ASEMI: Automated SEgmentation of Microtomography Imaging

Marc Tanti,[1] Camille Berruyer,[2,3] Paul Tafforeau,[2] Adrian Muscat,[1] Reuben Farrugia,[1] Kenneth Scerri,[1] Gianluca Valentino,[1] V. Armando Solé,[2] and Johann A. Briffa[1*]

[1]University of Malta, Msida MSD 2080, Malta; [2]ESRF, 71 Avenue des Martyrs, CS-40220, 38043 Grenoble cedex 09, France; [3]ArchéOrient (UMR 5133, Maison de l'Orient et de la Méditerranée Jean Pouilloux), Lyon, France
[*]Corresponding author: johann.briffa@um.edu.mt

ABSTRACT

Propagation Phase Contrast Synchrotron Microtomography is becoming a golden standard for a non-invasive and non-destructive access to internal structures of archaeological remains. However, the manual segmentation of complex biological samples can require weeks of work, even for small volumes. Machine learning techniques can automate this process. We describe the ASEMI segmenter, a tool for automatic segmentation of arbitrarily large volumes, and compare with commercially available software. Results are promising, and demonstrate the feasibility of these techniques. Further work would scale this tool to facilitate the segmentation of collections of scans, and to integrate within existing workflows.

Keywords: tomography; segmentation; archaeology.

1. INTRODUCTION

Archaeologists have always been interested in applying new technologies to their research to better understand our past. Propagation Phase Contrast Synchrotron Microtomography is becoming a golden standard for a non-invasive and non-destructive access to internal structures of archaeological remains. This technique has been recently applied to archaeozoological studies of mummified animal remains from the Ptolemaic and Roman periods of ancient Egypt (around 3rd century BC to 4th century AD). Researchers performed virtual autopsies and virtual unwrapping, uncovering information about animal life and death in past civilisations, as well as revealing the processes used to make these mummies [1], [2]. However, this can be a long process. After microtomographic data processing and reconstruction, the virtual specimen has to be segmented to separate the different parts or different materials of the sample. For biological samples, segmentation is usually done semi-manually, and can require weeks of human effort, even for small volumes. Effective automatic segmentation based on Artificial Intelligence (AI) could drastically reduce this effort, while computation time ranges from hours to days, depending on the size and complexity of the volumetric image. This is particularly relevant as the number and sizes of volumes increases. For example, the European Synchrotron Radiation Facility (ESRF) is currently generating huge amounts of data of human organs (up to 2 TiB for a single scan); only AI approaches can handle segmentation at this scale.

We developed a fully automatic segmenter based on classical machine learning, requiring the user to manually segment only a small sample of the volume. Our approach has a number of advantages over commercial solutions. Specifically, our segmenter is not limited by the computer's main memory, allowing it to segment arbitrarily large volumes. Its parameters are automatically tuned for the given volume, minimising user input. It also works directly with three dimensional features, rather than reducing the problem to an independent segmentation of two-dimensional slices, without increasing computational complexity from quadratic to cubic. This allows our system to scale well, particularly for larger volumes. Finally, our system can use interpretable machine learning models, making it possible to inspect the reasons behind segmentation errors.

Our output segmentations have an overall accuracy of 94–98% when compared with manually segmented slices. This approaches the results of off-the-shelf commercial software using deep learning (97–99%) at much lower complexity. A qualitative analysis of the outputs shows that our results are close in term of usability to those from deep learning. Some postprocessing is necessary to clean up segmentation boundaries, for both our system and the deep learning approach. Preliminary work using a Markov Random Field gave promising results, and we plan to implement this in a scalable way.

2. STATE OF THE ART

The state of the art in segmentation is the use of deep learning models such as U-Net [3], designed to work on 2D images. Unfortunately, a direct extension of U-Net to three dimensions dramatically increases memory and computational complexity. As a result, the technique of 3D U-Net does not scale well, and its use has been limited to relatively small volumes, such as those from medical scanners [4], [5]. An alternative approach is to independently segment every slice in a given volume. Although this makes the process more practical, it also results in ridged artifacts where the labelling from one slice to the next is inconsistent.

Since this project started, deep learning segmentation algorithms are available in Dragonfly [6], a commercial software for volumetric image editing and analysis. To avoid the scalability problem, Dragonfly independently segments every slice in the volume; the latest version, recently released, also introduces the ability to use adjacent slices as context. ASEMI then appears as a timely work in a very rapidly developing field.

