ZEBRAFISH
Volume 00, Number 00, 2019

DOI: 10.1089/zeb.2019.1754

Machine Learning Methods for Automated Quantification of Ventricular Dimensions

Mark Schutera,[1] Steffen Just,[2] Jakob Gierten,[3,4] Ralf Mikut,[1]
Markus Reischl,[1] and Christian Pylatiuk[1]

Abstract

Medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*) contribute substantially to our understanding of the genetic and molecular etiology of human cardiovascular diseases. In this context, the quantification of important cardiac functional parameters is fundamental. We have developed a framework that segments the ventricle of a medaka hatchling from image sequences and subsequently quantifies ventricular dimensions.

Keywords: zebrafish, fractional shortening, segmentation, biomedical imaging, deep learning, medaka

Over the past two decades, medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*) have significantly contributed to our understanding of the genetic and molecular etiology of human cardiovascular diseases.[1,2] In addition, zebrafish and medaka proved to be a valuable whole organism-based *in vivo* model system.[3,4] In this context, screening pipeline automation, including reliable and standardized image and video acquisition, as well as automated measurement and quantification of important cardiac functional parameters such as heart rate, systolic/diastolic volume, ejection fraction, stroke volumes, and fractional shortening as values for drug effects, is fundamental.[5,6] For these reasons, we have developed a framework that segments the ventricle of a medaka hatchling in image sequences from a ventral or lateral view and subsequently quantifies ventricular dimensions such as heart rate, systolic/diastolic volume, ejection fraction, stroke volumes, and fractional shortening. State-of-the-art approaches to estimate ventricular dimensions in medaka require manual intervention and labeling and involve a correspondingly large amount of engineering effort.[7] To some extent, existing approaches cover only part of the ventricular dimensions.[8] The reluctance to make annotated data publicly available renders it difficult to evaluate and assess present methods. In data-driven approaches, the medical and biological domain knowledge is implicitly captured within the data and its labels, and consequently the initial annotation effort is increased.

This study proposes a deep learning approach that reduces the manual engineering effort while covering the same or even greater scope of functionality. The image segmentation algorithm, and core of our framework, is based on the U-net architecture,[9] a symmetric convolutional neural network, which is successfully applied to a wide range of biomedical segmentation tasks.[10] The model was trained on 1226

[1]Institute for Automation and Applied Informatics (IAI), Karlsruhe Institute of Technology (KIT), Eggenstein, Germany.
[2]Department of Internal Medicine II, University of Ulm, Ulm, Germany.
[3]Department of Pediatric Cardiology, University Hospital Heidelberg, Heidelberg, Germany.
[4]Centre for Organismal Studies Heidelberg, Heidelberg University, Heidelberg, Germany.

pixel-wise annotated images and evaluated on 500 pixel-wise annotated images. During inference, the proposed framework follows a feed forward concept:

- Extract frames (red, green, blue channel [RGB], 640×480 pixels, *.tif*) from the image sequence (*.avi*, A1–A7 in Fig. 1)
- Predict the ventricle segment with U-net (B1–B7 in Fig. 1)
- Extract ventricular features from the segments generated by U-net (segment area, equivalent diameter, ellipse minor, and major axis)
- Determine the systolic and diastolic frames (C–E in Fig. 1)
- Estimate medaka hatchlings' ventricular dimensions from a ventral or lateral view (heart rate, ejection fraction, etc).

The individual steps are outlined in detail in the Supplementary Data; Supplementary Figures S1–S3, Supplementary Tables S1–S4. During inference, the segmentation of a single frame takes 2.32 s on average on a CPU (Intel Core i7-2670QM, 4×2.20 GHz). Evaluated on 125 frames of 5 ventral test set sequences, our segmentation algorithm achieves a mean accuracy of 88.16% (standard deviation [SD]: 9.23%) for the ellipse minor axis, and a mean accuracy of 86.68% (SD: 12.48%) for the ellipse major axis (D1 and E1 in Fig. 1). Evaluated on 125 frames of 5 lateral test set sequences, our segmentation algorithm achieves a mean accuracy of 78.59% (SD: 13.56%) for the ellipse minor axis, and a mean accuracy of 86.85% (SD: 12.79%) for the ellipse major axis. When calculating the ventricular dimensions (for the most part systolic–diastolic ratios), the predictions are evaluated on 150 frames of the ventral test set sequence R0004 and the lateral test set sequence N0092. The ventricular dimension estimation is robust over multiple cardiovascular cycles, such as heart rate SD of 0.08 beats per second (bps), heart rate mean of 1.59

