Detection of Knots in Oak Wood Planks

Instance versus Semantic segmentation (AT22-543)

Evalds Urtans (Riga Technical University), Karlis Bumanis (SIA Meza un koksnes produktu petniecības un attistibas instituts),
Valters Vecins (RTU), Maris Ancans (RTU), Aiga Andrijanova (RTU), Marcis Teodors Upenieks (RTU), Kristofers Volkovs (Ventspils University)

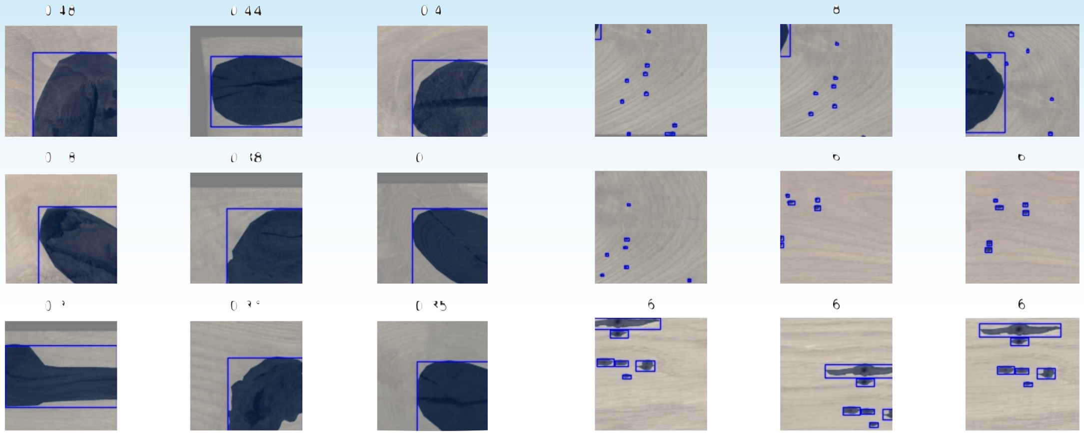
Summary

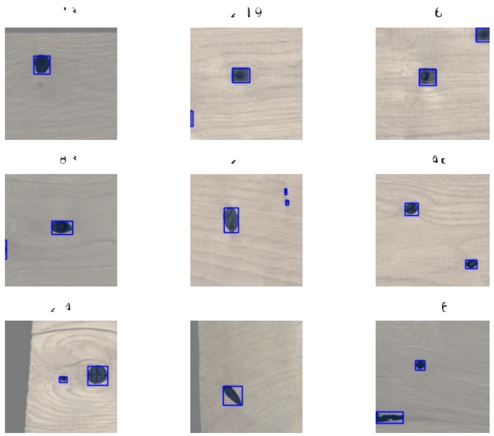
- Novel, publicly available dataset "FSCC oak knots" introduced
- Comparitve analysis between Instance and Semantic segmentation models
- Open-source implementation (https://github.com/evaldsurtans/FSCC-Oak-Knots-Dataset)

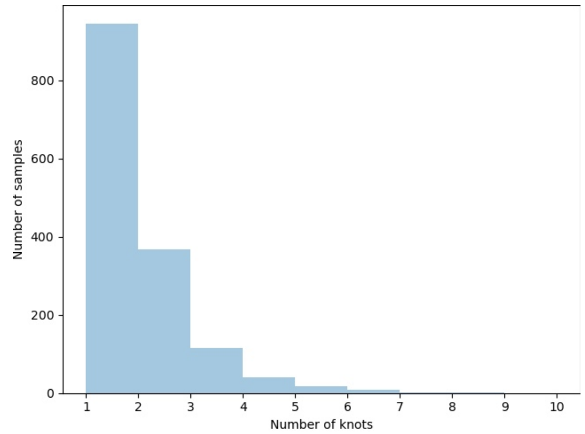
Novel Dataset

FSCC oak knots

- 1500 samples
  (1200 train, 300 test)
- 500x500 pixels RGB
- Labelled, 1 to 11 knots per image (mask & box)
- Alternatives: LSIDWS, WOOD, Wood Spieces and Kaggle Wood Textures



