Markus Fjellheim, Nils Magne Fossåen, Renny Octavia Tan

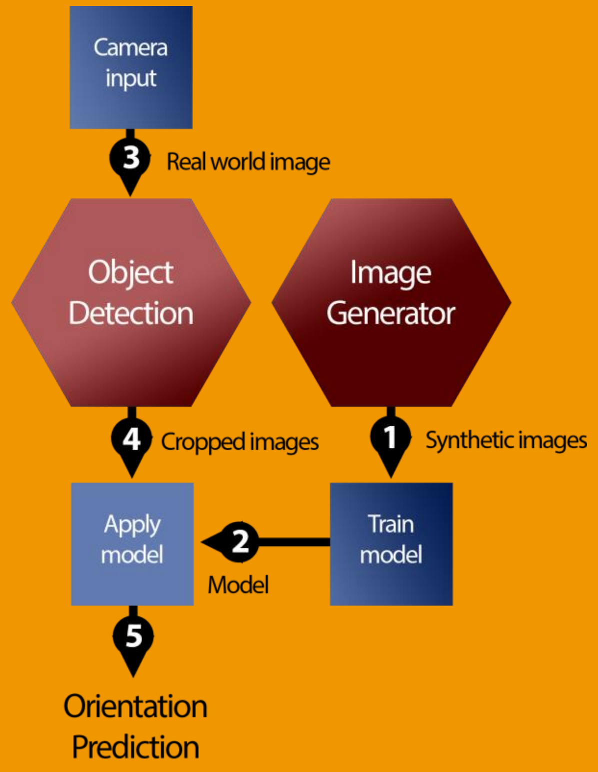
Synthetic Images for Training and Object Orientation Prediction

A concept for raising spatial awareness from 2D-images

4/24/2019

Introduction

- our goal is to detect the position and orientation of objects in regular images
- generate synthetic images labeled with known spatial data
- train our models on the synthetic images
- apply the model to real world images

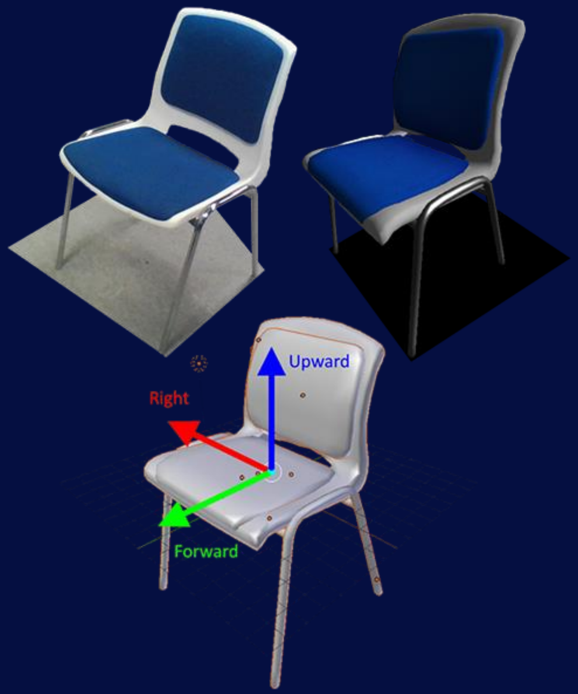


Technology

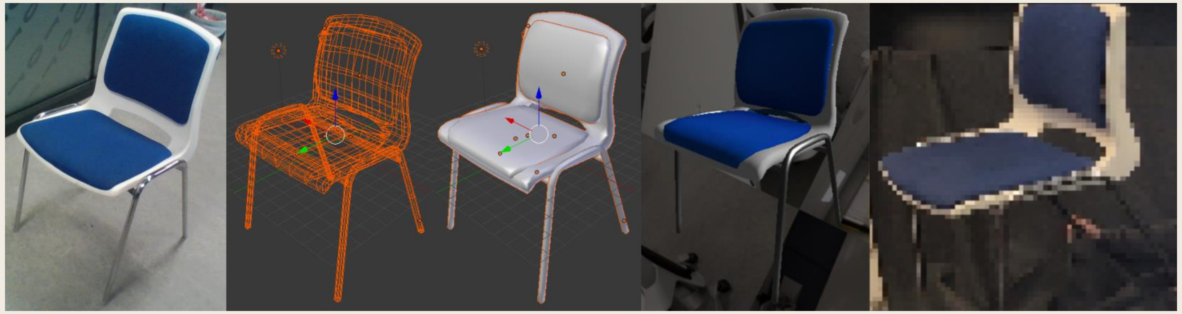
- Python
- Modeling and rendering
  - Blender
  - ModernGL
- Orientation prediction
  - Keras
- Object detection
  - OpenCV
  - YOLO v3

