

PROCEEDINGS OF SPIE

***Signal Processing,
Sensor/Information Fusion,
and Target Recognition XXX***

**Ivan Kadar
Erik P. Blasch
Lynne L. Grewe**
Editors

**12–16 April 2021
Online Only, United States**

Sponsored and Published by
SPIE

Volume 11756

Proceedings of SPIE 0277-786X, V. 11756

SPIE is an international society advancing an interdisciplinary approach to the science and application of light.

Signal Processing, Sensor/Information Fusion, and Target Recognition XXX, edited by Ivan Kadar,
Erik P. Blasch, Lynne L. Grewe, Proc. of SPIE Vol. 11756, 1175601 · © 2021 SPIE
CCC code: 0277-786X/21/\$21 · doi: 10.1117/12.2598593

Proc. of SPIE Vol. 11756 1175601-1

The papers in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. Additional papers and presentation recordings may be available online in the SPIE Digital Library at SPIDigitalLibrary.org.

The papers reflect the work and thoughts of the authors and are published herein as submitted. The publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

Please use the following format to cite material from these proceedings:

Author(s), "Title of Paper," in *Signal Processing, Sensor/Information Fusion, and Target Recognition XXX*, edited by Ivan Kadar, Erik P. Blasch, Lynne L. Grewe, Proc. of SPIE 11756, Seven-digit Article CID Number (DD/MM/YYYY); (DOI URL).

ISSN: 0277-786X
ISSN: 1996-756X (electronic)

ISBN: 9781510643499
ISBN: 9781510643505 (electronic)

Published by
SPIE
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone +1 360 676 3290 (Pacific Time)
SPIE.org
Copyright © 2021 Society of Photo-Optical Instrumentation Engineers (SPIE).

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by SPIE subject to payment of fees. To obtain permission to use and share articles in this volume, visit Copyright Clearance Center at copyright.com. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher.

Printed in the United States of America by Curran Associates, Inc., under license from SPIE.

Publication of record for individual papers is online in the SPIE Digital Library.

SPIE. DIGITAL LIBRARY
SPIDigitalLibrary.org

Paper Numbering: A unique citation identifier (CID) number is assigned to each article in the Proceedings of SPIE at the time of publication. Utilization of CIDs allows articles to be fully citable as soon as they are published online, and connects the same identifier to all online and print versions of the publication. SPIE uses a seven-digit CID article numbering system structured as follows:

- The first five digits correspond to the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. These two-number sets start with 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 0A, 0B ... 0Z, followed by 10-1Z, 20-2Z, etc. The CID Number appears on each page of the manuscript.

Contents

vii *Invited Panel Slides*

MULTISENSOR FUSION, MULTITARGET TRACKING, AND RESOURCE MANAGEMENT I

- 11756 03 **Hybrid probabilistic information matrix fusion** [11756-1]
- 11756 04 **Resolving two point targets with unequal and unknown intensities in a FPA** [11756-2]
- 11756 05 **Estimation for a feedback system with a desired final state and intermittent stochastic inputs**
[11756-3]
- 11756 06 **Launch point estimation with a single fixed passive sensor without trajectory state estimation**
[11756-4]
- 11756 07 **Application of the sliding innovation filter for fault detection and diagnosis of an
electromechanical system** [11756-5]
- 11756 08 **A multiple model-based sliding innovation filter** [11756-6]

MULTISENSOR FUSION, MULTITARGET TRACKING, AND RESOURCE MANAGEMENT II

- 11756 09 **The two-pass sliding innovation smoother** [11756-7]
- 11756 0C **Application of machine learning for drone classification using radars** [11756-10]

INFORMATION FUSION METHODOLOGIES AND APPLICATIONS I

- 11756 0H **2D point set registration via stochastic particle flow filter** [11756-15]

INFORMATION FUSION METHODOLOGIES AND APPLICATIONS II

- 11756 0K **The application of machine learning to signal processing for detection and identification of
signals of interest and anomalies** [11756-18]
- 11756 0L **Adversarial machine learning and adversarial risk analysis in multi-source command and
control** [11756-19]

11756 OM **Risk-based security: from theory to practice** [11756-20]

INFORMATION FUSION METHODOLOGIES AND APPLICATIONS III

11756 ON **Anomaly detection of unstructured big data via semantic analysis and dynamic knowledge graph construction** [11756-21]

11756 OO **Learning intent and behavior models from motion trajectories for unsupervised semantic labeling** [11756-22]

11756 OP **Weighted certainty grids for dynamic search** [11756-23]

INFORMATION FUSION METHODOLOGIES AND APPLICATIONS IV

11756 OR **Anomaly detection with noisy and missing data using a deep learning architecture** [11756-25]

11756 OS **Fairness-by-design Dempster-Shafer reasoning system** [11756-26]

11756 OT **OCULUS iCrowd: an integrated C2I and simulation environment for security management, anomaly detection, and risk assessment** [11756-27]

11756 OU **OCULUS wayGoo integrated C2I system with geocoding, path planning and navigation** [11756-28]

SIGNAL AND IMAGE PROCESSING, AND INFORMATION FUSION APPLICATIONS I

11756 OX **Camera placement optimization for sports filming** [11756-31]

11756 OY **Effect of the short time Fourier transform on the classification of complex-valued mobile signals** [11756-32]

11756 OZ **Sliding window study of brain connectivity dynamics based on energy landscape analysis.** [11756-33]

11756 IO **Energy landscape analysis of fMRI data from schizophrenic and healthy subjects** [11756-34]

SIGNAL AND IMAGE PROCESSING, AND INFORMATION FUSION APPLICATIONS II

11756 I2 **Real-time location fingerprinting for mobile devices in an indoor prison setting** [11756-35]

11756 I3 **Health crisis situation awareness using mobile multiple modalities** [11756-36]

- 11756 14 **Low-resolution infrared temperature analysis for disease situation awareness via machine learning on a mobile platform [11756-37]**
- 11756 15 **Adversarial indoor signal detection [11756-38]**
- 11756 17 **Ship formation detection based on spatial distribution and attribute information [11756-40]**
- 11756 19 **Computed extended depth of field photoacoustic microscopy using ratio of low-pass pyramid fusion [11756-42]**

SIGNAL AND IMAGE PROCESSING, AND INFORMATION FUSION APPLICATIONS III

- 11756 1D **CalibDNN: multimodal sensor calibration for perception using deep neural networks [11756-46]**
- 11756 1E **Demonstration of hybrid lidar and passive polarimetric infrared imaging [11756-47]**
- 11756 1F **Vibroacoustic monitoring of technological processes employing electrophysical phenomena [11756-48]**

POSTER SESSION

- 11756 1G **Comparison of DFT calculated and measured IR absorption spectra [11756-49]**
- 11756 1H **Simulation study of compressed sensing photoacoustic tomography based on k-space pseudospectral method [11756-50]**

INVITED PANEL DISCUSSION “JOINT DATA LEARNING”

Joint Data Learning

(Panel: SPIE Defense & Commercial Sensing)

12 April 2021 • 12:30 PM - 2:00 PM PDT

SPIE Panel

Panel Organizers: Erik P. Blasch, MOVEJ Analytics / Air Force Research Lab.
Ivan Kadar, Interlink Systems Sciences, Inc., USA;

Panel Moderators: Lynne Grewe, California State Univ., USA
Chee-Yee Chong, Independent Consultant

Panelists: Erik Blasch – MOVEJ Analytics
Genshe Chen - Intelligent Fusion Technology, Inc.
Ivan Kadar, Interlink Systems Sciences, Inc., USA;
Uttam Majumder – Air Force Research Laboratory
Ali Khalid Raz – George Mason University
Andreas Savakis – Rochester Institute of Technology
Yufeng Zheng - University of Mississippi Medical Center



30 years

Conference: Signal Processing, Sensor/Information Fusion, and Target Recognition

1

Panel Members

- **Panel Organizers:** Erik P. Blasch, MOVEJ Analytics / Air Force Research Lab.
Ivan Kadar, Interlink Systems Sciences, Inc., USA;

- **Panel Moderators:** Lynne Grewe, California State Univ., USA
Chee-Yee Chong, Independent Consultant

- **Panelists:**

- Erik P. Blasch, MOVEJ Analytics
- Genshe Chen - Intelligent Fusion Technology, Inc.
- Ivan Kadar - Interlink Systems Sciences, Inc., USA
- Uttam Majumder – Air Force Research Laboratory
- Ali Khalid Raz – George Mason University
- Andreas Savakis – Rochester Institute of Technology
- Yufeng Zheng - University of Mississippi Medical Center

2

viii

SPIE Panel: Joint Data Learning (1)



- **ABSTRACT:** AI techniques are based on learning a model based on a large available data set. The data sets typically are from a single modality (e.g., imagery) and hence the model is based on a single modality.
- **Multiple models** are each built for a common scenario (e.g., video and natural language processing of text describing the situation). There are issues of robustness, efficiency, and explainability needed. A second modality can improve **efficiency** (e.g., cueing), **robustness** (e.g., results cannot be fooled such as adversary systems), and **explainability** from different sources help. The challenge is how to organize the data needed for joint data training and model building. For example, what is needed
 - (1) **structure for indexing data** as an object file,
 - (2) **recording of metadata** for effective correlation, and
 - (3) **supporting models** and analysis for model interpretability for users. There are a variety of questions to be discussed, explored, and analyzed for fusion-based AI tool.

