REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
03-11-2022	Final Report	21-May-2021 - 21-Aug-2022

4. TITLE AND SUBTITLE
Final Report: Information-Theoretical Transfer Learning for Bridging the Gap between Simulated and Real-World Data

5a. CONTRACT NUMBER
W911NF-21-1-0290

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Mississippi State University
133 Etheredge Hall, 449 Hardy Road

Mississippi State, MS 39762 -9662

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
78331-MI-II.3

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Bo Tang
UU	UU	UU	UU		19b. TELEPHONE NUMBER 662-325-8757

RPPR Final Report
as of 13-Mar-2023

Agency Code: 21XD

Proposal Number: 78331MIII **Agreement Number: W911NF-21-1-0290**
INVESTIGATOR(S):

Name: Bo Tang
Email: tang@ece.msstate.edu
Phone Number: 6623258757
Principal: Y

Organization: **Mississippi State University**
Address: 133 Etheredge Hall, 449 Hardy Road, Mississippi State, MS 397629662
Country: USA
DUNS Number: 075461814 EIN: 646000819

Report Date: 21-Nov-2022 Date Received: 03-Nov-2022
Final Report for Period Beginning 21-May-2021 and Ending 21-Aug-2022
Title: Information-Theoretical Transfer Learning for Bridging the Gap between Simulated and Real-World Data
Begin Performance Period: 21-May-2021 **End Performance Period:** 21-Aug-2022
Report Term: 0-Other

Submitted By: Bo Tang Email: tang@ece.msstate.edu
Phone: (662) 325-8757

STEM Degrees: 1 **STEM Participants:** 1

Major Goals: The major goals of this Short-Term Innovative Research (STIR) project are to develop fundamental machine learning algorithms that are effective and robust for domain adaptation and domain fusion in autonomous driving.

Accomplishments: During this one-year STIR project, the team has contributed two fundamental machine learning algorithms in image-to-image translation and data fusion, as well as collecting both simulated and real-world datasets, which are of great significance in autonomous driving.

1. Simulated data and real-world data collection for off-road autonomous driving: During this project, we have collected simulated autonomous driving data sets using the MSU Autonomous Vehicle Simulator (MAVS). This simulator is a physics-based sensor simulator for ground vehicle robotics that includes high-fidelity simulations of LiDAR, cameras, and several other sensors. Using MAVS, we simulated different off-road autonomous driving scenes with ground, trees, and vegetables under different weather conditions at different time of the day. In addition to simulated data, we also have collected real-world data from open-source benchmarks for off-road autonomous driving. The availability of these datasets are of great importance for PI's future research in image-to-image translation.

2. Image-to-image translation: In this image-to-image translation research, we have focused on the application of image deraining which enables the autonomous vehicles driving in heavy raining days. Most of existing image deraining algorithms require exact pairs of rainy and clean images. However, obtaining such paired images for the same natural scene is very cumbersome task. In this research, we propose a generative adversarial networks (GANs) based approach, called BiRelDerainGAN, which do not require any paired rainy and clear images during the training process. More specifically, the proposed approach is based on the relative information of two random input and output images which is determined on the basis of presence or absence of rain. Thus, our proposed method is able to generate clear images from rainy images and vice-versa, thus being bidirectional. Moreover, BiRelDerainGAN can generate the controlled derained images from rainy images based on interpolation concept. The qualitative and quantitative results of the proposed method are analyzed in detail and compared against state-of-the-art algorithms where our method performs quite favorably.

