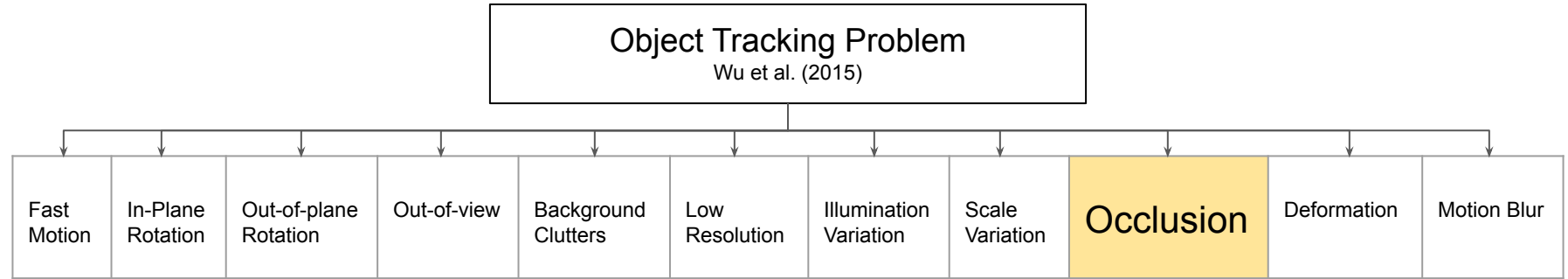


Laporan Progress Tesis

1906337053 - Siti Hadiyan Pratiwi

Background Overview



Occlusion merupakan keadaan dimana objek utama terhalang secara sebagian ataupun secara menyeluruh.

Background Overview

Depth Image



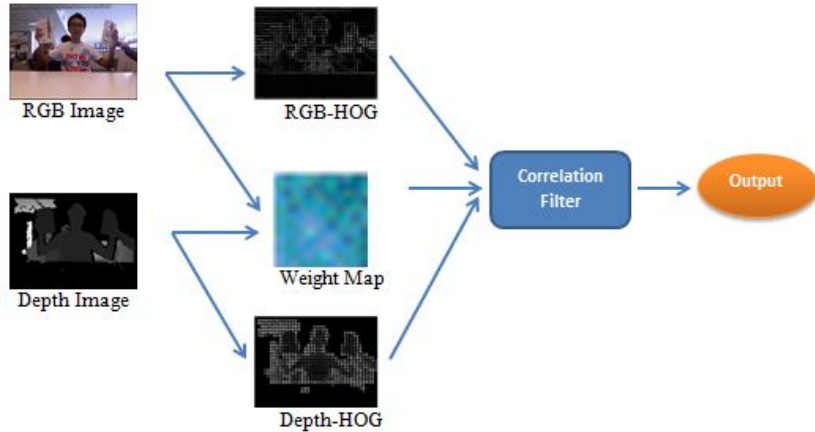
Memiliki nilai
depth berbeda



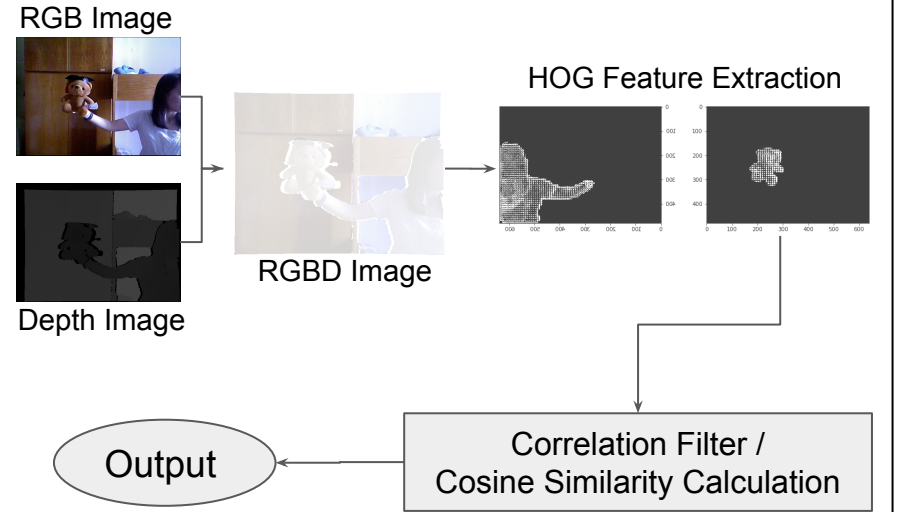
- Depth Image merupakan gambar yang merepresentasikan jarak pada benda terhadap ruang.
- Semakin jauh benda maka akan memiliki warna semakin terang (putih). Begitu pula sebaliknya benda yang lebih dekat akan memiliki warna lebih gelap (hitam).
- Image depth dimanfaatkan untuk mendapatkan bentuk 3 dimensi dari objek
-  Jika dilihat potongan gambar disamping, image depth memiliki potensi untuk dapat mengidentifikasi benda tertutup. Terlihat dengan adanya perbedaan pada kedua objek sehingga masing-masing objek dapat diidentifikasi secara terpisah.

