



HAL
open science

An end-to-end pipeline based on open source deep learning tools for reliable analysis of complex 3D images of Medaka ovaries

Manon Lesage, Jérôme Bugeon, Manon Thomas, Thierry Pécot, Violette Thermes

► To cite this version:

Manon Lesage, Jérôme Bugeon, Manon Thomas, Thierry Pécot, Violette Thermes. An end-to-end pipeline based on open source deep learning tools for reliable analysis of complex 3D images of Medaka ovaries. 2022. hal-03759490

HAL Id: hal-03759490

<https://hal.science/hal-03759490>

Preprint submitted on 24 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **An end-to-end pipeline based on open source deep learning tools**

2 **for reliable analysis of complex 3D images of Medaka ovaries**

3 Manon Lesage^{1,#}, Jérôme Bugeon¹, Manon Thomas¹, Thierry Pécot², Violette Thermes^{1,#}

4 ¹INRAE, Fish Physiology and Genomics Institute, 16 Allee Henri Fabre, Rennes 35000, France.

5 ²BIOSIT, UAR 3480 US 018, Université de Rennes 1, 2 rue Professeur Leon Bernard, Rennes 35042,
6 France

7
8 # Corresponding authors

9 10 **Keywords:**

11 Fish, reproduction, N2V, Cellpose, confocal microscopy, 3D imaging, optical tissue clearing,
12 Deep learning segmentation, Artificial Intelligence

13 14 **Short title:**

15 A deep-learning based workflow to assess the ovarian oocyte content in Medaka

16 17 **Summary statement**

18 An accessible image analysis method for biologists, which includes easy-to-use deep learning
19 algorithms, designed for accurate quantitative measurement of ovarian content from complex
20 3D fluorescent images.

21 22 **Grants support :**

23 The DYNAMO project (Agence National de la Recherche, ANR-18-CE20-0004).

24 The IMMO project (grants from the INRAE Metaprogramme DIGIT-BIO).

25 26 **Corresponding authors:**

27 manon.lesage@inrae.fr

28 violette.thermes@inrae.fr

29 INRAE, Laboratoire de Physiologie et Génomique des poissons, Campus de Beaulieu, 35042

30 Rennes cedex, France

31

32 **ABSTRACT**

33 Computational analysis of bio-images by deep learning (DL) algorithms has made
34 exceptional progress in recent years and has become much more accessible to non-
35 specialists with the development of ready-to-use tools. The study of oogenesis
36 mechanisms and female reproductive success in fish has also recently benefited from the
37 development of efficient three-dimensional (3D) imaging protocols on entire ovaries.
38 Such large datasets have a great potential for the generation of new quantitative data on
39 oogenesis but are, however, complex to analyze due to imperfect fluorescent signals and
40 the lack of efficient image analysis workflows. Here, we applied two open-source DL tools,
41 Noise2Void and Cellpose, to analyze the oocyte content of medaka ovaries at larvae and
42 adult stages. These tools were integrated into end-to-end analysis pipelines that include
43 image pre-processing, cell segmentation, and image post-processing to filter and combine
44 labels. Our pipelines thus provide effective solutions to accurately segment complex 3D
45 images of entire ovaries with either irregular fluorescent staining or low autofluorescence
46 signal. In the future, these pipelines will be applicable to extensive cellular phenotyping
47 in fish for developmental or toxicology studies.

48

49 INTRODUCTION

50 As imaging methods for thick biological samples improve and become more widespread in
51 various fields of life sciences, the volume of image data keeps growing and their analysis
52 becomes even more complex. Biologists are therefore facing a rising need for computational
53 tools to analyze large bio-image datasets and extract reproducible and meaningful biological
54 information.

55 The fish ovary is a complex organ that shows important structural and functional changes
56 during reproductive cycles. It contains different types of cells, including oocytes (*i.e.*, female
57 gametes) and numerous surrounding somatic supporting cells that form, together with each
58 oocyte, the functional units known as ovarian follicles (Lubzens et al., 2010; Nakamura et al.,
59 2009). During oogenesis, each follicle grows and differentiates until finally giving rise to eggs
60 that are ultimately released during spawning. One of the greatest challenges facing research on
61 the development of ovarian dynamics and functions is the lack of an effective method to
62 accurately count growing oocytes regardless of their stage. Studies have indeed traditionally
63 been limited to automatic or manual oocyte counting on two-dimensional (2D) ovarian sections
64 and extrapolation of the data to the whole organ or to manual counting of dissociated follicles
65 (Gay et al., 2018; Iwamatsu, Takashi, 1978; Iwamatsu, 2015). Some studies have also focused
66 on the development of complex stereological approaches to limit the biases induced by 2D
67 approaches (Charleston et al., 2007). Recently, the emergence of optical tissue clearing methods
68 and powerful microscopes have opened new perspectives with the possibility of imaging whole
69 ovaries in three dimensions (3D), notably for mice and fishes (Fiorentino et al., 2021; Lesage
70 et al., 2020; Soygur and Laird, 2021). It is thus now possible to generate 3D image data,
71 generally of very large size, that ideally allows direct and comprehensive access to all structures
72 and to achieve a precise 3D image reconstruction of the whole ovary. However, tools for 3D
73 image analyses are still too inaccurate and tedious, especially for image segmentation, partly

74 because of an irregular contrast signal in depth and the presence of oocytes of heterogenous
75 sizes, as reported previously for the adult Medaka ovary (Lesage et al., 2020). Ovarian 3D
76 imaging therefore has a promising future, but its widespread use still relies on the availability
77 of more efficient and easier-to-use computerized analytical tools.

78 In recent years, artificial intelligence (AI) has developed considerably and is proving to be
79 highly effective for digital image analysis in biology, which has recently led to a deluge of
80 publications in this field. Various algorithms based on deep learning (DL) have emerged and
81 provide many applications in microscopy allowing to overcome classical limitations such as
82 image segmentation. They allow to increase object recognition accuracy, segmentation
83 reproducibility and enable to save a considerable amount of time for the analysis of large
84 datasets by limiting manual interventions of users (Moen et al., 2019). Some specific methods
85 have thus been proposed to automatically segment follicles in the mammalian ovary from
86 histological 2D sections using a convolutional neural network (CNN) (Inik et al., 2019; Sonigo
87 et al., 2018). Other more generalist tools have recently emerged to democratize the use of DL
88 technology with few prerequisites in computed coding, by providing either DL trained models
89 accessible from public databases (<https://bioimage.io/#/>), notebooks accessible from any
90 computer (von Chamier et al., 2021), or other open-source plugins such as CSBDeep (Weigert
91 et al., 2018) or DeepImageJ (Gómez-de-Mariscal et al., 2021). Among the available models for
92 cell segmentation, Cellpose is a particularly versatile one, providing a generalist pre-trained
93 model for segmentation that can perform on various cell types in a great variety of acquisition
94 modalities (Stringer et al., 2021). Cellpose has recently proven to be very effective in
95 segmenting muscle fibers from 2D images of histological sections (Waisman et al., 2021).
96 Noise2Void (N2V) is another approach that stands out for its image denoising performance.
97 N2V does not require noisy image pairs nor clean target images, therefore allowing training
98 directly on the corpus of data to be denoised (Krull et al., 2019). In the era of deep learning, it

99 thus appears that some of the routine limitations for bio-image analysis are now solved. All that
100 remains for the biologist is the delicate task of integrating deep learning steps into the various
101 analytical procedures for 2D and for 3D images in particular.

102 The aim of this study was to test the possibility of using a pre-trained open-source model to
103 improve the critical step of segmentation of Medaka ovary 3D images without undergoing the
104 fastidious and complex task of neural network training. We generated 3D fluorescent images
105 of the adult ovarian follicle boundaries, by using the Methyl Green nuclear dye. We also
106 generated 3D images of ovaries at the larvae stage, by using the autofluorescence signal in
107 oocyte cytoplasm. For 3D segmentation of both types of images, we applied the generalist
108 Cellpose model for oocyte 3D segmentation, which was even more efficient after image pre-
109 processing steps and N2V denoising. A post-processing step after Cellpose was also set up to
110 eliminate any remaining error and to combine labels when necessary. N2V and Cellpose have
111 thus been integrated into a complete pipeline that allows an accurate estimation of the oocyte
112 content from complex 3D images of the whole Medaka ovary.

