COMP 2200/6200 Data Science Unit

machine learning engineering in data science

Guest Lecturer
Dr Donghao Zhang

ABOUT ME

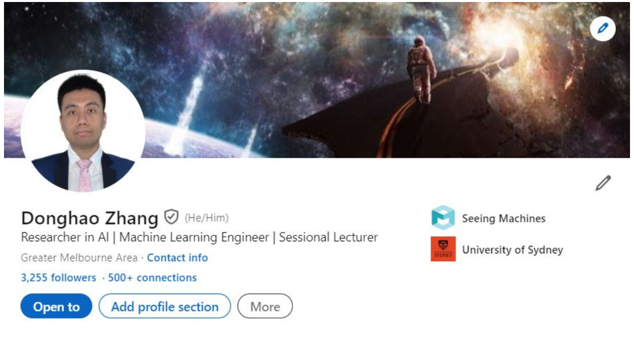
Education Background

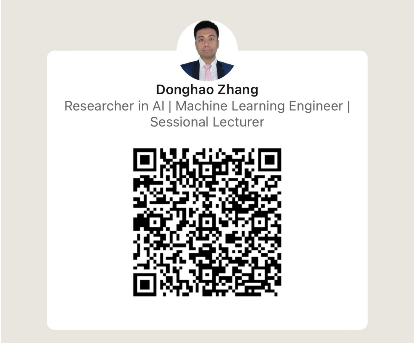
- PhD, Mphil in computer science, The University of Sydney, 2015-2020/09
- Bachelor with honours: Australian National University 2011-2014

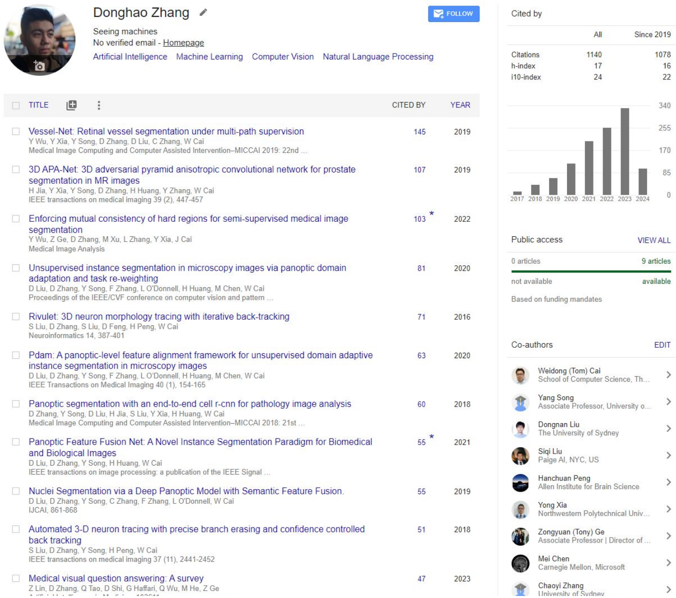
Work Experience (Selected)

- Machine Learning Scientist at Seeing Machines July 2023 - present
- Sessional Lecturer at La Trobe University 2022/09 - July 2023
- Postdoctoral Research Fellow (**Level B**), Monash University, Melbourne, 2020/09-2022/09
- Machine learning Engineer (**Intern**) Sydney Neuroimaging Analysis Centre
- Research Scientist (**Intern**) Siemens Healthineer, USA
- Tutors and Teaching Assistant at USYD

Happy to be connected







Project (Selected)

- Medical Report Generation
- Agilent Project
- Computational Histopathology
- Curvilinear Morphology Reconstruction
- Biomedical Image Segmentation

Publications Summary

- Top-tier conference and journal papers in biomedical image processing, computer vision, machine learning and AI (CVPR, IJCAI, AAAI, ISBI, MICCAI, TMI, MIA, and Neuroinformatics).
- H-index: 17 with 1140 citation

Medical Report Generation

MACQUARIE University

Significance:

- Challenging to control the reports' qualities due to the experience variations of medical professionals

- Reduce the labor-intensive workload observing medical images and typing findings of diseases and lesions into the computer

Motivation:

- Conventional scaled-dot product attention in the Transformer maps three input vectors: query, key, and value to the weighted sum of values but ignored higher-order interaction of query $=>$ **requiring better language and vision fusion**

- object location variations in natural images result in a notable region of interest, while majorities of diseases and lesions occupy relatively small regions. $=>$ **requiring better language and vision fusion**

- Taking advantage of the unlabeled data is a potential solution for image encoder pre-training

- pre-training the image encoder from the same domain can improve the performance