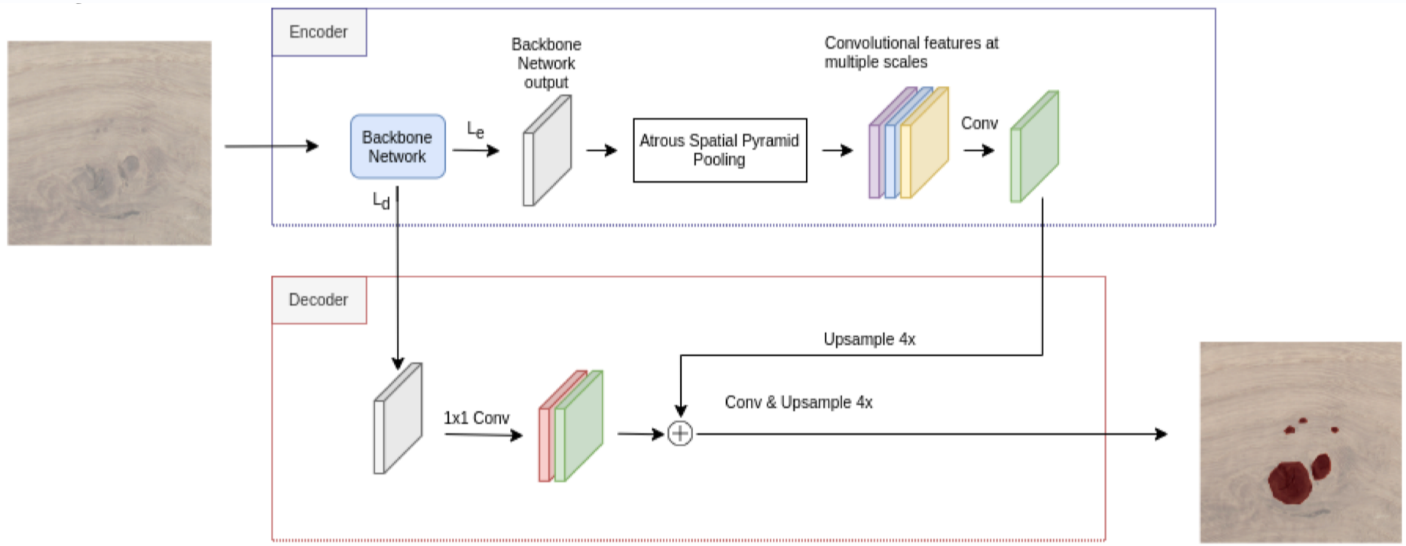




Semantic segmentation

DeepLab V3, UNet

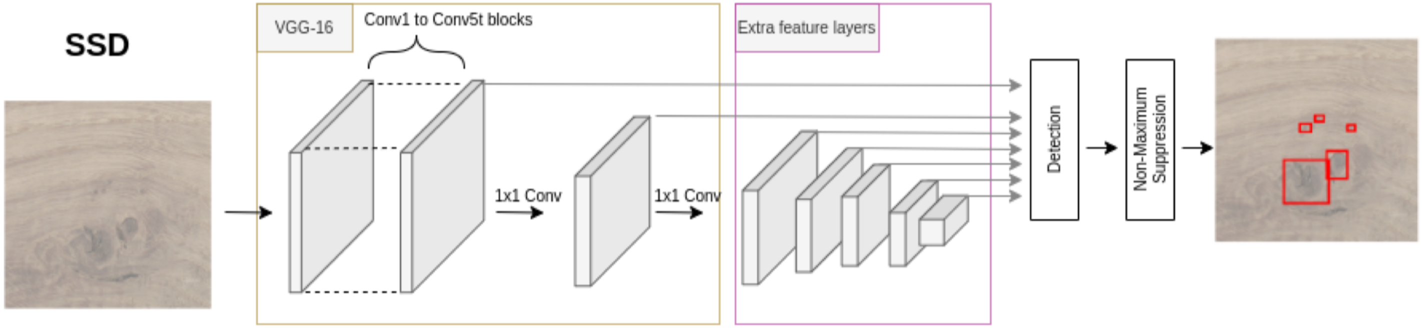
- Binary Cross-Entropy



$$\mathcal{L}_{BCE}(y, y') = -\frac{1}{N}\sum(y\ log(y' + \epsilon) + (1 - y)\ log((1 - y') + \epsilon))$$