Dragonfly already gives usable results in a range of situations when the training slices are selected to be representative of a single complex specimen, but its algorithms come with a number of caveats. In particular, Dragonfly is limited to volumes that fit in main memory, and recommends having $4\times$ as much RAM as the largest volume size. It also appears that for specimens as complex as our mummies, a trained model can hardly be applied to a different specimen, even if it was scanned in the same conditions. To obtain a more general model would require much more human time for the manual segmentation of relevant slices from several samples, as well as much longer training time for the machine.

3. BREAKTHROUGH CHARACTER OF THE PROJECT

Our system resolves the fundamental scalability problem using traditional machine learning techniques, while still approaching the results obtainable with deep learning when this is applied independently on every slice of the volumetric image. Normally, the computation of a voxel's features scales cubically with the size of the neighbourhood considered. However, we can take advantage of the fact that the neighbourhoods of adjacent voxels have significant overlap, and can express the computation of a voxel's features as a change from that of its neighbours. This reduces the complexity of the feature computation to one that scales quadratically with neighbourhood size, for a whole class of features. This allows us to work with three dimensional information without sacrificing performance. Since the segmentation uses adjacent slices as context, the segmentation is continuous across slices and ridged artifacts are reduced.

Furthermore, our implementation is designed specifically for very large volumes, dividing a given volume into blocks and working on these independently. This means it is not limited by the main memory or GPU memory available, and also facilitates parallel processing on multiple compute nodes.

Where necessary, we can also use white box classifiers, allowing us to interpret what the classifier has learned and to investigate the source of segmentation errors.

4. PROJECT RESULTS

A comparison of the per-label and overall accuracy between the U-Net implementation in Dragonfly (operating independently on each 2D slice) and the ASEMI segmenter with random forest (RF) and neural network (NN) classifiers is shown in Table 1. These results are computed against a ground truth consisting of a selection of manually segmented slices that were not used in the tuning or training processes. It can be seen that in general U-Net outperforms the ASEMI segmenter, and that the NN classifier often outperforms the RF classifier.

A visual inspection of the ASEMI segmenter output shows that this is still useful and that a lot of the errors are due to indecision near object boundaries. This exhibits as boundary fuzziness and small islands of mislabelled regions, as shown in Fig. 1. Preliminary work using a Markov Random Field to clean up the decision boundaries gave promising results. An alternative is the use of morphological operations. In this example there is also evidence of mislabelling in the manual segmentation which is corrected by ASEMI but not by Dragonfly. Observe how, in the manual segmentation, a region of dense textile was not marked and the bone cross-section was only partially marked; U-Net faithfully replicates these errors, while ASEMI correctly extends the dense textile region and fully labels the bone cross-section.

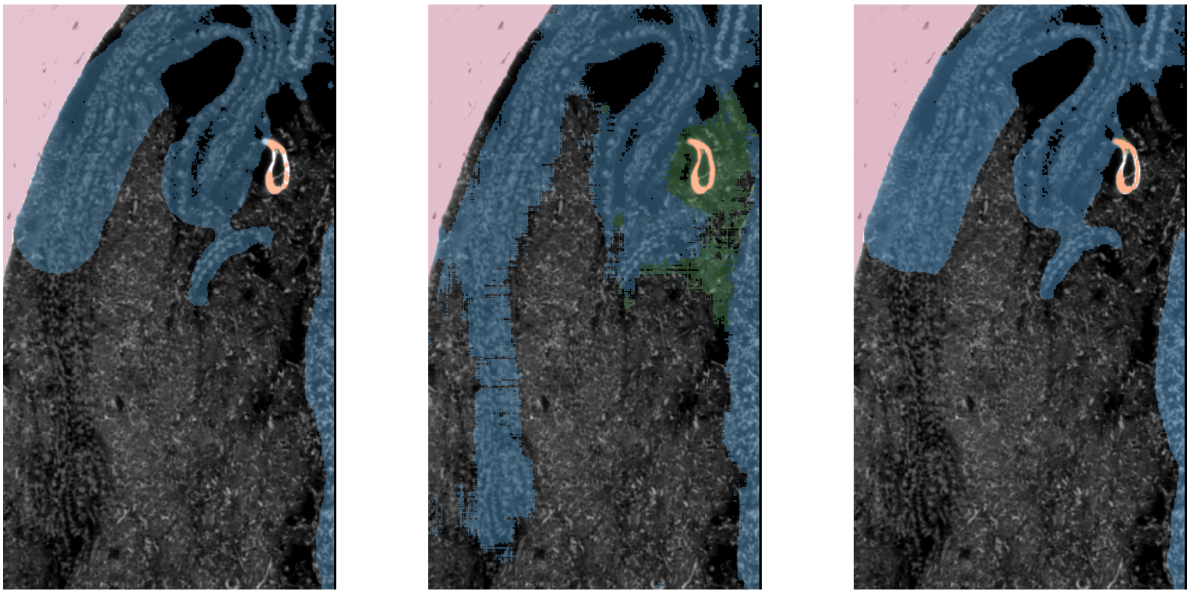
A 3D rendering of a segmented ibis mummy is shown in Fig. 2, comparing the U-Net implementation in Dragonfly with the initial ASEMI implementation and the tuned ASEMI segmenter with RF and NN classifiers. One can observe how the ASEMI output is comparable to the Dragonfly output, particularly when using the NN classifier. While Dragonfly seems to perform better in some regions (e.g. the tail section), ASEMI is somewhat better in others (e.g. the wing bones).