Conference: Signal Processing, Sensor/Information Fusion, and Target Recognition

3

Joint Data Learning (2)



- **Questions**
 - 1) Summarize a unique **data fusion success** story within the last six months
 - 2) For joint data fusion (beyond multimedia and medicine) what are the **next emerging applications**?
 - 3) How to **train a joint classifier** - issues and preparations
 - 4) Who are the **users** that the DF/SF can support and subsequently the metrics of interest?
 - 5) what is the **future** - such as a standard or evaluation method

4

Panel

SPIE. DEFENSE+
COMMERCIAL
SENSING



Ivan Kadar



Erik P. Blasch



Lynne Grewe



Chee-Yee Chong



Genshe Chen



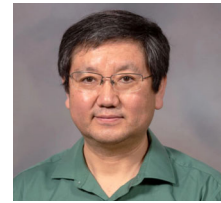
Uttam Majumder



Ali Khalid Raz



Andreas Savakis



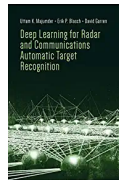
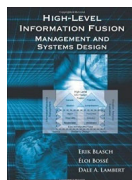
Yufeng Zheng



Joint Data Learning

(Panel: SPIE Defense & Commercial Sensing)

12 April 2021 • 12:30 PM - 2:00 PM PDT
SPIE Panel



Erik Blasch
MOVEJ Analytics



Conference: Signal Processing, Sensor/Information Fusion, and Target Recognition



Sensor Data Fusion

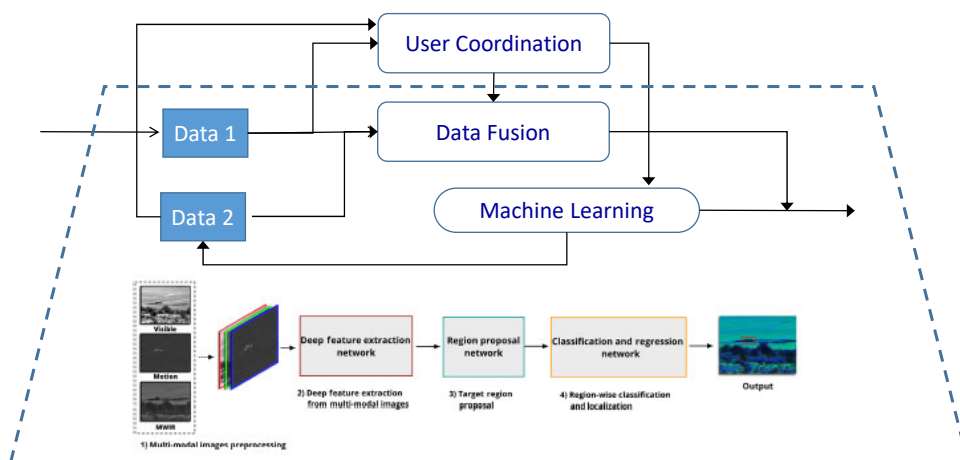
Physics-based /Human-derived Information Fusion (PHIF)

1



Joint Data Learning :Q1 – Data Fusion

- 1) Summarize a recent unique data fusion success stor



Shuo Liu, Huan Liu, Vijay John, Zheng Liu, Ying Huang, E. Blasch, "Enhanced Situation Awareness through CNN-based Deep MultiModal Image Fusion," *Optical Engineering*, 59(5): 053103, April 2020.

2

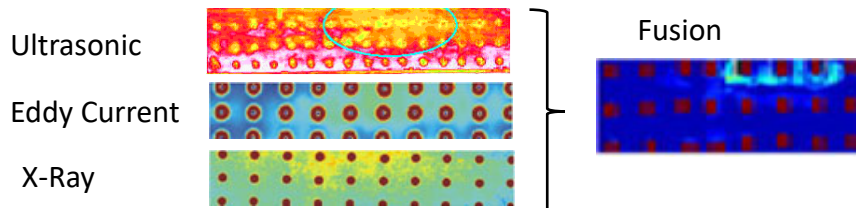
xi

Joint Data Learning Q2: Emerging



- 2) For joint data fusion (beyond multimedia and medicine) what are the next emerging applications? **Inspection, Targeting**

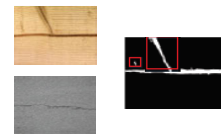
Y. Zheng, E. Blasch, Z. Liu, *Multispectral Image Fusion and Colorization*, SPIE Press, 2018.



- **Evidential Neural Networks (ENN)**
- **WGGAN – Wavelet Guided Generative Adversarial Network**
- Domain Adaptation/Transfer Learning

W. Zhai, J. Zhu, Y. Cao and Z. Wang, "A Generative Adversarial Network Based Framework for Unsupervised Visual Surface Inspection," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.

R. Zhang, J. Bin, Z. Liu, E. Blasch, "WGGAN: A **Wavelet-Guided Generative Adversarial Network** for Thermal Image Translation," *Generative Adversarial Networks for Image-to-Image Translation*, Elsevier, edited by M Naved, 2021

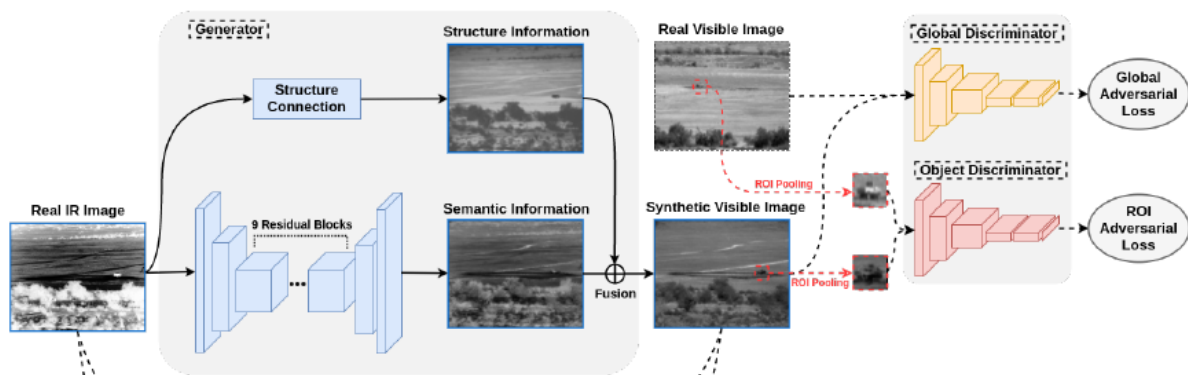


3

Joint Data Learning :Q3: Preparation



- 3) How to train a joint classifier - issues and preparations
- **Data Alignment**



S. Liu, M. Gao, V. John, Z. Liu, E. Blasch, "Deep Learning Thermal Image Translation for Night Vision Perception," *ACM Transactions on Intelligent Systems and Technology*, 12(1):1-18, Dec. 2020.

4

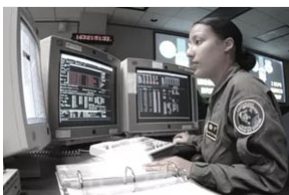
xii

Joint Data Learning :Q4: Users

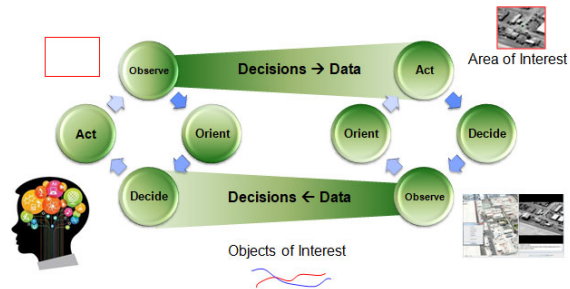
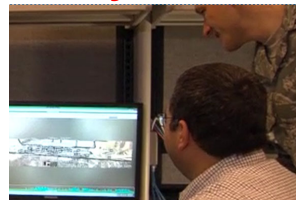


- 4) Who are the users that the DF/SF can support and subsequently the metrics of interest?

User/Operational Analyst



Career as a USAF Space Systems Operator
(thebalancecareers.com)



Human-Machine Teaming - Observe-Orient-Decide-Act (OODA) loops
Metrics – Trust, Explainability, Interpretability, Usability, Understandability,

E. Blasch, A. Steinberg, S. Das, J. Llinas, C.-Y. Chong, O. Kessler, E. Waltz, and F. White, "Revisiting the JDL model for information Exploitation," *Int'l Conf. on Info Fusion*, 2013.

