ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Generation of Simulated Dataset of Computed Tomography Images of Eggs and Extraction of Measurements Using Deep Learning

Jean Pierre B. L. Vargas[1], Davi D. de Paula[1], Denis H. P. Salvadeo[1], Emílio B. Júnior[1]*

1- São Paulo State University (UNESP) - Institute of Geosciences and Exact Sciences
Rio Claro, São Paulo - Brazil

Abstract. This paper extracts morphometric measurements of the different volumes of chicken egg components (shell, yolk, albumen and air chamber) by evaluating the segmentation algorithms, U-Net and Fully Convolutional Network (FCN). It also presents a new data set of 3D CT images of chicken eggs, simulating the different densities of a real one in the Digital Imaging and Communications in Medicine (DICOM) format and its labeled masks. The 3D models trained end-to-end showed high generalization even in the presence of variations in egg size and internal structures, achieving state-of-the-art segmentation performance with 99.4% accuracy.

1 Introduction

Consumption of chicken eggs is increasing worldwide and is being recommended as part of balanced nutrition[1]. Therefore, the relevance of doing an analysis for the processing of the characteristics of this food is very high. In the agriculture and food science the automated analysis of chicken eggs components throught computed tomography(CT) imaging presents a novel avenue for advancing both scientific understanding and practical applications. Typically, these analyzes are performed using conventional methods, which are destructive approaches using precision metrological equipment to make the physical measurements and bring valuable insights, but integration with deep learning techniques, particularly the approach using U-Net 3D[2] and Fully Convolutional Network(FCN) 3D[3], offers a promising solution for the tasks of segmentation and morphometric measurements.

Computed Tomography images assist in modern healthcare on a daily basis, allowing the visualization of internal structures, adding important information to diagnose diseases and guide treatment decisions. Segmentation is an image analysis task, delineating anatomical structures and regions of interest in images, each region representing different tissues, organs or pathologies. Segmentation contribuite to advancements in diagnostics, treatment planning, disease monitoring and in many others areas.

There are a variety of papers in the area of 3D segmentation of medical CT images using deep learning techniques as in [4] and also papers on conventional

*This work was funded in part by São Paulo State Research Support Foundation (FAPESP) through processes 2022/07446-3, 2022/14804-3, 2023/06422-6, 2024/02575-5.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

analysis of chicken eggs as in [5, 6, 7, 8], but there are no previous works on applications or studies on the analysis of chicken eggs using deep learning techniques with semantic segmentation to estimate morphometric measurements from CT or on analysis of the quality of chicken eggs with these techniques.

Our research seeks to make significant advances in the automation of egg analysis, with potential implications for quality control, breeding programs and nutritional studies. The objective of this work is to obtain morphometric measurements of the structures of the chicken egg from 3D computed tomographic images.

The main contributions of this paper are the creation of a synthetic dataset of computerized tomographic images of chicken eggs, the identification of the parts of the egg (shell, yolk, albumen, air chamber) through Deep Learning techniques and their morphometric measurements.

The rest of the paper is organized as follows: The dataset created for this research and the methodology used to identify the parts and estimate morphometric measurements are described in Section 2. The minute details of the experiments performed using these segmentation algorithms and the results are shown in Section 3. Finally in Section 4, we present some discussions and concluding remarks.

2 Methodology

This section describes the considerations for preparing the set of simulated CT images of chicken eggs. The architectures of the selected models are also shown, as well as the hyperparameters used for their training. The Pytorch framework was used in the Ubuntu OS and the Nvidia RTX 4090 GPU.

2.1 Simulator of Computed Tomographic Images of Chicken Eggs.

The egg is composed of 4 regions of interest: (i) eggshell; (ii) albumen; (iii) yolk and (iv) air chamber. The simulator generates these 4 regions. To simulate the eggshell (i) the 3D ellipse equation was used. The height, width and length of the ellipse was set to a random value between 110mm to 130mm, 50mm to 70mm and 50mm to 70mm, respectively. Eggshell thickness was also set to a random range between 1 mm and 5 mm. To simulate the albumen (ii), a fill with a fixed color was used inside the ellipse, a different color from the shell. To simulate the yolk (iii), a semi-sphere was used in the upper part of the egg. To simulate the air chamber (iv), a 3D parabola was used overlapping the upper part of the yolk. All egg components were filled with a single fixed color, so that detection of each component could be performed correctly.