113

114 **MATERIAL AND METHODS**

115 **Ethical Statement**

116 All fish were reared in the INRAE ISC-LPGP fish facility, which hold full approval for animal
117 experimentation (C35-238-6). All fish were handled in strict accordance with French and
118 European policies and guidelines of the INRAE LPGP Institutional Animal Care and Use
119 Committee (no.M-2020-126-VT-ML, no.M-2019-48-VT-SG).

120

121 **Medaka breeding and sample collection**

122 Medaka fish (*Oryzias latipes*) from the CAB strain were raised at 26°C under artificial
123 photoperiod dedicated to growth phase (16 h light/ 8 h dark) or reproductive cycles (14 h light/

124 10 h dark). Female fish were sampled either at larvae stage (20 days post-hatch, dph) or adult
125 stage (5 months old). Fish were euthanized by immersion in a lethal dose of MS-222 at 300mg/L
126 supplemented with NaHCO₃ at 600mg/L and fixed overnight at 4°C in 4% paraformaldehyde
127 (PFA) diluted in 0.01 M phosphate buffer saline (PBS) pH 7.4. Larvae were then dehydrated
128 gradually in methanol and stored at -20°C. Adult ovaries were dissected after fixation and
129 directly stored at 4°C in PBS + 0.5% (w/v) sodium azide (S2002, Sigma-Aldrich).

130

131 **Fluorescent staining and clearing**

132 Larvae were progressively rehydrated in PBS and ovaries were dissected. Ovaries were then
133 permeabilized and immunostained following the iDISCO protocol with some modifications
134 (Renier et al., 2014). Samples were successively incubated in PBS/0.2% Triton X-100 (PBSTx)
135 for 30 min twice, PBSTx/20% DMSO for 30 min at 37°C and in PBSTx/0.1% Tween-20/20%
136 DMSO/0.1% deoxycholate/0.1% NP40 at 37 °C for 3 h. Ovaries were washed in PBSTx for 15
137 min twice, then blocked in PBS/0.1% Triton X-100/20% DMSO/6% Sheep Serum for 2H30-
138 3H at 37°C. Samples were immunolabelled with anti-phospho-Histone H3 (Ser10) primary
139 antibody (1:500, 06-570 Merck millipore), washed for 0.5 day in PBS/0.1% Tween-20/10µg/ml
140 heparin (PBSTwH) under gentle agitation, and incubated with Alexa-Fluor 546 secondary
141 antibody (1:500, A11035, ThermoFisher). Antibodies incubations were conducted for 2.5 days
142 at 37°C in PBSTwH/5% DMSO/3% Sheep serum. Finally, stained larvae ovaries were
143 embedded in low-melting agarose 1% before proceeding to clearing. Adult ovary samples were
144 stained and cleared according to the C-ECi method with few modifications (Lesage et al., 2020).
145 For staining, adult ovaries were incubated with the Methyl Green dye (MG) (40 µg/mL, 323829,
146 Sigma-Aldrich) in PBS/0.1% Triton X-100 at 37°C for 2.5 days. After Staining, both adult
147 ovaries and embedded larvae ovaries were dehydrated in serial methanol/H₂O dilution series
148 supplemented with Tween-20 (2% and 0.1%, respectively), then immersed in 100% ethyl-3-

149 phenylprop-2-enoate (ethyl cinnamate [ECi]) (W243000, Sigma-Aldrich) and finally kept at
150 room temperature until subsequent imaging step.

151

152 **Samples mounting and imaging**

153 Image acquisitions were performed with a Leica TCS SP8 laser scanning confocal microscope
154 equipped with a 16x/0.6 IMM CORR VISIR HC FLUOTAR objective (ref. 15506533, Leica,
155 Wetzlar, Germany). For larvae ovaries, samples embedded in agarose blocks were glued on a
156 coverslip and placed in a glass Petri dish filled with ECi. Adult ovaries were successively placed
157 with ventral side up or down for complete imaging despite the objective working distance
158 limitation, and mounted as described previously (Lesage et al., 2020). Mosaic z-stack tiles were
159 stitched in Leica software using 11,72% overlap. Larvae ovaries were acquired in 1024x1024
160 pixels, 400Hz (unidirectional) with an optical zoom of 1.3 and a z-step of 1.63 μm (voxel size
161 0.52 x 0.52 x 1.6264 μm). PH3 fluorescent signal was acquired using 552 nm laser excitation
162 slightly above optimal intensity (3-4%), and frame average was set to 2. Acquisitions took
163 between 1.5 and 5.5 hrs according to ovary size and generated 1 to 2 GB of data. Adult ovaries
164 were acquired in 512x512 pixels, 600Hz (bidirectional), optical zoom 0.75, z-steps 6 μm (voxel
165 size 1.80 x 1.80 x 6.00 μm), line accumulation 2 and frame average 2. Ventral and dorsal z-
166 stacks were acquired in about 10 hrs each and generated 8 to 10 GB of data. MG staining was
167 detected with 638 nm laser and excitation gain compensation was used along Z axis (5 to 10%
168 intensity).

169

170 **Image processing**

171 A schematic overview of image treatment workflows is shown on Figure 1. All steps were
172 conducted on the open-source FIJI software, unless otherwise specified.

173

174 ***Image intensity and contrast enhancement***

175 Before image enhancement, adult z-stacks were downscaled in order to reduce computation
176 time. A resampling factor of 3 on X and Y axes was used, which resulted in images sizes of
177 1214 x 970 pixels. Progressive intensity and gamma correction plugin was applied along the Z
178 axis to compensate fluorescence loss in depth (Fig. 1B). For larvae, exponential or linear
179 interpolation method were used with default parameters and intensity enhancement was set
180 between 150 and 400% depending on samples. For adults, linear interpolation method was used,
181 intensity set between 200 and 800% and normalization was selected (modifying range of pixel
182 intensity values by linear scaling method). A linear gamma correction was also performed
183 (factor 1.5) to enhance mid tones pixels on adult images. Image contrast was then enhanced by
184 applying Contrast Limited Adaptive Histogram Equalization (CLAHE) with following
185 parameters: block size 128, bins 256, slope 3 and fast mode, for larva; block size 512, bins 256,
186 slope 30 and fast mode, for adult. A Fiji macro was used to apply this function on Z-stacks by
187 batches, which is available on our GitHub page: [https://github.com/INRAE-](https://github.com/INRAE-LPGP/ImageAnalysis_CombineLabels)
188 [LPGP/ImageAnalysis_CombineLabels](https://github.com/INRAE-LPGP/ImageAnalysis_CombineLabels).

189

190 ***3D registration***

191 Adult ventral and dorsal 3D stacks were registered, aligned and combined with the FijiYama
192 plugin using the “two images registration mode (training)” (Fig. 1B). A manual registration was
193 first performed to roughly superimpose the two volumes. Automatic registration was then
194 applied for linear image transformation with block-matching alignment method. Linear
195 transformations included rigid transformations (translation and rotation) and, if necessary,
196 similarities transformations (rigid and isotropic homothetic factor). The two registered stacks
197 were fused with Image calculator (Max operator), resulting in image size of 1324 x 1108 x 713
198 pixels.

199

200 ***Signal-to-noise ratios enhancement***

201 Three-dimensional images were denoised using Noise2Void (N2V) deep-learning based tool
202 available on Fiji, using a model trained on a few selected 3D stacks snippets. A 2D model was
203 trained on folder containing ~15 Z-stacks snippets (512x512, from 50 to 115 z-steps) cropped
204 from different larvae samples. Training patch shape was set at 96x96 pixels and N2V
205 automatically used data augmentation (90, 180 and 270 rotations and flipping), thereby
206 multiplying total patches amount by 8. The resulting pool of 2D patches were used for training
207 (90%) and validation (10%). Training was performed with 250 epochs, 150 steps/epoch and
208 batch size set to 128, resulting in ~13 hrs of training with our computer specifications.
209 Denoising prediction duration was estimated to ~12 min for 1GB of data with our stated
210 parameters (batch size 2). For adult, similar strategy was used for training, using 10 z-stack
211 snippets (256x256, 100-200 Z-steps), patch shape 64x64 pixels. Training was performed with
212 300 epochs, 200 steps/epoch, batch size 128, for a total of ~9 hrs of training. Denoising
213 prediction duration was estimated to ~8 min for 1GB of data with our hardware specifications
214 and stated parameters (batch size 2). For image edges enhancement, stacks were subjected to a
215 3D median filter (x,y,z radius 1,1,1 for larvae and 2,2,2 for adult). Filtered image was then
216 subjected to external morphological gradient computation (shape: ball; x,y,z radius 3,3,3 for
217 larvae and 2,2,2 for adult) with Morphological filters (3D) function of MorpholibJ plugin.
218 External gradient image was then subtracted from original pre-treated stack (without median
219 filtering). An internal morphological gradient was also computed on larvae stacks (element
220 shape: ball, x,y,z radius 4,4,4) and added to image data. For 3D visualization of data, volume
221 reconstructions were performed on the Amira software using Volren rendering (Fig 2A, E) or
222 Volume-rendering (Fig 3A, F and 4A, F).