5

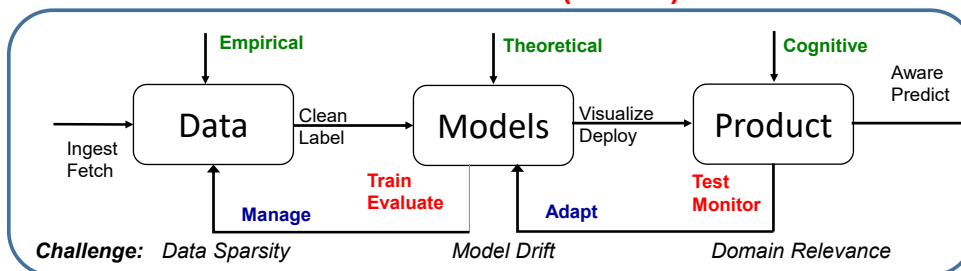
Joint Data Learning :Q5: Standards



- 5) what is the future - such as a standard or evaluation method

Evaluation (e.g., STANAG 2511, 4162)

Multisource AI Scorecard Table (MAST)



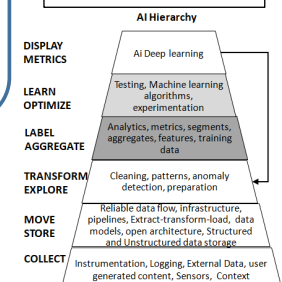
• Intelligence Community Directive 203 – Analytic Standards

- Test and Evaluation of AI/ML systems

E. Blasch, J. Sung, T. Nguyen, "Multisource AI Scorecard Table for System Evaluation," *AAA/ FSS-20: Artificial Intelligence in Government and Public Sector*, 2020. [arXiv:2102.03985](https://arxiv.org/abs/2102.03985)

ICD203 Standards

1. Sourcing
2. Uncertainty
3. Distinguishing
4. Analysis of Alternatives
5. Customer Relevance
6. Logical Argumentation
7. Consistency
8. Accuracy
9. Visualization



6

xiii

Joint Data Learning - Summary



- **Questions**

- 1) **Data fusion success** – AI/ML Active Learning
- 2) **Emerging applications** – Real-time Inspection
- 3) **Train a joint classifier** – Data Alignment
- 4) **Users** – Operators, but need to test AI/ML in the work domain
- 5) **Future** – AI/ML Data Fusion standards

Erik Blasch
MOVEJ Analytics

Joint Multisource Data Learning Standards

Erik Blasch

MOVEJ Analytics, Dayton, OH 45324

Abstract: Joint multisource data learning (JMDL) includes that of incorporating multiple data sources within a deep learning (DL) technique. The challenges of utilizing multisource data in DL include data association, embeddings, and resolution. Data association is critical for a data fusion technique such as to align entities of information from different sensor types. In this position paper, the areas of interest include: (1) benefits of active learning, (2) emerging methods of multisource deep learning for inspection, (3) preparation of data for alignment and training, (4) teaming for an operational analyst to enhance performance, and (4) future needs of JMDL standards including metrics and evaluation methods. Key enablers are an extensible ontology, joint embeddings, and JMDL architectures.

Keywords: Multimodal Deep Learning, Machine Learning, Test and Evaluation, Joint Data Embeddings, Ontology

Information fusion has a long history of development including that of low-level sensor data fusion [1] to that of Systems-Level information fusion [2]. Within the last decade, there has been an interest in the coordination between information fusion [3] and deep learning [4, 5]. The foundation of sensor data fusion supports physics-based and human-derived information fusion (PHIF) [6] which leverages machine learning (ML), DL, and artificial intelligence (AI) methods. However, much of the DL literature is focused on single modalities, such as video; while some examples are for multimedia (video and text). Going beyond the imagery to such elements of synthetic aperture radar [7] as well as natural language processing with contextual information [8] requires more thought in the DL architectures.

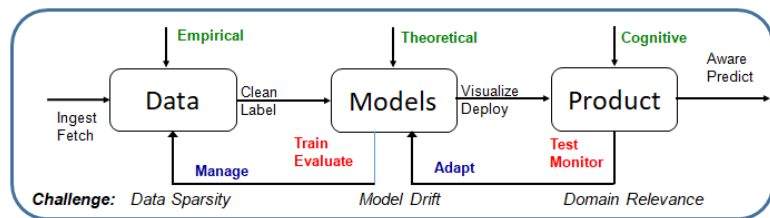


Figure 1: Multisource AI/ML Pipeline

Multiple models from data are each built for a common scenario (e.g., video and natural language processing of text describing the situation) as shown in Figure 1. There are issues of robustness, efficiency, and explainability needed, especially with multisource data. A second modality can improve efficiency (e.g., cueing), robustness (e.g., results cannot be fooled such as adversary systems), and explainability from different sources. The challenge is how to organize the data needed for joint data training and model building. For example, some directions needed are:

- (1) **structuring of data** such as indexing an object file and processing in a DL architecture;
- (2) **recording of metadata** for effective data correlation, embedding, and processing; and
- (3) **supporting knowledge** for model analysis explainability, and interpretability.

To answer future needs, there are issues associated with data analytics, model construction, and user products for which the community has just started to explore.

Element 1: Data Analytics

Data analytics focuses on the use of data for such things as data fusion. For sensor data fusion with user involvement such as active learning (Figure 2) or machine focused, the correspondence of data with the decision need requires joint data embedding for data association, alignment, and coordination during training. There is a need for embeddings for joint multisource data to translate high-dimensional information into a low-dimensional space affording ML over sparse data, semantic representations, and model reuse. For the joint data, it allows for mapping concepts in a similar space such as for data association, efficient training, effective results, and user relevance.

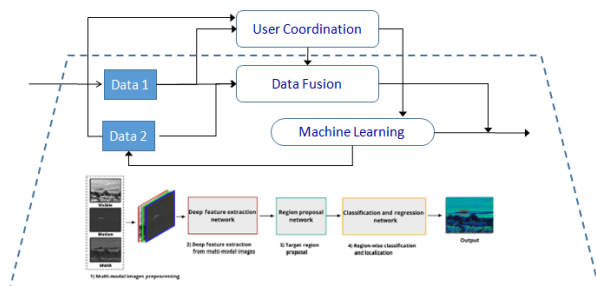


Figure 2: Active Learning

Element 2: Model Construction

Model construction with ML/DL over joint data requires an architecture (Figure 3) from which to combine the data [9]. Using standard information methods, the challenge is how to construct the architecture for processing signals, features, and decisions. Hence, the DL challenge typical resides in the methods and strategies for processing the joint data.

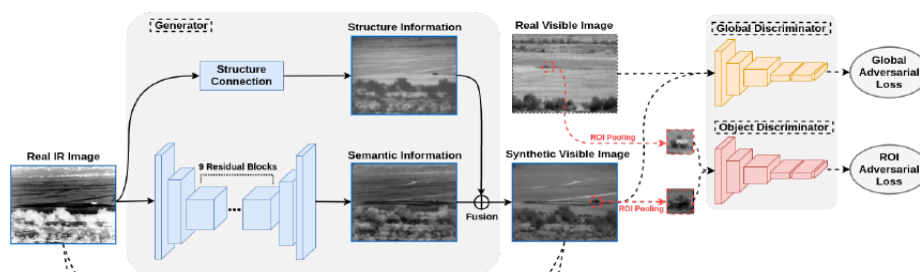
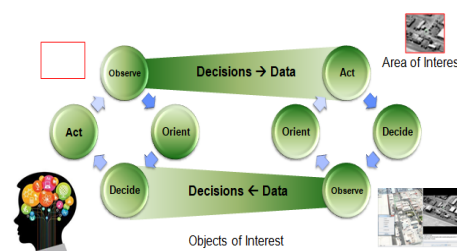


Figure 3: Data Alignment to allow one mode in the “resolution” of the other [9].

Element 3: User of Products

Users are critical to the performance of qualitative and quantitative evaluation of systems [10] and standards are needed to support decisions –to-data and data-to-decisions (Figure 4) for human machine teaming. Recent efforts in ML/DL standards [11] would assist in the verification and validation of multisource DL systems. As with previous discussions [12], the specific operator has to be considered for the end-product use to determine the value of the multisource information, DL, and system design for improved performance.



Conclusions

Deep Learning has the potential to focus the model development based on a corpus of multisource data. Before the entire corpus is a reality, there is a need for physics-based and human-derived information fusion approaches to utilize the data. For the physics, it is a foundational scalable alignment (space, time) for the data (such as EO, IR), while the human-derived semantic information requires an extensible ontology, joint embedding, and preferred architecture. Such developments in JMDL are needed to direct the extensible domain data fusion to go beyond text-video multimedia DL to that of other sources of data.

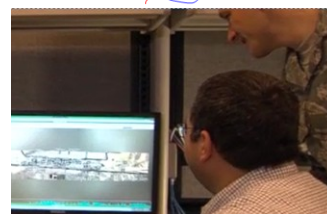


Figure 4. Human-Machine Teaming

References

- [1] Zheng, Y., Blasch, E., Liu, Z., [Multispectral Image Fusion and Colorization], SPIE Press, (2018).
- [2] Blasch, E. P., Bosse, E., D. A. Lambert, [High-Level Information Fusion Management and Systems Design] Artech House, (2012).
- [3] Blasch, E., Steinberg, A., Das, S., Llinas, J., Chong, C.-Y., Kessler, O., Waltz, E., White, F., "Revisiting the JDL model for information Exploitation," *Int'l Conf. on Info Fusion*, (2013).
- [4] Liu, S., Liu, H., John, V., Liu, Z., et al., "Enhanced Situation Awareness through CNN-based Deep MultiModal Image Fusion," *Optical Engineering*, 59(5): 053103, April (2020).
- [5] Blasch, E., Cruise, R., et al., "Methods of AI for Multi-modal Sensing and Action for Complex Situations," *AI Magazine*, (2019).
- [6] Blasch, E., Aved, A. J., "Physics-Based and Human-derived Information Fusion Video Activity Analysis," *Int'l. Conf. on Info. Fusion*, (2018).
- [7] Majumder, U., Blasch, E., Garren, D., [Deep Learning for Radar and Communications Automatic Target Recognition], Artech House, (2020).
- [8] Snidaro, L., Garcia, J., Llinas, J., et al. (eds.), [Context-Enhanced Information Fusion: Boosting Real-World Performance with Domain Knowledge], Springer, (2016).
- [9] Liu, S., Gao, M., John, V., Z. Liu, E. Blasch, "Deep Learning Thermal Image Translation for Night Vision Perception," *ACM Transactions on Intelligent Systems and Technology*, 12(1):1-18, Dec. (2020).
- [10] Zheng, Y., Dong, W., et al., "Qualitative and quantitative comparisons of multispectral night vision colorization techniques," *Optical Engineering*, Vol. 51, Issue 8, Aug. (2012).
- [11] Blasch, J. Sung, T. Nguyen, "Multisource AI Scorecard Table for System Evaluation," *AAAI FSS-20: Artificial Intelligence in Government and Public Sector*, (2020).
- [12] Blasch, E., Kadar, I., Grewe, L. L., Stevenson, G., Majumder, U. K., Chong, C.-Y., "Deep learning in AI and information fusion panel discussion," *Proc. SPIE, 11018*, (2019).