Instance segmentation
Single-Shot Detector (SSD)

- Binary Cross-Entropy
- L1 Smooth loss
- Non Maximum Suppression algorithm



$$\mathcal{L}_{L1\ smooth} = \frac{1}{N}\sum \begin{cases} 0.5(y_n - y_n')^2/\beta, & \text{if } |y_n - y_n'| < \beta \\ |y_n - y_n'| - 0.5\beta, & \text{otherwise} \end{cases}$$

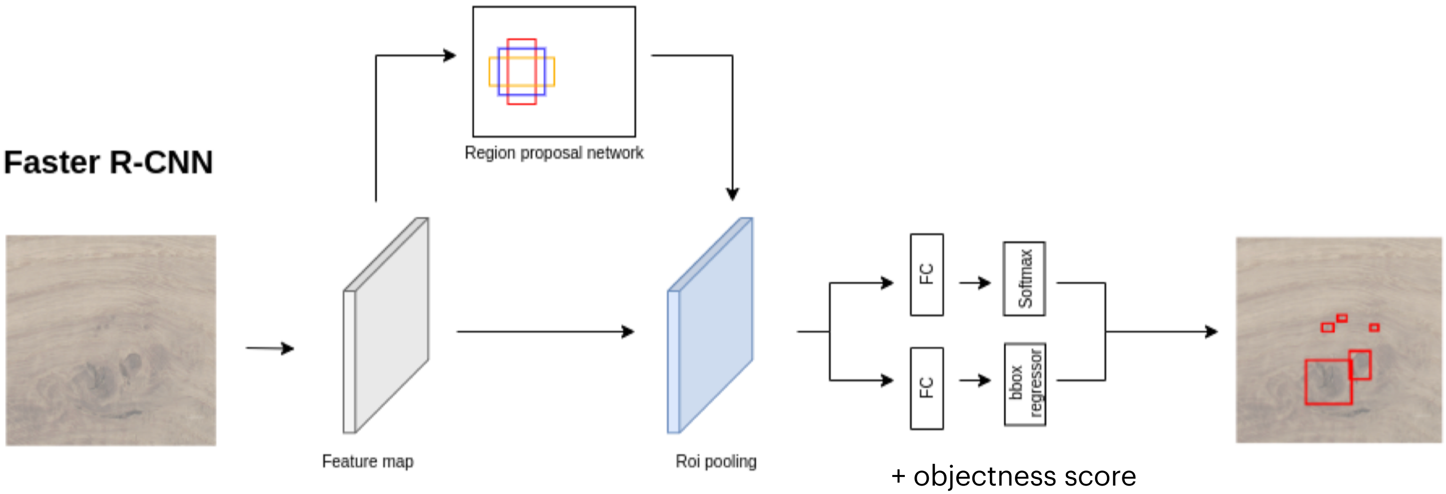
$$\mathcal{L}_{SSD} = \mathcal{L}_{L1\ smooth} + C\,\mathcal{L}_{BCE}$$

Instance segmentation

Faster Region-based Convolutional Neural Network (Faster R-CNN)

- Binary Cross-Entropy
- L1 Smooth loss
- Objectness loss
- Non Maximum Suppression algorithm