223

224 ***Deep learning 3D segmentation***

225 Follicle segmentation was performed using Cellpose algorithm with local environment
226 installation, launched from Anaconda command prompt (Fig. 1C). For larvae, X and Y scale
227 were first reduced by half so that mean follicle diameter approach ~30 pixels, which is the
228 optimum diameter for Cellpose cell segmentation (final image size 1194 x 610 pixels). Cellpose
229 was then run in 3D with “cyto” pre-trained model, setting parameters as follows: diameter 30,
230 cellprob threshold -2, anisotropy 1,6, min size 10. A batch size of 2 was used, depending on
231 GPU memory allocation, resulting in ~50 min for segmentation prediction of ~250 Mb of data.
232 Resulting masks were saved in TIFF format for subsequent data treatment. For adults, the same
233 process was used except anisotropy was set to 1.7. 3D segmentation took ~4 hrs for ~1 Gb of
234 data. To segment out-of-range follicles, adult stacks were downscaled once more by applying
235 a resampling factor of 2 in X, Y and Z (no interpolation, final images size of 662 x 554 x 357
236 pixels). Downscaled stacks were subjected to Cellpose segmentation with diameter size set to
237 30 and 60 pixels. 3D segmentation took ~35 min and ~11 min for ~125 Mb of downsized data,
238 for 30 and 60 pixel diameter respectively.

239

240 ***Post-processing and data extraction***

241 For post-processing of segmented follicles, data were first slightly narrowed. For that operation,
242 label boundaries were computed with MorpholibJ plugin and subtracted from the original
243 Cellpose results. For larvae ovary images, labels were then post-processed on AMIRA software.
244 Labels were subjected to an opening morphological operator (3 pixels, precise) and then filtered
245 based on their size (Equivalent Diameter $\geq 1.5e-5m$) and shape (ShapeVAa3d ≤ 3.5). Few
246 remaining errors were manually corrected. For adult ovary images, label shrinkage, filtration
247 and combination were performed automatically or semi-automatically using a Fiji macro.
248 Labels were filtered based on volume and sphericity parameters. When necessary, segmentation

249 images were rescaled to match 3D registered image size (1324 x 1108 x 713px). The
250 combination strategy consisted in adding largest segmented labels from downscaled images
251 (using Cellpose diameter 30) on original scale label segmentation (Cellpose diameter 30) where
252 largest follicles were over-segmented (Supplemental Fig. 1). Briefly, labels >650 μ m volume-
253 equivalent diameter (EqDiameter) were filtered from the downscaled image and added to the
254 original scale label image after selection and deletion (using morphological reconstruction
255 operation) of wrong labels resulting from over-segmentation. For combination of missing
256 largest labels, a similar strategy was used with downscaled label image obtained with Cellpose
257 diameter 60, but with semi-automatic method. Missing labels were manually selected with
258 multi-point tool and then processed as presented before. Macro “CombineLabels” was
259 developed in IJ1 Macro language and can be downloaded from the Github page:
260 https://github.com/INRAE-LPGP/ImageAnalysis_CombineLabels. The volumes of all
261 segmented follicles were exported and equivalent diameters were calculated. For adults,
262 EqDiameter were subjected to a correction factor of 1.12 to compensate the volume shrinkage
263 due to sample clearing, as described in Lesage *et al.* (Lesage et al., 2020). Data analysis was
264 performed on labels above 25 and 50 μ m in diameter, for larvae and adult samples respectively.
265 Label 3D reconstructions were generated on Amira using volume-rendering object.

266

267 **Hardware and software**

268 Data were analyzed on a 64-bit Windows 10 Pro computer equipped with a 2x Intel Xeon Silver
269 4110 (8 Cores, 3.0GHz) processor, a Nvidia Geforce GTX 1080 graphic card, and 384 Go of
270 RAM. We used the Amira 2020.2 software with the XLVolume extension (Thermo Fisher
271 Scientific, Waltham, Massachusetts, United States), Anaconda3-2021.11 python distribution,
272 Python 3.7.9, CUDA toolkit 10.0, PyTorch 1.6.0 and Cellpose v0.6.1 (Stringer et al., 2021).
273 We also used FIJI (Schindelin et al., 2012) and the following plugins: CLAHE (Pizer et al.,

274 1987; Zuiderveld, 1994), Progressive intensity and gamma correction (Murtin, 2016), Fijiyama
275 (fiji-yama-4.0.0) (Fernandez and Moisy, 2021), CLIJ2 (clij2-2.5.3.0, Haase et al., 2020),
276 MorpholibJ (morpholibJ-1.4.3, Legland et al., 2016), Noise2Void (n2v-0.8.6)(Krull et al.,
277 2019) and CSBDeep (csbdeep-0.6.0)(Weigert et al., 2018).

278

279 **RESULTS**

280 **3D imaging of the ovaries**

281 To detect oocytes within the ovary at both adult and larvae stages, sample were
282 fluorescently stained and optically cleared to allow full imaging by confocal fluorescence
283 microscopy (Fig. 1A). For adult ovaries, nuclei of supporting cells surrounding the oocytes
284 were stained with the fluorescent nuclear dye Methyl-Green (MG) identified as a convenient
285 marker for delineating follicle boundaries (Lesage et al., 2020). For larvae ovaries (20 dph),
286 which are composed of small early developing oocytes flanked by only a few supporting
287 somatic cells, we took advantage of the cytoplasmic autofluorescence generated by
288 immunostaining (here anti-phospho-histone H3 antibody, PH3). Resulting images displayed a
289 very low signal-to-noise ratio (SNR) and a rapid loss of signal recovery in depth for larvae
290 ovaries (Fig. 2A-D). Signal intensity was twice as low at 440 μm in depth compared to the top
291 (150 μm depth, Fig. 3B, B'). In addition, it is noteworthy that smaller oocytes were less
292 distinguishable than larger ones having thicker cytoplasm, especially in very compact regions
293 (Fig. 2D and 3B). Images stacks of adult ovaries displayed a higher fluorescence signal with a
294 high SNR that was recovered up to 1152 μm in depth, although some heterogeneity in
295 fluorescence intensity was observable (Fig. 2F, G). At a greater depth (2 000 μm), images
296 display a substantial loss of signal intensity (Fig. 4B, B').

297

298 **Image enhancement and 3D visualization**

299 Given the uneven signal intensity of the images, and especially the very low SNR observed
300 with the non-specific staining of larval ovaries, we applied successive processing steps to
301 enhance the fluorescent signal throughout Z-stacks prior segmentation. For larvae ovary,
302 fluorescence intensity of image stacks was progressively enhanced along Z-axis to increase the
303 signal in depth, and mean grey values were increased and homogenized to enhance contrasts
304 (Fig. 3B, B', C, C'). To minimize the noise potentially introduced by intensity and contrasts
305 adjustment and to avoid potential aberrant enhancement of noisy structures, image stacks were
306 denoised using the self-supervised N2V deep-learning-based algorithm (Krull et al., 2019) (Fig.
307 3D, D'). Finally, edges were refined using morphological gradients (Fig. 3E, E'). XY views
308 from Z-stacks and fluorescent intensity profiles through adjacent oocytes show the progressive
309 signal recovery over the different steps at both 150 and 440 μm in depth. It is noticeable that
310 while normalizing grey values distribution, N2V denoising preserves oocytes edges with
311 limited blurring effect, thus minimizing any feature loss (Fig. 3D, D'). In addition, it is
312 noteworthy that overexposure was created in some cases as a side effect of edge refinement.
313 The challenge here was therefore to find a compromise between the loss of detection of
314 underexposed oocytes and the overexposure generated in order to achieve the greatest
315 difference between light and dark levels. Image pre-processing steps thus enabled to increase
316 the overall fluorescent signal intensity, to better define edges of the oocytes and to homogenize
317 the fluorescence intensity across the Z-stack, thereby allowing a better 3D reconstruction of the
318 larval ovary (Fig. 3A, F).