Medical Report Generation

MACQUARIE University

Contribution:

- the novel weighted query-key interacting linear attention module to increase the capability of expressing a complex multi-modal relationship between the visual feature space and the semantic feature space

- the first to introduce contrastive pre-training to the medical report generation

- collected and processed Retina ImBank and Retina Chinese, which will be released to serve as benchmarking datasets to encourage further research on generating reports with retina images

- evaluated on two retina datasets and two Chest X-ray datasets, the proposed method achieved state-of-the-art performances in majorities of natural language evaluation metrics

Medical Report Generation

MACQUARIE University

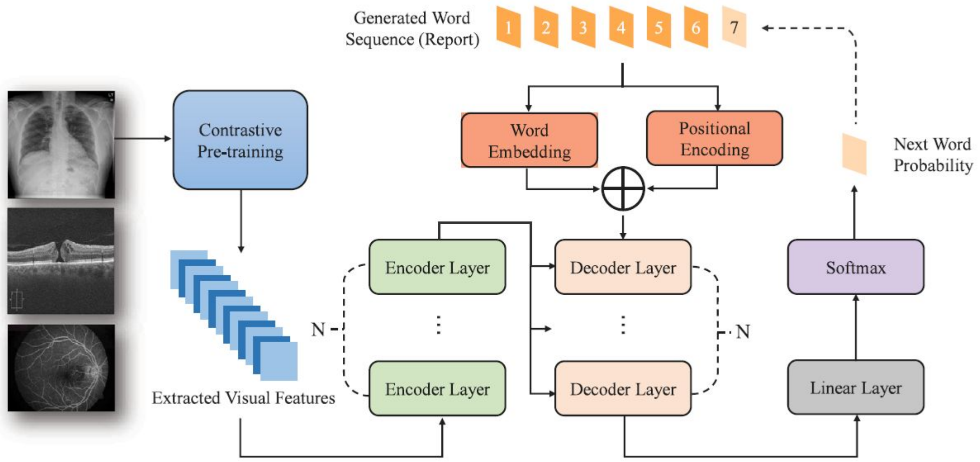
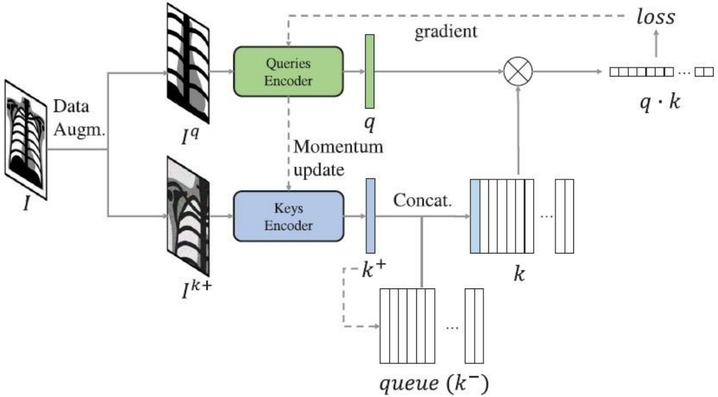


Illustration of the overall architecture. + denotes the add operation, and positional encoding introduces relative location information of the individual feature token in the whole sequence

Medical Report Generation
- Contrastive Pretraining



The detailed MoCo (v2) framework. The *Augm.* stands for *augmentation*. The *Concat.* stands for *concatenation*. The solid arrow lines mean the forward operations, and the dashed arrow lines mean the backward operations.

InfoNCE loss is computed as:

$$L_{q,k^+,\{k^-\}} = -\log \frac{\exp(q \cdot k^+/\tau)}{\exp(q \cdot k^+/\tau) + \sum_{k^-} \exp(q \cdot k^-/\tau))}$$

where q is a query representation,
k+ is positive (similar) key representation,
{k−} are a set of negative (dissimilar) key representations and
τ is a hyperparameter of temperature.