Joint Data Learning Panel Discussion

Learning with EO and SAR data

Andreas Savakis

Rochester Institute of Technology, USA

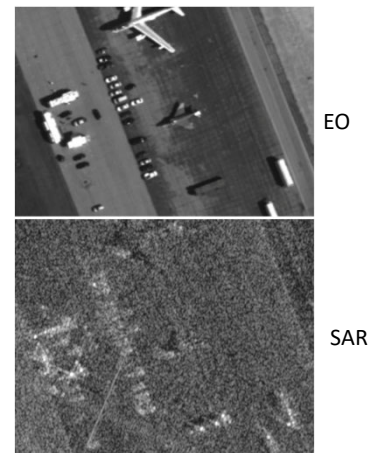
Andreas.Savakis@rit.edu

SPIE DCS 2021

1

Motivation for Joint Data Learning

- Deep Learning
 - Significant success with large labelled datasets in the EO domain
 - Limited SAR data available for training
- Synthetic Aperture Radar (SAR) vs. Electro-Optical (EO) imagery
 - SAR Imagery is not affected by weather
 - SAR works at night
 - Labelling SAR data is expensive and time consuming



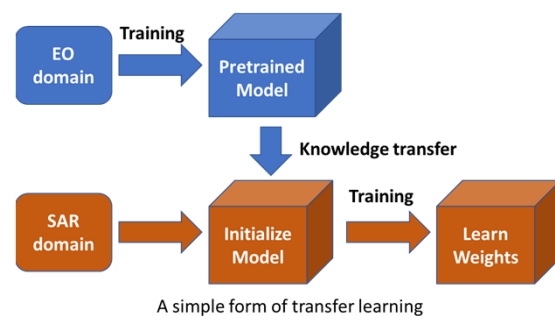
[1] Colin Leong et al. "Unified Coincident Optical and Radar for Recognition (UNICORN) 2008 Dataset" <https://github.com/AFRL-RY/data-unicorn-2008>, 2019.

2

- Transfer Learning
 - Large datasets available for EO but not for SAR
 - Requires labels in both domains
- Adaptation
 - Unsupervised or semi-supervised
 - Labels available in EO, but not in SAR
- Joint Classification
 - Classification using data from both domains
 - Requires data correspondence across domains

3

- Transfer knowledge from EO to SAR domain [2]
- Evaluate performance under various SAR data collection parameters such as sensor type, polarization, and incidence angle
- Visualize the network activation maps for explainability.

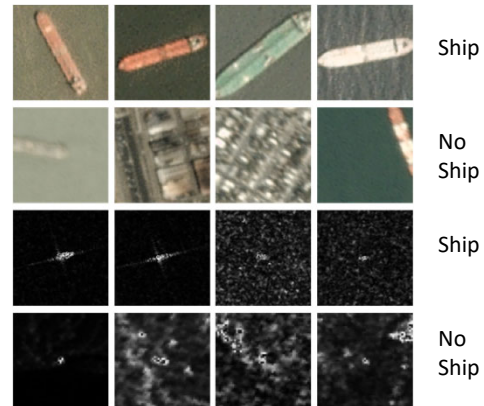


[2] A. M. N. Taufique, N. Nagananda, A. Savakis, "Visualization of Deep Transfer Learning in SAR Imagery," IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2020

4

xviii

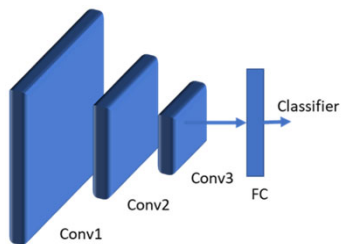
- EO ships in satellite imagery dataset [3]
 - Taken using Planet satellites
 - San Francisco and San Pedro Bay areas
 - 1000 ships chips
 - 3000 no ships chips
- SAR ships dataset [4]
 - Sintel-1 extended wide swath and RADARSAT-2 ScanSAR images
 - South African Exclusive Economic Zone
 - 1596 ships chips
 - 7890 no ships chips



[3] R. Hammell "Ships in Satellite Imagery, ver.~9" <https://www.kaggle.com/rhammell/ships-in-satellite-imagery>, 2018

[4] C. P. Schwegmann et al. "Very deep learning for ship discrimination in Synthetic Aperture Radar imagery" IGARSS, 2016

5



Training dataset	Class	SAR Percent Accuracy
Only EO	Ship	27%
	No ship	72%
	Overall	49%
TL to SAR	Ship	91%
	No ship	95%
	Overall	93%

- CNN trained on EO data with cross entropy loss
- Mean per class accuracy classification metric
- Results significantly improve after Transfer Learning (TL)



6

xix

RIT SAR Data Attributes



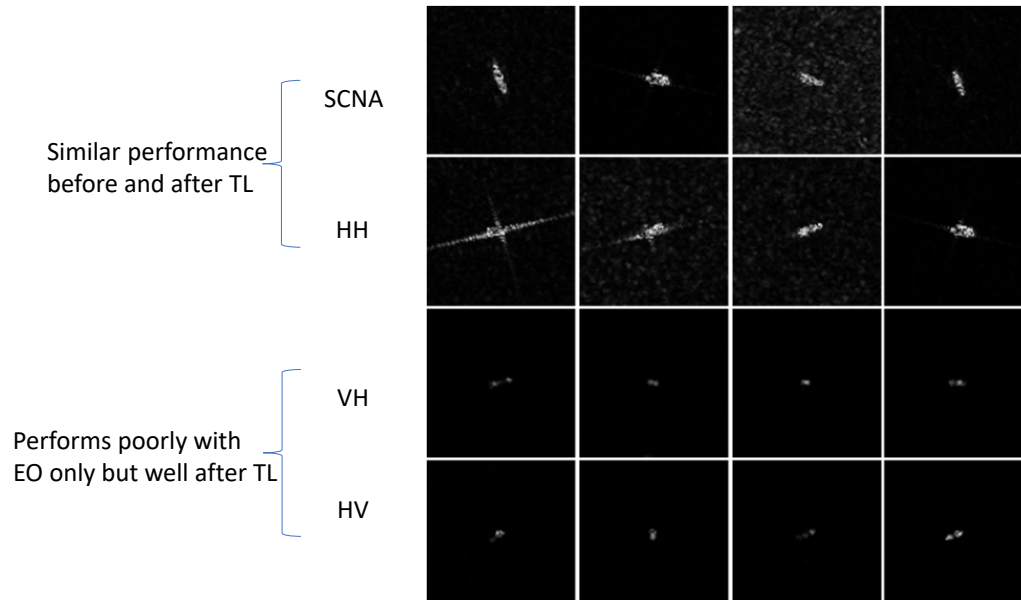
- Sensors
 - Sintel-1
 - GRDH: Ground Range Detected imagery with High resolution
 - GRDM: Ground Range Detected imagery with Medium resolution
 - RADARSAT2-ScanSAR
 - SCNA
- Polarization
 - HH, HV, VV, and VH
- Incidence angle of the transmitted wave
 - Small, Medium, Large

Training	Class	GRDH	GRDM	SCNA
Only EO	Ship	34%	19%	83%
	No ship	75%	71%	62%
	Overall	55%	45%	73%
TL to SAR	Ship	89%	92%	83%
	No ship	99%	94%	96%
	Overall	94%	93%	90%

Training	Class	HH	HV	VV	VH
Only EO	Ship	74%	6%	0.47%	0%
	No ship	62%	93%	0.54%	95%
	Overall	68%	49%	51%	48%
TL to SAR	Ship	76%	94%	93%	93%
	No ship	98%	97%	97%	91%
	Overall	87%	96%	95%	92%

7

RIT Results Analysis

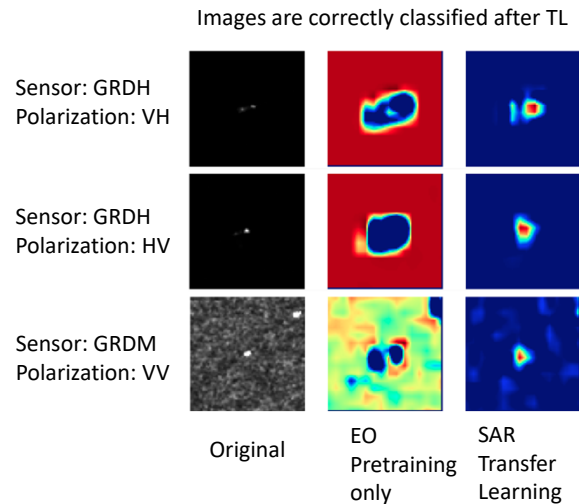


8

xx

Explainability Visualization Results

- GradCAM visualization shows the salient image region [5]
- Images incorrectly classified before transfer learning when
 - Blob is very small
 - There is significant noise
- Images correctly classified after SAR transfer learning
 - Activation maps shift to object region after TL



[5] R. R. Selvaraju et al. "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization" ICCV, 2017