Perbandingan proposal dan usulan metode

Proposal



Usulan Metode



Checklist Overview

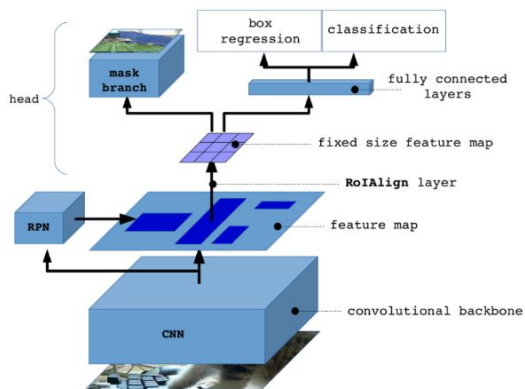
Literature Review	✓ Literature Overview ✓ Initial Experiment
Persiapan Dataset	✓ Persiapan dataset RGB dan Depth ✓ Pembentukan groundtruth dan anotasi gambar ✓ Pembentukan image 4-Channel RGBD dengan melakukan merge image RGB dan Depth
Pembentukan Learning Model	✓ Pembentukan model ✓ Training Model ✓ Hasil Evaluasi
Aplikasi Algoritma Tracking: Cosine Similarity / Correlation Filter	✓ Ekstraksi Fitur ☐ Perhitungan Cosine Similarity ☐ Aplikasi Metode Evaluasi
Output	☐ Visualisasi ☐ Evaluasi ☐ Pembentukan kesimpulan

Literature Review

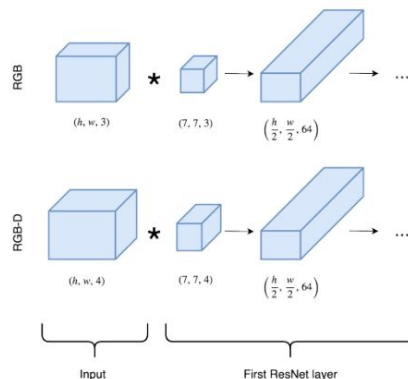
✓ Literature Overview

☐ Initial Experiment

- Paper: **Does Depth Matter? RGB-D Instance Segmentation with Mask R-CNN** (Orestis Zambounis, Furrer Fadri, Novkovic Tonci, Grinvald Margarita)



Mask RCNN



Modified Mask RCNN
4 Channel Layer

Literature Review



Literature Overview



Initial Experiment

Experiment Result:

- Pre-training on SceneNet:
 - 400 epochs
 - 2000 steps / epoch
- Fine-tuning on NYU-Depth V2
 - 40 Epochs
 - 1000 steps / epoch