Medical Report Generation
- Contrastive Pretraining

MACQUARIE University

Algorithm of MOCO

- Generate a positive pair by sampling data augmentation functions
- No gradient to positive keys
- Use the running queue of keys as the negative samples
- Compute the InfoNCE loss
- Update the query encoder using momentum f_k.params = m*f_k.params+(1-m)*f_q.params m
- Update the FIFO negative sample queue with k the positive key (k+) is pushed into the queue and replaced with the earliest key in a FIFO (first-in, first-out) manner

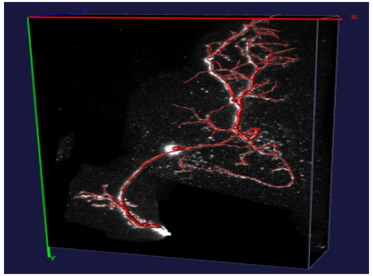
Improvement of MOCOv2

- Non-linear projection head and strong data augmentation

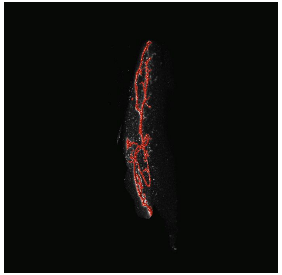
Curvilinear Morphology Reconstruction

-Automatic Neuron Reconstruction

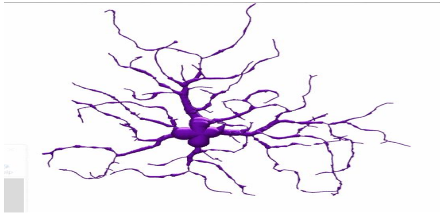
- **Significance**:
  - Easy to generalise large datasets
  - Reduce the manual time-consuming reconstruction
- **Contribution**:
  - Propose a novel and accurate back-tracking method based on the gradient of time-crossing map
  - Back-tracking with sub-voxel gradient interpolation
  - Longest branch is found iteratively
  - Terminates at previous traced branch or soma location



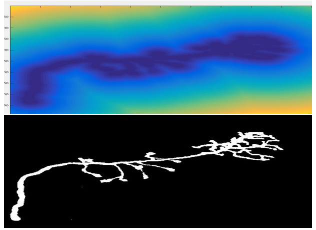
Ground Truth



Proposed



Back-tracking



Time-crossing Map

Curvilinear Morphology Reconstruction

-Liver Vessel Reconstruction

Method

- Train a vessel enhancement CNN
- Initialize the vessel graph tracing with high sensitivity and low specificity
- Graph attention network with graph attention layers to estimate the confidence of sub-branch in the initial reconstruction
- We proposed a graph attention network with input graph nodes interpolating the CNN features

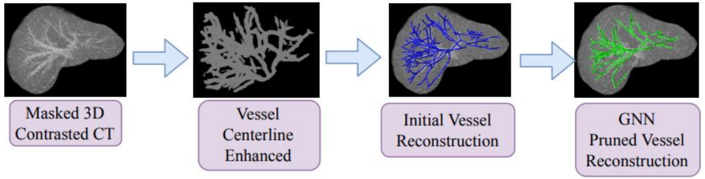


Fig: The illustration of the overall liver vessel reconstruction framework.

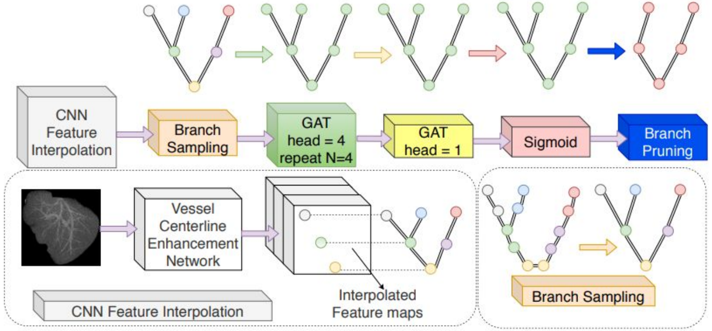
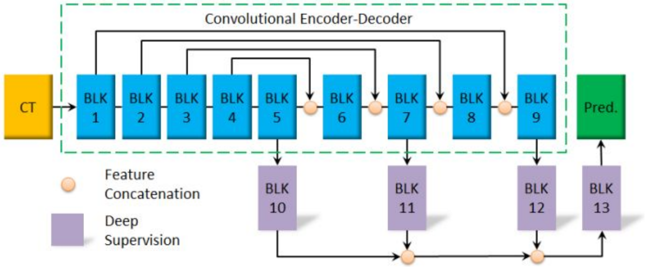


Fig: The illustration of the overall liver vessel reconstruction framework.