In the case of generating the simulated images, a slice thickness was established for each slice of the 3D egg model, directly generating the slices. Because the reconstruction process of the real 3D CT image contains noise[9], additive Gaussian noise of mean = 0 and variance = 0.1 was added for pixels between 0-1, being then normalized between 0-255, these slices were saved in a DICOM file.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2.
Available from http://www.i6doc.com/en/.

In the preprocessing stage, we crop the images and masks from 150 slices x 150 height x 150 width to 80 slices x 140 height x 90 width. This process accelerated network training and reduced memory consumption on the Graphics processing unit(GPU), allowing the number of batches to be increased. The result of cropping the image has only one channel and enters the network with dimension 10 batches x 1 channel x 80 slices (depth) x 140 height x 90 width.

The chicken egg CT image simulator generated a data set of 100 3D images separated by slices. In addition to these images, masks with the same dimensions were generated for each image, these were created as a matrix with pixel values from 0 to 4 corresponding to the classes to be segmented (Background, Shell, Yolk, Album, Air Chamber). The Figure 1 illustrates a slices of a samples of the simulated tomographic images.

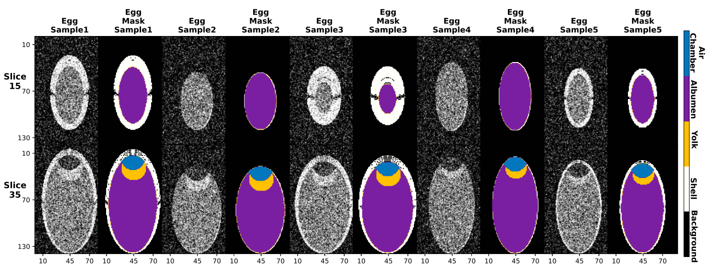


Fig. 1: Samples of a slices and masks of the Simulated Tomographic Images

2.2 Architectures and Hyperparameters

The hyperparameters of U-Net 3D/FCN 3D: Loss function = Cross Entropy, Optimizer = Adam, Learning rate = 0.01/0.0001, Epochs = 150/1000, Dropout = 0.1/0.5 as the result of different tests. The arquitecture is shown in Figure 2.

3 Experiments and Results

The database was divided into three sets: training 80%, validation 10% and test 10% following literature recommendations[10]. The Table 1 shows the values of

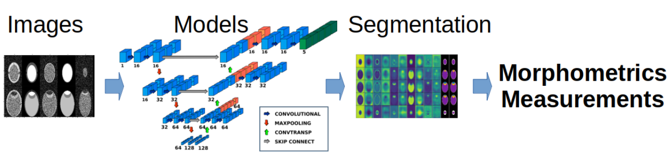


Fig. 2: Morphometric measurement pipeline

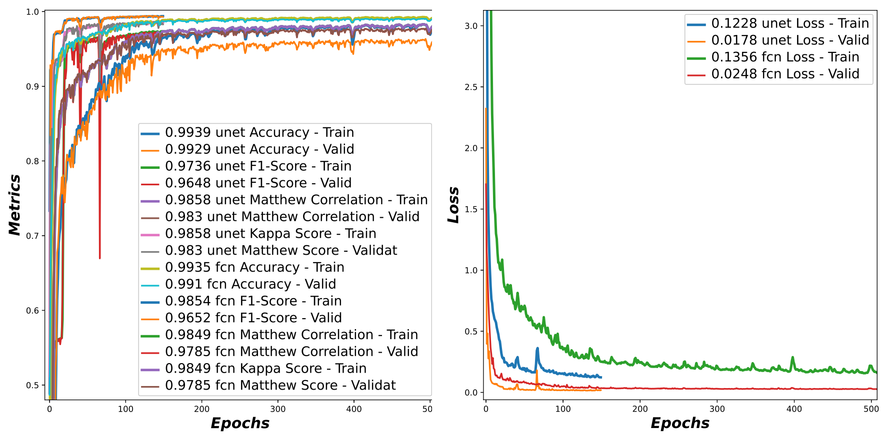
ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2.
Available from http://www.i6doc.com/en/.

the accuracy metrics achieved with the networks trained.

