Ongoing Master Thesis (40ECs):

Using Deep Learning to Integrate Thermal and Visual Images for Human Detection in Autonomous Driving System

- **Objective:** to find out the best sensor fusion architecture with the approach of convolutional neural network (CNN) applied in human detection.
- **Methodology:** The architecture takes thermal and visual images as input (figure 1). It detects human in these images. The research aims at testing three different architectures and provide the most accurate architecture in human detection. The three different architectures are early fusion, halfway fusion and late fusion (figure 2). They are all based on the state-of-art **Convolutional Neural Network RetinaNet**.
- **Why I choose RetinaNet?** Because RetinaNet has a powerful loss function to solve class imbalance problem. Furthermore, RetinaNet uses a feature pyramid network to solve the other problem: usually high-resolution maps have low-level features. Retinanet architecture is in Figure 3.
- **Training Dataset**: Kaist dataset. Human have been annotated in both thermal and visual images.
- **Evaluation approach**: the best CNN model is determined by the best tradeoff between precision and recall and the largest average intersection over union.
- **Programming language**: Python, with Keras and Tensoflow.
- I enjoy working on this thesis!

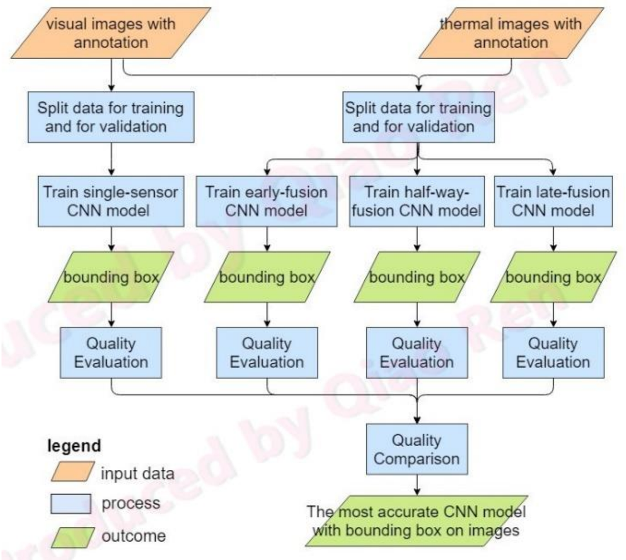


Figure 1 methodology

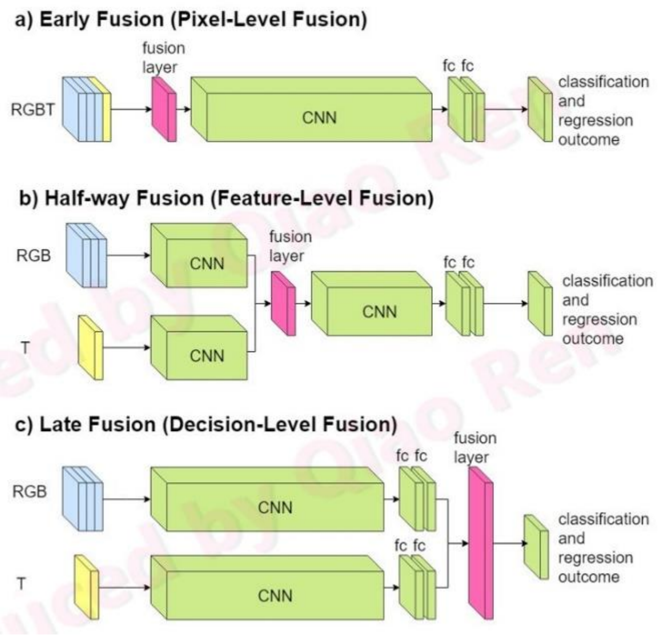


Figure 2 early, halfway and late fusion (RGB: visual images, T: Thermal images, fc: fully connected layer)

- **How Far I am:** I have successfully built and tested the following models
  - ✓ An early fusion architecture based on RetinaNet
  - ✓ A late fusion architecture based on RetinaNet
  - ✓ two single-sensor models finetuned by ImageNet dataset (figure 3): KaistT, KaistRGB
  - ✓ two single-sensor models finetuned by cocokitti dataset (figure 3): CocoKittikaistT, CocoKittiKaistRGB
  - ✓ have completed evaluation: precision-recall and average intersection over union, for all the above models
  - ✓ have completed trade-off analysis on precision and recall, for all the above models
  - ✓ have got a conclusion on which model has the best performance.

