International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

GAN-Based Day-to-Night Image Style Transfer for Nighttime Vehicle Detection

DENNIS MICHAEL D , KESAVAN R

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTING

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited with Grade "A" by NAAC
JEPPIAAR NAGAR, RAJIV GANDHISALAI,
CHENNAI - 600119

ABSTRACT

THE major cause of traffic accidents is mainly due to improper following distance and distracted driving. The most critical function in the advanced driver assistance systems (ADAS) and autonomous vehicles is vehicle detection. One expects that vehicles around host driver could be detected as accurately as possible by an ADAS all day, including day and night. However, vehicle's appearance at daytime is quite different from its counterpart at nighttime. The existing was AugGAN, a GAN based data augmenter which could transform on-road driving images to a desired domain while image-objects would be well preserved. We quantitatively evaluate different methods by training Faster R-CNN and YOLO with datasets generated from the transformed results and demonstrate significant improvement on the object detection accuracies by using the proposed AugGAN model. In this paper, for images in source domain to be properly translated to the target one while image-objects are well-preserved. We will try to explicitly vector in order to gain multi-modality in performing unpaired image - to – image translation

CHAPTER 1

INTRODUCTION

THE major cause of traffic accidents is mainly due to improper following distance and distracted driving. The most critical function in the advanced driver assistance systems (ADAS) and autonomous vehicles is vehicle detection. One expects that vehicles around host driver could be detected as accurately as possible by an ADAS all day, including day and night. However, vehicle's appearance at daytime is quite different from its counterpart at nighttime. The existing was AugGAN, a GAN based data augmenter which could transform on-road driving images to a desired domain while image-objects would be well preserved. We quantitatively evaluate different methods by training Faster R-CNN and YOLO with datasets generated from the transformed results and demonstrate significant improvement on the object detection accuracies by using the proposed AugGAN model. In this paper, for images in source domain to be properly translated to the target one while image-objects are well-preserved. We will try to explicitly vector in order to gain multi-modality in performing unpaired image - to – image translation.

CHAPTER 2

LITERATURE SURVEY

Che-Tsung Lin received the B.S. degree in mechanical engineering from the National Taiwan University of Science and Technology, Taipei, Taiwan, in 2003, and the M.S. degree from the Institute of Applied Mechanics, National Taiwan University, Taipei, in 2005. He is currently pursuing the Ph.D. degree with the Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan. He was a Visiting Scholar with the Department of Computer Science, University of California, Santa Barbara, CA, USA, in 2013. He has been working with the Mechanical and Mechatronics System Research Laboratory, Intelligent Vehicle Division, Safety Sensing and Control Department, Industrial Technology Research Institute, Hsinchu, since 2006, where he is currently a Researcher in the field of intelligent transportation system, intelligent vehicle, and advanced driver assistance system. His research interests include computer vision, machine learning, deep learning, and generative adversarial networks and their applications in on-road object detection.

Sheng-Wei Huang received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2017. He was with Prof. Shang-Hong Lai as a Research Assistant at the Department of

Computer Science, National Tsing Hua University. His research interests in computer vision include semantic segmentation and generative adversarial networks.

Yen-Yi Wu received the B.S. degree in computer science from the University of Taipei, Taipei, Taiwan, in 2017, and the M.S. degree in computer science from National Tsing Hua University, Hsinchu, Taiwan, in 2019. Her research interests in computer vision include generative adversarial networks, domain adaption, and object detection.

Shang-Hong Lai received the Ph.D. degree in electrical and computer engineering from the University of Florida, Gainesville, in 1995. Then, he joined the Siemens Corporate Research, Princeton, New Jersey, as a member of technical staff. Since 1999, he became a faculty member with the Department of Computer Science, National Tsing Hua University (NTHU), Taiwan, where he is currently a Professor. Since 2018, he has been on leave from NTHU to join the Microsoft AI R&D Center, Taipei, as a Principal Researcher. He has authored more than 200 articles published in the related international journals and conferences. His research interests include computer vision, image processing, and machine learning. He has been a member of program committee of several international conferences, including CVPR, ICCV, ECCV, ACCV, ICPR, PSIVT, and ICME. He has been an Associate Editor or a Guest Editor for several international journals, including Pattern Recognition, Journal of Signal Processing Systems, Journal of Visual Communication and Image Representation, and IPSJ Transactions on Computer Vision and Applications.