Algorithm	Metrics	Training	Validation	Testing
U-Net	Accuracy	0.9925	**0.9938**	**0.9933**
	F1-Score	0.9669	0.9686	0.9681
	M. Correlation	0.9826	0.9853	0.9680
	Kappa Score	0.9826	0.9853	0.9678
	Loss	0.1228	0.0178	-
FCN	Accuracy	0.9935	0.9910	0.9901
	F1-Score	0.9854	0.9652	0.9582
	M. Correlation	0.9849	0.9785	0.9580
	Kappa Score	0.9849	0.9785	0.9578
	Loss	0.1356	0.0248	-

Table 1: Accuracy values in training, validation and testing sets for 150 epochs for U-Net and 1000 for FCN.

The Figures 3a and 3b shows by epoch of the values of the different accuracy metrics selected and the behavior of the loss function respectively.



(a) Accuracies (b) Losses

Fig. 3: The accuracy and losses metrics of the U-Net 3D and FCN 3D models for each epoch are shown for both the training and validation data set.

The extraction of morphometric measurements of the egg, from the estimation of the masks, considers the total volume of a class as a count of the number of voxels of each class, which was then multiplied by the volume of a voxel. The volume of a voxel is $1mm^3$.

In Table 2 shows the number of voxels for each mask class (ground truth), also shows the number of estimated voxels and the total volume per class. These values correspond to the total of the test set. Percentage values greater than 100% in the fourth column indicate that more voxels corresponding to that class were detected than there actually were.

In Figure 4, a comparison is shown between the slices of a simulated sample

Model	Class	Number of Voxels	Estimated Number of Voxels/ %mean	Mean Precision per class	Mean Recall per class	Mean Total Estimated Volume(mm^3)
U-Net	Shell	556250	536158 (94.66%)	95.46%	90.44%	53615.8
	Yolk	91239	88358 (94.57%)	96.13%	90.95%	8835.8
	Albumen	1888047	1919788 (101.74%)	97.87%	99.57%	191978.8
	Ar chamber	78850	79780 (101.04%)	96.37%	97.33%	7978.0
FCN	Shell	556250	551415 (99.21%)	91.29%	90.54%	55141.5
	Yolk	91239	91070 (103.09%)	92.37%	94.41%	9107.0
	Albumen	1888047	1885569 (99.92%)	99.12%	99.04%	188556.9
	Ar chamber	78850	78835 (99.91%)	95.89%	95.80%	7883.5

Table 2: Estimated volume results for each egg component in the Test set.

from the test set, the resulting slices of the network as probabilities for each class and the result of the extraction of *argmax* of probabilities (final segmentation).