Curvilinear Morphology Reconstruction

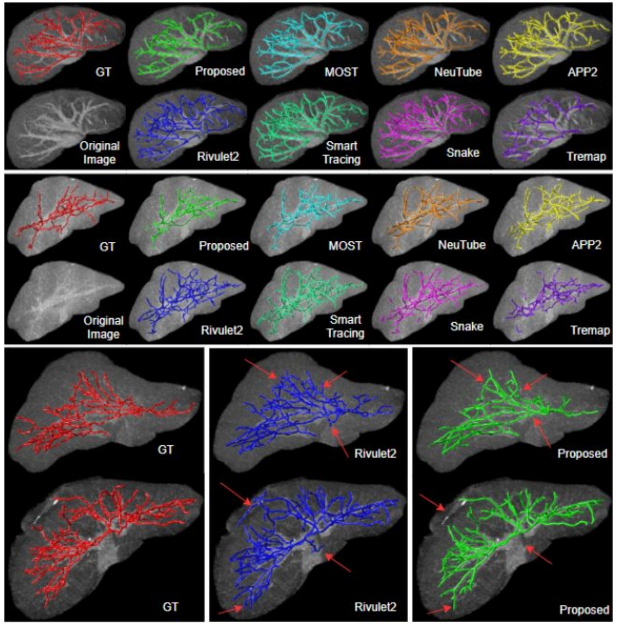
-Liver Vessel Reconstruction



CNN network architecture

Method	Precision	Recall	F1	SD	SSD	pSSD
MOST [12]	0.8625 ± 0.0775	0.6044 ± 0.1330	0.6984 ± 0.0957	2.97 ± 1.196	9.90 ± 3.280	0.167 ± 0.0511
NeuTube [27]	0.7622 ± 0.1089	0.2937 ± 0.0703	0.4158 ± 0.0737	3.86 ± 1.116	$\mathbf{9.28 \pm 2.517}$	0.279 ± 0.0587
APP2 [23]	0.8051 ± 0.1531	0.1890 ± 0.1265	0.2921 ± 0.1629	9.93 ± 8.470	16.44 ± 9.567	0.348 ± 0.1473
Smart Tracing [2]	0.7233 ± 0.1487	0.7780 ± 0.2048	0.7521 ± 0.0771	6.23 ± 6.873	14.87 ± 8.926	0.255 ± 0.1381
Snake [18]	0.7949 ± 0.0870	0.7719 ± 0.1119	0.7738 ± 0.0616	2.99 ± 0.883	10.34 ± 2.459	0.174 ± 0.0475
TreMap [28]	0.7844 ± 0.1610	0.2287 ± 0.1114	0.3422 ± 0.1313	5.22 ± 2.343	9.40 ± 2.974	0.337 ± 0.0920
Rivulet2 [11]	0.7634 ± 0.0961	$\mathbf{0.8823 \pm 0.0678}$	0.8124 ± 0.0514	3.02 ± 1.008	10.27 ± 3.204	0.184 ± 0.0501
Proposed	$\mathbf{0.9280 \pm 0.0497}$	0.8354 ± 0.0849	$\mathbf{0.8762 \pm 0.0549}$	$\mathbf{2.46 \pm 0.668}$	9.72 ± 3.172	$\mathbf{0.136 \pm 0.0433}$