319 For 3D images of adult ovaries, a similar strategy was applied except an extra step of
320 automatic 3D registration that was performed for the reconstruction of the whole ovary (Fig.
321 4B-E and 4B'-E'). As a result of the combination of images in the overlap region, 3D
322 registration led to a slight increase in fluorescence intensity in this region in the final stack (Fig.
323 4A, F). XY views of Z-stacks and fluorescent intensity profiles through adjacent oocytes

324 showed a significant increase of the SNR, especially at 2 000 μm in depth. Similar to the larvae
325 ovary, the pre-processing allowed to improve the fluorescence signal, and especially to
326 homogenize the fluorescence intensity through the Z-stack for a better 3D reconstruction of the
327 adult ovary (Fig. 4A, F).

328

329 **Cellpose efficiently identifies oocytes and follicles on 3D images**

330 For 3D oocyte and follicle segmentation on larvae and adult images, we selected the open-
331 source Cellpose deep-learning algorithm because of its generalist nature for cell segmentation
332 (Stringer et al., 2021). We compared the efficiency of Cellpose for 3D segmentation before and
333 after image pre-treatment. In both cases, Cellpose could detect either internal fluorescent
334 staining (oocyte cytoplasm) or external fluorescent staining (somatic follicular cells), on larvae
335 and adult ovary images respectively (Fig. 5A-D and 5E-H). Notably, Cellpose was much more
336 efficient on pre-treated images than raw images. Although XZ views of larvae stacks revealed
337 accurate segmentation along the Z axis, several undetected oocytes and some Z-label fusions
338 were detectable in the absence of preprocessing (Fig. 5B and 5D, insets). For adult ovaries,
339 segmentation of raw images leads to many cases of over-segmentation in conjunctive tissues or
340 in large follicles, as well as fewer detected follicles, compared to segmentation of pre-processed
341 images (Fig. 5F and 5H, insets).

342

343 **Post-processing of label images after Cellpose 3D segmentation**

344 Cellpose output images were post-processed to adjust the label sizes to that of the oocytes (label
345 shrinkage) and to remove outliers (label filtration) (Fig. 6). Label shrinkage was performed by
346 automatically subtracting the label boundaries to the original Cellpose labels. For adult ovaries
347 that have the unique feature of containing heterogeneous follicle sizes (ranging from about 20
348 to 1 000 μm in diameter), different Cellpose label images were generated by modulating image

349 resolution of the input image (Fig.6B, C). If necessary, a 60 pixel diameter was used for
350 Cellpose segmentation to detect largest follicles. The different resulting label images were
351 combined in an additional post-processing step by using a Fiji Macro named “CombineLabels”
352 (Fig. 6D). Images of larvae ovary labels show that, after post-processing, the majority of labels
353 perfectly fit to the shape and size of the oocytes and that aberrant labels with elongated shapes
354 or very small sizes were removed. In few cases, some inaccuracies still persisted, mainly under-
355 segmentation of small oocyte clusters (Fig. 6A, arrows) or non-segmented oocytes (Fig. 6A,
356 arrowhead). Similar to larvae ovary images, results of segmentation and post-processing of
357 adult ovary images were highly accurate, both in terms of follicle detection, label shape and
358 size fitting (Fig. 6B-D). After post-processing, remaining segmentation errors were limited to
359 a few outlier labels located outside the relevant structures.

360

361 **Oocyte content analyses**

362 To assess the ovarian oocyte content at both larvae and adult stages, ovaries were imaged at
363 each of these stages and 3D computational analyses were performed following our deep
364 learning-based pipeline. Three-dimensional reconstructions after data pre-processing revealed
365 the thin oval-shape of larvae ovaries oriented along the anteroposterior axis, which then evolves
366 into a thicker rounded shape at the adult stage (Fig. 3F, 4F, 7A and 7D). Ventrally, larvae
367 ovaries exhibited lateral folds and a marked central depression, likewise adult ovaries displayed
368 two lateral folds as well as a ventro-median bulge, giving the ovary a wheat grain appearance.
369 Diameters of segmented oocytes or follicles were computed, classified into different size
370 classes and merged to the 3D ovary reconstructions (Fig. 7A', B, D', E). The ventral and dorsal
371 3D views of the larvae ovary, revealed that small oocytes were preferentially visible from the
372 ventral views, whereas larger oocytes were only observable from dorsal views, while no
373 obvious regionalization was observable in the adult ovary (Fig. 7B and 7E). To analyze the

374 relative abundance of the different size classes, the developmental stage of oocytes/follicles
375 was determined according to their diameter and as described in the oocyte developmental table
376 of Iwamatsu *et. al.* (Iwamatsu et al., 1988). In the larvae ovary, a total of 1231 ± 182 (n=2)
377 oocytes were detected. The mean size distribution showed a high predominance of small
378 previtellogenic follicles ranging from 25 to 60 μm in diameter (chromatin-nucleolar stage, stage
379 I), which suggests a synchronized oocyte growth during larval development (Fig. 7C). By
380 contrast, all follicular developmental stages were found at the adult stage. A total of 1275
381 follicles were counted with a large predominance of pre-vitellogenic follicles (from stage II to
382 IV, 50-150 μm) and of early vitellogenic follicles (stages V and VI, 150-400 μm , Fig. 7F).
383 Proportion of follicles then progressively decrease as they progress through late vitellogenesis
384 (stages VII and VIII, 400-800 μm). The pool of post-vitellogenic follicles (maturation stage IX,
385 > 800 μm) is clearly distinguishable and reflects upcoming egg laying with a consistent number
386 of about 23 follicles measuring more than 950 μm in diameter.

387

388 **DISCUSSION**

389 Three-dimensional imaging of whole fish ovaries typically generates large image data sets that
390 are particularly complex to analyze. In this study, we generated two types of 3D images. On
391 the one hand, we generated images of adult fish ovaries with low-contrast follicle outline signal
392 at great depths, which usually greatly impairs the final segmentation efficiency, as described
393 previously (Lesage et al., 2020). On the other hand, we generated images of larvae ovaries with
394 a low-contrast signal inside the oocytes throughout image stacks, which makes segmentation
395 otherwise impossible with conventional approaches. Here, we applied the generic Cellpose pre-
396 trained algorithm that allows cell segmentation without any manual annotation and neural
397 network training. To optimize 3D segmentation results and maximize accuracy of
398 oocyte/follicle content analyses, Cellpose was integrated into an end-to-end analysis pipeline.

399

400 **Enhancement and homogenization of input dataset**

401 The first part of our pipeline was dedicated to signal quality improvement in depth of raw image
402 stacks. Such image pre-processing steps allowed improving segmentation efficiency by
403 Cellpose. To some extent, the decrease in fluorescence level in depth on raw images should not
404 be a major issue for predicting feature boundaries with Cellpose as it uses vector gradients
405 representation of objects to accurately predict complex cell outlines with non-homogenous cell
406 marker distribution (Stringer et al., 2021). However, our result indicates that the SNR is an
407 important prerequisite for image analysis with Cellpose, in line with previous observations (Kar
408 et al., 2021). Along with an enhanced visualization of the structures of interest across the
409 sample, the pre-processing of 3D images therefore allows for homogenization of the data set
410 and much more efficient 3D segmentation with Cellpose, thus increasing the reproducibility
411 and quality of analysis.