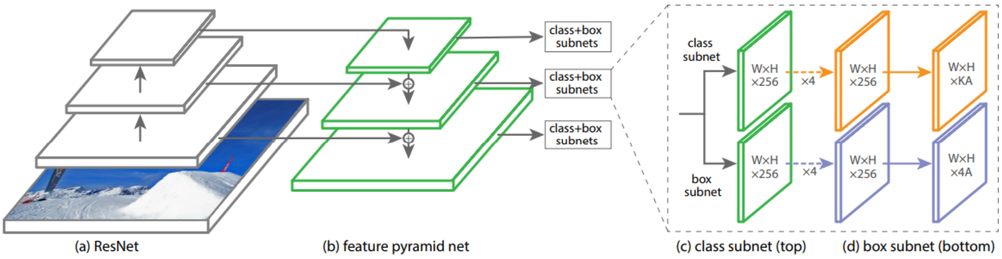


Figure 3 architecture of RetinaNet

- **Intermediate result**: (figure 4) Each rectangle represents a bounding box of a detected human. Blue bounding boxes are predicted by RetinaNet. Red bouding boxes are ground truth.
    - Click here to see human detection trained on thermal images using RetinaNet
    - Click here to see human detection trained on visual images using RetinaNet
    - Click here to see midterm presentation

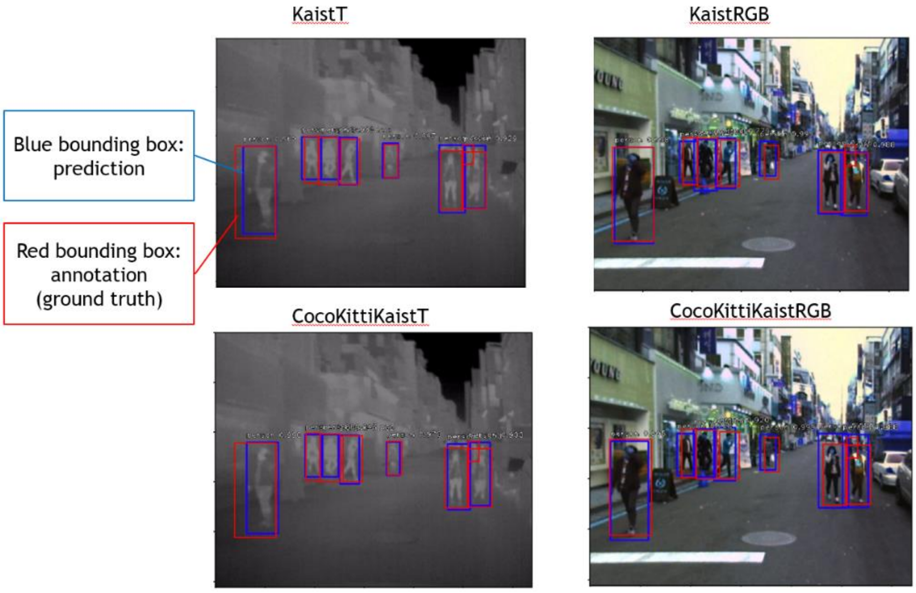


Figure 4 my intermediate result. Human detection predicted by four single-sensor models

- **Literature research**: I have written a very good literature review on traditional approaches and deep learning approaches for human detection in images. My supervisor likes it! See Figure 5. Important approaches: R-CNN, fast R-CNN, faster R-CNN, histogram of gradient (HoG), aggregated channel feature detector (ACF).