9

EO/SAR Classification Challenge

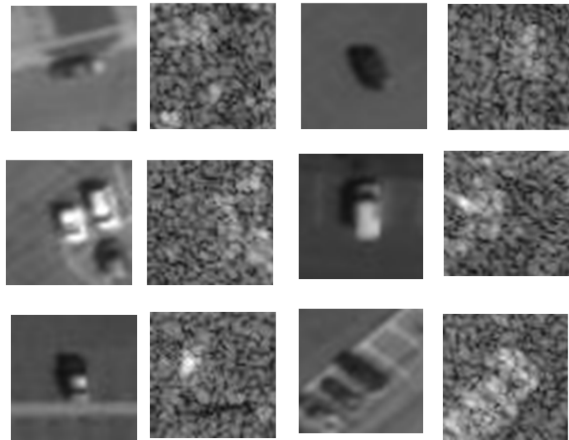
- CVPR 2021 NTIRE Workshop
- Multi-modal Aerial View Object Classification
 - <https://competitions.codalab.org/competitions/28095>
 - SAR classification with EO/SAR training
 - SAR/EO joint classification
- New EO/SAR Dataset
 - 10 classes of vehicles
 - Low resolution crops: EO 31x31, SAR 55x55
 - noisy
 - highly imbalanced

10

xxi

EO/SAR Challenge Dataset

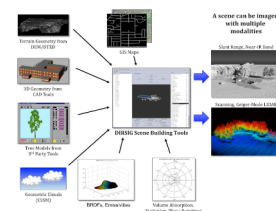
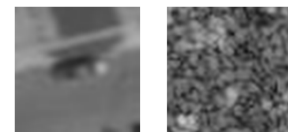
- 10 classes
 - Sedan (~234,000)
 - SUV (~28,000)
 - Pickup truck (~15,000)
 - Van (~10,000)
 - Box truck (~1,800)
 - Motorcycle (~850)
 - Flatbed truck (~800)
 - Bus (~600)
 - Pickup truck with trailer (~800)
 - Flatbed truck with trailer (~600)



11

Discussion and Future Directions

- Transfer Learning
 - Useful when labels are available
 - Feature generalization
- Adaptation
 - Can help if few or no labels are available
 - Unsupervised or semi-supervised
- Joint Classification
 - Feature fusion: early vs. late fusion
 - Class Balance
- Synthetic data generation
 - DIGSIG <http://www.dirsig.org/>
 - Generate synthetic data → Domain transfer
 - Train on synthetic data → Transfer Learning to real



<http://www.dirsig.org/>

12

xxii

Thank you!

- Panel Organizers and Moderators
 - Erik Blasch
 - Ivan Kadar
 - Lynne Grewe
 - Chee-Yee Chong
- RIT Vision Lab PhD students
 - Navya Nagananda
 - Abu Taufique

Joint Data Learning with EO and SAR data

Andreas Savakis*

Rochester Institute of Technology, Rochester, New York 14623, USA

ABSTRACT

Joint learning with data from different modalities offers research opportunities on feature learning, information fusion, domain adaptation and joint classification. We discuss transfer learning as a motivating example for joint data learning using Electro-Optical (EO) and Synthetic Aperture Radar (SAR) imagery. Transfer learning involves training a convolutional neural network in the EO domain and fine-tuning the network in the SAR domain, where labeled data may be limited. We further discuss joint classification of EO and SAR data with a multi-modal dataset used in the CVPR 2021 NTIRE workshop challenge. Challenging multi-modal datasets can motivate new research that leads to advances in joint data learning.

Keywords: Joint Data Learning, Transfer Learning, Domain Adaptation, Multi-modal Classification, EO and SAR data

1. INTRODUCTION

Deep learning methods for image classification¹ have been widely successful for Electro-Optical (EO) data, including RGB and panchromatic images in a variety of applications. These advances require the availability of large labeled datasets for training, which are generally available in the EO domain. However, large labeled datasets may not be widely available with other sensor modalities, such as Synthetic Aperture Radar (SAR) data². Furthermore, labelling SAR data is expensive and time consuming. SAR offers certain advantages over EO, such as the ability to operate at night or under different weather conditions, and joint data learning is a way to leverage the strengths of both modalities for classification or other tasks. Examples of EO and SAR images of ships are shown in Figure 1.

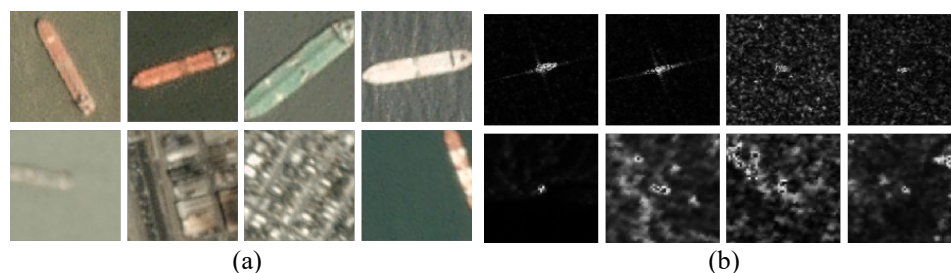


Figure 1. Examples of EO and SAR images from (a) Ships in Satellite Imagery dataset and (b) SAR Ships dataset.

Top row shows images with ships and bottom row shows images without ships.

Joint learning with convolutional neural networks (CNNs) using EO and SAR data can take the form of transfer learning, domain adaptation or joint classification. Transfer learning is useful in situations where labeled data are available in both EO and SAR, so that initial training is performed in the EO domain and is followed by fine-tuning in the SAR domain. Domain adaptation begins with training in the source EO domain, where labeled data are available. Adaptation takes place in an unsupervised manner in the target SAR domain when target labels are not available. Semi-supervised adaptation utilizes a small number of labeled samples in the target domain to facilitate the adaptation process.

Joint classification incorporates data from both modalities, e.g. EO and SAR samples. The joint data learning process may involve fusion of information from EO and SAR data. Fusion can take place early, during the feature extraction process, or late, during the classification process. We next discuss transfer learning from the EO domain to the SAR domain for classification of SAR images and a challenge dataset of EO and SAR images that is designed for multi-modal classification.

*andreas.savakis@rit.edu; <https://www.rit.edu/directory/axseec-andreas-savakis>

2. TRANSFER LEARNING FROM EO TO SAR DOMAIN

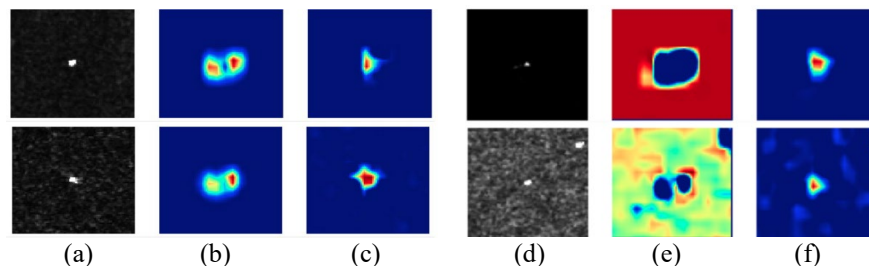


Figure 2. Examples of Grad-CAM activations before and after transfer learning. SAR images in columns (a) and (d); Grad-CAM activations before transfer learning in columns (b) and (d) and after transfer learning in columns (c) and (f).

A recent study of transfer learning from EO to SAR³ demonstrated the effects of transfer learning across modalities. The EO dataset was used to pre-train a small CNN and then SAR data were used to perform transfer learning to the SAR domain. The datasets considered were the EO Ships in Satellite Imagery⁴ dataset and the SAR Ships⁵ dataset. Representative samples from these datasets are shown in Figure 1. The Ships in Satellite Imagery samples were collected using Planet satellites in the San Francisco and San Pedro Bay areas. This EO dataset consists of 1,000 chips of ships and 3,000 background chips. The SAR Ships dataset contains Sintel-1 Extended Wide Swath images and RADARSAT-2 ScanSAR images from the South African Exclusive Economic Zone. It consists of 1,596 positive samples that contain a ship and 7,890 negative samples. Various data attributes related to sensor resolution (High and Medium for Sintel-1 and SCNA for RADARSAT-2), polarization (HH, HV, VV, and VH), and incidence angle of the radar signal (Small, Medium, Large) are included in the dataset as annotations for the SAR image acquisition.

We tested the performance of our network before and after transfer learning and found that transfer learning significantly increased performance in most cases. These gains depend on the SAR data attributes, since there are significant differences in ship size and brightness among various attributes. Training in the EO domain without transfer learning was not sufficient for the classification of SAR data, as only 27% of SAR ships were correctly classified and the overall accuracy on the SAR dataset was 49% before transfer learning. Transfer learning to the SAR dataset dramatically improved performance to 91% for ship images and 93% overall for the SAR Ships dataset.

We utilized Grad-CAM⁶ visualization to highlight the salient image region that activated the network during classification. Figure 2 shows representative examples of the Grad-CAM activation heatmaps before and after transfer learning. These results illustrate that SAR ship images were incorrectly classified before transfer learning when the blob is very small or when there is significant noise. After transfer learning the activation heatmaps shift to the correct location of the ship indicating a correct network response.

3. JOINT CLASSIFICATION CHALLENGE WITH EO AND SAR DATA

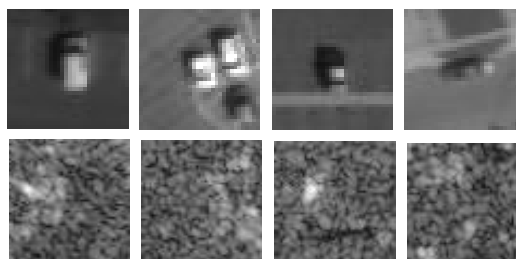


Figure 3. Samples of EO images (top row) and corresponding SAR images (bottom row) from the Multi-modal Aerial View Object Classification dataset.