412

413 **Improvement of Cellpose output label images**

414 Despite its high efficiency, Cellpose led to some substantial errors, including slightly oversized
415 or aberrant labels, and it also failed to segment oocytes of highly heterogeneous sizes. To
416 overcome these limitations and refine labels produced by Cellpose, we performed post-
417 segmentation corrections. The size of 3D labels was adjusted following an automated boundary
418 subtraction strategy. Our strategy differs from other methods that use the pixel-by-pixel label
419 erosion operation, such as in LabelsToROIs Fiji plugin designed on 2D myofiber sections, and
420 is likely to be faster when dealing with large 3D data (Waisman et al., 2021). The combination
421 of multiple Cellpose segmentation images, implemented with a Fiji macro “CombineLabels”,
422 also allows identification of highly heterogeneous objects sizes, that was previously not
423 possible with Cellpose algorithm alone. It is however worth noting that there are still few

424 inaccuracies that could not be fixed. Under-segmentations or unsegmented objects were
425 sometimes detected mostly with larvae image datasets. Albeit minor, these errors occur in
426 highly oocyte-dense regions or with non-optimal signal levels. Such observation is in agreement
427 with some studies that do not recommend Cellpose for highly overlapping masks or that
428 describe lower accuracy with over- or underexposed images (Kar et al., 2021). This could be
429 attributed to the 2D averaging process for the 3D Cellpose extension that may have lower
430 accuracy than a model trained with 3D data, especially for highly dense regions (Lalit et
431 al.,2022; Stringer et al., 2021). Obviously, one can assume that better accuracy could be
432 achieved by using a dedicated specialized DL model, and in particular with 3D trained model
433 on our data, as shown by D.Eschweiler et *al* (Eschweiler et al., 2022). It would thus be
434 interesting in the future to use our segmentation results for Cellpose algorithm fine-tuning. This
435 could indeed limit the need for image pre-processing as well as post-processing corrections of
436 segmentation results. But in this case, we would somewhat lose the advantage of versatile
437 generalist models like Cellpose and different models would have to be trained for each type of
438 data. Another solution could therefore be to improve the input images quality, by using a
439 suitable oocyte marker to avoid sharp signal enhancement and possibly in combination with a
440 membrane marker for better boundary discrimination. Alternatively, and in absence of such
441 specific staining, another denoising process, either trained in three dimensions, with noisy/non-
442 noisy paired images (CARE) or combining deconvolution process (DecoNoising), could also
443 help objects recognition accuracy (Weigert et al., 2018; Goncharova et al., 2020).

444

445 **An accurate and comprehensive content analysis of larvae and adult medaka** 446 **ovaries**

447 Implementation of Cellpose for oocytes/follicle 3D segmentation eventually enabled unbiased,
448 reproducible and comprehensive studies for meaningful biological information, which offers

449 great possibilities for a complete description of fish ovarian growth and development. From a
450 morphological point of view, we could clearly distinguish the oval shape of the ovary
451 thickening over time and shaping a bulge in the ventro-median position that connects the
452 mesentery and attaches to the gut (Iwamatsu, 2015; Lesage et al., 2020). *In situ* follicular size
453 measurements by our 3D imaging and DL-based segmentation approaches allowed producing
454 size distribution profiles for both larvae and adult ovaries. Our results are consistent with those
455 obtained previously from dissociated follicles measured manually for the larvae ovary
456 (Iwamatsu, 2015) or semi-automatically from 3D images by classical watershed segmentation
457 approaches for the adult ovary as shown in our previous study (Lesage et al., 2020). However,
458 greater confidence can be attributed to the present study, particularly for the pre- and post-
459 vitellogenic stages in the adult ovary for which we achieved fewer segmentation errors. In
460 general, we also achieved a better estimation of follicle size due to the accurate shape detection
461 enabled by the Cellpose algorithm. Interestingly, we also noticed that the spatial distribution of
462 oocytes between 30 and 70 μm in diameter tended to be regionalized along the ventro-dorsal
463 axis in the larvae ovary, suggesting an oriented follicular growth through this axis in
464 consistency with observation of Nakamura *et al.* (Nakamura, 2018). In the future, the ovarian
465 morphogenesis and spatial organization of follicles according to their size should however be
466 further characterized during the ovarian development by using refined 3D spatial analysis
467 approaches.

468

469 **Conclusion**

470 Overall, the use of the generic Cellpose algorithm has been successful for 3D ovary images and
471 has allowed ovarian segmentation of unprecedented quality. Cellpose significantly accelerated
472 and improved the efficiency and the quality of ovarian follicles 3D segmentation in adults,
473 leading to an accurate count and measurement of all oocyte diameters. Even more remarkably,

474 this generalist model also allowed the successful segmentation of images of larvae ovaries with
475 weak fluorescent signal, otherwise not exploitable with conventional methods, and quite
476 certainly even after image pre-processing. This possibility challenges the dogma that a good
477 raw image is necessary for an accurate object segmentation and thus significantly increases
478 further analysis opportunities. Furthermore, thanks to its ease of use, implementation of
479 Cellpose avoids the tedious and complex step of setting up an AI segmentation method and is
480 therefore largely accessible to non-specialist biologists with limited coding and hardware
481 knowledge. In the deep learning era, it is thus now clearly possible to apply such a cutting-edge
482 technology for tissue 3D phenotyping with relative ease. To our knowledge, our pipeline is the
483 first application using developer-to-user deep learning solutions for 3D image analysis of the
484 ovary in vertebrates, thus opening the way for further innovative in-depth morphometric studies
485 within the framework of developmental or toxicological studies.

486

487 **ACKNOWLEDGMENTS**

488 We thank the INRAE ISC-LPGP fish facility staff and especially Amélie Patinote and
489 Guillaume Gourmelen for fish rearing and husbandry.

490

491 **ADDITIONAL INFORMATION AND DECLARATIONS**

492 **Funding**

493 This work was funded by The DYNAMO project (Agence Nationale de la Recherche, ANR-
494 18-CE20-0004 to V.T.). This work has also been supported by the IMMO project (grants from
495 the INRAE Metaprogramme DIGIT-BIO to V.T.). The funders had no role in study design, data
496 collection and analysis, decision to publish, or preparation of the manuscript.

497

498 **Competing Interests**

499 The authors declare there are no competing interests.

500

501 **Author Contributions**

502 M.L. performed the experiments, the computational analyses and wrote the manuscript.

503 J.B. participated to the setup of the computation bio-image analyses. M.T. participated in

504 the setup of the clearing protocol and to the image acquisition. T.P. participated in the

505 choice and implementation of the tool for the 3D registration of adult ovary images. V.T.

506 conceived the study, participated in data analyses and manuscript writing. All authors

507 reviewed drafts of the article and approved the final manuscript.

508 REFERENCES

- 509
- 510 **Charleston, J. S., Hansen, K. R., Thyer, A. C., Charleston, L. B., Gougeon, A., Siebert, J.**
511 **R., Soules, M. R. and Klein, N. A.** (2007). Estimating human ovarian non-growing follicle
512 number: the application of modern stereology techniques to an old problem†. *Human*
513 *Reproduction* **22**, 2103–2110.
- 514 **Eschweiler, D., Smith, R. S. and Stegmaier, J.** (2022). Robust 3D Cell Segmentation:
515 Extending the View of Cellpose.
- 516 **Fernandez, R. and Moisy, C.** (2021). FijiYama: a registration tool for 3D multimodal time-
517 lapse imaging. *Bioinformatics* **37**, 1482–1484.
- 518 **Fiorentino, G., Parrilli, A., Garagna, S. and Zuccotti, M.** (2021). Three-dimensional
519 imaging and reconstruction of the whole ovary and testis: a new frontier for the reproductive
520 scientist. *Molecular Human Reproduction* **27**, gaab007.
- 521 **Gay, S., Bugeon, J., Bouchareb, A., Henry, L., Montfort, J., Le Cam, A., Bobe, J. and**
522 **Thermes, V.** (2018). MicroRNA-202 (miR-202) controls female fecundity by regulating
523 medaka oogenesis.
- 524 **Gómez-de-Mariscal, E., García-López-de-Haro, C., Ouyang, W., Donati, L., Lundberg,**
525 **E., Unser, M., Muñoz-Barrutia, A. and Sage, D.** (2021). DeepImageJ: A user-friendly
526 environment to run deep learning models in ImageJ. *Nat Methods* **18**, 1192–1195.
- 527 **Goncharova, A. S., Honigmann, A., Jug, F. and Krull, A.** (2020). Improving Blind Spot
528 Denoising for Microscopy.
- 529 **Haase, R., Royer, L. A., Steinbach, P., Schmidt, D., Dibrov, A., Schmidt, U., Weigert,**
530 **M., Maghelli, N., Tomancak, P., Jug, F., et al.** (2020). CLIJ: GPU-accelerated image
531 processing for everyone. *Nat Methods* **17**, 5–6.
- 532 **İnik, Ö., Ceyhan, A., Balcioglu, E. and Ülker, E.** (2019). A new method for automatic
533 counting of ovarian follicles on whole slide histological images based on convolutional neural
534 network. *Computers in Biology and Medicine* **112**, 103350.
- 535 **Iwamatsu, T.** (2015). Growth of the Medaka (IV) - Dynamics of Oocytes in the Ovary
536 During Metamorphosis. *Bulletin of Aichi Univ. of Education* **64**, 37–46.
- 537 **Iwamatsu, T., Ohta, T., Oshima, E. and Sakai, N.** (1988). Oogenesis in the Medaka
538 *Oryzias latipes* : Stages of Oocyte Development : Developmental Biology. *Zoological Science*
539 **5**, 353–373.
- 540 **Iwamatsu, Takashi, T.** (1978). Studies on Oocyte Maturation of the Medaka, *Oryzias latipes*
541 VI. RELATIONSHIP BETWEEN THE CIRCADIAN CYCLE OF OOCYTE
542 MATURATION AND ACTIVITY OF THE PITUITARY GLAND. *J. Exp. Zool.* **206**, 355–
543 364.
- 544 **Kar, A., Petit, M., Refahi, Y., Cerutti, G., Godin, C. and Traas, J.** (2021). *Assessment of*
545 *deep learning algorithms for 3D instance segmentation of confocal image datasets.*
546 *Bioinformatics.*
- 547 **Krull, A., Buchholz, T.-O. and Jug, F.** (2019). Noise2Void - Learning Denoising from
548 Single Noisy Images. *arXiv:1811.10980 [cs]*.
- 549 **Lalit, M., Tomancak, P. and Jug, F.** Embedding-based instance segmentation in
550 microscopy. *PMLR* 399–415.
- 551 **Legland, D., Arganda-Carreras, I. and Andrey, P.** (2016). MorphoLibJ: integrated library
552 and plugins for mathematical morphology with ImageJ. *Bioinformatics* btw413.
- 553 **Lesage, M., Thomas, M., Bugeon, J., Branthonne, A., Gay, S., Cardona, E., Bobe, J. and**
554 **Thermes, V.** (2020). *C-Eci: A Cubic-Eci Combined Clearing Method For 3D Follicular*
555 *Content Analysis In The Fish Ovary.* Developmental Biology.
- 556 **Lubzens, E., Young, G., Bobe, J. and Cerdà, J.** (2010). Oogenesis in teleosts: how eggs are
557 formed. *Gen. Comp. Endocrinol.* **165**, 367–389.