2.1 INFERENCES FROM LITREATURE SURVEY

In this review, we present results from published studies that applied machine learning to the diagnosis and differential diagnosis of PD. Since the number of included papers was relatively large, we focused on a high-level summary rather than a detailed description of methodology and direct comparison of outcomes of individual studies. We also provide an overview of sample size, data source and data type, for a more in-depth understanding of methodological differences across studies and their outcomes. Furthermore, we assessed (a) how large the participant pool/dataset was, (b) to what extent new data (i.e., unpublished, raw data acquired from locally recruited human participants) were collected and used, (c) the feasibility of machine learning and the possibility of introducing new biomarkers in the diagnosis of PD. Overall, *methodology* studies that proposed and tested novel technical approaches (e.g., machine learning and deep learning models, data acquisition devices, and feature extraction algorithms) have repetitively shown that features extracted from data modalities including voice recordings and handwritten patterns could lead to high patient-level diagnostic performance, while facilitating accessible and non-invasive data acquisition. Nevertheless, only a small number of studies further validated these technical approaches in *clinical* settings using local human participants recruited specifically for these studies, indicating a gap between model development and their clinical applications.

2,2 OPEN PROBLEMS IN EXISTING SYSTEM

In this existing method, with the advent of R-CNN, a sequence of two-stage detectors including Fast R-CNN, Faster R-CNN, R-FCN, MS-CNN, etc., continuously achieve better accuracy. YOLO regards object detection as a single regression problem as how CNNs are applied for image classification. Then, a multi-scale version of the one-stage detector, SSD, demonstrates significant improvements. The contribution of this work is three fold: we design a structure-unpaired image to image translation network which learns the latent data transformation across different domains while artifacts in the transformed images are reduced.

Disadvantages / Problem Statement:

- Keep pushing the limits of object detection.
- Ill-posed unsupervised image-to-image translation problem.

Proposed Method:

In this proposed method, we will try explicitly encode random noise vector to our structure-aware latent vector in order to gain multi-modality in performing unpaired image-to-image translation, such as day-to-night, while imageobjects are still well-preserved. This way, a nighttime vehicle detector could learn to better detect vehicles under different degrees of ambient light in the same domain.

Advantages / Solution Statement:

- Accuracy improvement by introducing segmentation in regularizing the image translation phase.

Best quantitative result is achieved by CNN

CHAPTER 3
REQUIREMENT ANALYSIS

3.1 FEASIBILITY STUDIES/RISK ANALYSIS OF THE PROJECT

3.1 FEASIBILITY STUDIES

The feasibility of the project is analyzed in this phase and business proposal is put forth N with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential. Three key considerations involved in the feasibility analysis are

i. Economical Feasibility

ii. Technical Feasibility

iii. Social Feasibility

i. Economic Feasibility

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

ii. Technical Feasibility

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

iii. Social Feasibility

The aspect of study is to check the level of acceptance of the system by the user. This includes the

process of training the user to use the system efficiently. The user must not feel
threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

3.2 SOFTWARE REQUIREMENTS SPECIFICATION DOCUMENT

3.2.1 REQUIREMENT ANALYSIS

Requirements are a feature of a system or description of something that the system is capable of doing in order to fulfil the system's purpose. It provides the appropriate mechanism for understanding what the customer wants, analyzing the needs assessing feasibility, negotiating a reasonable solution, specifying the solution unambiguously, validating the specification and managing the requirements as they are translated into an operational system.

3.2.2 DEEP LEARNING:

Deep learning (also known as **deep structured learning**) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.[2]

Deep-learning architectures such as deep neural networks, deep belief networks, deep reinforcement learning, recurrent neural networks, convolutional neural networks and Transformers have been applied to fields including computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis, climate science, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.

Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems. ANNs have various differences from biological brains. Specifically, artificial neural networks tend to be static and symbolic, while the biological brain of most living organisms is dynamic (plastic) and analogue.

The adjective "deep" in deep learning refers to the use of multiple layers in the network. Early work showed that a linear perceptron cannot be a universal classifier, but that a network with a nonpolynomial activation function with one hidden layer of unbounded width can. Deep learning is a modern variation which is concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically informed connectionist models, for the sake of efficiency, trainability and understandability, hence the "structured" part.