Faster R-CNN



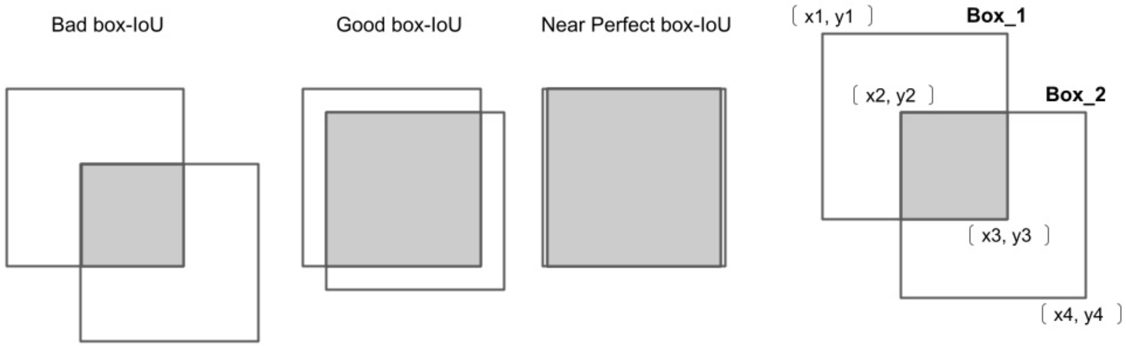
$$\mathcal{L}_{L1\ smooth} = \frac{1}{N}\sum \begin{cases} 0.5(y_n - y_n')^2/\beta, & \text{if } |y_n - y_n'| < \beta \\ |y_n - y_n'| - 0.5\beta, & \text{otherwise} \end{cases}$$

$$\mathcal{L}_{objectness} = \frac{1}{N}\sum (IoU - IoU')^2$$

$$\mathcal{L}_{RCNN} = \mathcal{L}_{L1\ smooth} + C_1\ \mathcal{L}_{BCE} + C_2\ \mathcal{L}_{objectness}$$

Metrics

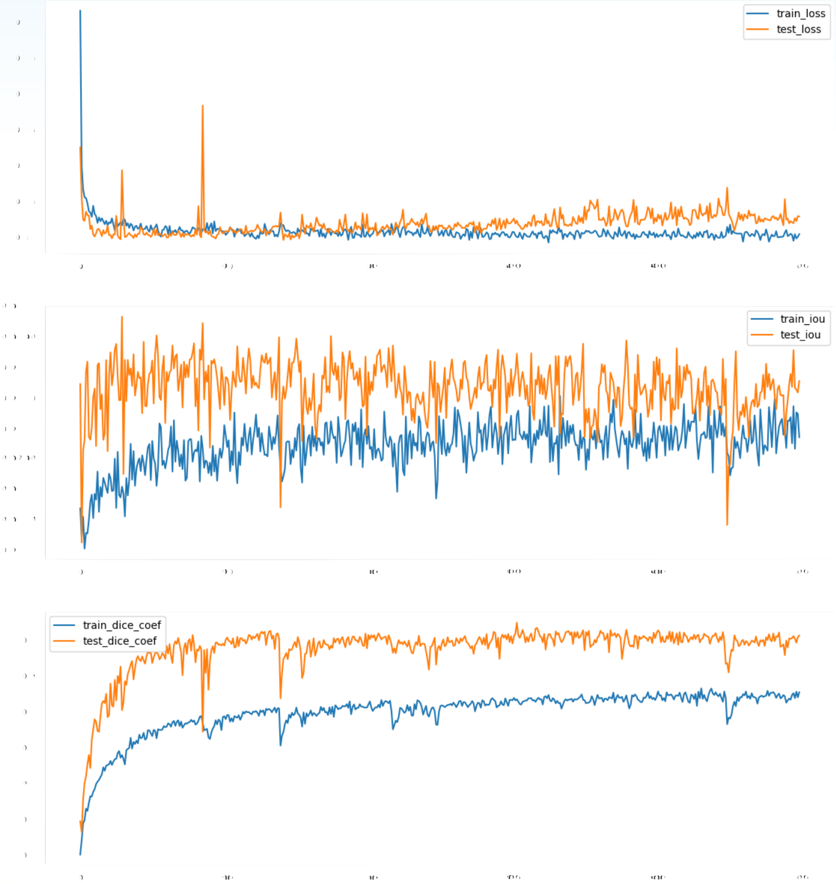
- Pixel-based IoU (Intersection over Union Jaccard index)
- Box-based IoU
- Accuracy (did it detected a kont at all in given image that contains a knot)