	AP	AP ₅₀	AP ₇₅
RGB	44.4	62.9	46.6
RGB-D	54.5	71.4	57.6

(a) Binary segmentation

	AP	AP ₅₀	AP ₇₅
RGB	28.2	41.2	30.0
RGB-D	37.0	49.5	38.9

(b) Class-specific segmentation

Table 1: Mask AP on SceneNet RGB-D

	AP	AP ₅₀	AP ₇₅
RGB	25.3	46.1	24.5
RGB-D	27.8	49.1	28.1

(a) Binary segmentation

	AP	AP ₅₀	AP ₇₅
RGB	17.6	30.7	18.1
RGB-D	18.4	31.5	19.0

(b) Class-specific segmentation

Table 2: Mask AP on NYU-Depth V2



Literature Review

- ✓ Literature Overview
- ✓ Initial Experiment

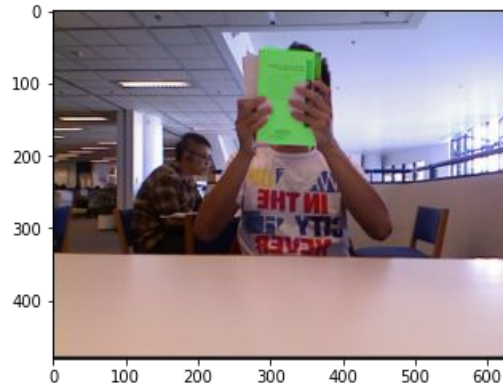
- Data:
 - Train: 7 Images, Test: 189 Images (including 7 train images)
 - Epoch: 5, LR: 0.001



Literature Review

- ✓ Literature Overview
- ✓ Initial Experiment

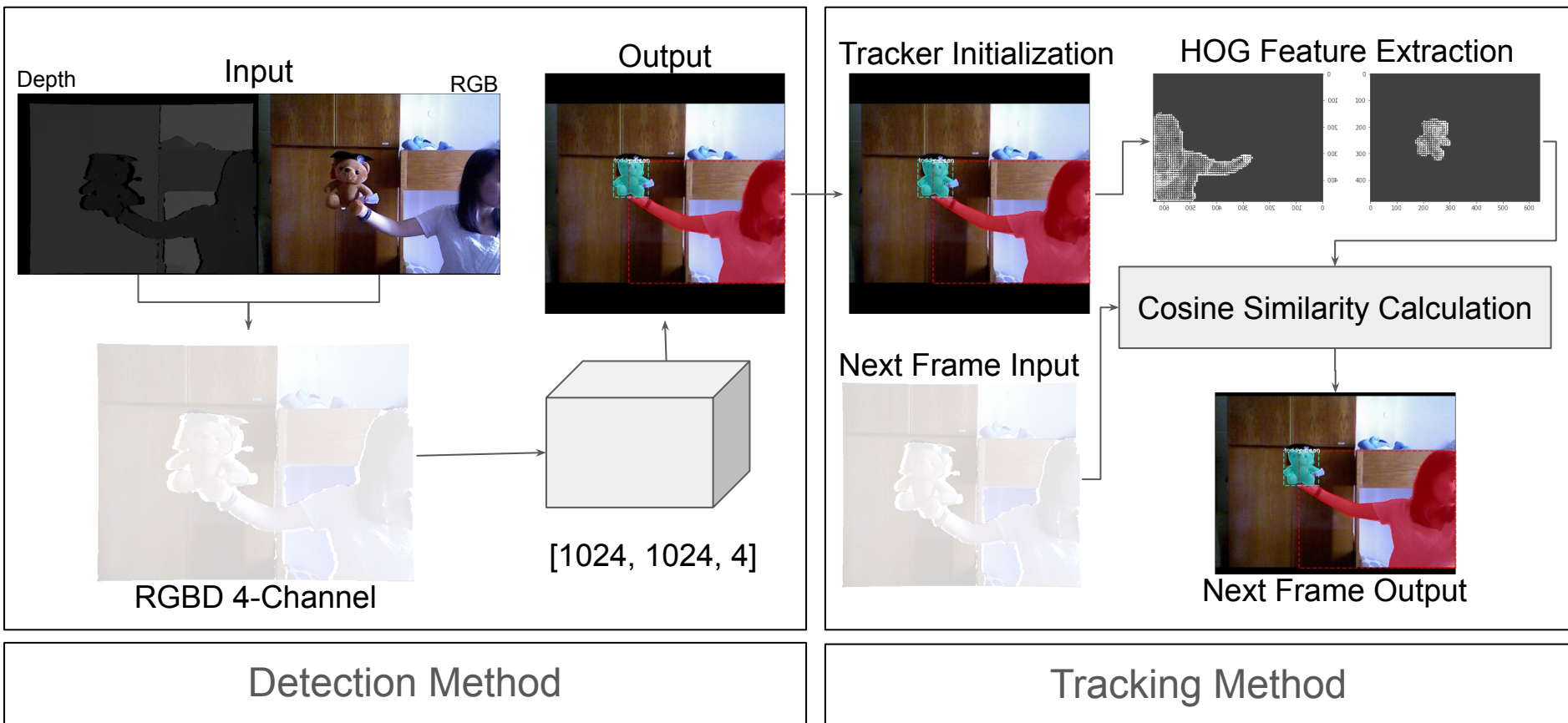
Detection result (1 image random)