Our case

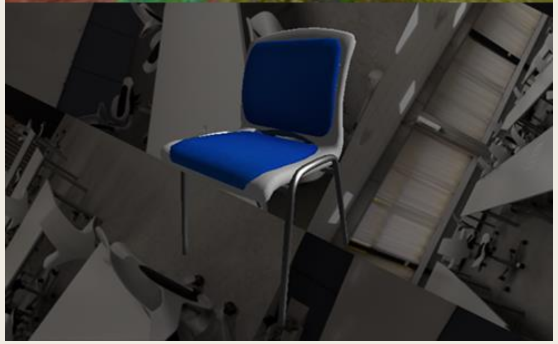
- We choose a chair as object to train on
- Recreated the chair with 3D-modeling software



Rendering

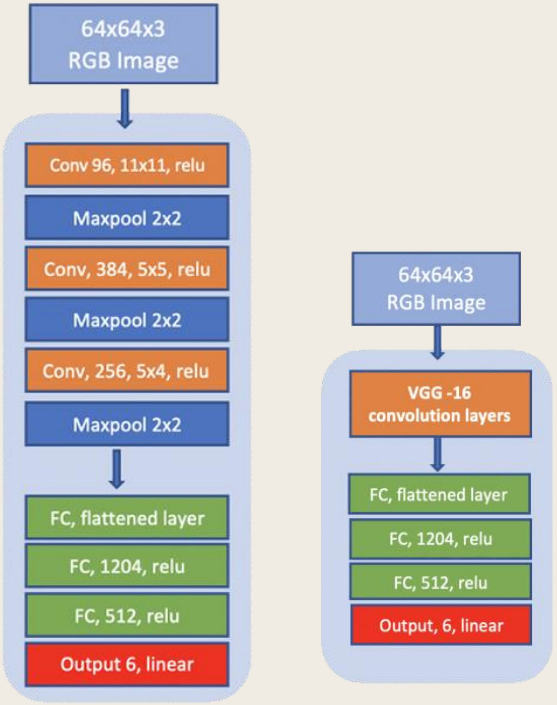


- Obtaining the training data
- ModernGL
- Blender
- Rendereing
- Describing orientation
- Background
- Visualization



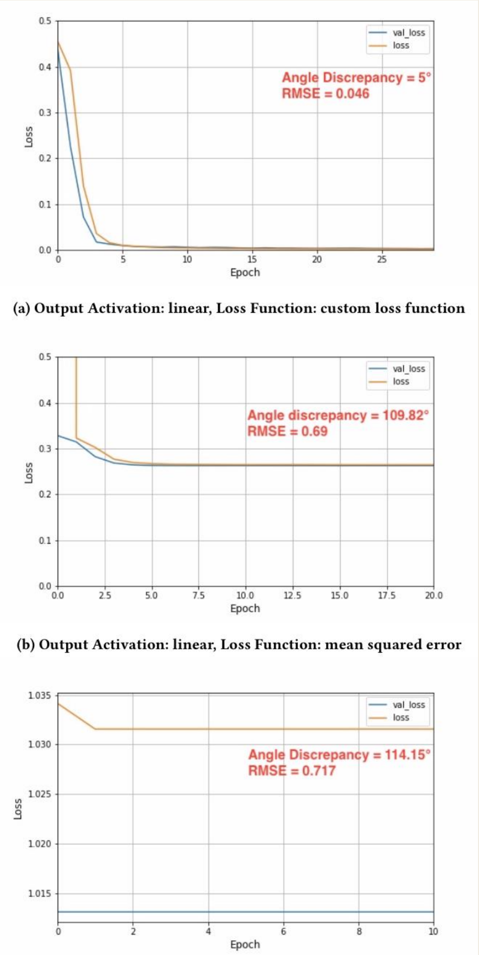
Orientation prediction

- Architechture
  - Simple CNN
  - VGG-16
  - ResNet-50



Customized Loss Function

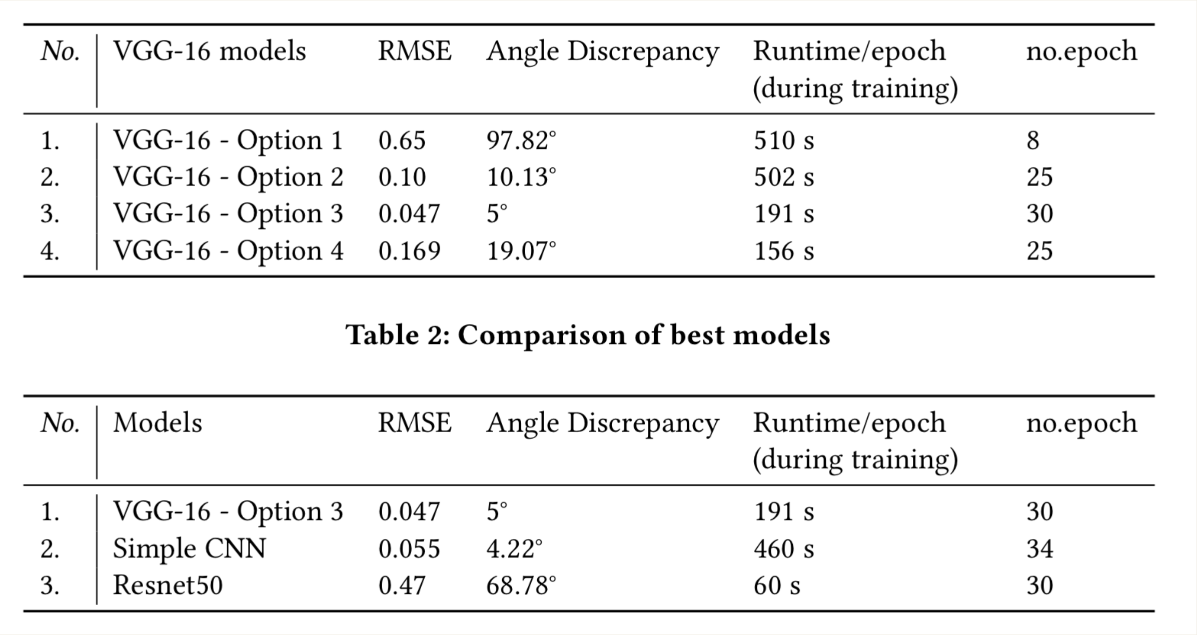
- Modification of a standard mean squared error function implemented on Tensorflow