Quantitative comparison with state-of-the-art methods

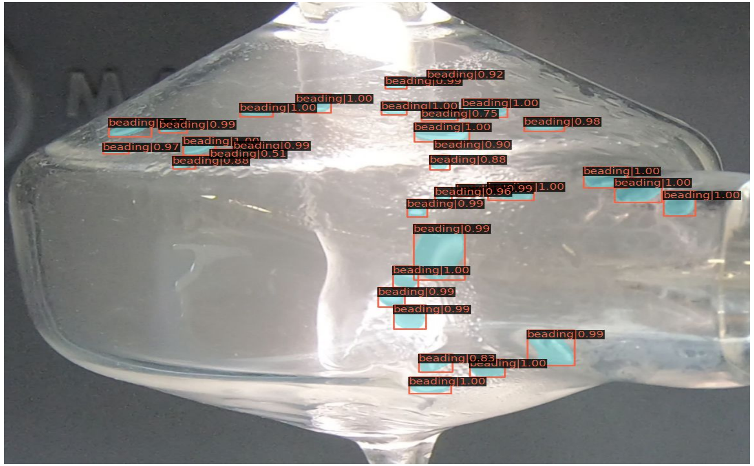


Visual comparison with state-of-the-art methods

Agilent Project

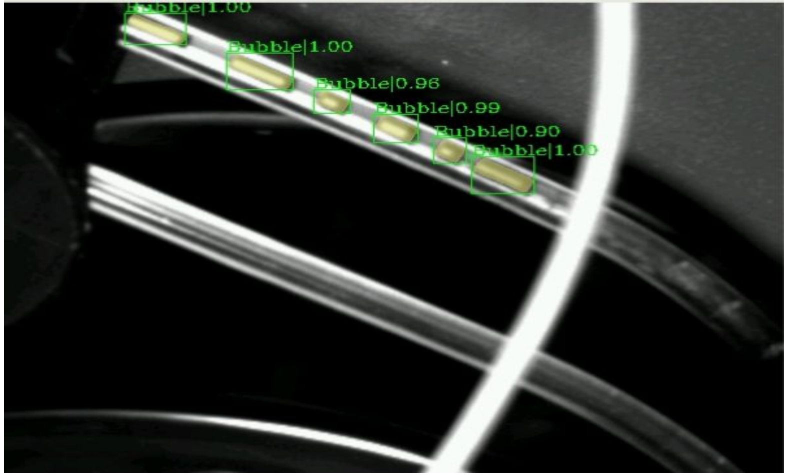
1) Improved Bead Detection Models

We choose the optimal combination of the backbone and the detector, with hard negative mining, and paste and copy augmentation to boost the accuracy.



2) Bubble Speed Estimation

We use object tracking to estimate the distances bubbles moved between two consecutive frames. We see a decrease in bubble speed overtime

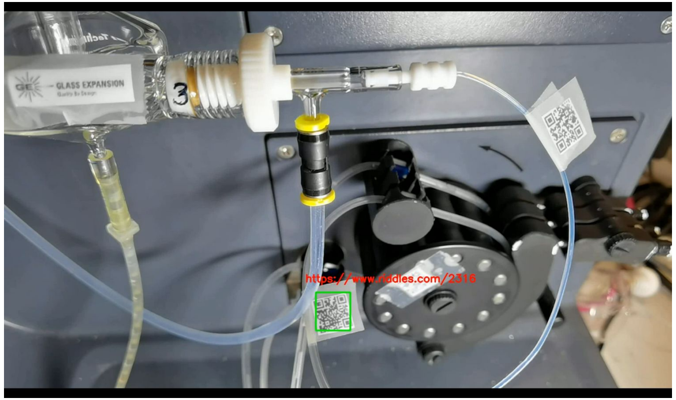


Agilent Project

MACQUARIE University

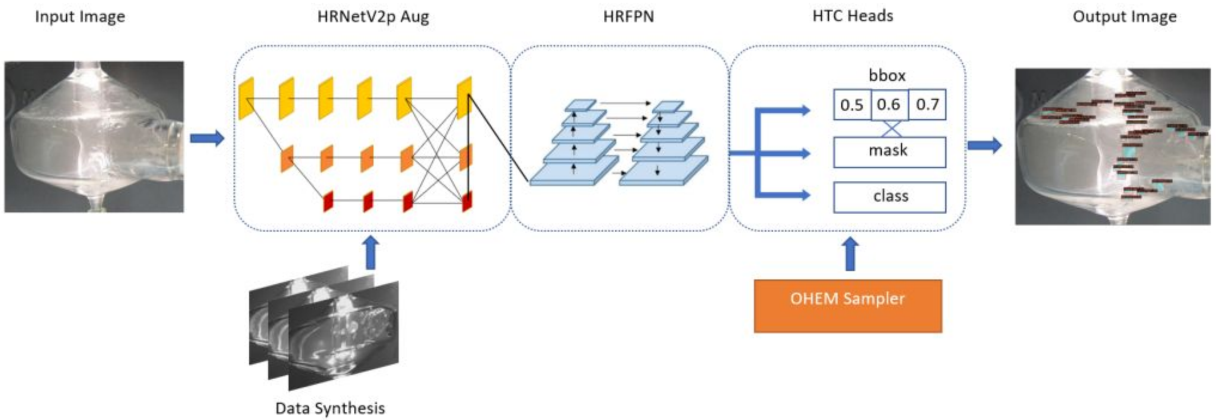
3) QR Code Detection and Decoding

We combine different QR code detection techniques to detect and decode QR codes at different locations



- Super resolution
- Convert image to grayscale
- Use bilateral filter and gaussian blur to remove noise
- Use canny edge detection
- Find contours in images
- Determine if a contour is a square
- Find squares with similar size and check if they are close each other
- Determine the location of the QR code and crop it out
- Use pyzbar and cv2 wechat to decode the QR code

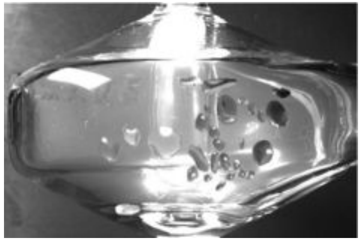
Bead Detection Pipeline



- **Backbone**: HRNetV2p is the HighResolution Network, which works to maintain the same resolution of an image in parallel.
- **HRFPN: High Resolution Feature Pyramids**
- **Hybrid Task Cascade**: inspired by high-quality detection from Cascade Mask R-CNN, via the cascade of three localization heads with increasing IoU thresholds of 0.5, 0.6, and 0.7
- **Hard Negative Mining**: To deal with the false positive, we include more training examples that do not contain any beading.
- **Data Synthesis and Copy Paste**

