Completed Master Thesis (40ECs):

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

- **Keywords:** *deep learning, CNN, multi sensor integration, human detection, thermal images, RGB images*
- **Introduction and Applications**: Human detection is important in a wide range of applications nowadays. It has been applied in autonomous driving, human-robotics interaction, victim rescuing and automated surveillance.
- **Objective:** This research aims at finding out the best architecture with the approach of **convolutional neural network** applied in human detection by integrating thermal and visual images.
- **Methodology:** The convolutional neural network RetinaNet is a state-of-art approach. This research implements two multi-sensor fusion models and four single sensor models (Figure 1), based on RetinaNet. Multi-sensor fusion models are early-fusion and late-fusion (Figure 2). The early fusion model integrates thermal and visual images on pixel level. The late fusion model integrates the predictions that are provided by single-sensor models.
- **Results:** The result shows that late fusion model has significantly improves the performance of single sensor models. The improvement made by late fusion model is 7.1 %, 3.8%, 11.1% and 8.1%, compared with non-finetuned thermal model kaistT, finetuned thermal model cocokittikaistT, non-finetuned RGB model kaist RGB and finetuned RGB model cocokittikaistRGB respectively. Early fusion model does make improvement. Therefore, it is concluded that late-fusion model based on RetinaNet is the best approach in human detection.
- **Result**: Figure 4 shows the example of bounding boxes predicted by six RetinaNet models in human detection. Red bounding boxes are ground truth. Non-red bounding boxes are predicted by models.
  - Watch this video to see human detection trained on thermal images using fine-tuned CNN RetinaNet
  - Watch this video to see human detection trained on visual images using fine-tuned CNN RetinaNet
  - Watch this video to see human detection using fine-tuned late fusion CNN RetinaNet. White bounding boxes are predicted by late fusion CNN model. Blue and yellow bounding boxes are predicted by single-sensor CNN.
  - See Presentation
  - See Proposal
  - Click here to see evaluation results of single sensor models.

- **Publications:**
  - Paper of multi-sensor fusion models has been submitted to **a conference: the 22nd International Conference On Information Fusion.** Ottawa, Canada.
  - Paper of single-sensor fusion models has been accepted in **a conference: the 13th Intelligent Transport Systems European Congress.** Eindhoven, Netherlands. It will be published in April, 2019.

- **Programming language**: Python, with Keras and Tensoflow.
- **Training Dataset**: Kaist dataset. Coco dataset. Kitti dataset. Human have been annotated in both thermal and visual images.
- **Evaluation approach**: precision, recall, F1 score (which indicates the balance between precision and recall), miss rate, false positive per image.
- **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.
- I enjoy working on this thesis!