Detection result (sequence 189 image)



Metodologi Penelitian

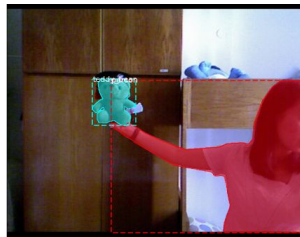


Milestone

Input: 4-Channel
RGBD Image



Detection Method:
4-Channel Mask RCNN

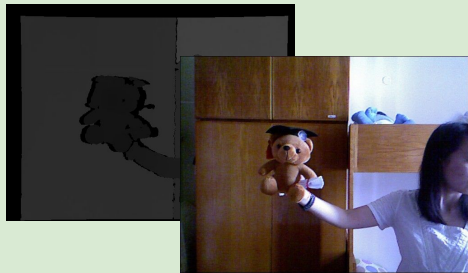


Tracking Method:
Cosine Similarity /
Correlation Filter

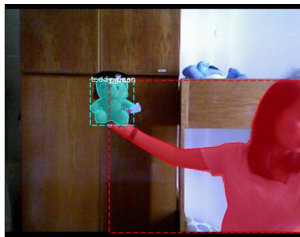
Output

Progress Milestone

Input: 4-Channel
RGBD Image



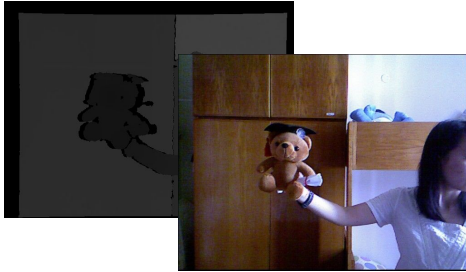
Detection Method:
4-Channel Mask RCNN



Tracking Method:
Cosine Similarity /
Correlation Filter

Output

Input: 4-Channel
RGBD Image



Persiapan Dataset

- ❑ Persiapan dataset RGB dan Depth
- ❑ Pembentukan groundtruth dan anotasi gambar
- ❑ Pembentukan image 4-Channel RGBD dengan melakukan merge image RGB dan Depth
- ❑ Visualisasi hasil masking

Persiapan Dataset

- ✓ Persiapan dataset RGB dan Depth
- ❑ Pembentukan groundtruth dan anotasi gambar
- ❑ Pembentukan image 4-Channel RGBD dengan melakukan merge image RGB dan Depth
- ❑ Visualisasi hasil masking

Spesifikasi Dataset:

- Source: Princeton Tracking Benchmark (PTB) Dataset
- Sequence:
 - Bear_front (30 fps) : 297 Images RGB + 297 Images Depth
- Preview:

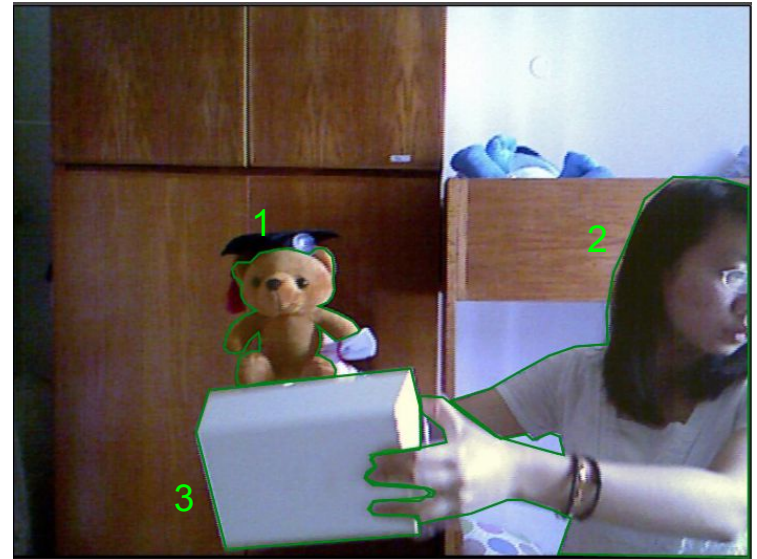


Persiapan Dataset

- ✓ Persiapan dataset RGB dan Depth
- ✓ Pembentukan groundtruth dan anotasi gambar
- ❑ Pembentukan image 4-Channel RGBD dengan melakukan merge image RGB dan Depth
- ❑ Visualisasi hasil masking

Anotasi dan pemberian label terhadap 3 objek:

1. Teddy_bear
2. Person
3. Box



Total gambar: 295

Persiapan Dataset

- ✓ Persiapan dataset RGB dan Depth
- ✓ Pembentukan groundtruth dan anotasi gambar
- ✓ Pembentukan image 4-Channel RGBD dengan melakukan merge image RGB dan Depth
- ☐ Visualisasi hasil masking

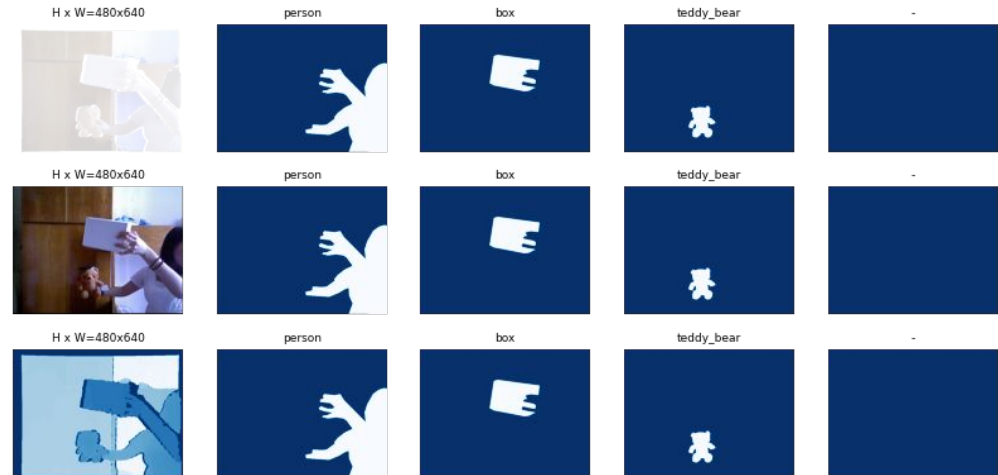
- Pembentukan image 4-Channel RGBD dilakukan dengan cara menggabungkan 3-Channel image RGB dengan 1-Channel image Depth.
- Preview:



Persiapan Dataset

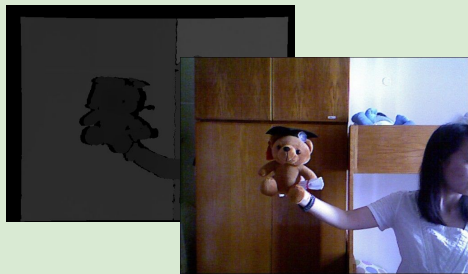
- ✓ Persiapan dataset RGB dan Depth
- ✓ Pembentukan groundtruth dan anotasi gambar
- ✓ Pembentukan image 4-Channel RGBD dengan melakukan merge image RGB dan Depth
- ✓ Visualisasi hasil masking