A new multimodal dataset with EO and SAR data was released for the Multi-modal Aerial View Object Classification Challenge⁷ in the CVPR 2021 NTIRE Workshop. Examples of EO and SAR images are shown in Figure 3. This dataset contains chips of the following 10 classes of vehicles with approximate number of samples in each class shown in parentheses: Sedan (234,000), SUV (28,000), Pickup truck (15,000), Van (10,000), Box truck (1,800), Motorcycle (850),

Flatbed truck (800), Bus (600), Pickup truck with trailer (800), Flatbed truck with trailer (600). This dataset is highly imbalanced with samples per class ranging from 234,000 for sedans to only 600 for buses. Additionally, the images have low resolution (31x31 pixels for EO and 55x55 pixels for SAR). SAR samples are very noisy, as shown in Figure 3. These conditions make joint data learning with this dataset quite challenging.

Challenging multi-modal datasets are useful for motivating research that leads to new methods and advances in the field of joint data learning. In cases where labeled or registered data are not easy to obtain, synthetic data generation is an alternative. Synthetic data may be generated using DIRSIG⁸ or Blender and used for initial training of deep networks followed by transfer learning using real data.

4. DISCUSSION

Joint data learning offers opportunities for transfer learning, domain adaptation and joint classification using datasets from various sensor modalities, such as EO and SAR. In some cases, it is useful to leverage knowledge from one modality, e.g. EO, to another modality, e.g. SAR. Joint classification with multi-modal data motivates research in feature fusion and methods that overcome the class imbalance problem. Future advances in the field may be motivated by challenging datasets consisting of a mix of synthetic and real data.

ACKNOWLEDGEMENTS

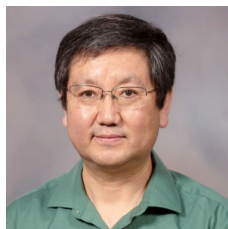
This research was supported in part by the Air Force Office of Scientific Research (AFOSR) under Dynamic Data Driven Applications Systems (DDAS) grant FA9550-18-1-0121.

REFERENCES

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
- [2] M. Rostami, S. Kolouri, E. Eaton, and K. Kim, "Deep transfer learning for few-shot SAR image classification," *Remote Sensing*, vol. 11, pp. 1374, 2019.
- [3] A. M. N. Taufique, N. Nagananda, A. Savakis, "Visualization of Deep Transfer Learning in SAR Imagery," *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2020.
- [4] R. Hammell, "Ships in Satellite Imagery," <https://www.kaggle.com/rhammell/ships-in-satellite-imagery>, 2018.
- [5] C. P. Schwegmann, W. Kleynhans, B. P. Salmon, L. W. Mdakane, and R. G. V. Meyer, "Very deep learning for ship discrimination in Synthetic Aperture Radar imagery," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 104–107, 2016.
- [6] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *IEEE International Conference on Computer Vision (ICCV)*, pp. 618–626, 2017.
- [7] CVPR 2021 NTIRE Workshop Multi-modal Aerial View Object Classification Challenge. <https://competitions.codalab.org/competitions/28095>
- [8] DIRSIG Remote Sensing Data and Image Simulation Framework <http://www.dirsig.org>

Invited Panel Discussion

Joint Data Learning



Yufeng Zheng, UMMC
yzheng@umc.edu

Joint Data Learning at SPIE DCS

1

How to train a CNN to Colorize

Night Vision Images

Joint Data Learning

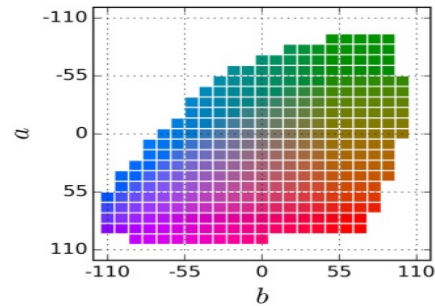
SPIE. The international society
for optics and photonics

2

xxvii

Color Transferring for Night Vision Colorization Using Convolutional Neural Network (CNN)

- ▶ Zhang et al. presented a CNN model for colorizing *grayscale* images.
- ▶ A *grayscale* image can be thought as the *L*-channel of the image in the **Lab** color space and the objective is to find the *a* and *b* components.
- ▶ The *ab* space is quantized into 313 bins (squared pixels) as shown below. So the *color prediction* task is turned into a *multinomial classification* problem, where for every gray pixel there are 313 classes in 1D space (corresponding to the 2D *ab* space) to choose from.
- ▶ Zhang's CNN architecture is a VGG-style network with multiple convolutional blocks.
- ▶ Each block has 2-3 convolutional layers followed by a *Rectified Linear Unit* (ReLU) and terminating in a *Batch Normalization* layer.



R. Zhang, P. Isola, et al., **Colorful Image Colorizations**, 14th European Conference on Computer Vision (ECCV), Amsterdam (2016).

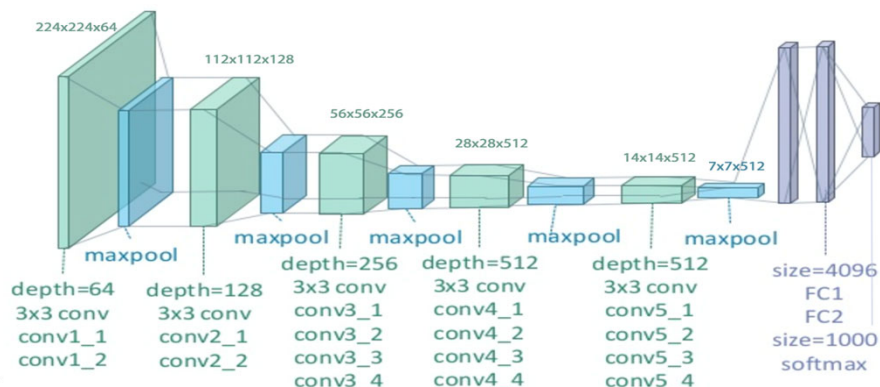
Joint Data Learning

SPIE The international society for optics and photonics

3

VGG-19 model for Classification

- 19-layer (16 conv., 3 fully-connected) CNN
- 3×3 filters with stride and pad of 1,
- 2×2 max-pooling layers with stride 2.
- Input image 224×224×3, 138M parameters, placed **2nd** in classification, and 1st in localization in ILSVRC 2014



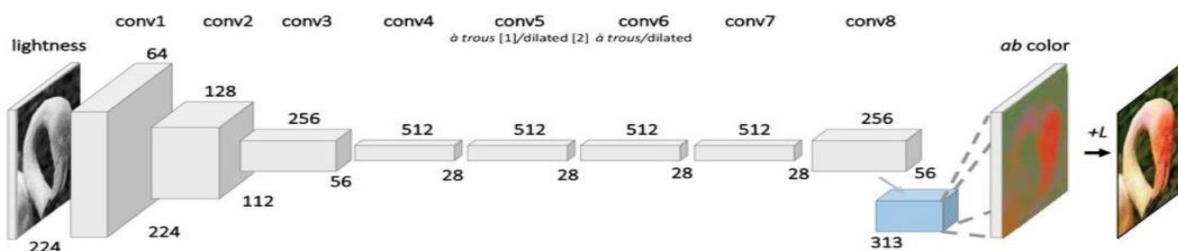
Joint Data Learning

4

xxviii

Color Transferring for Night Vision Colorization Using Convolutional Neural Network (CNN)

- ▶ Unlike the VGG net, there are no *pooling* or *fully connected* layers.
- ▶ The input image is rescaled to 224×224 , and the output dimensions of the last convolution layer are $56 \times 56 \times 313$.
- ▶ Each of 56×56 pixels contains a vector of 313 probability values.
- ▶ The predicted **ab** image also has the dimension 56×56 . To obtain the color image, it is **upsampled** to the original image size and then added to the lightness channel, **L**, to produce the final color image.
- ▶ The network was trained with 1.3M images from the *ImageNet* training set.



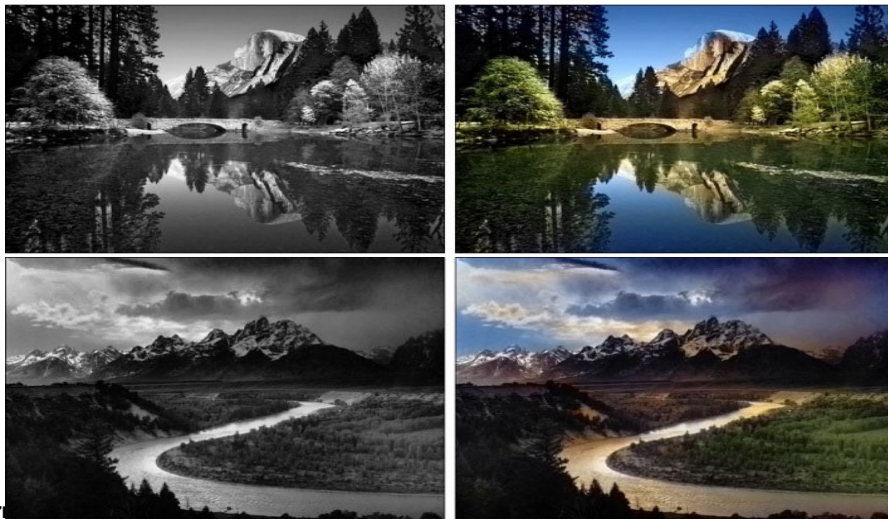
Joint Data Learning

SPIE The international society for optics and photonics

5

Color Transferring for Night Vision Colorization Using Convolutional Neural Network (CNN)

- ▶ Illustration of colorizing *grayscale* images using Zhang's VGG-Net (Grayscale photos by Ansel Adams of Yosemite)



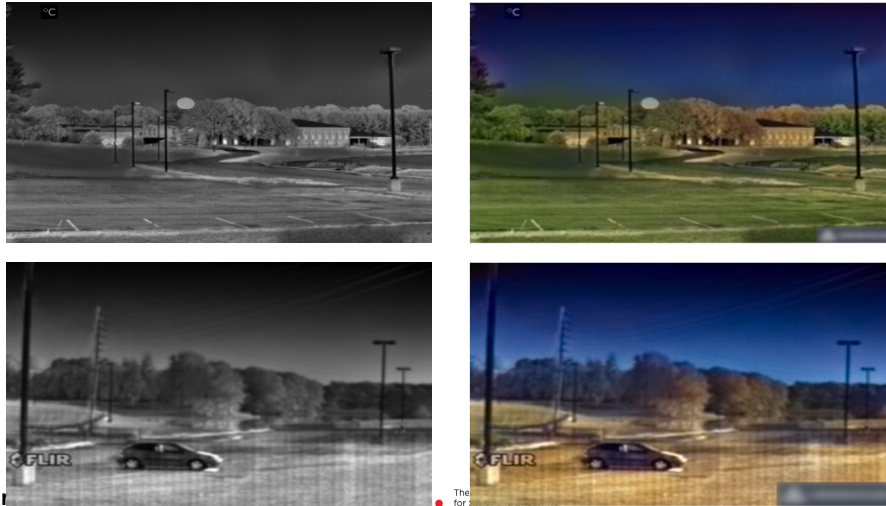
Joint Data Learning

6

xxix

Color Transferring for Night Vision Colorization Using Convolutional Neural Network (CNN)

- ▶ Illustration of colorizing the fused images (NIR + LWIR) from *night-vision* images using Zhang's VGG-Net



Joint Data Learning

7

Color Transferring for Night Vision Colorization Using Convolutional Neural Network (CNN)

- ▶ **Color transferring** uses a convolutional neural network (CNN) to map colors.
- ▶ The VGG-Net is one of commonly-used CNN models due to its great tradeoff between accuracy and complexity.
- ▶ Adding more NV samples (including multispectral images and corresponding daylight pictures) to the training → more realistic colorization results.
- ▶ A new VGG-net model can be trained by using only LWIR images because the LWIR (thermal) images have **consistent contrast** in both daytime and nighttime.

Joint Data Learning

SPIE The international society for optics and photonics

8

xxx

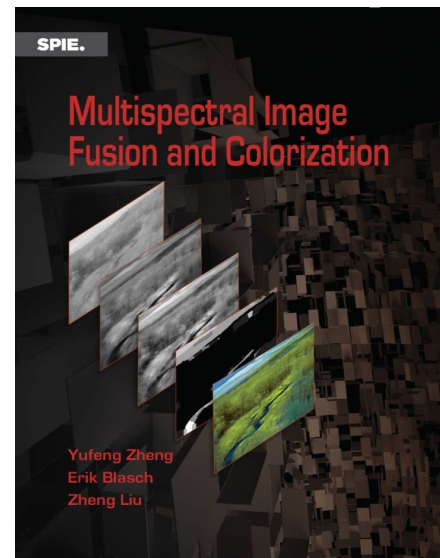
Future Trends

- With the development of network sensors, the Internet of Things, and Big Data, the coordination of information **fusion** and image fusion is **growing** quickly.
- **Multimodal** data and **multichannel** data fusion provide proven benefits, and as a result they can improve the performance of pattern recognition.
- Recent technological developments in *convolutional neural networks*, *deep learning*, and computational cognition will also facilitate the progress of image fusion research. *Vice versa*.

New Textbook Released

Multispectral Image Fusion and Colorization

- ▶ provides an overview of image fusion, associated evaluation methods, and its range of applications.
- ▶ presents an overview of fusion metrics and a comparison of objective metrics and subjective evaluations.
- ▶ includes 4 parts (~ 400 pages, \$70)
 - 1) historical background and basic concepts.
 - 2) image fusion theory.
 - 3) quantitative and qualitative evaluation.
 - 4) fusion applications: face matching, biomedical imaging, and night vision.
- ▶ **Link to the book:** [http://](http://spie.org/Publications/Book/2316453)
- ▶ spie.org/Publications/Book/2316453



Colorizing Night Vision Images a Trained CNN

Yufeng Zheng

Dept. of Data Science, University of Mississippi Medical Centre, Jackson, MS, USA

Emails: yzheng@umc.edu

ABSTRACT

Imagine an aircraft equipped with a multispectral imaging system to execute a nighttime object assessment task. In such a context, analyzing colorized multisensory images definitely improves observer situational awareness, reaction time, and perceptual analysis (human vision). The existing colorization methods have deficits in color resemblance and coloring efficiency. Therefore we propose to use deep learning neural network to realistically colorize night vision imagery. The objectives of the proposed research are (a) to transfer colors from daylight pictures to nighttime images, and (b) to train a convolutional neural network (CNN) model via deep learning.

Keywords: Color transferring, Night vision Colorization, Convolutional neural network (CNN), Deep learning.

1. INTRODUCTION

Multispectral images enable target recognition in low-light conditions (e.g., night vision), but it is inconvenient to directly observe and analyze multiple images of a scene. Instead, the analysis of the synthesized (fused or colorized) multisensory image proves more informative and more efficient for target recognition.¹ The fused multispectral imagery (in grayscale) can increase the reliability of interpretation^{2,3} and supports machine analysis (computer vision), whereas the colorized multispectral imagery (in colors) improves observer situational awareness,⁴ reaction time,⁵ and perceptual analysis (human vision).⁶

Colorized multispectral night vision (NV) images can enhance human vision by improving observer object classification and reaction times especially for low light conditions. NV colorization techniques can produce the colorized images that closely resemble natural scenes. Two categories of coloring methods, color fusion and color mapping, are briefly reviewed hereby. *Color fusion* directly combines multispectral NV images into a color-version image by mixing pixel intensities at different color planes. *Color mapping* usually maps the color properties of a false-colored NV image (source) onto that of a true-color daylight target picture (reference). Four coloring mapping methods, statistical matching, histogram matching, joint histogram matching, and lookup table (LUT) are reported in literatures. Comparable experiments^{7,8} showed that: (i) the segmentation-based colorization method produces very impressive and realistic colors but it requires intense computations; (ii) color fusion and LUT-based methods run very fast but their results are less realistic; (iii) the statistic matching method always provides acceptable results (i.e., never fails); (iv) histogram matching and joint-histogram matching can generate impressive and vivid colors when the color distributions between source and target are similar; and (v) the statistic matching then joint-histogram matching (SM-JHM) method is a reliable and efficient method recommended from both qualitative and quantitative evaluations.

The proposed CNN NV colorization performs color transferring from daylight pictures to nighttime images.

2. NIGHT VISION COLORIZATION WITH CNN

2.1 Colorizing grayscale images using VGG-Net

Zhang et al.⁹ presented a CNN model for colorizing *grayscale* images. Their CNN model is a fully automatic approach that produces vibrant and realistic colorizations. They embrace the underlying uncertainty of the problem by posing it as a classification task and use class-rebalancing at training time to increase the diversity of colors in the result.

The grayscale image to be colored can be thought as the *L*-channel of the image in the *Lab* color space and the objective is to find the *a* and *b* components. To simplify calculations, the *ab* space of the *Lab* color space is quantized into 313 bins as shown in Fig. 1a. So the *color prediction* task is now turned into a *multinomial classification* problem where for

every gray pixel there are 313 classes in the ab space to choose from.

The architecture proposed by Zhang et al.⁹ is a VGG-style network (Fig. 1b) with multiple convolutional blocks. Each block has two or three convolutional layers followed by a Rectified Linear Unit (ReLU) and terminating in a Batch Normalization layer. Unlike the VGG net, there are no pooling or fully connected layers. The input image is rescaled to 224×224 , and the output dimensions of the last convolution layer are $56 \times 56 \times 313$. Each of 56×56 pixels contains a vector of 313 values where each value represents the probability of the pixel belonging to that class. The goal is to find a single pair of ab channel values for each probability distribution. Notice that when the image passes through the CNN, its size decreases to 56×56 . Therefore the predicted ab image also has the dimension 56×56 . To obtain the color image, it is upsampled to the original image size and then added to the lightness channel, L , to produce the final color image.

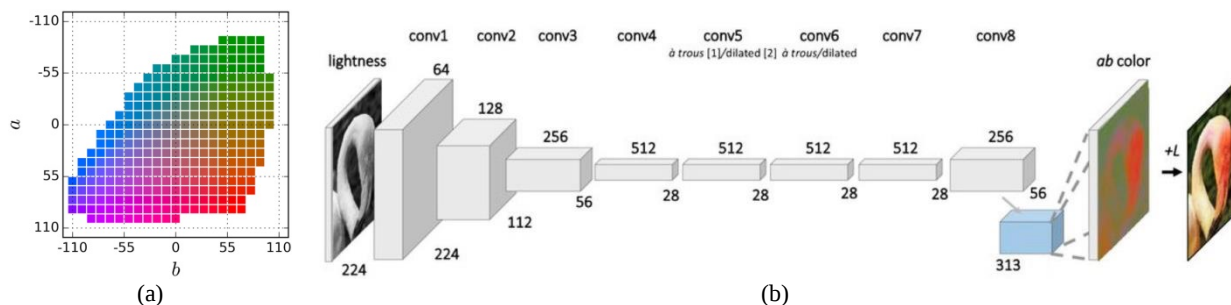


Fig. 1 Illustration of colorizing grayscale images using VGG-Net⁹: (a) 313 quantized ab colors in gamut. (b) Network architecture. Each conv layer refers to a block of 2 or 3 repeated conv and ReLU layers, followed by a BatchNorm layer. The net has no pool layers. All changes in resolution are achieved through spatial downsampling or upsampling between convolution blocks.