3.2.3 MATLAB

MATLAB (an abbreviation of "MATrix LABoratory") is a proprietary multi-paradigm programming language and numeric computing environment developed by MathWorks. MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages.

Although MATLAB is intended primarily for numeric computing, an optional toolbox uses the MuPAD symbolic engine allowing access to symbolic computing abilities. An additional package, Simulink, adds graphical multi-domain simulation and model-based design for dynamic and embedded systems.

As of 2020, MATLAB has more than 4 million users worldwide.[They come from various backgrounds of engineering, science, and economics.

MATLAB was first released as a commercial product in 1984 at the Automatic Control Conference in Las Vegas. MathWorks, Inc. was founded to develop the software and the MATLAB programming language was released. The first MATLAB sale was the following year, when Nick Trefethen from the Massachusetts Institute of Technology bought ten copies.

By the end of the 1980s, several hundred copies of MATLAB had been sold to universities for student use. The software was popularized largely thanks to toolboxes created by experts in various fields for performing specialized mathematical tasks. Many of the toolboxes were developed as a result of Stanford students that used MATLAB in academia, then brought the software with them to the private sector.

Over time, MATLAB was re-written for early operating systems created by Digital Equipment Corporation, VAX, Sun Microsystems, and for Unix PCs. Version 3 was released in 1987. The first MATLAB compiler was developed by Stephen C. Johnson in the 1990s.

In 2000, MathWorks added a Fortran-based library for linear algebra in MATLAB 6, replacing the software's original LINPACK and EISPACK subroutines that were in C.] MATLAB's Parallel Computing Toolbox was released at the 2004 Supercomputing Conference and support for graphics processing units (GPUs) was added to it in 2010.

3.3.1 SOFTWARE REQUIREMENTS:

Operating System	Windows 7or later
Simulation Tool	MATLAB
Documentation	Ms – Office

3.3.2 HARDWARE REQUIREMENTS:

CPU type	I5
Ram size	4GB
Hard disk capacity	80 GB
Keyboard type	Internet keyboard
Monitor type	15 Inch colour monitor

CHAPTER 4

DESCRIPTION OF PROPOSED SYSTEM

In this proposed method, we will try explicitly encode random noise vector to our structure-aware latent vector in order to gain multi-modality in performing unpaired image-to-image translation, such as day-to-night, while imageobjects are still well-preserved. This way, a nighttime vehicle detector could learn to better detect vehicles under different degrees of ambient light in the same domain.

Advantages / Solution Statement:

- Accuracy improvement by introducing segmentation in regularizing the image translation phase.
- Best quantitative result is achieved by CNN.

MODULES DESCRIPTION

Module 1: Preprocessing

The images which are collected are subjected to pre- processing. In Pre-processing stage basic steps are image resizing and applying Median filters for a perfect input clear image for easy identification of an image. Pre-processed images will be segmented digitally into various pixels. We do this segmentation for an image is to modify its representation to have more clarity to analyze the images.

Module 2: Segmentation and feature extraction

Image segmentation is a method in which a digital image is broken down into various subgroups called Image segments which helps in reducing the complexity of the image to make further processing or analysis of the image simpler. Segmentation in easy words is assigning labels to pixels. Feature extraction is a part of the dimensionality reduction process, in which, an initial set of the raw data is divided and reduced to more manageable groups. So when you want to process it will be easier. The most important characteristic of these large data sets is that they have a large number of variables.