Using single sensor	Traditional approaches		HOG
			THOG
			Standard ACF
	Deep learning approaches	Two stage	R-CNN
			Fast R-CNN
			Faster R-CNN
		One stage	RatinaNet
			SSD[12]
			YOLO[13]
Using multiple sensors	Traditional approaches		ACF+T
			ACF+T+TM+TO
			ACF+T+THOG
			ACF+C+T
	Deep learning approaches	Multimodal applications	Applications in object detection
			Applications in human detection
		Fusion types	Pixel-level fusion
			Feature-level fusion
			Decision level fusion

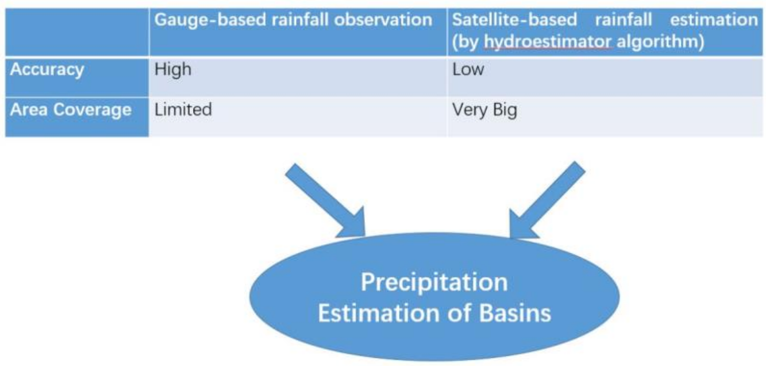
Figure 5 my literature research

- **Theory learning in Master Thesis**: convolutional layer, activation functions, learning process, loss function, test and validation, overfitting and underfitting, learning rate, batch size, Back propagation, finetuning, data augmentation, prediction with a neural network, supervised, unsupervised and semi-supervised learning, zero padding, max pooling, backpropagation, weight initialization, bias in an artificial neural network, learnable parameters in a convolutional neural network, RNN (recurrent neural network) and LSTM (long short term memory)

Completed Bachelor Thesis (19ECs):

Statistical Estimation of Daily Precipitation by Integrating a Satellite-Based Algorithm and Rain Gauges Network

- **Objective**: to generate daily precipitation estimations by integrating gauge precipitation observation and satellite precipitation estimates.
- **Background:** Use Rain Gauge measurement to correct the rainfall value that has been estimated by satellite algorithm. Because rain gauge measurement is more accurate than satellite estimation (Figure 5). The problem is what is the best approach to correct the bias. This is where the Combined Scheme plays a role.
- **Methodology:** Combined Scheme **compares** the rainfall value corrected by two different approaches, **ratio-bias correction and additive bias correction**, and **makes decision** on the final precipitation value (Figure6). Two cases have been implemented: Latin America case and Africa case. Click here to see methodology.
- **Programming language**: python with ILWIS 4 extension
- **Output**: the Latin America and Africa cases both have the following output:
    - Four precipitation maps:
        - Satellite-based precipitation corrected by addictive -bias correction
        - Satellite-based precipitation corrected by ratio -bias correction
        - Satellite-based precipitation corrected by combined-algorithm correction, without using filter
        - Satellite-based precipitation corrected by combined-algorithm correction, using filter
    - Statistics result for basin management
- **Results:** here



	Gauge-based rainfall observation	Satellite-based rainfall estimation (by hydroestimator algorithm)
Accuracy	High	Low
Area Coverage	Limited	Very Big

Figure 5 advantage and disadvantage of gauge observation and satellite estimation

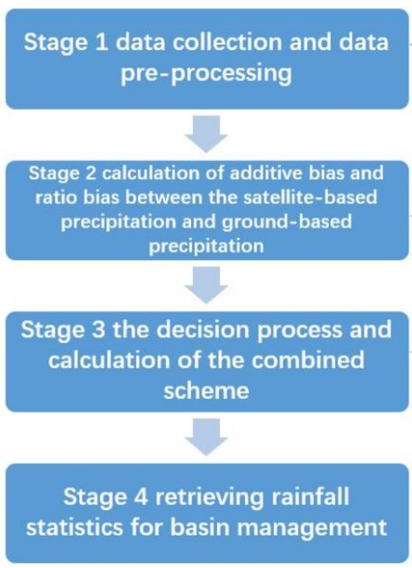


Figure 6 methodology