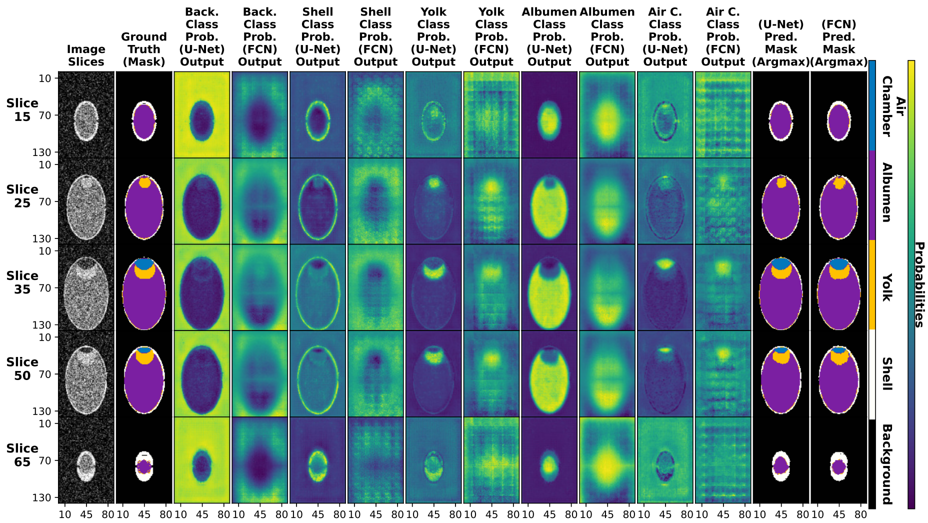


Fig. 4: Comparison of simulated tomographic image slices with the corresponding mask slices, the networks probability outputs for each class, showing in yellow the most probable class and the final predictions for U-Net and FCN Networks in the last two columns.

4 Discussions and Conclusions

The results obtained in this paper are satisfactory and, in a way, expected, since it is an application on a set of simulated and well-structured data, despite the effort to recreate the nature of the real data by adding noise to the samples.

In summary, the results achieved to date in the design, implementation and training using simulated 3D computed tomographic images for the problems of segmentation and extraction of morphometric measurements have been highly satisfactory, showing that the model is robust to noise.

This paper contributes to the state of the art of applications with a focus on Segmentation of Tomographic Images in Chicken Eggs, being the first of its kind. The creation of the tomographic image simulator for this study is another great contribution, opening the possibility of carrying out studies related to the food and production industry. The simulated base used in this paper and the simulator is available for free use by the community[11].

Future work includes consideration of the egg equation to generate more realistic models, simulation of the complete CT acquisition process (ray tracing + backprojection algorithms), application of the model trained with simulated data and tested on real images of eggs acquired by CT.

References

[1] H. Ritchie and P. Rosado. Per capita consumption of eggs. Technical report, Our World in Data, 2023.

[2] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[3] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440, Los Alamitos, CA, USA, jun 2015. IEEE Computer Society.

[4] M. H. Asnawi, A. A. Pravitasari, G. Darmawan, T. Hendrawati, I. N. Yulita, J. Suprijadi, and F. A. L. Nugraha. Lung and infection ct-scan-based segmentation with 3d unet architecture and its modification. *Healthcare (Switzerland)*, 11(2), 2023.

[5] V. G. Narushin, M. N. Romanov, and D. K. Griffin. Non-destructive measurement of chicken egg characteristics: Improved formulae for calculating egg volume and surface area. *Biosystems Engineering*, 201:42–49, 2021.

[6] Y. Liu, Y Ying, A. Ouyang, and Y. Li. Measurement of internal quality in chicken eggs using visible transmittance spectroscopy technology. *Food Control*, 18(1):18–22, 2007.

[7] G. Milisits, T. Donkó, Z. Süto, P. Bogner, and I. Repa. Applicability of computer tomography in the prediction of egg yolk ratio in hen's eggs. *Italian Journal of Animal Science*, 8(SUPPL. 3):234–236, 2009.

[8] M. R. Tataram, A. Charuta, W. Krupski, I. Åuszczewska Sierakowska, A. Korwin-Kossakowska, K. Sartowska, M. Szpetnar, and J. O. HorbaÅczuk. Interrelationships between morphological, densitometric and mechanical properties of eggs in japanese quails (coturnix japonica). *Journal of Poultry Science*, 53(1):51–57, 2016.

[9] G. L. Zeng. *3D Image Reconstruction*, pages 87–123. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[10] V. R. Joseph. Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 15(4):531–538, 2022.

[11] D. D. de Paula. Chicken egg simulator. https://github.com/daviduarte/ct-chicken-egg-simulator, 2024.