558 **Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M. and Van Valen, D.** (2019). Deep
559 learning for cellular image analysis. *Nat Methods* **16**, 1233–1246.
560 **Murtin, C. I.** (2016). Three-dimensional image analysis of high resolution confocal
561 microscopy data of the *Drosophila melanogaster* brain. *Image Processing [eess.IV]*.
562 **Université de Lyon.**, 1–166.
563 **Nakamura, Y. T.** (2018). All Oocytes Attach to the Dorsal Ovarian Epithelium in the Ovary
564 of Medaka, *Oryzias latipes*. *Zoological Science* **35**, 306–313.
565 **Nakamura, S., Kurokawa, H., Asakawa, S., Shimizu, N. and Tanaka, M.** (2009). Two
566 distinct types of theca cells in the medaka gonad: germ cell-dependent maintenance of
567 *cyp19a1*-expressing theca cells. *Dev. Dyn.* **238**, 2652–2657.
568 **Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T. and**
569 **Zuiderveld, K.** (1987). Adaptive Histogram Equalization and Its Variations. *COMPUTER*
570 *VISION, GRAPHICS, AND IMAGE PROCESSING* 355–368.
571 **Renier, N., Wu, Z., Simon, D. J., Yang, J., Ariel, P. and Tessier-Lavigne, M.** (2014).
572 iDISCO: A Simple, Rapid Method to Immunolabel Large Tissue Samples for Volume
573 Imaging. *Cell* **159**, 896–910.
574 **Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,**
575 **Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al.** (2012). Fiji: an open-source
576 platform for biological-image analysis. *Nature Methods* **9**, 676–682.
577 **Sonigo, C., Jankowski, S., Yoo, O., Trassard, O., Bousquet, N., Grynberg, M., Beau, I.**
578 **and Binart, N.** (2018). High-throughput ovarian follicle counting by an innovative deep
579 learning approach. *Sci Rep* **8**, 13499.
580 **Soygur, B. and Laird, D. J.** (2021). Ovary Development: Insights From a Three-
581 Dimensional Imaging Revolution. *Front. Cell Dev. Biol.* **9**, 698315.
582 **Stringer, C., Wang, T., Michaelos, M. and Pachitariu, M.** (2021). Cellpose: a generalist
583 algorithm for cellular segmentation. *Nat Methods* **18**, 100–106.
584 **von Chamier, L., Laine, R. F., Jukkala, J., Spahn, C., Krentzel, D., Nehme, E., Lerche,**
585 **M., Hernández-Pérez, S., Mattila, P. K., Karinou, E., et al.** (2021). Democratising deep
586 learning for microscopy with ZeroCostDL4Mic. *Nat Commun* **12**, 2276.
587 **Waisman, A., Norris, A. M., Elías Costa, M. and Kopinke, D.** (2021). Automatic and
588 unbiased segmentation and quantification of myofibers in skeletal muscle. *Sci Rep* **11**, 11793.
589 **Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B.,**
590 **Schmidt, D., Broaddus, C., Culley, S., et al.** (2018). Content-aware image restoration:
591 pushing the limits of fluorescence microscopy. *Nature Methods* **15**, 1090–1097.
592 **Zuiderveld, K.** (1994). Contrast Limited Adaptive Histogram Equalization. In *Graphics*
593 *Gems*, pp. 474–485. Elsevier.
594
595

596 **FIGURE LEGENDS**

597 **Figure 1: Pipeline overview for 3D image analysis of the whole ovary at larvae and** 598 **adult stages.**

599 (A) Fluorescent staining strategies for whole ovary imaging. Cytoplasmic
600 autofluorescence from Histone H3 phosphorylation immunofluorescence (PH3) is used
601 for larvae (left panel). Methyl-green (MG) nuclear staining delineating follicles contour is
602 used for adult stage (right panel). Z-projection of raw data are shown (standard deviation
603 method. Raw stacks sizes are indicated. (B) Image pre-processing steps used for
604 reconstruction and enhancement, listed from top to bottom. (C) 3D segmentation step is
605 performed with Cellpose algorithm. Adult images are subjected to several segmentations
606 runs before and after image downscaling. (D) Image post-processing is performed for
607 segmentation correction, label filtering and final quantitative analysis. Opensource tools
608 are indicated in white boxes, deep-learning opensource tools in blue, commercially
609 available software in red (AMIRA). Relative computation time for one sample is indicated
610 (B-D). Voxel size is indicated in brackets (A-C). Scale bars 200 μm (for larvae), 1000 μm
611 (for adult), 100 μm (for larvae inset), 500 μm (for adult inset)

612

613 **Figure 2: 3D reconstruction of whole medaka ovaries.**

614 (A) Larvae ovary reconstruction with raw data. Total ovary size approach 1240 μm in
615 length (x), 630 μm in width (y) and 515 μm in height (z). XY plane at 188 μm and XZ plane
616 are shown with dotted lines. (B) XY plane showing PH3 staining and cytoplasmic
617 background in oocytes at 188 μm in depth, magnified in inset (D). (C) XZ orthoslice of
618 larvae PH3 staining. A decrease in fluorescence intensity is observable near 400 μm in
619 depth. (E) Adult ovary reconstruction with raw data. Only back stack (dorsal face) is
620 shown, with a size of 7645 μm in length (x), 6033 μm in width (y) and 2778 μm in height

621 (z). XY and YZ virtual slices are shown with dotted lines. (F) XY plane at 1152 μ m depth
622 shows MG staining resulting in delimited follicular contours, magnified in inset (H). (G)
623 YZ orthoslice of dorsal face of adult ovary. Heterogeneity of MG staining is observable
624 through depth. Scale bars 200 μ m (B, C), 50 μ m (D), 1000 μ m (F, G), 40 μ m (H).

625

626

627 **Figure 3: Image pre-processing for enhancement of features detection through**
628 **larva ovary depth.**

629 (A) Representative 3D reconstruction of 20 dph (days post-hatching) larvae ovary before
630 image processing. (B-E) Effect of successive image processing steps at 150 μ m depth and
631 (B'-E') at 400 μ m depth assessed on XY cropped planes. A profile line intensity is used to
632 assess fluorescence intensities nearby relevant objects to be segmented. Fluorescence
633 intensity and signal to noise ratio are progressively enhanced. (F) 3D reconstruction of
634 20 dph larvae after image pre-processing showing signal homogenization. Color gradient
635 is representative of grey levels (1-255). Scale bar 100 μ m, Grid square size 50 μ m.

636

637 **Figure 4: Image pre-processing for enhancement of features detection through**
638 **adult ovary depth.**

639 (A) Volume reconstruction of front and back adult z-stacks before image processing. (B-
640 E) Effect of successive image processing steps at 780 μ m deep and (B'-E') at 2000 μ m deep
641 assessed on XY cropped planes extracted on front stack. A profile line intensity is used to
642 assess fluorescence intensities nearby relevant objects to be segmented. Fluorescence
643 intensity loss in depth is greatly recovered and resolution of follicles contours is
644 improved. (L) Final adult ovary reconstruction after 3D registration and image

645 enhancements. Color gradient representative of grey levels. Scale bar 300 μm , Grid square
646 size 500 μm .

647

648

649 **Figure 5: Effect of image pre-processing on Cellpose 3D segmentation efficiency**

650 (A) XZ orthoslice of larvae ovary showing raw data and (C) image data after pre-
651 processing, magnified on insets. (B) Cellpose segmentation output using raw data as input
652 or (D) pre-processed image data. Results are shown after label erosion to correctly
653 visualize labels individualization. Insets show more segmentation errors without image
654 pre-processing, including unsegregated labels (arrows) or over-segmentation
655 (arrowhead). (E) XY plane of adult ovary on raw data and (G) after image processing,
656 magnified on insets. (F) Cellpose segmentation output for 30 pixel diameter using either
657 raw data input or (H) pre-processed image. Segmentation results show high error number
658 without image pre-processing, including many missing labels (arrowhead) especially in
659 locations with heterogenous staining and high over-segmentation in medium and large
660 size follicles. Scale bars: 200 μm (A-D), 100 μm (insets in A-D), 1000 μm (E-H), 500 μm
661 (insets in E-H).

662

663 **Figure 6: Post-processing corrections of segmented labels.**

664 (A, top) XY planes of 20 dph larvae ovary at $\sim 330 \mu\text{m}$ depth showing pre-treated image
665 data, Cellpose output, labels boundaries, eroded labels (subtracted boundaries) and final
666 results after semi-automatic label filtration in AMIRA superimposed onto image data. (A,
667 bottom) YZ planes for qualitative assessment of segmentation along Z-axis. Segmentation
668 results after post-processing show good accuracy and shape fitting in XY or YZ planes with
669 only few labels fusion for smallest oocytes (arrows) or missing label (arrowhead). (B, top)

670 XY planes of adult ovary at ~2500 μm depth before and after Cellpose 30 segmentation.
671 (B, bottom) YZ planes of adult ovary showing segmentation accuracy along Z axis. Largest
672 follicles are over-segmented but other labels are correctly fitting follicles size. (C)
673 Downscaled image data, Cellpose 30 segmentation results and post-processed images of
674 adult data on XY plane (top) or YZ orthoslices (bottom). Large labels (stars) are combined
675 to original scale segmentation result in (B) to replace segmentation errors of largest
676 follicles. (D) Combined and filtered adult segmentation data superimposed onto image
677 data. Labels show a good fit in size and shape of follicles at various sizes either on XY or
678 YZ plane. Scale bars 100 μm (larvae panels), 1000 μm (adult panels).