- Preview masking anotasi manual

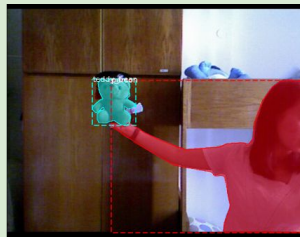


Progress Milestone

Input: 4-Channel
RGBD Image



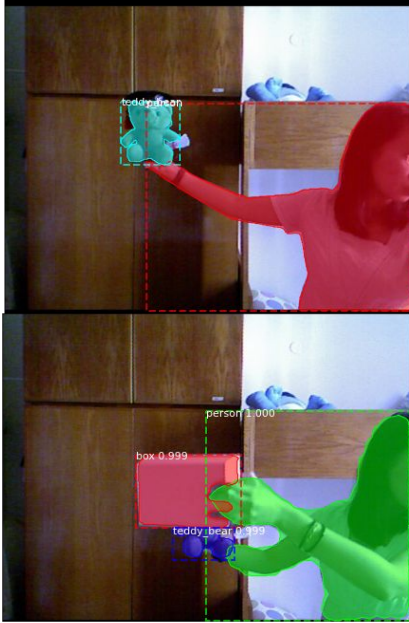
Detection Method:
4-Channel Mask RCNN



Tracking Method:
Cosine Similarity /
Correlation Filter

Output

Detection Method:
4-Channel Mask
RCNN



Pembentukan Learning Model

- ❑ Pembentukan model
- ❑ Training Model
- ❑ Hasil Evaluasi

Pembentukan Learning Model



Pembentukan Model



Training Model



Hasil Evaluasi

- Spesifikasi Model:

```
Configurations:
BACKBONE                resnet50
BACKBONE_STRIDES        [4, 8, 16, 32, 64]
BATCH_SIZE              1
BBOX_STD_DEV            [0.1 0.1 0.2 0.2]
COMPUTE_BACKBONE_SHAPE  None
DETECTION_MAX_INSTANCES 100
DETECTION_MIN_CONFIDENCE 0.7
DETECTION_NMS_THRESHOLD 0.3
FPN_CLASSIF_FC_LAYERS_SIZE 1024
GPU_COUNT               1
GRADIENT_CLIP_NORM      5.0
IMAGES_PER_GPU          1
IMAGE_CHANNEL_COUNT      4
IMAGE_MAX_DIM           1024
IMAGE_META_SIZE         16
IMAGE_MIN_DIM           800
IMAGE_MIN_SCALE         0
IMAGE_RESIZE_MODE        square
IMAGE_SHAPE              [1024 1024 4]
LEARNING_MOMENTUM       0.9
LEARNING_RATE            0.001
LOSS_WEIGHTS             {'rpn_class_loss': 1.0, 'rpn_bbox_loss': 1.0, 'mrcnn_class_loss': 1.0, 'mrcnn_bbox_loss': 1.0, 'mrcnn_mask_loss': 1.0}
MASK_POOL_SIZE          14
MASK_SHAPE               [28, 28]
MAX_GT_INSTANCES        100
MEAN_PIXEL               [123.7 116.8 103.9 127.5]
MINI_MASK_SHAPE          (56, 56)
MODE                     RGBD
NAME                     rgbd
NUM_CLASSES              4
POOL_SIZE                7
POST_NMS_ROIS_INFERENCE 1000
POST_NMS_ROIS_TRAINING   2000
PRE_NMS_LIMIT            6000
ROI_POSITIVE_RATIO       0.33
RPN_ANCHOR_RATIOS        [0.5, 1, 2]
RPN_ANCHOR_SCALES        (32, 64, 128, 256, 512)
RPN_ANCHOR_STRIDE        1
RPN_BBOX_STD_DEV         [0.1 0.1 0.2 0.2]
RPN_NMS_THRESHOLD        0.7
RPN_TRAIN_ANCHORS_PER_IMAGE 256
STEPS_PER_EPOCH          200
TOP_DOWN_PYRAMID_SIZE    256
TRAIN_BN                 False
TRAIN_ROIS_PER_IMAGE     200
USE_MINI_MASK            True
USE_RPN_ROIS             True
VALIDATION_STEPS         50
WEIGHT_DECAY             0.0001
```

