# 5 RELATED WORK

For object verification, association or re-identification tasks, deep learning approaches and Siamese architectures were first explored in face domain instead of vehicle domain. DDML [9] used a Siamese network with a generalized logistic loss function to learn a Mahalanobis metric to obtain robust feature representations for different face images. Yi et al [19] proposed a Siamese architecture with Cosine layer to compute similarity between two features and adopts binomial deviance cost function for network training. Similarly, Ahmed et al [1] adopted a Siamese network followed by a cross-input neighborhood difference layer and a summarize layer to describe the differences of two feature vectors for person re-identification problems. Other variations on the Siamese architectures have also been extensively explored recently. In FaceNet [15], a triplet loss function was proposed for robust similarity metric learning between a triplet of images. Ding et al [5] proposed a similar approach where a 3-branch deep architecture was adopted along with the triplet loss function to address person re-identification tasks.

More recently, researchers have begun to explore deep learning methods in the vehicle domain. DRDL [11] proposed a two-branch deep neural network model with coupled cluster loss function that is inspired by triplet loss and incorporate each vehicle's attribute labels to learn a robust metric. Furthermore, studies have shown that besides pixels values, spatial or temporal information is also valuable for vehicle verification tasks. Wang et al [18] used deep neural networks to extract features from 20 keypoints from a vehicle and aggregated them to construct a discriminative feature vector and refer to spatial and temporal information of each vehicle to further verify the vehicles' similarity. Liu et al [12] proposed PROVID, in which appearance features resulted from a deep convolutional neural network (CNN), license plate difference measured by a Siamese network and spatial-temporal information are fused to address vehicle verification problems. Shen et al [16] adopted a chain MRF model with a deeply learned pair-wise potential function to generate visual-spatial-temporal path proposals, which are further evaluated by a Siamese-CNN+Path-LSTM model to obtain similarity scores between pairs of query vehicle images.

# 6 CONCLUSIONS

In this paper we propose an approach for vehicle verification under the framework of connected vehicle systems. The proposed deep learning architectures generalize well and achieve high verification accuracy on real world dataset of reasonable complexity. The results shown in this paper is only preliminary and more experiments are expected to fully utilized the kinematic information besides pixel information of vehicle's profile image to further improve the association accuracy under more complicated circumstances. The high accuracy and low dimensionality of the feature representation indicate the potential deployment in the real time systems. For future work, it is worth exploring in real time scenarios by deploying the deep learning architectures to connected vehicle systems and evaluating communication performances in terms of transmission latency and bandwidth requirement.

# ACKNOWLEDGMENTS

I would like to thank my advisor Marco Gruteser and anonymous reviewers for their valuable comments. This material is based in part upon work supported by the National Science Foundation under Grant No. CNS-1329939.

# REFERENCES

[1] Ejaz Ahmed, Michael Jones, and Tim K Marks. 2015. An improved deep learning architecture for person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 3908–3916.
[2] Sumit Chopra, Raia Hadsell, Yann LeCun, et al. 2005. Learning a similarity metric discriminatively, with application to face verification. In *CVPR (1)*. 539–546.
[3] Soteris Demetriou, Puneet Jain, and Kyu-Han Kim. 2018. Codrive: Improving automobile positioning via collaborative driving. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 72–80.
[4] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 248–255.
[5] Shengyong Ding, Liang Lin, Guangrun Wang, and Hongyang Chao. 2015. Deep feature learning with relative distance comparison for person re-identification. *Pattern Recognition* 48, 10 (2015), 2993–3003.
[6] Sae Fujii, Atsushi Fujita, Takaaki Umedu, Shigeru Kaneda, Hirozumi Yamaguchi, Teruo Higashino, and Mineo Takai. 2011. Cooperative vehicle positioning via V2V communications and onboard sensors. In *Vehicular Technology Conference (VTC Fall), 2011 IEEE*. IEEE, 1–5.
[7] Raia Hadsell, Sumit Chopra, and Yann LeCun. 2006. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, Vol. 2. IEEE, 1735–1742.
[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.
[9] Junlin Hu, Jiwen Lu, and Yap-Peng Tan. 2014. Discriminative deep metric learning for face verification in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1875–1882.
[10] Hansi Liu, Pengfei Ren, Shubham Jain, Mohannad Murad, Marco Gruteser, and Fan Bai. 2019. FusionEye: Perception Sharing for Connected Vehicles and its Bandwidth-Accuracy Trade-offs. In *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON) (to appear)*. IEEE, 1–9.
[11] Hongye Liu, Yonghong Tian, Yaowei Yang, Lu Pang, and Tiejun Huang. 2016. Deep relative distance learning: Tell the difference between similar vehicles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2167–2175.
[12] Xinchen Liu, Wu Liu, Tao Mei, and Huadong Ma. 2016. A deep learning-based approach to progressive vehicle re-identification for urban surveillance. In *European Conference on Computer Vision*. Springer, 869–884.
[13] Hang Qiu, Fawad Ahmad, Fan Bai, Marco Gruteser, and Ramesh Govindan. 2018. AVR: Augmented Vehicular Reality. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services (Mobisys)* (2018-01-01) *(MobiSys '18)*. ACM, Munich, Germany, 81–95. https://doi.org/10.1145/3210240.3210319
[14] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 779–788.
[15] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 815–823.
[16] Yantao Shen, Tong Xiao, Hongsheng Li, Shuai Yi, and Xiaogang Wang. 2017. Learning deep neural networks for vehicle re-id with visual-spatio-temporal path proposals. In *Proceedings of the IEEE International Conference on Computer Vision*. 1900–1909.
[17] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. 2014. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1701–1708.
[18] Zhongdao Wang, Luming Tang, Xihui Liu, Zhuliang Yao, Shuai Yi, Jing Shao, Junjie Yan, Shengjin Wang, Hongsheng Li, and Xiaogang Wang. 2017. Orientation invariant feature embedding and spatial temporal regularization for vehicle re-identification. In *Proceedings of the IEEE International Conference on Computer Vision*. 379–387.
[19] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z Li. 2014. Deep metric learning for person re-identification. In *2014 22nd International Conference on Pattern Recognition*. IEEE, 34–39.