05151 0.62285 0.03574 0.24124 0.03540 0.33400 <	Class Dice Score But Accuracy Precision Sensitivity Specificity Boundary F1 Score AUC Kappa C1 0.7749 0.23283 0.03048 0.83048 0.83048 0.89878 0.99876 C2 0.7739 0.22583 0.03048 0.84699 0.76712 0.67279 0.23638 0.03048 0.84699 0.76712 0.87719 0.23263 0.03048 0.84699 0.76712 0.87719 0.23247 0.84699 0.76712 0.82387 0.011366 0.52379 0.03643 0.82487 0.98976 0.82387 0.04493 0.82445 0.91819 C5 0.00977 0.12544 0.98926 0.11134 0.45383 0.82425 0.91819 C7 0.00001 0.00020 0.03141 0.23265 0.00004 0.00000 0.00000 0.00000 0.00000 0.00000 0.00004 0.990530 0.03746 0.12344 1 0.01390 0.02247 0.03616 0.62248 0.0377 0.23445

В

Metric				F	-value spineC	T-ResNetSeg	vs MONAI-Verse			
Class	Dice Score	loU	Accuracy	Precision	Sensitivity	Specificity	Boundary F1 Score	AUC	Карра	Hausdorff Distance
C1	0.00009	0.00000	0.00999	0.26757	0.00000	0.04576	0.00014	0.00000	0.00009	0.00007
C2	0.01107	0.00219	0.00194	0.03920	0.00309	0.03918	0.00393	0.00304	0.01103	0.00000
C3	0.00162	0.00011	0.00301	0.21819	0.00011	0.77313	0.00054	0.00008	0.00162	0.00011
C4	0.00031	0.00005	0.00841	0.03238	0.00001	0.29519	0.00022	0.00001	0.00031	0.00051
C5	0.03262	0.01585	0.04449	0.02216	0.03347	0.01699	0.02030	0.03712	0.03256	0.01675
C6	0.00000	0.00000	0.01378	0.00000	0.00000	0.28034	0.00000	0.00000	0.00000	0.00000
C7	0.00000	0.00000	0.00285	0.00000	0.00000	0.64321	0.00000	0.00000	0.00000	0.00000
T1	0.00000	0.00000	0.04222	0.75315	0.00000	0.04798	0.00000	0.00000	0.00000	0.00000
T2	0.03197	0.00472	0.49916	0.04358	0.03495	0.60738	0.00550	0.06750	0.01881	0.00001
Т3	0.04869	0.03442	0.10117	0.01611	0.26632	0.06077	0.01845	0.15439	0.08519	0.01131
T4	0.00098	0.00093	0.17255	0.10940	0.00001	0.41258	0.00100	0.00000	0.00098	0.08011
T5	0.00004	0.00001	0.33749	0.00890	0.00000	0.70684	0.00001	0.00000	0.00004	0.00005
T6	0.00000	0.00000	0.16731	0.00313	0.00000	0.68572	0.00000	0.00000	0.00000	0.00000
T7	0.00000	0.00000	0.16460	0.00021	0.00000	0.97582	0.00000	0.00000	0.00000	0.00000
T8	0.00000	0.00000	0.11214	0.00102	0.00000	0.96214	0.00000	0.00000	0.00000	0.00000
T9	0.00000	0.00000	0.25478	0.02706	0.00000	0.87060	0.00000	0.00000	0.00000	0.00000
T10	0.00080	0.00026	0.35119	0.00841	0.00013	0.78478	0.00011	0.00014	0.00081	0.00009
T11	0.00752	0.00576	0.36233	0.14362	0.00040	0.65906	0.00195	0.00084	0.00414	0.01684
T12	0.00295	0.00248	0.14787	0.62042	0.00101	0.94077	0.00269	0.00051	0.00541	0.06433
L1	0.00221	0.00150	0.22585	0.07469	0.00034	0.44340	0.00194	0.00035	0.00220	0.11389
L2	0.00509	0.00474	0.46612	0.02918	0.00089	0.52267	0.00107	0.00095	0.00511	0.07658
L3	0.00171	0.00157	0.57122	0.04284	0.00006	0.74193	0.00083	0.00033	0.00051	0.04490
L4	0.00731	0.00483	0.35114	0.01723	0.00127	0.41694	0.00448	0.00130	0.00731	0.03846
15	0.00012	0 00004	0 11469	0.01287	0 00000	0 27333	0.00011	0 00000	0 00006	0.00686