Incremental Cross-Modality Deep Learning for Pedestrian Recognition

Key Words : Stereo Vision, Deep Learning, Pedestrian Recognition.

Supervisor : Pr A. Bensrhair (abdelaziz.bensrhair@insa-rouen.f)
LITIS laboratory, INSA Rouen Normandie
685 Av de l'Université
76800 Saint Etienne du Rouvray
France

In spite of the large number of existing methods, pedestrian detection remains an open challenge. In recent years, deep learning classification methods combined with multimodality images within different fusion schemes have achieved the best performance. It was proven that the late-fusion scheme outperforms both direct and intermediate integration of modalities for pedestrian recognition. Hence, in this thesis, we focus on improving the late-fusion scheme for pedestrian classification on the Daimler stereo vision data set.
Each image modality, Intensity, Depth and Flow, is classified by an independent Convolutional Neural Network (CNN), the outputs of which are then fused by a Multi-layer Perceptron (MLP) before the recognition decision.
In this Thesis, we propose different methods based on Cross-Modality deep learning of CNNs:
(1) a correlated model where a unique CNN is trained with Intensity, Depth and Flow images for each frame,
(2) an incremental model where a CNN is trained with the first modality images frames, then a second CNN, initialized by transfer learning on the first one is trained on the second modality images frames,
(3) and finally a third CNN initialized on the second one, is trained on the last modality images frames.

The purpose of this work is :

1. The showing that the incremental cross-modality deep learning of CNNs improves classification performances not only for each independent modality classifier, but also for the multi-modality classifier based on late-fusion.
2. The learning algorithms investigation.
3. The improvement of models with new CNN architecture by creating a new neural network, the performance of which will not be hindered by complex computing resources. We will concentrate on improving that model using optimal settings for different training modality.

References:

Miron, Alina; Rogozan, Alexandrina; Ainouz, Samia, Bensrhair Abdelaziz et Broggi Alberto "An Evaluation of the Pedestrian Classification in a Multi-Domain Multi-Modality Setup" SENSORS Volume: 15 Issue: 6 Pages: 13851-13873 Published: JUN 2015.

Pierre Sermanet, Koray Kavukcuoglu, Soumith Chintala, and Yann Lecun. Pedestrian detection with unsupervised multi-stage feature learning. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2013.

Jan Hosang, Mohamed Omran, Rodrigo Benenson, and Bernt Schiele. Taking a deeper look at pedestrians. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.

H. Fukui, T. Yamashita, Y. Yamauchi, H. Fujiyoshi, and H. Murase. Pedestrian detection based on deep convolutional neural network with ensemble inference network. In 2015 IEEE Intelligent Vehicles Symposium (IV), pages 223–228, June 2015.

Anelia Angelova, Alex Krizhevsky, and Vincent Vanhoucke. Pedestrian detection with a large-field-of-view deep network. In IEEE International Conference on Robotics and Automation, ICRA 2015, Seattle, WA, USA, 26-30 May, 2015, pages 704–711, 2015.

Rodrigo Benenson, Mohamed Omran, Jan Hosang, and Bernt Schiele. Ten Years of Pedestrian Detection, What Have We Learned?, pages 613–627. Springer International Publishing, Cham, 2015.

Piotr Dollar, Christian Wojek, Bernt Schiele, and Pietro Perona. Pedestrian detection: An evaluation of the state of the art. IEEE Trans. Pattern Anal. Mach. Intell., 34(4):743–761, April 2012.

M. Eisenbach, D. Seichter, T. Wengefeld, and H. M. Gross. Cooperative multi-scale convolutional neural networks for person detection. In 2016 International Joint Conference on Neural Networks (IJCNN), pages 267–276, July 2016.

Xiaogang Chen, Pengxu Wei, Wei Ke, Qixiang Ye, and Jianbin Jiao. Pedestrian Detection with Deep Convolutional Neural Network, pages 354–365. Springer International Publishing, Cham, 2015.

J¨orgWagner, Volker Fischer, Michael Herman, and Sven Behnke. Multispectral pedestrian detection using deep fusion convolutional neural networks. In 24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), pages 509–514, April 2016.

Danut¸ Ovidiu Pop, Alexandrina Rogozan, Fawzi Nashashibi, and Abdelaziz Bensrhair. Fusion of stereo vision for pedestrian recognition using convolutional neural networks. Proceedings of the European Sympoisum on Artificial Neural Networks (ESANN), 2017.

C. Karaoguz and A. Gepperth. Incremental learning for bootstrapping object classifier models. In 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pages 1242–1248, Nov 2016.