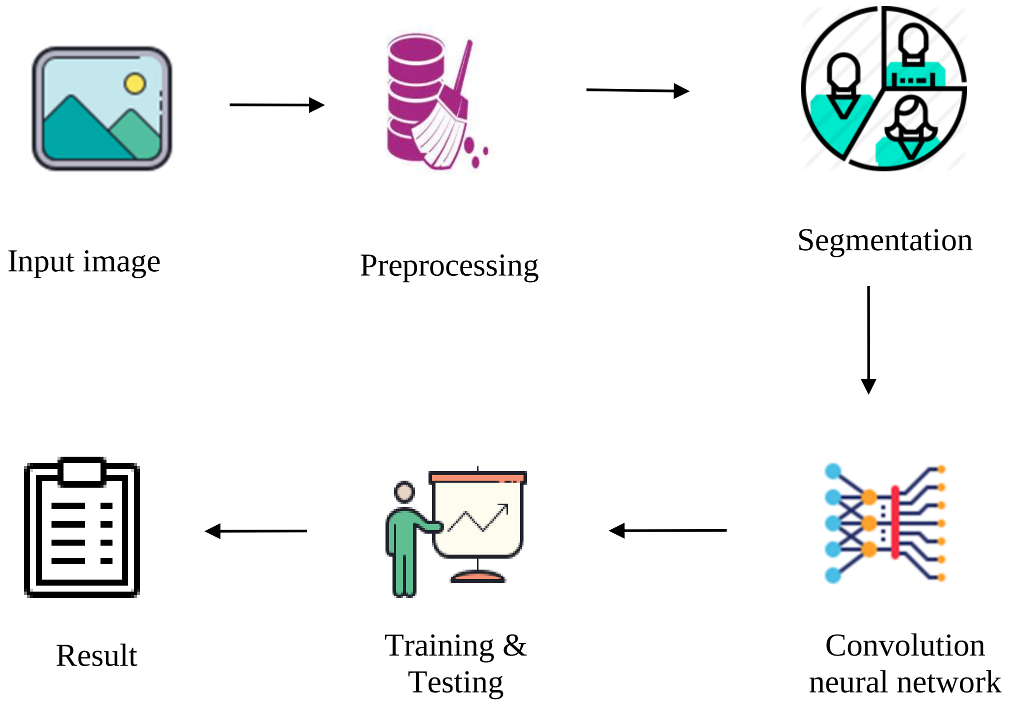
Module 3: Convolution neural network

A **Convolutional Neural Network (ConvNet/CNN)** is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

This module is used to establish the deep convolutional neural network concept for training the image and testing the image with the help of weight estimating classifier. The result image will compare with the dataset images and it will display whether it is detect or not.

4,1 SELECTED METHODOLOGY OR PROCESS MODEL



4,2 ARCHITECTURE / OVERALL DESIGN OF PROPOSED SYSTEM

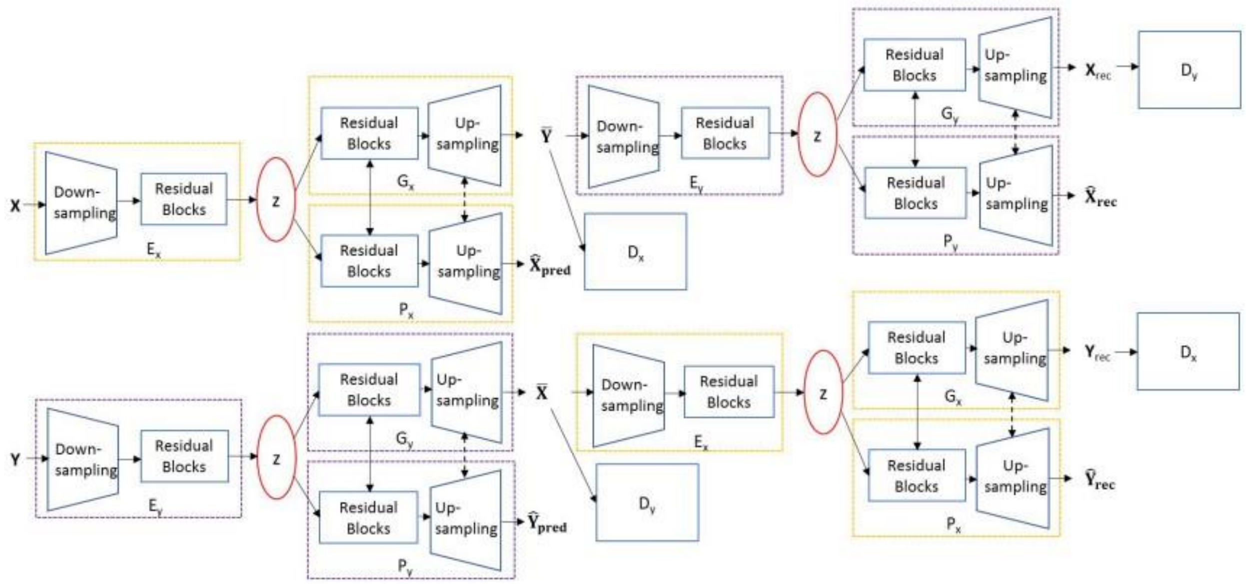


Fig 4.2: System Architecture for Drowsiness Judgment

The block diagram of the proposed system has been shown in the above figures. The camera captures the image of the person inside the car and sends that to the HOG model to train and it detects each feature from the face using facial landmark