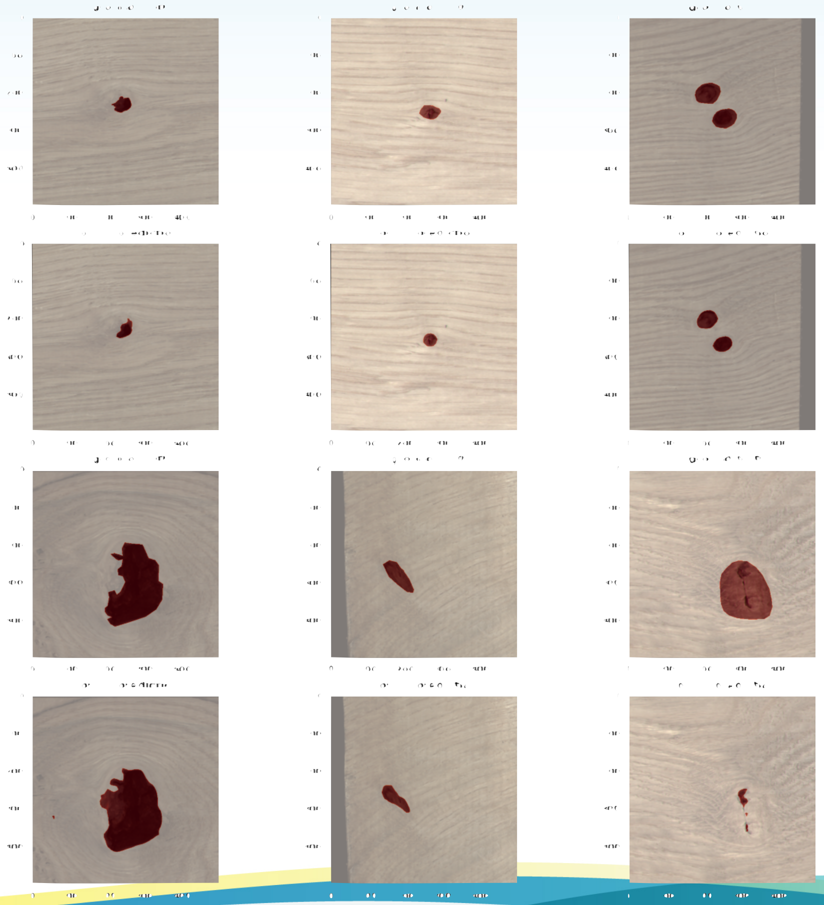


Training protocol

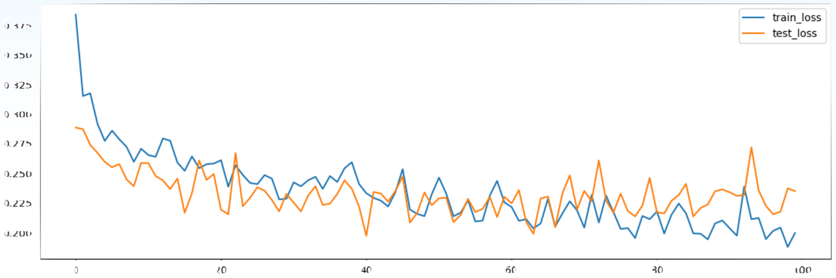
- Augmentations (horizontal flip, vertical flip, scale, rotation, and color jitter)
- Weighted Binary Cross-Entropy, based on pixel distribution
- Grid search of hyper-parameters (batch_size, learning_rate, model, confidence thresholds)