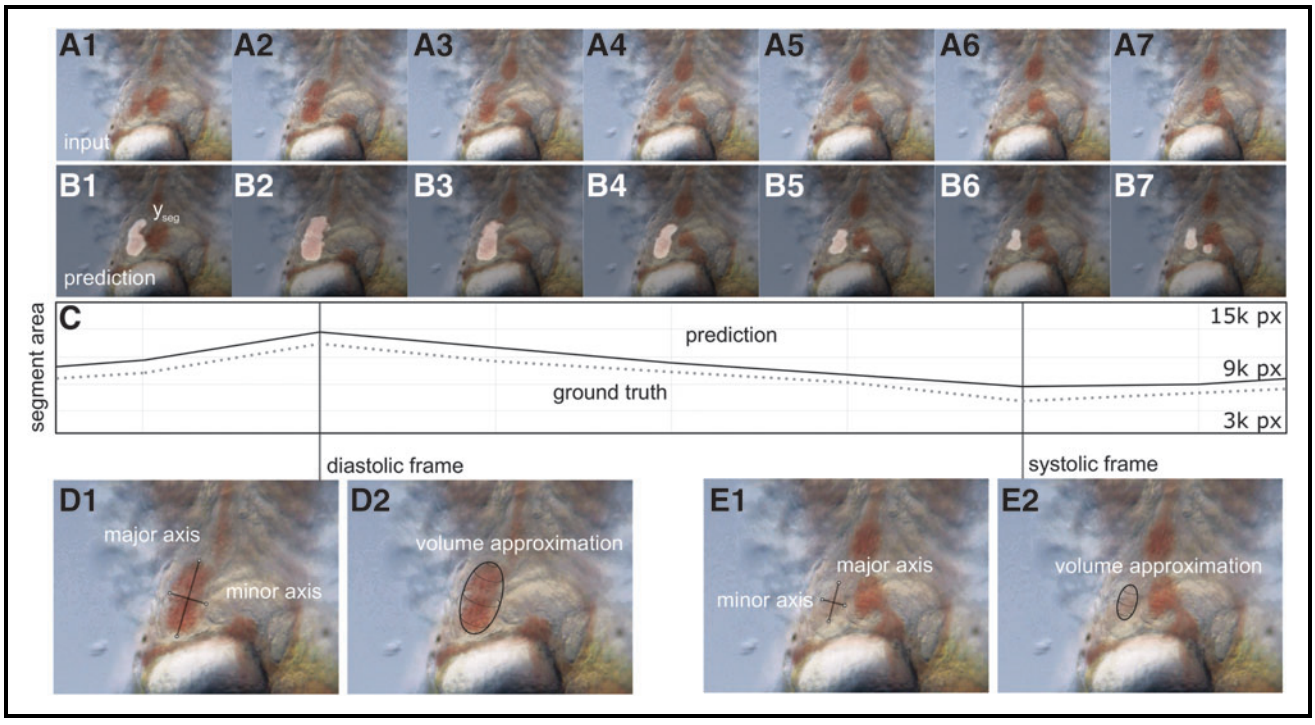

FIG. 1. **(A1–A7)** are input frames **x** of the ventral test set. **(B1–B7)** form an overlay of the corresponding binary heart segment predictions $\mathbf{y}_{seg}$ of the U-net. The overlay is binary, whereas the prediction is of a *white color*, which produces a *light reddish color* when superimposed on the input frames. **(C)** is the time series of the extracted segment area a, which is plotted as a function of consecutive frames, that is, over time. The *dotted curve* represents the segment area $\tilde{a}$ of the annotated ground truth segments; the *solid curve* depicts the predicted segment area. The two *vertical solid lines* mark the determined peaks, that is, frames with local segment extrema. These represent the diastolic frame (maximum, **D)** and the systolic frame (minimum, **E)**, which are displayed with the visualization of the ventricle's determined ellipse minor axis, ellipse major axis **(D1, E1)**, and prolate spheroidal approximation **(D2, E2)**.

bps, ellipse minor axis fractional shortening SD of 7.10%, ellipse major axis fractional shortening SD of 4.98%, ejection fraction SD of 4.46%. The point in time of the diastolic and systolic frames is detected robustly with a mean precision error of 0.53 frames and a maximal detection precision error of 1.00 frame (0.07 s). Increasing the frames per second of the video might thus lead to further improvement in detection precision. The recall of systolic–diastolic pairs is 100% on the test set.