Pembentukan Learning Model

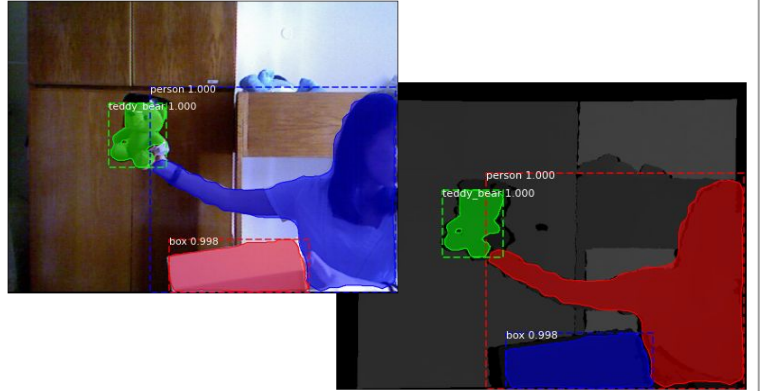
- ✓ Ekstraksi Fitur
- ✓ Training Model
- 📄 Hasil Evaluasi

- Training dilakukan dengan mengambil gambar secara acak menggunakan komposisi sebagai berikut:
 - Train : 118 Images (40%)
 - Validation : 29 Images (10%)
 - Test : 148 Images (50%)
- Spesifikasi training:
 - Epoch : 10
 - Learning Rate : 0.001

Persiapan Dataset

- ✓ Ekstraksi Fitur
- ✓ Training Model
- ✓ Hasil Evaluasi

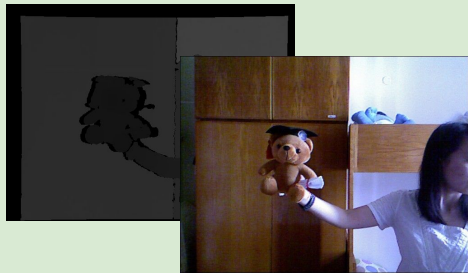
- Val_Loss: 0.1726
- Loss: 0.1218
- Preview:

Groundtruth	Detection Result
	

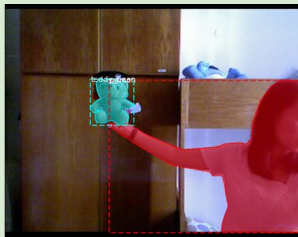
- mAP: 0.9740112999738273

Progress Milestone

Input: 4-Channel
RGBD Image



Detection Method:
4-Channel Mask RCNN



Tracking Method:
Cosine Similarity

Output

Tracking Method:
Cosine Similarity /
Correlation Filter

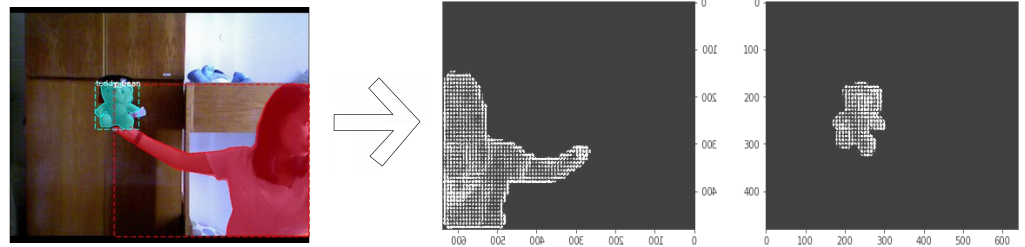
Aplikasi Algoritma Tracking

- ❑ Ekstraksi Fitur
- ❑ Perhitungan Cosine Similarity
- ❑ Hasil Evaluasi

Aplikasi Algoritma Tracking

- ❑ Ekstraksi Fitur
- ❑ Perhitungan Cosine Similarity
- ❑ Hasil Evaluasi

- Ekstraksi fitur menggunakan metode Histogram of Gradient (HOG).
- Ekstraksi dilakukan pada hasil deteksi objek
- Preview:



Aplikasi Algoritma Tracking

- ❑ Ekstraksi Fitur
- ❑ Perhitungan Cosine Similarity
- ❑ Hasil Evaluasi

- Preview tracker:

