U-net segmentation of tubular structures from live imaging confocal microscopy: Successes and challenges

Kasra Arnavaz, Pia Nyeng, Jelena Miskovic Krivokapic, Oswin Krause, Aasa Feragen

Summary

We study the problem of segmenting tubular structures in the pancreas from live imaging via confocal microscopy. The segmentation is difficult due to noise, making it challenging even for a human expert to distinguish structure from background. We present the results of the state-of-the-art U-Net and inspect its successes and failures. We discuss challenges in the manual segmentation of the structure, the impact of these challenges on our choice of experimental design, and future directions for improving the segmentation.
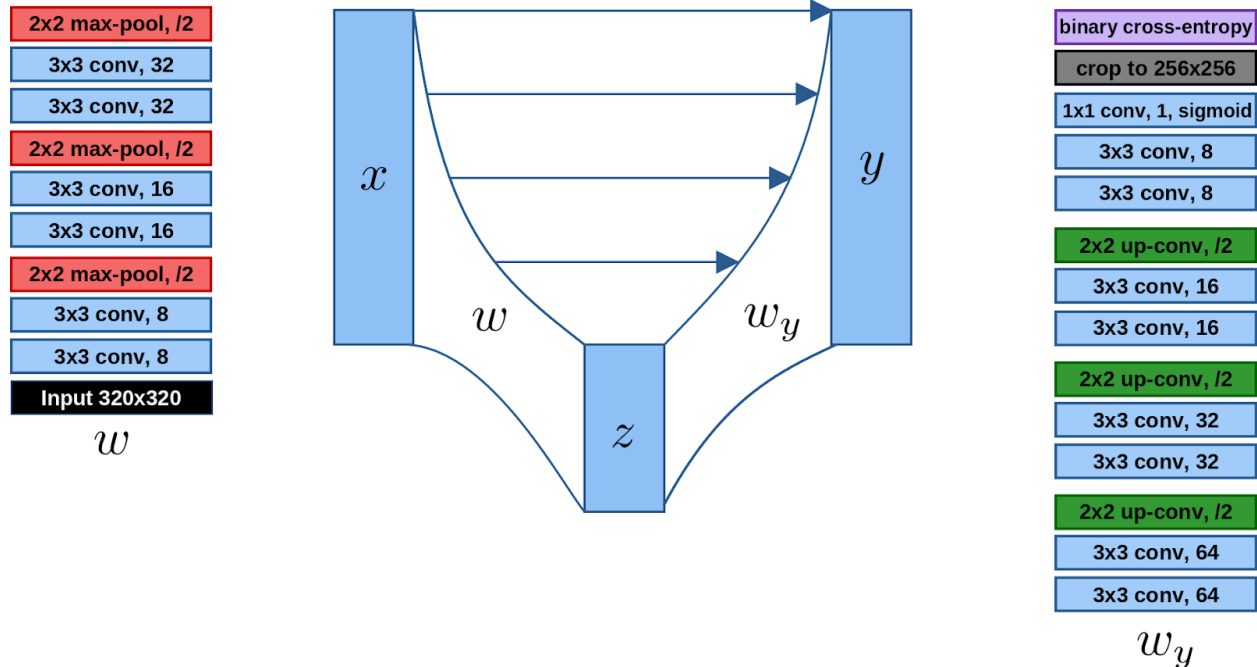
Introduction

During embryonic development, the pancreatic tubes (Figure 1) grow from a web-like network with loops to a tree-like structure without loops. A thorough understanding of the mechanisms governing this morphology has yet defied developmental biologists (Shih et al. 2013), but observations indicate that the topological events happening during development of the tubular structure may be related to differentiation of beta cells nearby. However, confirming this quantitatively requires a binary segmentation of the tubes.