In this study, we provide medaka researchers with a framework (in *Python*, *TensorFlow*, *Keras*, and *OpenCV)* for end-to-end automated estimation of ventricular dimensions through heart segmentation in image sequences of the heart region of medaka hatchlings from a ventral or lateral view. Our algorithm enables fully automated determination of cardiac ventricle dimensions in a large number of experiments, such as for small molecule drug discovery, and toxicological and genetic screenings. We are certain that this study can be widely used, adjusted to a wide range of high-throughput screens,[11] and thus increasing biomedical research. To this end and to ensure comparability, the framework—including the trained model, the majority of the annotated training and test data, the scripts for determining the ventricular dimensions, as well as the associated readme-file, log-files, a demonstration video, and inference results—is freely available and can be downloaded from the approved online repository (https://osf.io/snb6p).

Acknowledgments

The authors thank Katherine Quinlan-Flatter for proofreading the article.

Disclosure Statement

No competing financial interests exist.

Funding Information

This work was supported by the Helmholtz Association program ''BioInterfaces in Technology and Medicine - BIFTM'', the Research Center for Molecular Medicine (HRCMM), Career Development Fellowship (CDF), the Deutsche Herzstiftung e.V. (S/02/17), and by an Add-On Fellowship for Interdisciplinary Science of the Joachim Herz Stiftung.

Supplementary Material

Supplementary Data
Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3
Supplementary Table S4

References

1. Gut P, Reischauer S, Stainier DY, Arnaout R. Little fish, big data: Zebrafish as a model for cardiovascular and metabolic disease. Physiol Rev 2017;97:889–938.
2. Just S, Berger IM, Meder B, Backs J, Keller A, Marquart S, *et al.* Protein kinase D2 controls cardiac valve formation in zebrafish by regulating histone deacetylase 5 activity. Circulation 2011;124:324–334.
3. Ishikawa Y. Medakafish as a model system for vertebrate developmental genetics. Bioessays 2000;22:487–495.
4. MacRae CA, Peterson RT. Zebrafish as tools for drug discovery. Nat Rev Drug Discov 2015;14:721.
5. Hoage T, Ding Y, Xu X. Quantifying cardiac functions in embryonic and adult zebrafish. Cardiovasc Dev 2012;843:11–20.
6. Denvir MA, Tucker CS, Mullins JJ. Systolic and diastolic ventricular function in zebrafish embryos: influence of norepenephrine, ms-222 and temperature. BMC Biotechnol 2008;8:21.

7. Nasrat S, Marcato D, Hirth S, Reischl M, Pylatiuk C. Semi-automated detection of fractional shortening in zebrafish embryo heart videos. Curr Dir Biomed Eng 2016;2:233–236.
8. Pylatiuk C, Sanchez D, Mikut R, Alshut R, Reischl M, *et al.* Automatic zebrafish heartbeat detection and analysis for zebrafish embryos. Zebrafish 2014;11:379–383.
9. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Med Image Comput Comput Assist Interv 2015;9351: 234–241.
10. Scherr T, Bartschat A, Reischl M, Stegmaier J, Mikut R. Best practices in deep learning-based segmentation of microscopy images. Proc Workshop Comput Intell 2018;28:175–195.
11. Schutera M, Dickmeis T, Mione M, Peravali R, Marcato D, Reischl M, *et al.* Automated phenotype pattern recognition of zebrafish for high-throughput screening. Bioengineered 2016;7:261–265.

Address correspondence to:
Christian Pylatiuk, MD
Institute for Automation and Applied Informatics (IAI)
Karlsruhe Institute of Technology (KIT)
Hermann-von-Helmholtz-Platz 1
Eggenstein 76344
Germany

E-mail: pylatiuk@kit.edu