4.3 DESCRIPTION OF SOFTWARE FOR IMPLEMENTATION AND TESTING PLAN OF THE PROPOSED MODEL/SYSTEM

4.3.1 TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub – assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the

Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

4.3.2 TYPES OF TESTS

1. UNIT TESTING

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

2. INTEGRATION TESTING

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

3. FUNCTIONAL TEST

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

: identified classes of application outputs must be exercised

Procedures : interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

4. SYSTEM TEST

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

5. WHITE BOX TESTING

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

6. BLACK BOX TESTING

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot "see" into it. The test provides inputs and responds to outputs without considering how the software works.

7. UNIT TESTING:

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

i Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

ii Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

. CHAPTER 5

RESULT & CONCLUSION

In this work, we proposed AugGAN, an unpaired image-toimage translation network for realizing domain adaptation in vehicle detection. Our method quantitatively surpasses competing methods for achieving higher nighttime vehicle detection accuracy because of better image-object preservation. Therefore, most daytime vehicle datasets in public domain become valuable in nighttime vehicle detector development. AugGAN is general in that it could also deal with synthetic-tosynthetic, synthetic-to-real, real-to-real, and real-to-synthetic transformations across different domains ranging from day, night, sunset, rain, etc. Currently, the major limitation of AugGAN is its uni-modality. In the future, we will try to explicitly encode random noise vector to our structure-aware latent vector in order to gain multi-modality in performing unpaired image-to-image translation, such as day-to-night, while imageobjects are still well-preserved. This way, a nighttime vehicle detector could learn to better detect vehicles under different degrees of ambient light in the same domain.

REFERENCES:-

[1] Z. Sun, G. Bebis, and R. Miller, “On-road vehicle detection: A review,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 5, pp. 694–711, May 2006.

[2] X. Wen, L. Shao, W. Fang, and Y. Xue, “Efficient feature selection and classification for vehicle detection,” IEEE Trans. Circuits Syst. Video Technol., vol. 25, no. 3, pp. 508–517, Mar. 2015.

[3] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, pp. 1627–1645, Sep. 2010.

[4] A. Takeuchi, S. Mita, and D. McAllester, “On-road vehicle tracking using deformable object model and particle filter with integrated likelihoods,” in Proc. IEEE IV Symp., Jun. 2010, pp. 1014–1021.

[5] H. T. Niknejad, T. Kawano, M. Shimizu, and S. Mita, "Vehicle detection using discriminatively trained part templates with variable size," in Proc. IEEE IV Symp, Jun. 2012, pp. 766–771..

[6] H. T. Niknejad, T. Kawano, Y. Oishi, and S. Mita, "Occlusion handling using discriminative model of trained part templates and conditional random field," in Proc. IEEE IV Symp, Jun. 2013, pp. 750–755.

[7] S. Sivaraman and M. M. Trivedi, "Vehicle detection by independent parts for urban driver assistance," IEEE Trans. Intell. Transp. Syst., vol. 14, no. 4, pp. 1597–1608, Dec. 2013.

[8] H. T. Niknejad, S. Mita, D. McAllester, and T. Naito, "Vision-based vehicle detection for nighttime with discriminately trained mixture of weighted deformable part models," in Proc. IEEE ITSC, Oct. 2011, pp. 1560–1565.

[9] H. Tehrani, T. Kawano, and S. Mita, "Car detection at night using latent filters," in Proc. IEEE IV Symp., Jun. 2014, pp. 839–844.

[10] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2014, pp. 580–587

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 779–788.

[12] M. Braun, S. Krebs, F. Flohr, and D. M. Gavrila, "The EuroCity persons dataset: A novel benchmark for object detection," 2018, arXiv:1805.07193. [Online]. Available: https://arxiv.org/abs/1805.07193

[13] I. Goodfellow et al., "Generative adversarial nets," in Proc. Adv. Neural Inf. Process. Syst., Dec. 2014, pp. 2672–2680.

[14] G. Ros, L. Sellart, J. Materzynska, D. Vazquez, and A. M. Lopez, "The SYNTHIA dataset: A large collection of synthetic images for semantic segmentation of urban scenes," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 3234–3243. [15] S. R. Richter, V. Vineet, S. Roth, and V. Koltun, "Playing for data: Ground truth from computer games," in Proc. Eur. Conf. Comput. Vis. (ECCV), Oct. 2016, pp. 102–118.