2D microscopy slice with tubular structure | Manual annotation of tubular structure
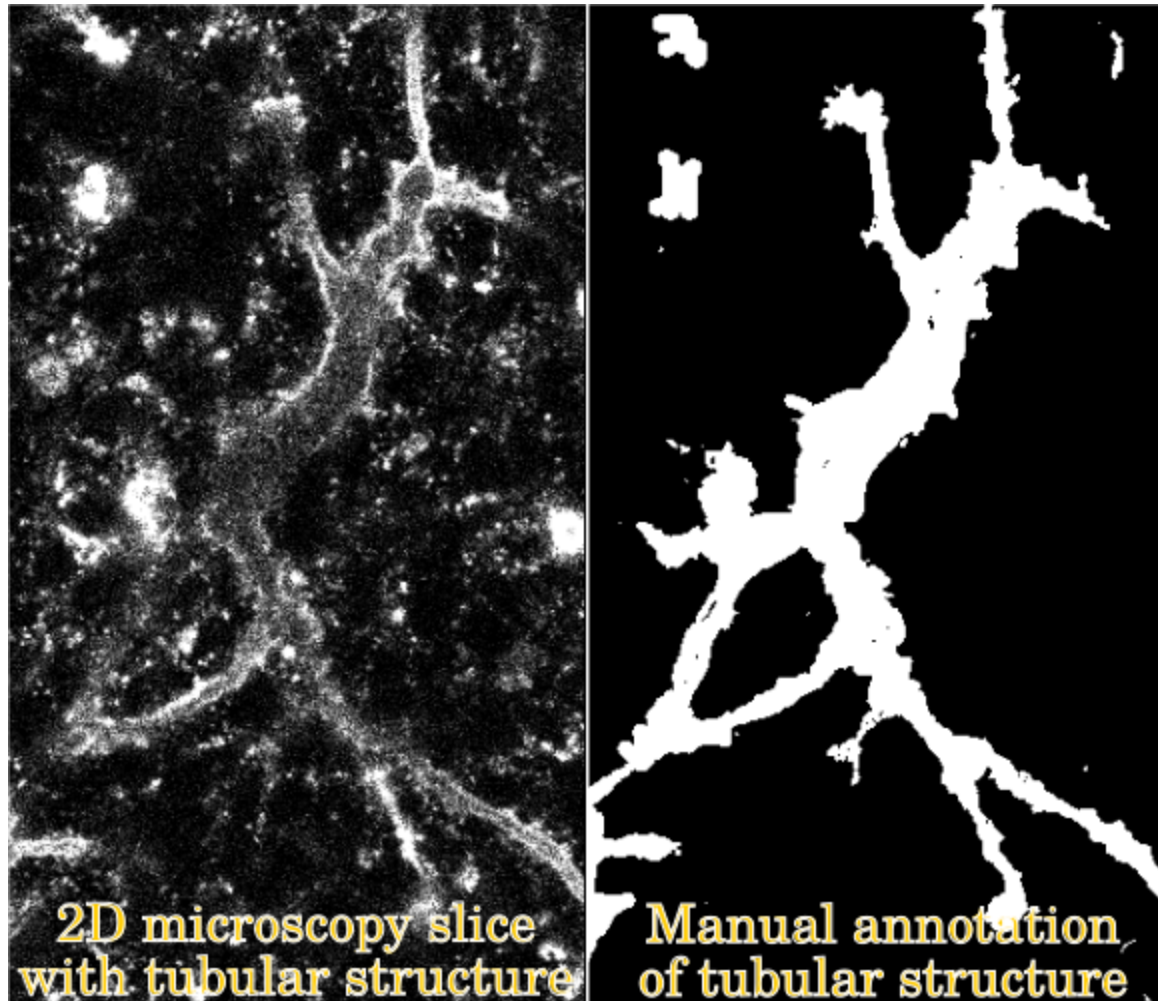
While segmentation of other tubular structures such as airways or vessels is well studied, the pancreatic tubes are different in the sense that we have no prior anatomical knowledge about their structure. Moreover, the images are recorded with low signal to noise ratio, creating a challenging segmentation task. We present results of the state-of-the-art U-net (Ronneberger et al, 2015) and discuss its successes and failures. In particular, we discuss challenges in the manual labeling of training data, as well as potential strategies for "optimally" labeling data.

Methods/Materials

We implement the U-net (Figure 2) in Keras, using 3 sets of down- and up-sampling layers and binary cross entropy loss, and train it using the Adam Optimizer (beta_1 = 0.9, beta_2 = 0.999, lr = 10^{-5}) for 50 epochs.

w

2x2 max-pool, /2
3x3 conv, 32
3x3 conv, 32
2x2 max-pool, /2
3x3 conv, 16
3x3 conv, 16
2x2 max-pool, /2
3x3 conv, 8
3x3 conv, 8
Input 320x320

w

x w y w_y z

w_y

binary cross-entropy
crop to 256x256
1x1 conv, 1, sigmoid
3x3 conv, 8
3x3 conv, 8
2x2 up-conv, /2
3x3 conv, 16
3x3 conv, 16
2x2 up-conv, /2
3x3 conv, 32
3x3 conv, 32
2x2 up-conv, /2
3x3 conv, 64
3x3 conv, 64

w_y

Our dataset consists of 2D image slices extracted from 3D image stacks obtained from live imaging using confocal microscopy of the developing pancreas of mouse embryos imaged for 40-48 hours in vitro (see abstract submission from P Nyeng).

The 2D slices are aggregated as follows: A total of 6 3D movies with approximately 220 time-points each are split randomly into two sets with 3 movies each, to be used for training and testing, respectively. As parts of the movies have very high noise levels with no discernible structure, a smaller set of images is manually subselected by first randomly sampling, from each movie, two 3D images at random. These images have a resolution of 1024 x 1024 x D, where D varies between 27 and 40. Each of these images is divided into a 4 x 4 grid of patches in the x-y plane, each of size 256 x 256 x D. From this set of patches, a trained laboratory assistant manually removes all patches which are too dark and noisy to be reliably annotated. The remaining patches typically form connected parts of the x-y plane, leading to 1024 x 1024 x D images with ``forbidden regions'' consisting of the removed patches.

Manual annotation: The training split consists of 6 3D images from 3 different movies with in total 174 2D image slices, where the tubular structure is fully annotated in a semi-manial fashion as follows: The 3D images were first processed in Imaris image processing software using Z-normalization, smoothing, volume rendering and filtering of the Muc1-mcherry signal. The resulting segmentations are manually edited by the laboratory assistant using the software Slicer. To validate on as diverse a dataset as possible, we select from the test split 10 3D patches at random, in which the tubular structure is annotated by the laboratory assistant, giving annotations for around 13\% of the test split.

Image preprocessing: While the mean intensity per movie is close to zero, the variance differs noticeably. Thus, every time-point is stadardized by its own mean and standard deviation, and intensities are clipped to be at most 3 standard deviations away from the mean (which is zero).

Training on patches: The U-net is trained on image patches from the fully annotated slices in the training set, avoiding the forbidden regions.

Results and Discussion

On the test set, the U-net segmentation produces:
- a sensitivity, giving the proportion of true foreground segmented pixels out of ground truth foreground pixels, of 0.648;
- a specificity/true negative rate giving the proportion of true background segmented pixels out of ground truth background pixels, of 0.987
- An accuracy, giving the proportion of correctly segmented pixels over all pixels, of 0.970
- A dice score, giving a similarity between the predicted segmentation and the ground truth segmentation, of 0.687 (optimal score is 1).

Note that the accuracy, which is often used to assess quality of segmentation, is high due to the large imbalance between foreground and background, where much of the dominant background class is easy to predict. As shown by the sensitivity and specificity, however, the U-net is undersegmenting the foreground.

This is problematic for two reasons. First, the expert annotation is already likely an undersegmentation, due to the fact that the expert annotator is unlikely to annotate foreground unless she is fairly confident. Second, undersegmentation easily leads to disruptions in the topology of the segmented tubular network, as a few missing pixels can be enough to cut a connected tube into two or more pieces.

| Original image | U-net segmentation | Expert annotation |
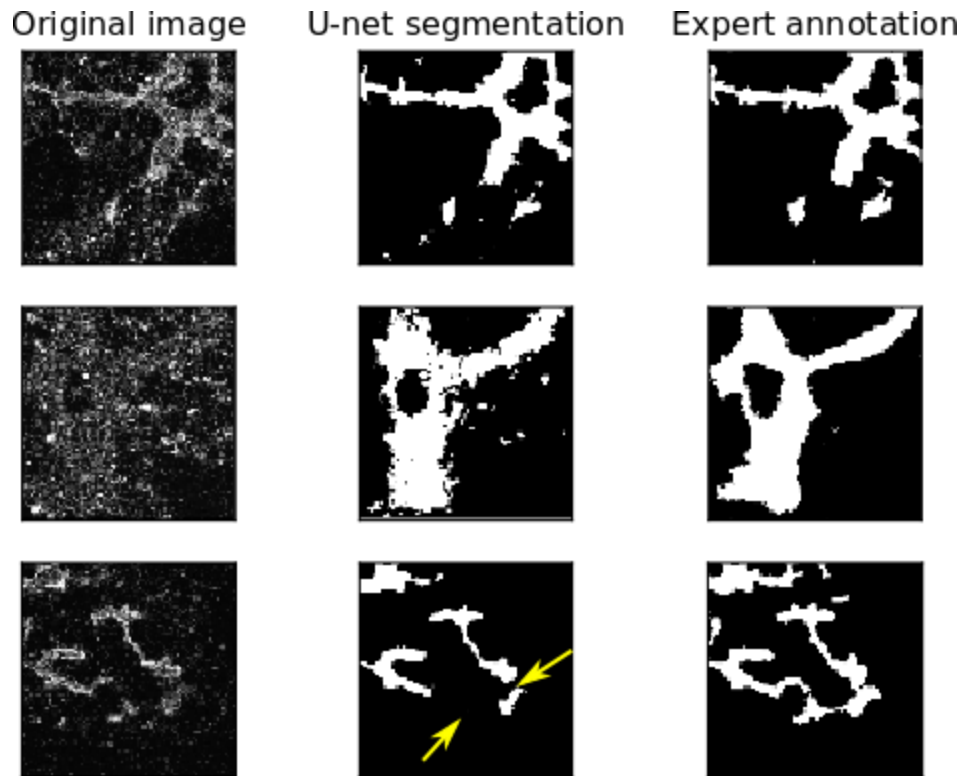
Conclusion

We have demonstrated how a U-net trained on patches from a limited set of 2D slices from a limited set of 4D microscopy live images obtains a qualitatively reasonable performance. In light of the limited amount of training data, with limited amount of variation, the results are positive, although they also present challenges: The U-net is shown quantitatively to undersegment the tubular structure, which raises some concern regarding the correctness of segmented network topology.

In theory, we expect that this could be handled by increasing the amount of hand-annotated training data. In practice, however, manual labeling is expensive, difficult and time consuming. This thus leads to multiple future research directions, including semisupervised learning to take advantage of our large amounts of unlabeled data, and including spatial and time information into the segmentation pipeline. Finally, we note that the segmentation problem is callenging, also for a human, who is able (more so than a neural network) to take contextual information into account. A final potential research problem is thus to specialize the network to search for tubular structures.

References

HP Shih, A Wang, M Sander, Annu Rev Cell Dev Biol 29 (2013), p.81-105
O Ronneberger, P Fischer, T Brox, MICCAI (2015), p.234-241

Submission Guidelines
  • Abstracts must contain original data and meet international ethical standards.
  • Authors may submit abstracts for either oral or poster presentations.
  • Abstracts are to be limited to a minimum of 100 words and a maximum of 1000 words including spaces .You may include supporting figures and tables and your abstract must contain adequate information to allow a sound referee review.
  • Abstracts should be a condensed version of the final presentation and include all significant findings. Write the text so that readers who are not specialists will be able to appreciate the purpose of the study and understand the procedures and conclusions. The text, entirely written in English, must include a short introduction and motivation of the study, including experimental procedures, main results and conclusions.
  • Abstracts should be structured as follows (please note you do not need to use the headings below in you abstract text):
  ◦ Summary - Brief description of the project/case study.
  ◦ Introduction - Includes the background, aim and/or objectives, and importance of the study.
  ◦ Methods/Materials - Summarises the approach, procedures or steps used in the study.
  ◦ Results and Discussion - Describe the observations or data in an orderly and logical sequence using both text and illustrative materials such as tables, figures. The discussion section should interpret the results and connect the key points or findings of the study with the objectives stated in the introduction.
  ◦ Conclusion - highlight the significant findings and outcomes.
  • All sections on this form should be completed:
  ◦ Abstract title.
  ◦ All authors' and co-authors' names and affiliations.
  ◦ Contact information (name, designation, email address, postal address, and phone and fax numbers) of corresponding author(s).
  ◦ Type of presentation (oral or poster presentation).
  ◦ Theme or sub-topic area.
  ◦ Detailed information with introduction, methods, results, and conclusions.
  • Abstracts will be reviewed according to the following criteria: (a) relevance to a specific symposium or keyword, (b) scientific content, quality and innovative proposals, (c) clarity of the text.
  • Accepted non-invited abstracts must be presented by an author who is registered and has paid to attend the conference. An exhibition visitor registration will not be considered a valid registration to present your work and failure to register for the conference prior to its start, will result in your work being excluded from the Congress.
  • Any major revisions may affect a symposium organiser's decision to accept your abstract, please review it carefully before submission.
  • You may express your preference about the form of presentation, however the Scientific Organising Committee reserves the right to make the final decision.
  • Abstracts will appear on the online programme on the emc2020 website. If you do not wish to have your abstract displayed, please email katejermey@rms.org.uk.

Abstract text *
Guidelines for Abstract Submission
Formatting:
  Main body of your abstract (1000 words)
        • If pasting copied text into the text box, please note that font type and font size will be automatically adjusted by the system. You do not need to enter the text in a particular font or font size.
        • If using a figure, be sure to include figure labels and scale markers. Colour figures may be used and will be reproduced as such in the electronic proceedings.
        • Figure captions should be placed below figures and tables. Figure captions should be formatted with the figure label in bold and followed by a concise and accurate description of the figure shown. e.g. Figure 3. Image comparison on a rough surface of sintered stainless steel.
If the figure shows more than one image, these should be listed under one figure heading and separated by alphabetical characters. e.g. (a), (b), (c).
        • To insert an image, please use the 'insert image ' at the appropriate point in your text. When inserting several images next to each other, please make sure your insert a space in the text field.
        • The required format for any images is jpeg, tiff or gif files (not to exceed 500KB). Please note that any files uploaded in a .doc (Word), .ppt (PowerPoint), .xls (Excel) or any other format not mentioned above, will not be accepted.
        • Please take into consideration the pixel width and height of your image in relation to your abstract text. Pictures that are too large will affect the presentation of your abstract on the online programme.