679

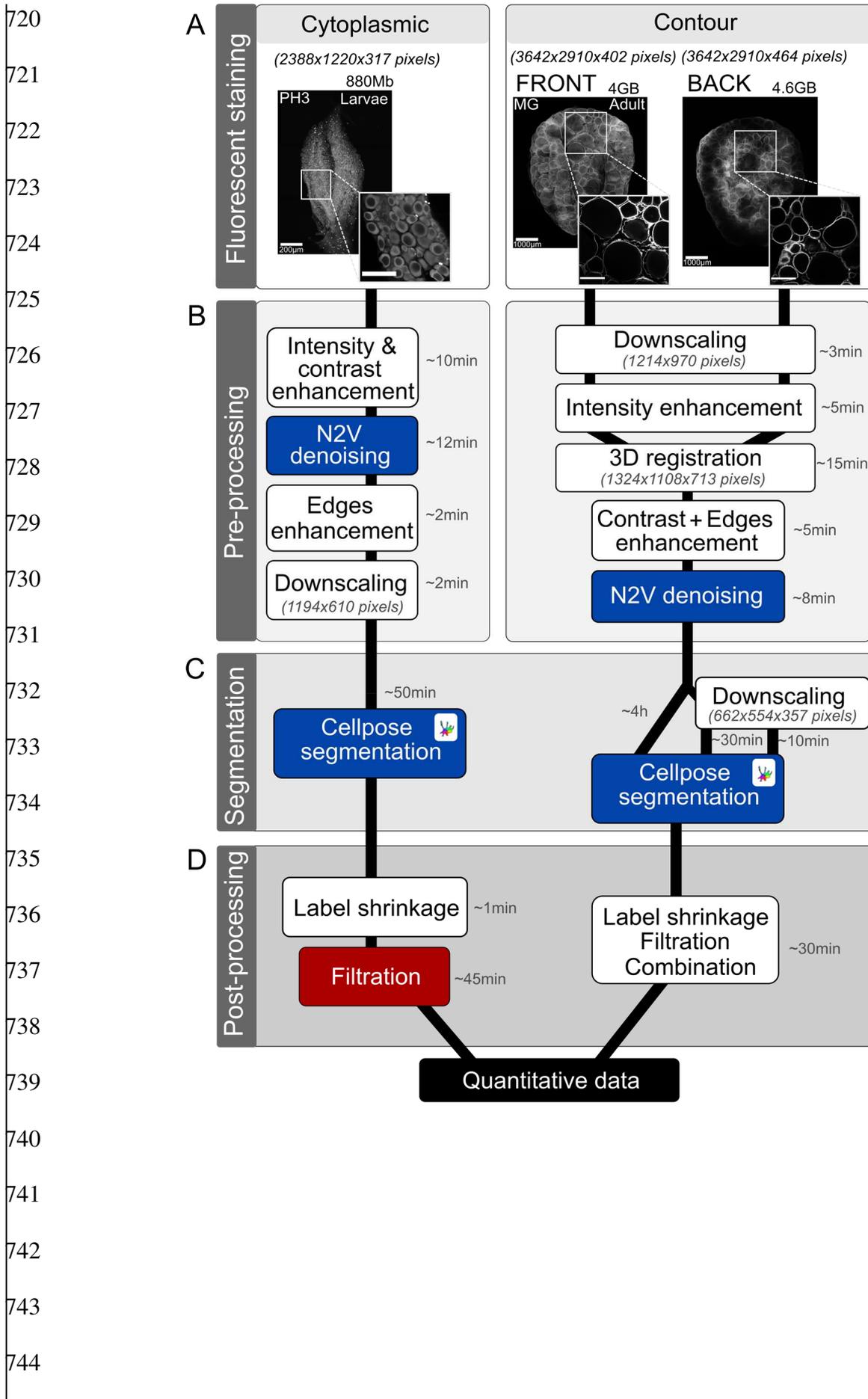
680

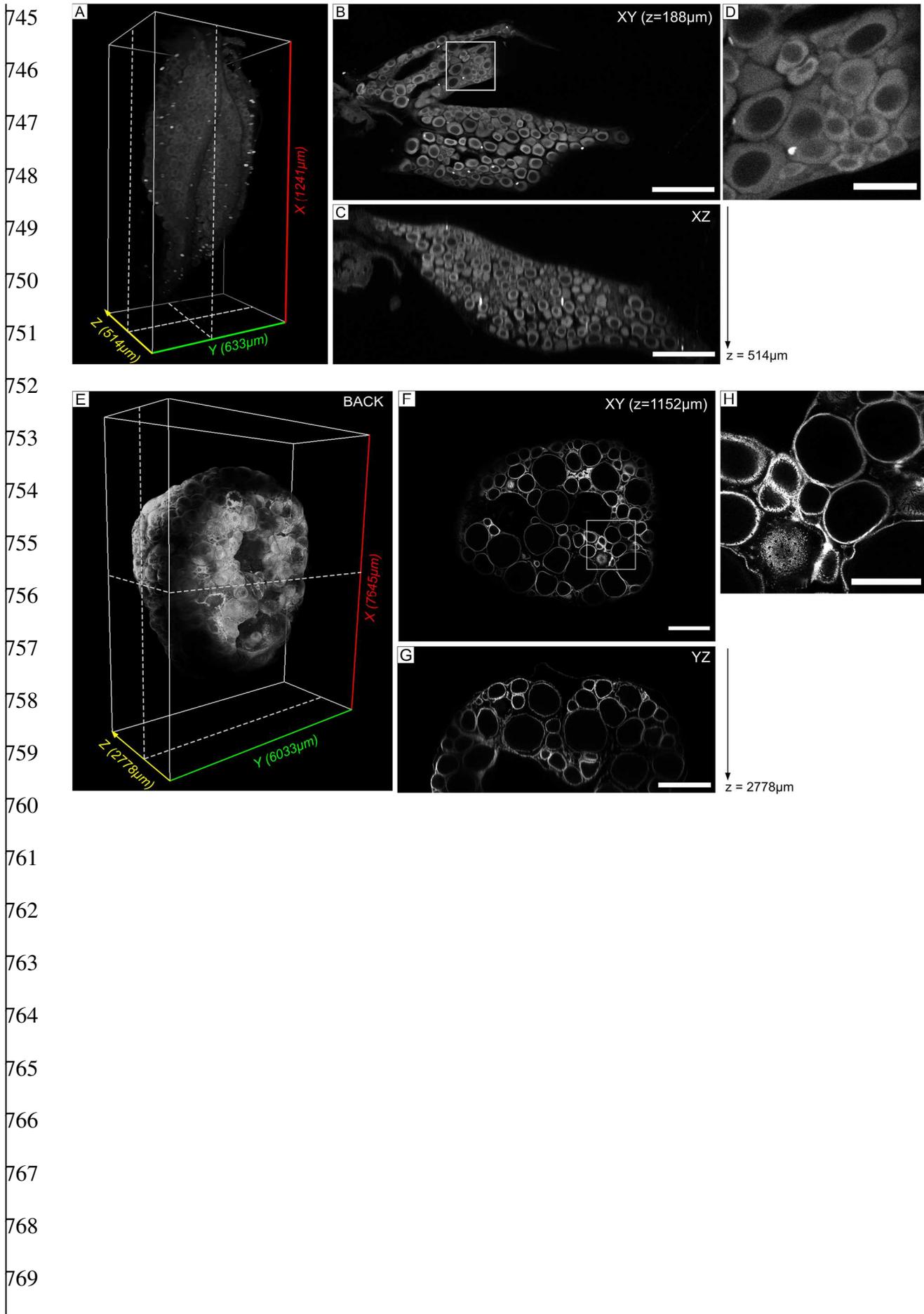
681 **Figure 7: 3D Qualitative spatial visualization and quantitative analysis of ovarian**
682 **content.**

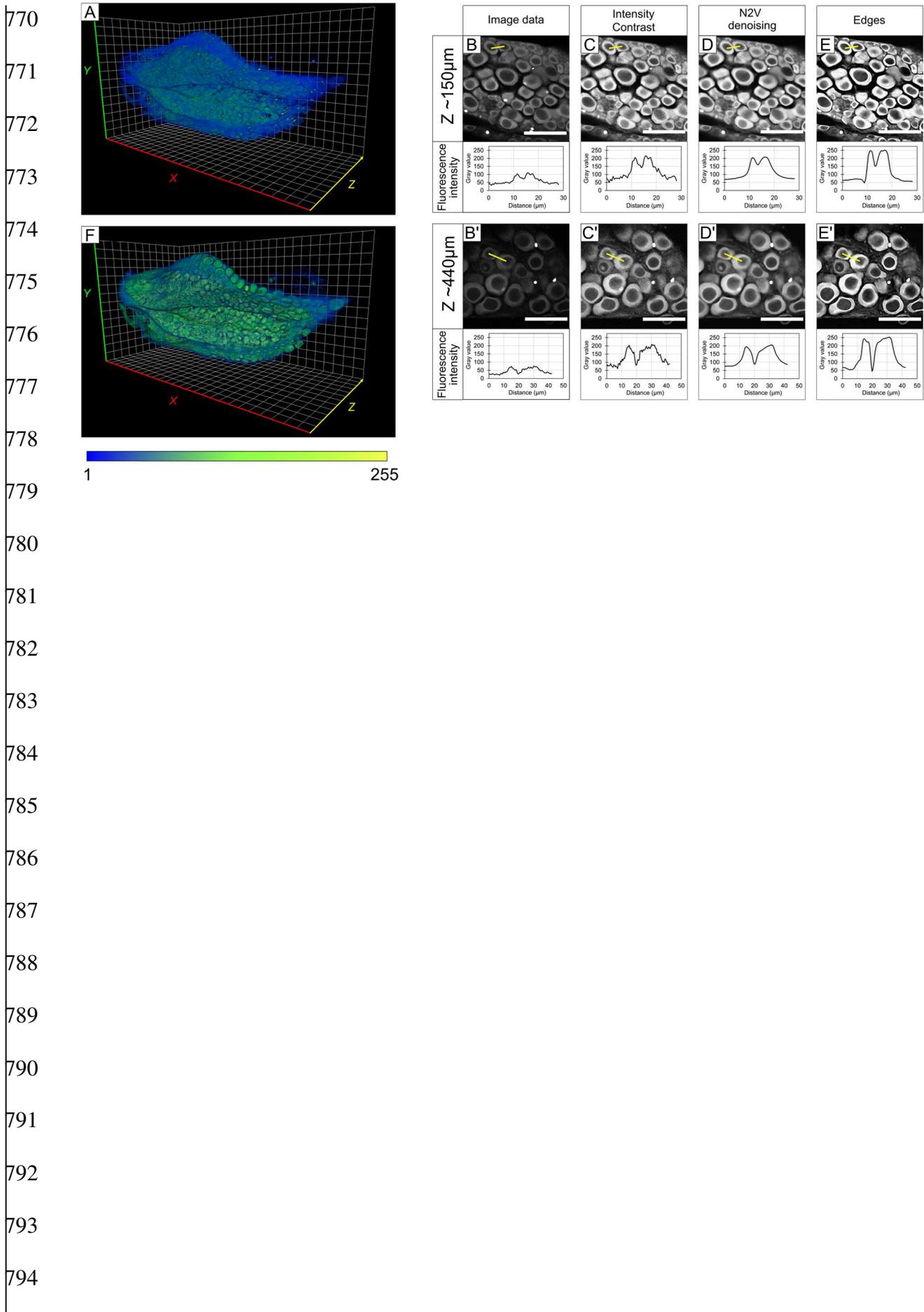
683 (A) 3D ventral and dorsal reconstruction views of 20 dph larvae ovary, and (A') merged
684 with segmented oocytes. Ovary size is approximately represented on bounding box
685 measuring 1240 μm in length (x, yellow), 630 μm in width (y, green) and 515 μm in height
686 (z, red). (B) Oocytes spatial distribution visualized by diameter range from ventral and
687 dorsal side of larvae ovary. Oocytes tend to localize dorsally through their growth. (C)
688 Oocyte distribution in entire larvae ovaries (mean \pm SD, n=2) depending on their
689 equivalent diameter. Diameter measure cut-off was applied at 25 μm . Corresponding
690 developmental stages of previtellogenesis are indicated: stage I, chromatin-nucleolar (25-
691 60 μm) and stage II, perinucleolar (60-90 μm). (D) 3D ventral and dorsal views of entire
692 adult ovary after registration and reconstruction, and (D') merged with 3D segmented
693 follicles. Bounding box size approximates whole ovary size with an antero-posterior
694 length of 7 mm (x, yellow), left to right width of 6 mm (y, green) and depth of 4,25 mm (z,

695 red). (E) Follicles spatial distribution within ovary based on their equivalent diameter
696 range. Follicle size classes show respective localization of various developmental stages,
697 namely previtellogenesis (50-150 μm), vitellogenesis (150-800 μm) and post-
698 vitellogenesis (>800 μm). (F) Total quantification of adult ovarian content distributed by
699 follicular diameter. Stages of development are listed, namely previtellogenesis (II-IV, 50-
700 150 μm), early vitellogenesis (V-VI, 150-400 μm), late vitellogenesis (VII-VIII, 400-800
701 μm) and post-vitellogenesis (IX, 800 μm and over). Oocyte and follicular distribution are
702 expressed as percentage of total objects counted within the ovaries.

703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719







795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819