[16] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in Proc. Int. Conf. Pattern Recognit. (CVPR), Jun. 2012, pp. 3354–3361.

[17] S.-W. Huang, C.-T. Lin, S.-P. Chen, Y.-Y. Wu, P.-H. Hsu, and S.-H. Lai, "Auggan: Cross domain adaptation with gan-based data augmentation," in Proc. Eur. Conf. Comput. Vis. (ECCV), Sep. 2018, pp. 718–731.

[18] M. Cordts et al., "The cityscapes dataset," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. Workshops, vol. 1, no. 2, Jun. 2015, p. 3.

[19] R. Girshick, "Fast R-CNN," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 1440–1448.

[20] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards realtime object detection with region proposal networks," in Proc. Adv. Neural Inf. Process. Syst., Dec. 2015, pp. 91–99.

[21] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object detection via regionbased fully convolutional networks," in Proc. Adv. Neural Inf. Process. Syst., Dec. 2016, pp. 379–387.

[22] Z. Cai, Q. Fan, R. S. Feris, and N. Vasconcelos, “A unified multi-scale deep convolutional neural network for fast object detection,” in Proc. Eur. Conf. Comput. Vis. (ECCV), Oct. 2016, pp. 354–370.

[23] W. Liu et al., “SSD: Single shot multibox detector,” in Proc. Eur. Conf. Comput. Vis. (ECCV), Oct. 2016, pp. 21–37.

[24] J. Redmon and A. Farhadi, “YOLO9000: Better, faster, stronger,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 6517–6525.

[25] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The PASCAL visual object classes (VOC) challenge,” Int. J. Comput. Vis., vol. 88, no. 2, pp. 303–338, Sep. 2009.

[26] T.-Y. Lin et al., “Microsoft COCO: Common objects in context,” in Proc. Eur. Conf. Comput. Vis. (ECCV), Sep. 2014, pp. 740–755.

[27] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 5967–5976.

[28] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 2242–2251.

[29] T. Kim, M. Cha, H. Kim, J. Lee, and J. Kim, “Learning to discover cross-domain relations with generative adversarial networks,” 2017, arXiv:1703.05192. [Online]. Available: https://arxiv.org/abs/1703.05192

[30] Z. Yi, H. Zhang, T. Gong, P. Tan, and M. Gong, “DualGAN: Unsupervised dual learning for image-to-image translation,” in Proc. Int. Conf. Comput. Vis. (ICCV), Jul. 2017, pp. 2868–2876.

[31] M.-Y. Liu and O. Tuzel, “Coupled generative adversarial networks,” in Proc. Adv. Neural Inf. Process. Syst., Dec. 2016, pp. 469–477.

[32] M.-Y. Liu, T. Breuel, and J. Kautz, “Unsupervised image-to-image translation networks,” in Proc. Adv. Neural Inf. Process. Syst., Dec. 2017, pp. 700–708.

[33] S. Sivaraman and M. M. Trivedi, “A general active-learning framework for on-road vehicle recognition and tracking,” IEEE Trans. Intell. Transp. Syst., vol. 11, no. 2, pp. 267–276, Jun. 2010.

[34] Y. Zhou, L. Liu, L. Shao, and M. Mellor, “DAVE: A unified framework for fast vehicle detection and annotation,” in Proc. Eur. Conf. Comput. Vis. (ECCV), Oct. 2016, pp. 278–293.

[35] L. Yang, P. Luo, C. Change Loy, and X. Tang, “A large-scale car dataset for fine-grained categorization and verification,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 3973–3981.

[36] S. J. Nowlan and G. E. Hinton, “Simplifying neural networks by soft weight-sharing,” Neural Comput., vol. 4, no. 4, pp. 473–493, 1992.

[37] K. Ullrich, E. Meeds, and M. Welling, “Soft weight-sharing for neural network compression,” 2017, arXiv:1702.04008. [Online]. Available: https://arxiv.org/abs/1702.04008

[38] R. Caruna, “Multitask learning: A knowledge-based source of inductive bias,” in Proc. ICML, Jun. 1993, pp. 41–48. 0.

[39] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 3431–3440.