(a) Output Activation: linear, Loss Function: custom loss function

(b) Output Activation: linear, Loss Function: mean squared error

Experiment results



No.	VGG-16 models	RMSE	Angle Discrepancy	Runtime/epoch (during training)	no.epoch
1.	VGG-16 - Option 1	0.65	97.82°	510 s	8
2.	VGG-16 - Option 2	0.10	10.13°	502 s	25
3.	VGG-16 - Option 3	0.047	5°	191 s	30
4.	VGG-16 - Option 4	0.169	19.07°	156 s	25

Table 2: Comparison of best models

No.	Models	RMSE	Angle Discrepancy	Runtime/epoch (during training)	no.epoch
1.	VGG-16 - Option 3	0.047	5°	191 s	30
2.	Simple CNN	0.055	4.22°	460 s	34
3.	Resnet50	0.47	68.78°	60 s	30

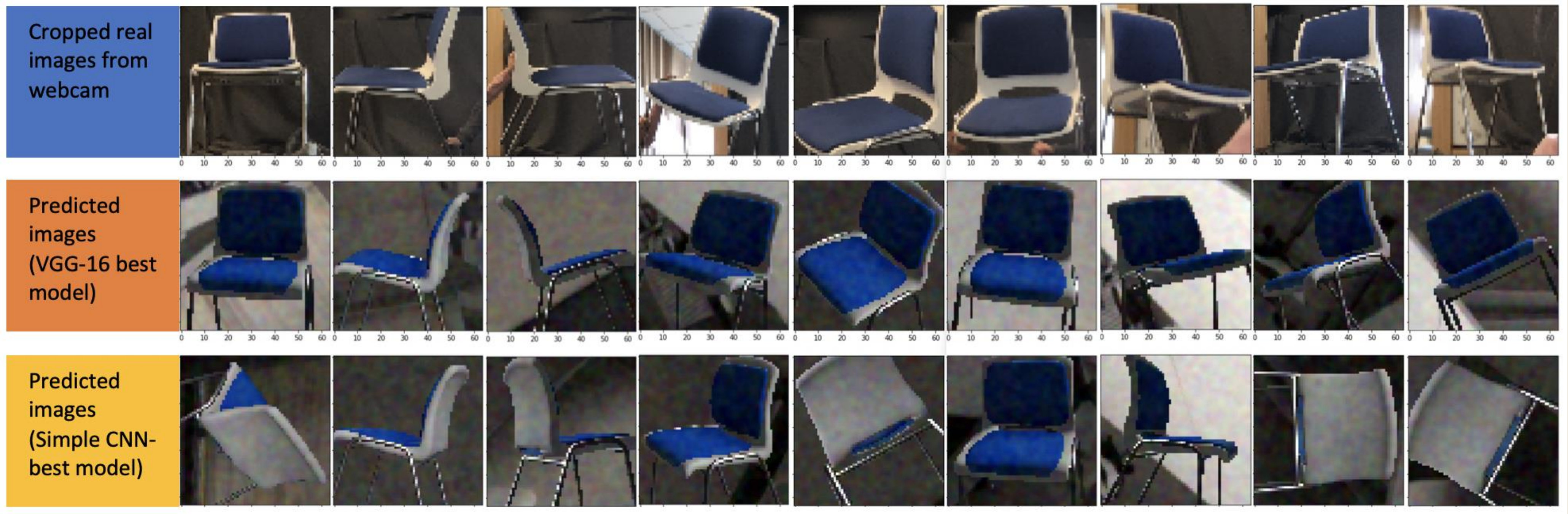
Object detection

- Background
- YOLO
- OpenCV

Conclusion

- To predict orientation using regression is possible, with modification of loss function
- Models sensitive to small differences between the synthetic images and the real test data
- The use of pre-trained weights is useful in our case to compensate for the lack of robustness of our model.

Comparison in Predicting With Real Rest Image



Cropped real images from webcam

Predicted images (VGG-16 best model)

Predicted images (Simple CNN-best model)