3. Data fusion: In this data fusion research, we have focused on the research question of how to effectively fuse camera images and LiDAR cloud points to improve the performance of semantic segmentation in autonomous driving. Mostly, the LiDAR and camera sensor-based data are being fused since they usually bear the

complementary information. Fusion of such sources helps to increase the segmentation performance. However, there still exists a performance gap due to their fusion mechanisms in terms of direct features concatenation or choice of architectures to learn the source specific features. In this research, we have developed a novel framework for the fusion of surface normal (SN) images calculated from depth/disparity maps from LiDAR scans with RGB camera images in order to learn effective features. While doing so, a two-step learning strategy is proposed such that the learning from one modal is continuously calibrated from the corresponding local and non-local attentions generated from the other modal. Furthermore, in order to overcome the degraded performance of segmentation due to unwanted artifacts around the class-boundaries, we propose to use the cycle spinning-based technique for training and testing of our proposed model. The use of such local and non-local attentions from one modal to another and cycle spinning-based scheme is the first time in the field of road segmentation. With extensive quantitative and qualitative results and ablation studies performed on KITTI dataset for road segmentation, we show that our proposed method performs very favorably against the state-of-the-art methods.

Training Opportunities: The PI has offered training sessions to graduate students in the topics of machine learning, tensorflow programming, data processing, and academic paper writing.

Results Dissemination: The PI has actively made efforts to disseminate our research results in top-tier IEEE journals:

1. Sharma, S., and Tang, B. BiRelDerainGAN: A Clear-Rain Relativeness-Based Bidirectional Unpaired Single Image Deraining Using Generative Adversarial Networks, IEEE Transactions on Image Processing, Under Review, 2022.

2. Sharma, S., Tang, B., Ball, J., Carruth, D., and Dabbiru, L. Local and Non-Local Cross Attention-Based Fusion for Semantic Segmentation with Cycle Spinning. IEEE Transactions on Circuits and Systems for Video Technology, Under Review, 2022.

Honors and Awards: The PI has received the prestigious NSF CAREER Award during this project period.

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: Graduate Student (research assistant)
Participant: Suvash Sharma
Person Months Worked: 10.00 **Funding Support:**
Project Contribution:
National Academy Member: N

ARTICLES:

RPPR Final Report
as of 13-Mar-2023

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 4-Under Review
Journal: IEEE Transactions on Circuits and Systems for Video Technology
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 8/16/22 12:00AM Date Published:
Publication Location:
Article Title: Local and Non-local Cross-Modal Attention-based Fusion for Semantic Road-Segmentation with Cycle Spinning
Authors: Suvash Sharma, Bo Tang, John Ball, Daniel Carruth, Lalitha Dabbiru
Keywords: Local attention, non-local attention, segmentation, cycle spinning, sensor fusion
Abstract: To improve the performance of semantic segmentation algorithms, information from multiple data sources are being used in-place-of a single data source. Mostly, the LiDAR and camera sensor-based data are being fused since they usually bear the complementary information. Fusion of such sources helps to increase the segmentation performance. However, there still exists a performance gap due to their fusion mechanisms in terms of direct features concatenation or choice of architectures to learn the source specific features. In this work, we propose a novel framework for the fusion of surface normal (SN) images calculated from depth/disparity maps from LiDAR scans with RGB camera images in order to learn effective features. While doing so, a two-step learning strategy is proposed such that the learning from one modal is continuously calibrated from the corresponding local and non-local attentions generated from the another modal.
Distribution Statement: 3-Distribution authorized to U.S. Government Agencies and their contractors
Acknowledged Federal Support: **Y**

Publication Type: Journal Article Peer Reviewed: Y **Publication Status:** 4-Under Review
Journal: IEEE Transactions on Image Processing
Publication Identifier Type: Publication Identifier:
Volume: Issue: First Page #:
Date Submitted: 11/3/22 12:00AM Date Published:
Publication Location:
Article Title: BiRelDerainGAN: A Clear-Rain relativeness-based bidirectional unpaired single image deraining using generative adversarial networks
Authors: Suvash Sharma, Bo Tang
Keywords: Image deraining, image generation, image-to-image translation, generative model
Abstract: Most of existing image deraining algorithms require exact pairs of rainy and clean images. However, obtaining such paired images for the same natural scene is very cumbersome task. In this research, we propose a generative adversarial networks (GANs) based approach, called BiRelDerainGAN, which do not require any paired rainy and clear images during the training process. More specifically, the proposed approach is based on the relative information of two random input and output images which is determined on the basis of presence or absence of rain. Thus, our proposed method is able to generate clear images from rainy images and vice-versa, thus being bidirectional. Moreover, BiRelDerainGAN can generate the controlled derained images from rainy images based on interpolation concept. The qualitative and quantitative results of the proposed method are analyzed in detail and compared against state-of-the-art algorithms where our method performs quite favorably.
Distribution Statement: 2-Distribution Limited to U.S. Government agencies only; report contains proprietary info
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 13-Mar-2023