Data Synthesis

Clean Spray Chamber



Copy and Paste

HTC	OHEM	Aug.	mAP	AP_{30}	AP_{50}	AP_{75}
✓			70.3	91.3	90.1	61.3
✓	✓		80.4	95.8	94.9	82.6
✓	✓	✓	83.9	96.0	95.1	84.7

hard-negative mining approach via OHEM
data synthesis pre-training(denoted by Aug. herein).
The implementation of HTC with the HRNetV2p backbone

Bead Detection Result

Backbone	mAP	AP_{30}	AP_{50}	AP_{75}	AP_s	AP_m
DarkNet [28]	62.1	86.7	86.5	66.6	58.1	68.9
x101 64-4d [55]	60.5	84.9	82.9	55.1	47.2	66.8
RegNet [56]	59.0	85.5	82.2	52.8	49.2	64.1
ResNeSt [42]	64.7	89.9	88.0	59.7	56.7	69.4
Res2Net [40]	58.7	83.3	80.4	53.1	46.3	64.8
Swin-96 [43]	56.3	82.4	78.9	46.5	49.3	60.1
TridentNet [41]	42.2	76.6	67.3	19.2	37.7	44.9
HRNetV2p [9]	**69.6**	**90.1**	**89.0**	**71.7**	**59.1**	**74.8**

- **Res2Net**: a network that deals with multi-scale feature maps by using multiple receptive field sizes at granular levels of an image rather than resolution.
- **ResNeSt**: a different approach to scale, using channel-wise attention on different network branches to improve cross-scale information and representation.
- **TridentNet**: a parallel multi-branch architecture, where each branch uses a different receptive field and is then only trained on instances of selective scales
- **Swin-96**: hierarchical vision transformer using shifted window

Bead Detection Result

Detector	mAP	AP_{30}	AP_{50}	AP_{75}	AP_s	AP_m
SSD-512 * [22]	27.2	59.8	46.9	7.50	9.70	36.8
RetinaNet * [44]	42.2	76.6	67.3	19.2	37.7	44.9
YOLOv3 * [28]	62.1	86.7	86.5	66.6	58.1	68.9
YOLOF * [57]	56.4	85.5	78.8	49.9	49	61.1
Faster [18]	56.3	82.4	78.9	46.5	49.3	60.1
Libra [19]	66.0	87.2	86.6	69.3	56.0	70.9
Cascade [45]	69.3	91.2	90.4	69.9	60.6	73.6
Cascade Mask [47]	69.6	90.1	89.0	71.7	59.1	74.8
HTC [10]	70.3	90.9	90.1	70.9	61.3	74.8
Proposed	83.9	96.0	95.1	84.7	65.6	71.4

- One-Stage detectors struggle to be competitive with the Two-Stage models in terms of mAP

- YOLOv3, which surpasses the results of Faster R-CNN

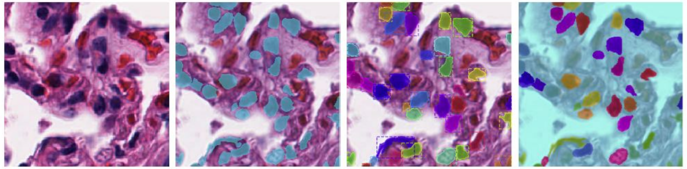
COMPUTATIONAL Histopathology

Significance

- Reduce workload for pathologists
- Important step of cell morphology analysis
- Critical Information for Cancer Diagnosis and Prognosis

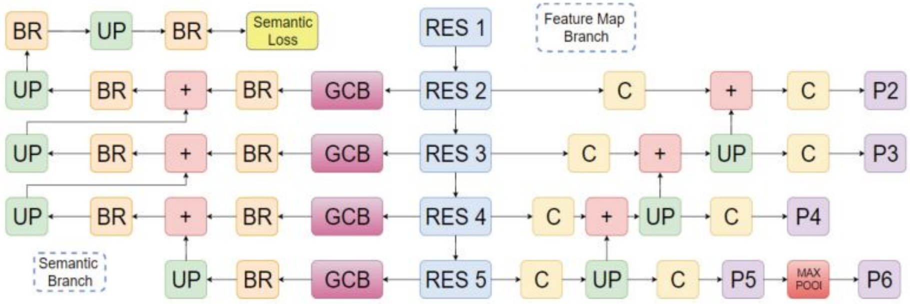
Challenges

- High level of heterogeneity in different types of organs or cells
- Similar structure such as cytoplasm and stroma