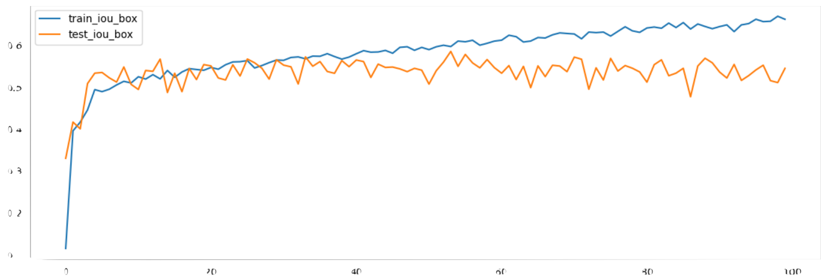
Semantic Segmentation

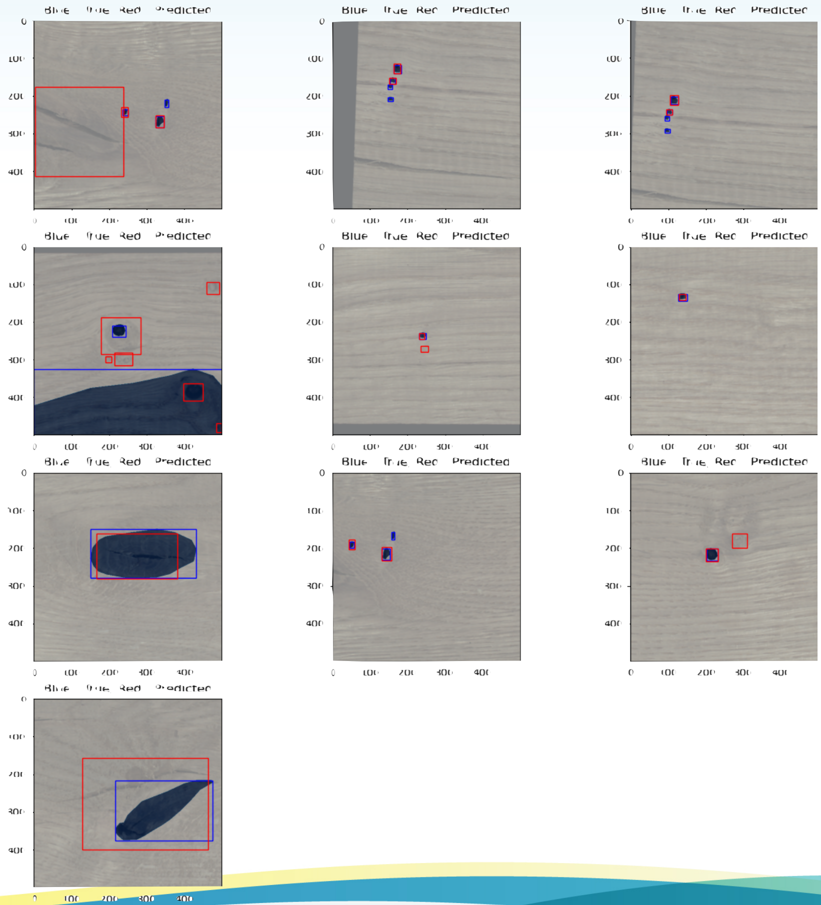




Instance Segmentation







Results
and conclusions

- Dataset is feasible for task
- Instance segmentation has superior **59% Box-IoU** and **90% acc.** versus **49% Box-IoU** and **89% acc.** using semantic segmentation
- UNet-based DeepLabV3 has superior **51% Pixel-IoU** versus **42% Pixel-IoU** using FCN

Model	Method	Pixel IoU	Box IoU	Acc.
Faster R-CNN	Instance segmentation	N/A	**0.59**	**0.90**
FCOS	Instance segmentation	N/A	0.51	0.77
SSD VGG-16	Instance segmentation	N/A	0.27	0.67
RetinaNet	Instance segmentation	N/A	0.49	0.72
DeepLab V3 ResNet-101	Semantic segmentation	**0.51**	**0.49**	**0.89**
DeepLab V3 ResNet-50	Semantic segmentation	0.45	0.40	0.86
DeepLab V3 MobileNet	Semantic segmentation	0.40	0.37	0.72
Lite R-ASPP MobileNet	Semantic segmentation	0.42	0.37	0.73
FCN ResNet-50	Semantic segmentation	0.42	0.43	0.81

Further research

- UNet-based models: UNet++, UNet 3+, etc.
- Combined models: Mask R-CNN, etc.
- Loss functions: Focal loss, Tversky loss, DICE loss and combinations
- Pre-training using other wood segmentation datasets