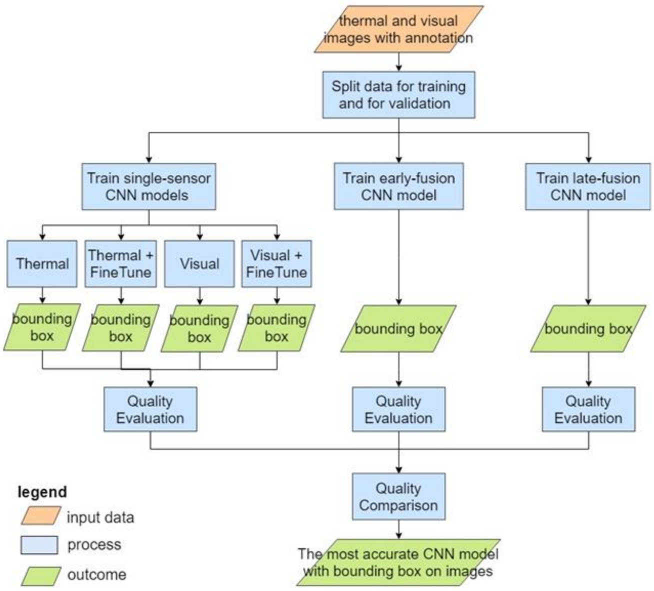

Figure 1 methodology

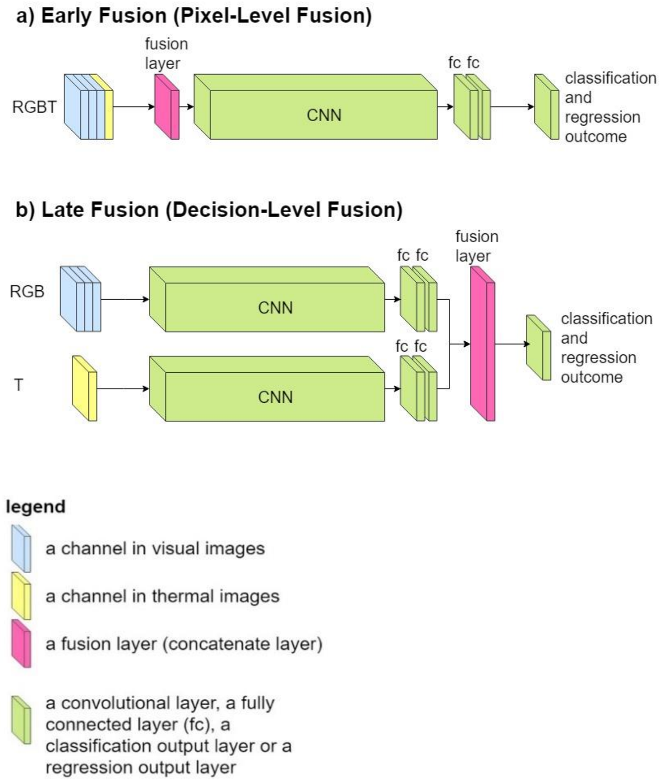

Figure 2 early and late fusion (RGB: visual images, T: Thermal images, fc: fully connected layer)

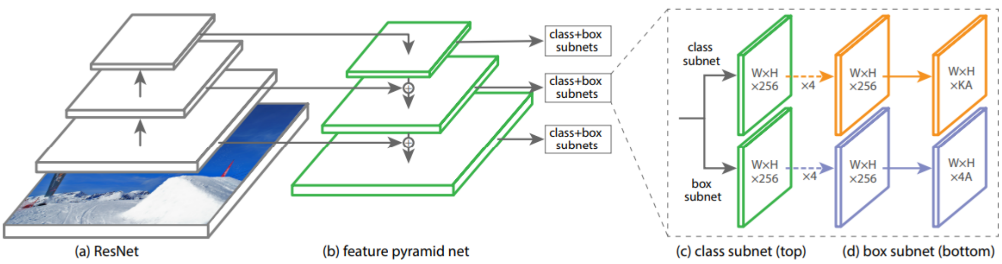

Figure 3 architecture of convolutional neural network RetinaNet

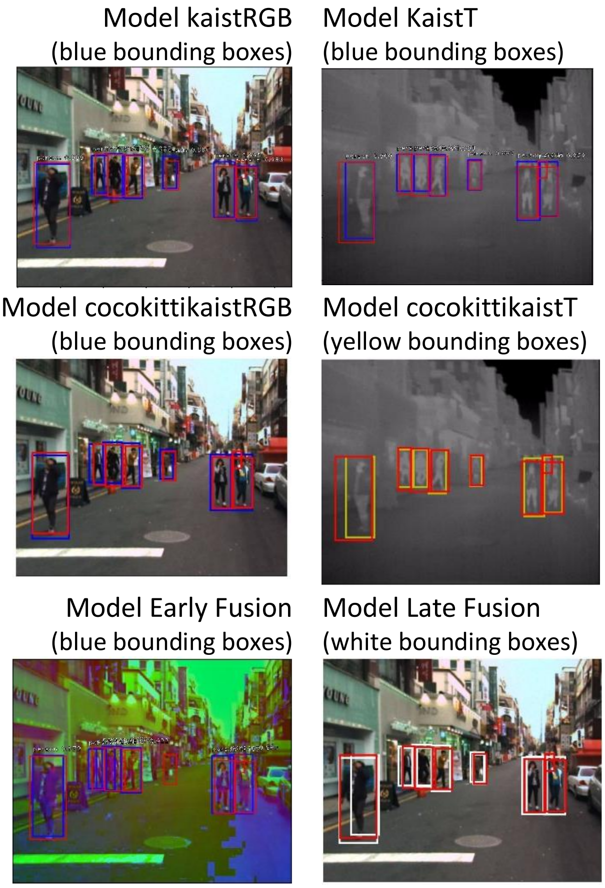

Figure 4 Human detection predicted by four single-sensor models and two fusion models (early fusion and late fusion). Red bounding boxes are ground truth. Non-red bounding boxes are predicted by models.

- **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

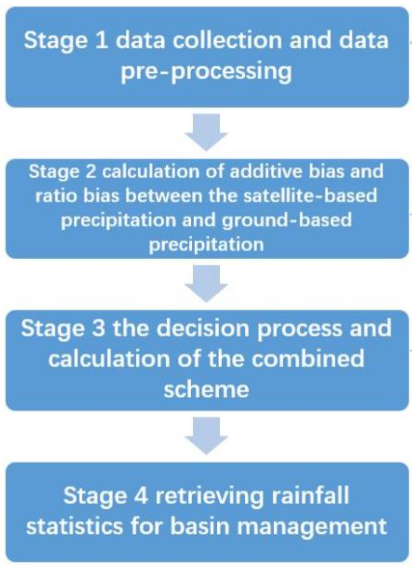

Figure 6 methodology