Partners

,

I certify that the information in the report is complete and accurate:
Signature: Bo Tang
Signature Date: 11/3/22 11:08PM

Abstract: Artificial intelligence (AI)-enabled autonomous vehicles is one of the critical and enabling components in the next-generation Army Modernization Strategy (AMS) which will transform the total Army of the United States into a multi-domain force. However, training advanced AI models usually requires a large amount of labeled data. This presents a great challenge in their practical applications in autonomous driving, where labeled data is typically scarce and easily outdated. The objective of this research project is to address this limited supervision challenge by investigating effective transfer learning approaches using information theory to leverage simulation data to facilitate the learning in real data for autonomous driving. The proposed methods in this project advance the foundation of machine learning and autonomous driving. They are also critical to many DoD related tasks where soldiers or autonomous vehicles are deployed in a new environment without being trained well for scene understanding. The development of advanced systems with strong cognitive capabilities will lead to significant breakthroughs for enhancing the Army's mobility and survivability.

Objectives: While our originally proposed research is a 3-year project, it has been awarded as a 1-year Short-Term Innovative Research (STIR) project and we have achieved the following new objectives which lay down foundations for PI's future ARO projects: i) simulated data and real-world data collection for off-road autonomous driving; ii) development of novel image-to-image translation or domain adaptation methods for image deraining; iii) development of novel domain fusion methods for road segmentation.

Findings:

Objective I: Simulated data and real-world data collection for off-road autonomous driving.

During this project, we have collected simulated autonomous driving data sets using the MSU Autonomous Vehicle Simulator (MAVS). This simulator is a physics-based sensor simulator for ground vehicle robotics that includes high-fidelity simulations of LiDAR, cameras, and several other sensors. Using MAVS, we simulated different off-road autonomous driving scenes with ground, trees, and vegetables under different weather conditions at different time of the day. Some typical simulated images are shown in Fig. 1 and Fig. 2. In addition to simulated data, we also have collected real-world data from open-source benchmarks for off-road autonomous driving. Some typical real-world images are shown in Fig. 3. The availability of these datasets are of great importance for PI's future research in image-to-image translation.

Fig. 1. Some typical images from a synthetic dataset for simple scenes only with ground and trees.

Fig. 2. Some typical images from a synthetic dataset for high-definition scenes.

Fig. 3. Some typical images from a real-world off-road autonomous driving dataset.

Objective II: Development of novel image-to-image translation or domain adaptation methods for image deraining.

In this image-to-image translation research, we have focused on the application of image deraining which enables the autonomous vehicles driving in heavy raining days. Most of existing image deraining algorithms require exact pairs of rainy and clean images. However, obtaining such paired images for the same natural scene is very cumbersome task. In this research, we propose a generative adversarial networks (GANs) based approach, called BiRelDerainGAN, which do not require any paired rainy and clear images during the training process. More

specifically, the proposed approach is based on the relative information of two random input and output images which is determined on the basis of presence or absence of rain. Thus, our proposed method is able to generate clear images from rainy images and vice-versa, thus being bidirectional, as shown in Fig. 4. Moreover, BiRelDerainGAN can generate the controlled derained images from rainy images based on interpolation concept. The qualitative and quantitative results of the proposed method are analyzed in detail and compared against state-of-the-art algorithms where our method performs quite favorably.

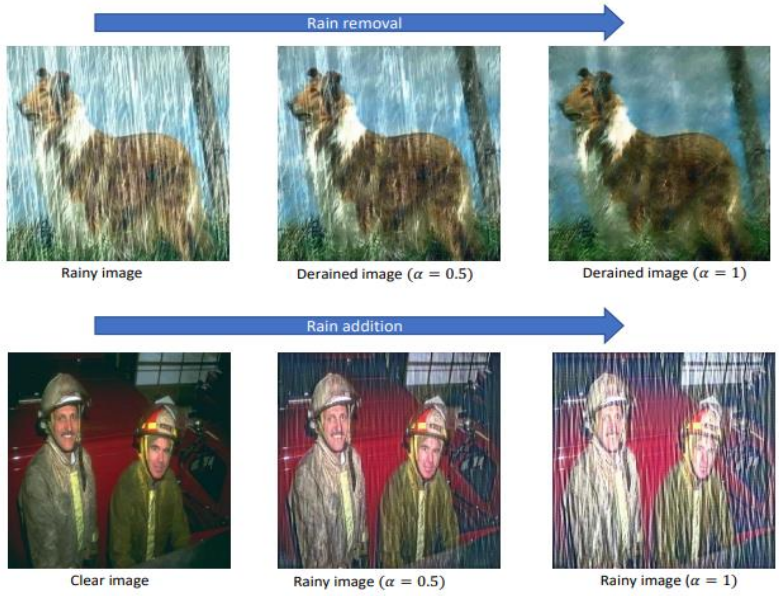
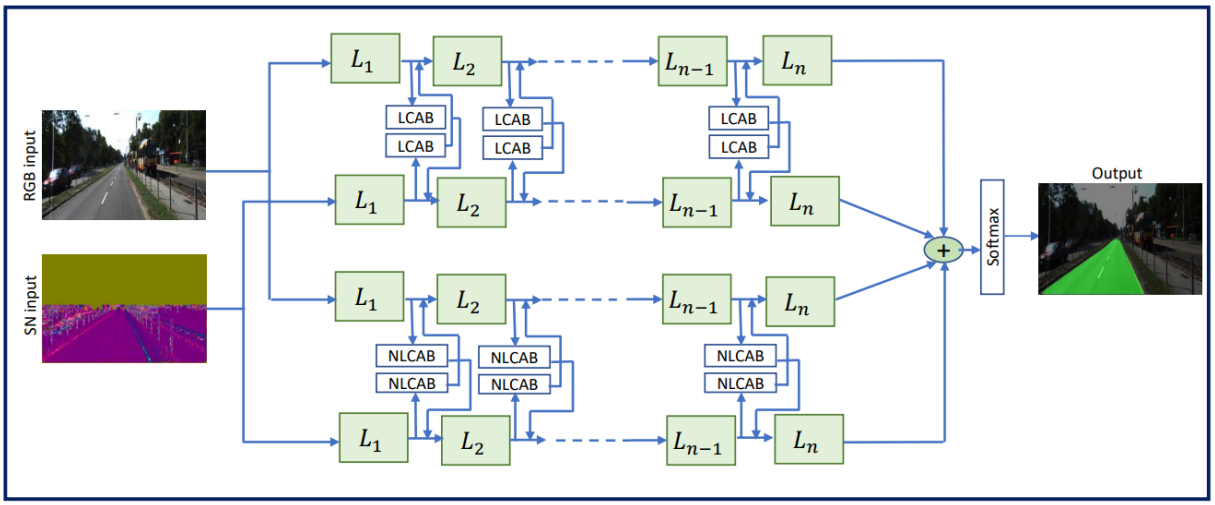


Fig. 4. Results of our proposed BiRelDerainGAN in terms of rain removal and rain addition. Towards both the direction our proposed method can generate the intermediate interpolated images, which demonstrates successive progression controlled by the factor α.

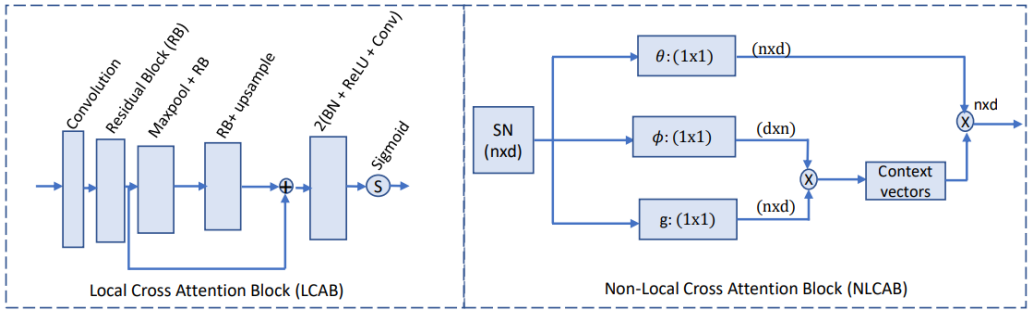
Objective III: Development of novel domain fusion methods for road segmentation.

In this data fusion research, we have focused on the research question of how to effectively fuse camera images and LiDAR cloud points to improve the performance of semantic segmentation in autonomous driving. Mostly, the LiDAR and camera sensor-based data are being fused since they usually bear the complementary information. Fusion of such sources helps to increase the segmentation performance. However, there still exists a performance gap due to their fusion mechanisms in terms of direct features concatenation or choice of architectures to learn the source specific features. In this research, we have developed a novel framework for the fusion of surface normal (SN) images calculated from depth/disparity maps from LiDAR scans with RGB camera images in order to learn effective features. While doing so, a two-step learning strategy, as shown in Fig. 5, is proposed such that the learning from one modal is continuously calibrated from the corresponding local and non-local attentions generated from the other modal.

Furthermore, in order to overcome the degraded performance of segmentation due to unwanted artifacts around the class-boundaries, we propose to use the cycle spinning-based technique for training and testing of our proposed model. The use of such local and non-local attentions from one modal to another and cycle spinning-based scheme is the first time in the field of road segmentation. With extensive quantitative and qualitative results and ablation studies performed on KITTI dataset for road segmentation, we show that our proposed method performs very favorably against the state-of-the-art methods.



(a) Proposed four branch fusion architecture. Note that each stream of the architecture is the separate instantiation of the same baseline network and l_n represents its' n^{th} layer. Further, RGB and SN represent the RGB camera inputs and SN inputs respectively. The arrow intersection with the network layer represents the attention injection at that point as per Eqs. 1 and 2. Best viewed in color.



(b) Local (left) and non-local (right) cross attention blocks. Note that θ, ϕ, and g are the $1x1$ convolutions used to linearly project the input feature.

Fig. 5. Illustration of the proposed domain fusion method.

In conclusion, during this one-year STIR project, we have investigated novel domain adaptation and domain fusion approaches under the goal of improving decision making performance in autonomous vehicles. Extensive experimental results have demonstrated the effectiveness of the proposed approaches in real-world environments. The outcomes of this project lay down the foundation of the PI's future work in this field, based on which the PI will develop a regular ARO proposal in the topics of advanced machine learning and autonomous agents.

Publications:

1. Sharma, S., and Tang, B. BiRelDerainGAN: A Clear-Rain Relativeness-Based Bidirectional Unpaired Single Image Deraining Using Generative Adversarial Networks, IEEE Transactions on Image Processing, Under Review, 2022.
2. Sharma, S., Tang, B., Ball, J., Carruth, D., and Dabbiru, L. Local and Non-Local Cross Attention-Based Fusion for Semantic Segmentation with Cycle Spinning. IEEE Transactions on Circuits and Systems for Video Technology, Under Review, 2022.