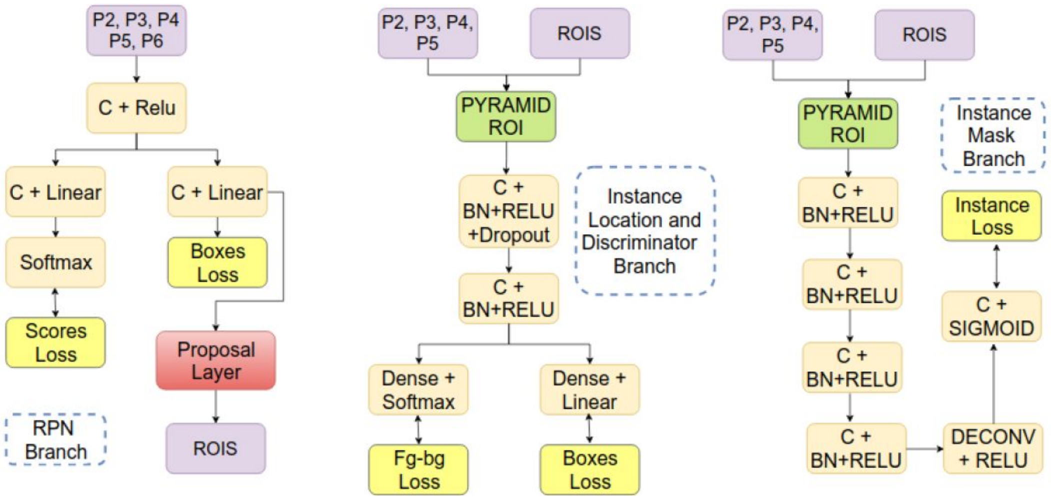
(a) Image (b) Semantic Segmentation (c) Instance Segmentation (d) Panoptic Segmentation

Cell R-CNN



The network architecture of semantic segmentation branch and feature map branch. C, BR, UP, GCB and RES X represent convolutional layer, boundary refine block, upsampling layer, global convolution block and specific ResNet layer respectively.

Region Proposal Network and Instance Branch



Region Proposal Network and Instance Branch. ROIS, PYRAMID ROI, and DECONV indicate regions of interest, pyramid ROI aligning layer and deconvolutional layer, respectively.

Biomedical Image Segmentation

Brian Tumor Segmentation

- **Significance**:
  - Difficult to annotate the fuzzy regions with binary masks
- **Challenges**
  - Diffusion of surrounding edema into the tumor region
  - Non-standardized voxel intensity values in MRI unlike the Computed Tomography and X-ray
  - irregular shapes and sizes of brain tumors
- **Contribution**:
  - Propose a novel lightweight 3D CNN model balancing efficiency and accuracy with 3D depthwise and separate convolutions

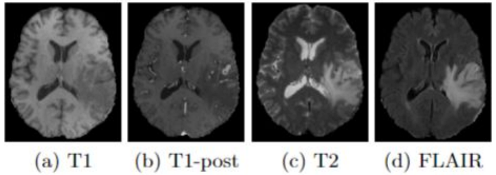


Fig: Visualization of different imaging modalities

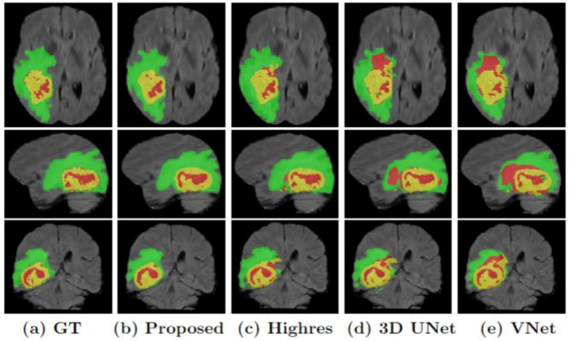


Fig: A visual example in the testing set compared with SOTA methods.

Publication (Main Research Topics)

MACQUARIE University

- ## Computational Histopathology
- D. Liu, D. Zhang, Y. Song, H. Huang, W. Cai, “Cell R-CNN V3: A Novel Panoptic Paradigm for Instance Segmentation in Biomed- ical Images.” IEEE Transactions on Image Processing, **TIP**, 2021. (**IF:10.86**)
- D. Liu, **D. Zhang**, Y. Song, F. Zhang, L. O’Donnell, H. Huang, M. Chen, W. Cai, “Unsupervised Instance Segmentation in Mi- croscopy Images via Panoptic Domain Adaptation and Task Re-weighting.” *Conference on Computer Vision and Pat- tern Recognition*, **CVPR**, ,2020. (**h5:389**)
- **D. Zhang**, Y. Song, D. Liu, H. Jia, S. Liu, Y. Xia, H. Huang, W. Cai, “Panoptic Segmentation with an End-to-end Cell R-CNN for Pathology Image Analysis.” *International Conference on Medical Image Computing and Computer Assisted Intervention*, **MICCAI**, 2018. (**h5:61**)
- **D. Zhang**, Y. Song, S. Liu, D. Feng, Y. Wang, W. Cai. “Nuclei Instance Segmentation With Dual Contour-enhanced Adversarial Network.” *IEEE International Symposium on Biomedical Imaging From Nano To Macro*, **ISBI**, 2018. (**h5:43**)
- ## Biomedical Image Segmentation
- L. Wang, X. Ye, **D. Zhang**, L. Ju, W. He, X. Wang, W. Feng, K. Song, X. Zhao, Z. Ge, “3D Matting: A Benchmark Study on Soft Segmentation Method Applied in Computed Tomography.” *Computers in Biology and Medicine*, 2022. (**IF:10.86**)
- Y. Wu, Z. Ge, **D. Zhang**, M. Xu, L. Zhang, Y. Xia, J. Cai, “Enforcing Mutual Consistency of Hard Regions for Semi-supervised Medical Image Segmentation.” *Medical Image Analysis*, **MIA**, 2022.(**IF: 8.55**)
- X. Yu, B. Lou, **D. Zhang**, D. Winkel, N. Arrahmane, M. Diallo, T. Meng, H. Busch, R. Grimm, B. Kiefer, D. Comaniciu, A. Ka- men, ProstateAI Clinical Collaborators, “Deep Attentive Panoptic Model for Prostate Cancer Detection Using Biparametric MRI Scans.” *International Conference on Medical Image Computing and Computer Assisted Intervention*, **MICCAI**, 2020. (**h5:61**)
- D. Liu, **D. Zhang**, Y. Song, F. Zhang, L. O’Donnell, W. Cai, “3D Large Kernel Anisotropic Network for Brain Tumor Segmentation.” *International Conference on Neural Information Processing*, 2018. (**h5:35**)
- **D. Zhang**, Y. Song, D. Liu, C. Zhang, Y. Wu, H. Wang, F. Zhang, Y. Xia, L. J. O’Donnell and W. Cai, “Efficient 3D Depthwise and Separable Convolutions with Dilation for Brain Tumor Segmentation.” *The Australasian Joint Conference on Artificial Intelligence*, 2019
- ## Curvilinear Morphology Reconstruction
- H. Wang, **D. Zhang**, Y. Song, S. Liu, Y. Wang, D. Feng, H. Peng, W. Cai, “Segmenting Neuronal Structure in 3D Optical Micro- scope Images via Knowledge Distillation with Teacher-Student Network.” *IEEE International Symposium on Biomedical Imaging*, **ISBI**, 2019, (**h5:43**)
- H. Wang, **D. Zhang**, Y. Song, S. Liu, H. Huang, M. Chen, H. Peng, W. Cai, “Multiscale Kernels for Enhanced U-shaped Network to Improve 3D Neuron Tracing.” *Computer Vision for Microscopy Image Analysis Workshop*, *In conjunction with IEEE Conference on Computer Vision and Pattern Recognition*, 2019
- **D. Zhang**, S. Liu, Y. Song, D. Feng, H. Peng, W. Cai. “Automated 3D Soma Segmentation with Morphological Surface Evolution for Neuron Reconstruction.” *Neuroinformatics*, 2018. (**IF: 4.09**)
- S. Liu, **D. Zhang**, Y. Song, H. Peng, W. Cai, “Automated 3D Neuron Tracing with Precise Branch Erasing and Confidence Con- trolled Back-Tracking.” *IEEE Transactions on Medical Imaging*, **TMI**, 2018. (**IF: 10.05**)
- S. Liu, **D. Zhang**, S. Liu, D. Feng, H. Peng, W. Cai. “Rivulet: 3D Neuron Morphology Tracing with Iterative Back-Tracking.” *Neuroinformatics*, 2016. (**IF: 4.09**)

Thank You