5. FUTURE PROJECT VISION

5.1. Technology Scaling

While we have demonstrated the feasibility of using machine learning for the 3D segmentation of large volumes, a number of further advances are necessary for an operational environment:

1) Apply our complexity reduction technique to estab-



(a) Manual (b) ASEMI (c) Dragonfly

Fig. 1. Segmentation detail from MG.2038 (ibis in a jar) mummy, showing fuzziness at the label boundaries for the ASEMI segmenter with random forest classifier, and correction of manual mislabelling, as compared to U-Net in Dragonfly. Legend: Blue – dense textile, Orange – bones, Green – soft parts, Pink – terracotta.

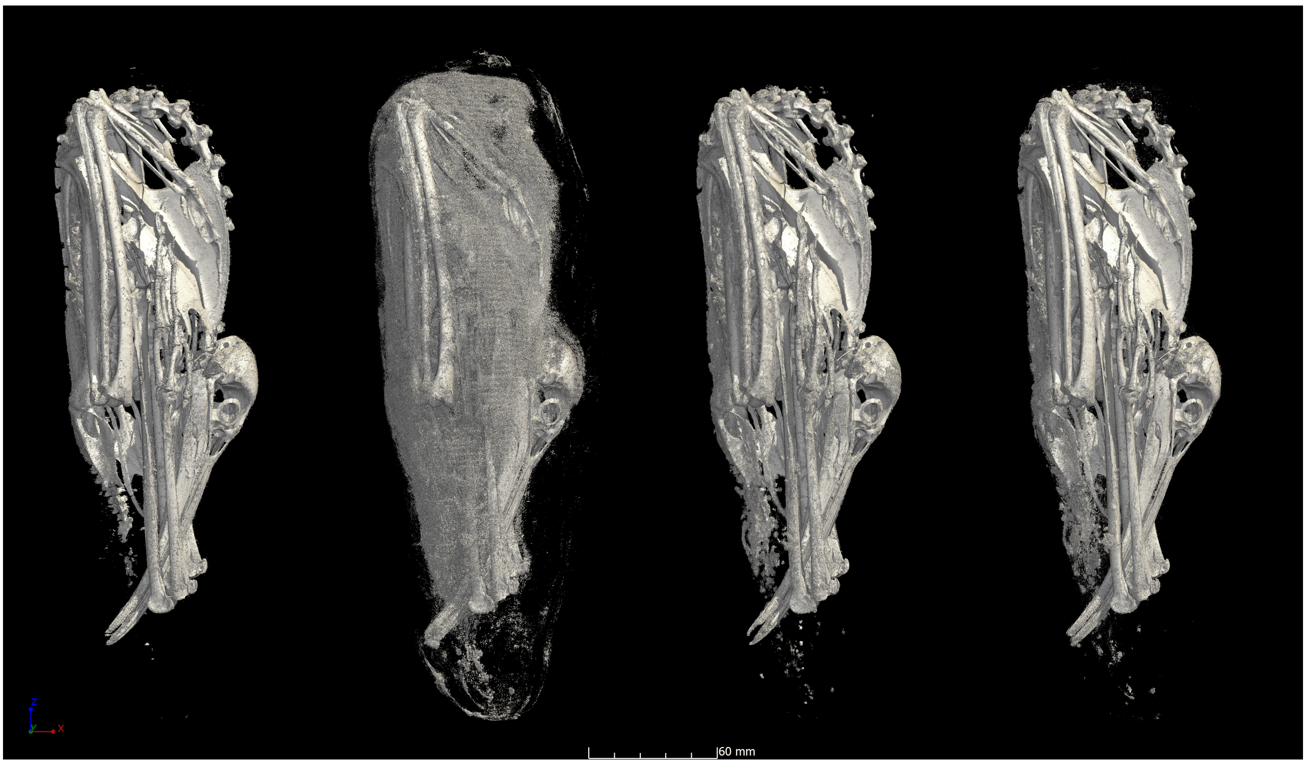


Fig. 2. 3D rendering of the segmented MHNGr.ET.1456 (ibis) mummy, showing the parts labelled as bones. Left to right: i) U-Net implementation in Dragonfly, ii) initial ASEMI implementation, iii) tuned ASEMI segmenter with RF classifier, iv) tuned ASEMI segmenter with NN classifier.

Tab. 1. Intersection-over-union results for individual labels and overall accuracy for four specimens. Legend: U-Net – Dragonfly implementation; RF – ASEMI segmenter with random forest classifier; NN – ASEMI segmenter with neural network classifier.

	MHNGr.ET.1023 (dog)			*MHNGr.ET.1017 (raptor)*			*MHNGr.ET.1456 (ibis)*			*MG.2038 (ibis in a jar)*		
Label	*U-Net*	*RF*	*NN*	*U-Net*	*RF*	*NN*	*U-Net*	*RF*	*NN*	*U-Net*	*RF*	*NN*
Bones	92.2%	76.2%	85.5%				88.7%	86.2%	81.9%	93.4%	96.0%	90.5%
Teeth	43.2%	14.5%	46.6%									
Feathers							69.3%	61.6%	64.9%	67.7%	21.5%	28.7%
Soft parts	87.4%	70.3%	72.9%	77.4%	66.1%	66.4%				92.8%	58.8%	81.6%
Soft powder							74.9%	58.3%	71.3%			
Stomach				95.9%	84.8%	80.4%						
Snails							87.2%	60.0%	54.2%			
Textiles	93.1%	83.6%	85.1%	96.5%	90.8%	91.4%	86.2%	82.7%	82.3%			
Balm textile				85.3%	81.1%	80.5%						
Dense textile				81.0%	66.5%	67.9%	67.0%	57.5%	62.8%	96.5%	78.0%	78.7%
Natron							60.4%	36.9%	22.0%			
Ceramics				78.8%	67.7%	65.3%						
Terracotta										99.8%	98.9%	99.3%
Cement										94.4%	74.8%	78.2%
Wood				84.4%	76.8%	94.8%						
Insects							25.4%	4.6%	4.5%			
Powder										96.8%	70.9%	71.8%
Unlabelled	99.2%	97.7%	98.7%	97.7%	95.0%	89.9%	99.0%	97.1%	98.8%	99.4%	97.0%	98.5%
Overall	98.9%	96.5%	97.7%	97.2%	94.2%	94.3%	97.4%	94.6%	96.8%	99.4%	96.0%	97.2%

lished deep learning models, making it feasible to use deep learning with a full 3D context. This should allow us to get the accuracy we have observed with deep learning models without the discontinuities that exist with current 2D-context approaches.

2) Extend our segmenter to use an incremental learning approach, allowing the iterative improvement of a learned model as further specimens are segmented. This would enable the use of a model trained on one or more specimens to segment a completely new specimen, with reduced user input. Such an extension is necessary for the automatic segmentation of collections of related mummies, which is currently not feasible.
3) Integrate our tool within a suitable software framework for volumetric image editing. This would allow seamless use of our segmenter from within the same environment already used to analyse these volumes, where memory allows.
4) Optimise our implementation for speed, to realise the complexity advantage of our method, also making use of high-performance computing (HPC) resources available in a user-friendly way.
5) Test our system on a much larger set of synchrotron microtomography scans, ideally in an operational environment, to ensure its readiness.

5.2. Project Synergies and Outreach

To implement the steps listed previously, the consortium would ideally be expanded to include an industrial member with background in developing tools for volumetric image editing. This would particularly facilitate steps 3 and 4. We believe the current consortium already has the necessary skill set for steps 1 and 5. However, for step 5 it may be useful to collaborate with other organisations that need to segment large volumes of a related nature.

During ATTRACT Phase 1, we actively disseminated our work and results in various ways:

- A scientific presentation at the Workshop on Cultural and Natural Heritage, ESRF-EBS, January 2020. The entire workshop was live streamed and remains available on the ESRF YouTube channel.
- A presentation at the Data Science Research Platform Seminar Series, University of Malta, with attendees from across University and external entities.
- Two interviews on national TV, during the live morning programme TVAM, and for the series Ras Imb Ras. Video clips are available on our website, courtesy of Public Broadcasting Services Ltd.
- Regular articles on the project website about ongoing activities.

These activities have strengthened our links with the local media, and we expect to promote our work in a similar way in Phase 2, should our project be chosen for funding.

5.3. Technology application and demonstration cases

The primary demonstration case in Phase 2 would remain the application to synchrotron microtomography of mummified remains. However, with the new BM18 beamline at ESRF-EBS, it is expected that much larger specimens, including human mummies, will be imaged with multi-resolution approach. We would also like to apply our automatic segmenter to collections of related mummies, enabling a statistical analysis which is at present impractical. These depend on the algorithm upgrades planned for Phase 2. A related demonstration case that we intend to include in Phase 2 is the application in paleontology, which is an important research topic at the ESRF, and where we have seen significant interest during our presentation at the Workshop on Cultural and Natural Heritage. Consortium members already have a sizable collection of synchrotron scans of paleontological samples, ready for further analysis. These samples are expected to be harder to automatically segment, due to the lower contrast inherent in the nature of the samples. Finally, we also intend to test our segmenter on the sizable collection of scans of human organs currently being generated at the ESRF. It is expected this would be of use in the fields of medicine and health sciences.

The European Research Infrastructure community is already represented in ASEMI through the ESRF. The technology application cases chosen are already active research areas at the ESRF, so that the results of this work are of direct relevance to the work done there.

5.4. Technology commercialization

So far we focused on the feasibility of our tool, and on the scientific application of its results. However, there is clear commercial interest in the integration of these technologies within existing tools for volumetric image editing and analysis. During Phase 1, in fact, we have seen the introduction of machine learning segmentation tools in existing commercial products. Hence, it is clear that the time could not be more opportune to commercialise the work done in this project. Concretely, we envisage two options for commercialising our tool. The simplest is to create a spin-off company to commercialise the tool as a plugin for existing volumetric image editors, implemented through their published Application Programming Interface (API). An alternative is to promote our tool with software houses that produce volumetric image editors, and if any show interest, to license our tool for integration within their software.

5.5. Envisioned risks

The biggest potential difficulty in our Phase 2 plans is the feasibility of reducing the complexity of deep learning models, in a way that allows their extension to 3D contexts without a step change in complexity. While we believe this is possible, similarly to the complexity reduction for conventional features as used in the ASEMI segmenter, we also have a contingency plan in case this proves problematic. An alternative approach is to apply deep learning with a 2D context independently to the three orthogonal planes, and fuse these outputs with that of our segmenter.

Other risks are easier to mitigate. If there is insufficient engagement by software companies with respect to integrating our tool within an existing volumetric image editor, we can always implement it ourselves as a plugin in software with a public API. We have already verified that this is possible with Dragonfly, for example. The small possibility of people leaving the team is mitigated by the knowledge overlap that exists within the team.

5.6. Liaison with Student Teams and Socio-Economic Study

With the collaboration with MSc student teams becoming an integral part of Phase 2, it is critical to have the necessary support within the project to make this work. In ASEMI, this requires input from at least two experts: an application domain expert to explain the context of the problem and the manual segmentation process, and an algorithm expert to explain how we tackle the problem in ASEMI. A small student liaison team, consisting of these experts, will be identified and its activities budgeted for, to ensure successful engagement with this action.

Similarly, a small project promotion team will be identified, with its activities budgeted for. This ensures that the necessary manpower is available for promotional activities related to ATTRACT. This also benefits activities such as media interviews, part of any successful project.

ACKNOWLEDGMENT

This project has received funding from the ATTRACT project funded by the EC under Grant Agreement 777222. Authors thank Alessandro Mirone for the initial GPU implementation of the 3D histogram algorithm, and Andy Götz for project management support.

REFERENCES

[1] C. Berruyer, S. M. Porcier, and P. Tafforeau, “Synchrotron “virtual archaeozoology” reveals how Ancient Egyptians prepared a decaying crocodile cadaver for mummification,” *PLoS ONE*, vol. 15, no. 2, p. e0229140, Feb. 2020.

[2] S. M. Porcier, C. Berruyer, S. Pasquali, S. Ikram, D. Berthet, and P. Tafforeau, “Wild crocodiles hunted to make mummies in roman egypt: Evidence from synchrotron imaging,” *Journal of Archaeological Science*, vol. 110, p. 105009, Oct. 2019.

[3] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, ser. LNCS, vol. 9351. Springer, 2015, pp. 234–241.

[4] Ö. Çiçek, A. Abdulkadir, S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: Learning dense volumetric segmentation from sparse annotation,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, ser. LNCS, vol. 9901. Springer, Oct. 2016, pp. 424–432.

[5] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” in *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, Oct. 2016.

[6] Dragonfly. Object Research Systems (ORS) Inc. 760 St-Paul West, Suite 101, Montreal, Canada H3C 1M4. [Online]. Available: https://www.theobjects.com/dragonfly/index.html