The network was trained with 1.3M images from the ImageNet training set. Two colorized images from daylight grayscale images are presented in Fig. 2.

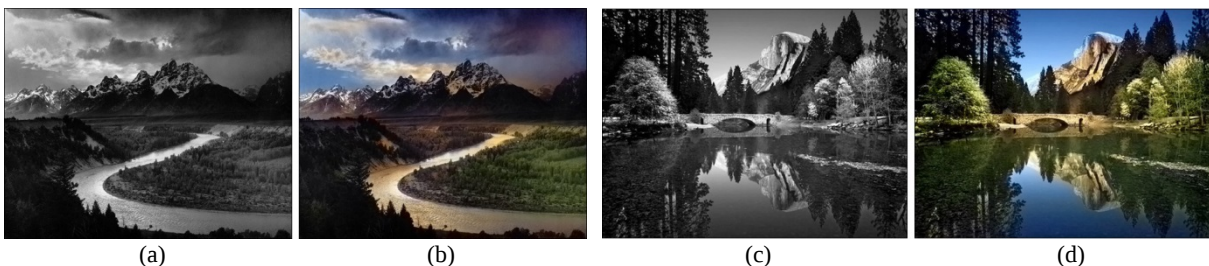


Fig. 2 Illustration of colorizing grayscale images using VGG-Net⁹: (a,c) Grayscale photos by Ansel Adams of Yosemite; (b,d) Colorized photos of (a,c).

2.2 Colorizing night vision images with CNN

The CNN model for colorizing grayscale images⁹ uses a modified VGG-Net (see Fig. 1b). The model training process takes a long time, which creates a mapping (weights) between the lightness (of a grayscale image, L) and its ab colors. The coloring process takes a (NV) grayscale image (L), passes it through the trained VGG-Net to get its ab components, then convert a Lab image to a RGB image. This coloring process is pretty fast and does not require extra memory. In addition, the trained model can be transferred.

Based on our reviews of the state-of-the-art CNN-based colorizing methods, we propose to use the VGG-Net (Fig. 1b) for color transferring from daylight pictures to NV images. In fact, the VGG-Net is one of commonly-used CNN models for color transferring due to its great tradeoff between accuracy and complexity. Using the pre-trained VGG-Net model (by ImageNet)⁹ to colorize two sets of NV grayscale images, the preliminary results are demonstrated in Fig. 3. In good illumination condition (daytime), both NIR and fused images resulted good colorizations (Fig. 3b, j), while in nighttime, the fused image led a better colorized image (Fig. 3l). The colored images (Fig. 3f, h) from thermal images are not realistic because no LWIR images were included in the training datasets (ImageNet). A new VGG-net model may be trained by using only LWIR images because the thermal images have consistent contrast in both daytime and nighttime.

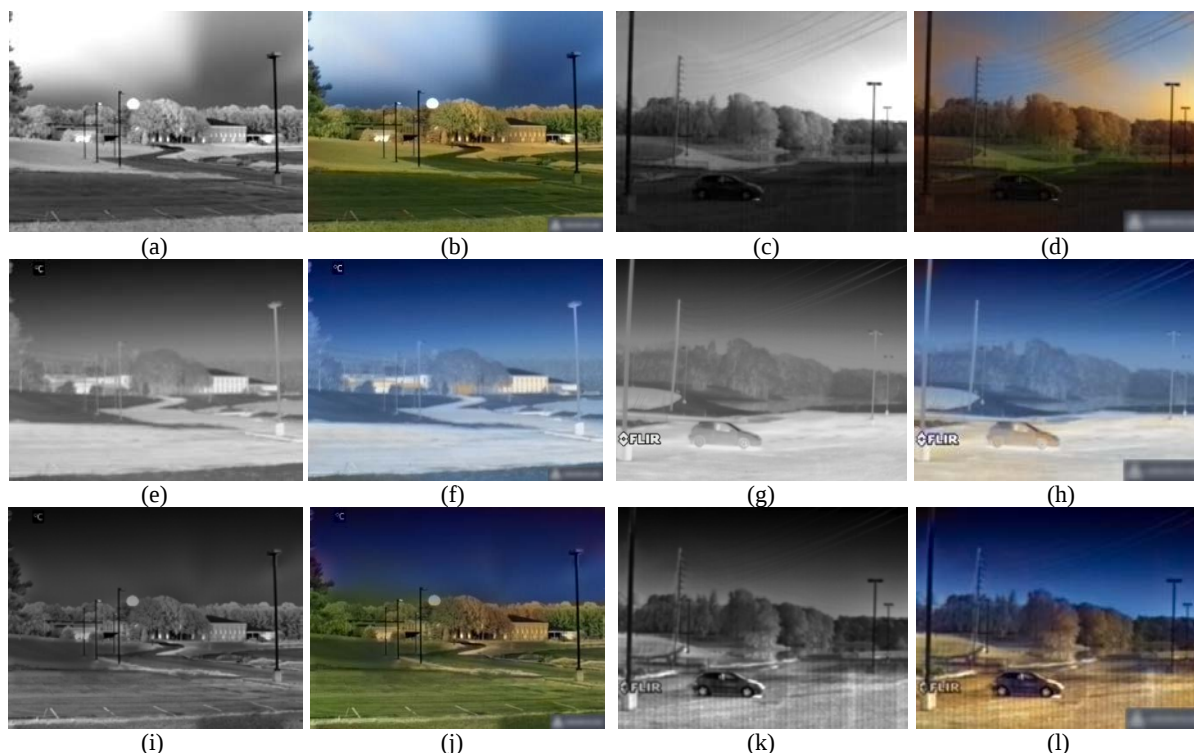


Fig. 3 Illustration of colorizing night-vision images using VGG-Net⁹: (a,e,i) NIR, LWIR, and fused (NIR+LWIR) images (#AT008, see Fig. 1); (b,f,j) Colorized images of (a,e,i); (c,g,k) NIR, LWIR, and fused (NIR+LWIR) images (#AT012, see Fig. 2); (d,h,l) Colorized images of (c,g,k).

3. CONCLUSIONS

The advantages of CNN-based color mapping include high color similarity, high robustness, and fast coloring process. Once a CNN model is trained, the coloring results do not depend on the contents of night vision images. The colorized images can be evaluated and compared with literature-reported results.

In some environmental conditions (e.g. low-light conditions), augmenting the imagery is of great interest to the military. Hence measurement data, combined with fast computer methods such as deep learning models and methods, could potentially offer a unique method for enhanced imagery presentation and target recognition.

4. REFERENCES

- [1] J. A. Ratches, "Review of Current aided/automatic target acquisition technology for military target acquisition tasks," *Opt. Eng.*, 50(7), 072001 (2011).
- [2] R. H. Rogers and L. Wood, "The history and status of merging multiple sensor data: an overview," *Technical Papers, ACSMASPRS Annual Conf. Image Processing and Remote Sensing* 4, pp. 352–360, (1990).
- [3] E. A. Essock, J. S. McCarley, M. J. Sinai, and J. K. DeFord, "Human perception of sensor-fused imagery," in *Interpreting Remote Sensing Imagery: Human Factors*, R. R. Hoffman and A. B. Markman, Eds., Lewis Publishers, Boca Raton, Florida, (2001).
- [4] A. Toet, J. K. IJpeert, A.M. Waxman, and M. Aguilar, "Fusion of visible and thermal imagery improves situational awareness," *Proc. SPIE* 3088, 177–188, (1997).
- [5] J. T. Varga, "Evaluation of operator performance using true color and artificial color in natural scene perception," (Report ADA363036), Naval Postgraduate School, Monterey, CA (1999).
- [6] A. M. Waxman et al., "Progress on color night vision: visible/IR fusion, perception and search, and low-light CCD imaging," *Proc. SPIE* 2736, 96–107 (1996).
- [7] Y. Zheng, W. Dong, E. Blasch, "Qualitative and quantitative comparisons of multispectral night vision colorization techniques" (Open Access), *Optical Engineering*, 51 (8), 087004, (2012).
- [8] Yufeng Zheng, Erik Blasch, Zheng Liu, "Multispectral Image Fusion and Colorization", SPIE Press, March 2018.
- [9] R. Zhang, P. Isola, A.A. Efros., *Colorful Image Colorizations*, 14th European Conference on Computer Vision (ECCV), Amsterdam 2016.

Invited Panel Discussion

“Joint Data Learning”

Frameworks of Maximum Entropy (MaxEnt) Based XAI in Multimodal Learning via Distributed Decision Fusion

Ivan Kadar

Interlink Systems Sciences, Inc.
Lake Success, NY, USA

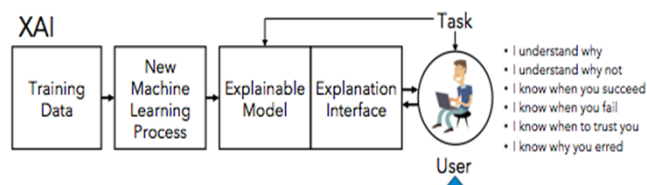
SPIE Conference 11756 “Signal Processing, Sensor Fusion and
Target Recognition XXX”, 12-16 April 2021, Online Digital Forum

1

Motivation and Approach

Motivation

Having introduced and demonstrated **the information theoretic model-based Maximum Entropy (MaxEnt) Machine Learning** Distributed Decision Fusion (DDF) “multimodal sensing” algorithm [1, 2, 3], achieving better detection performance than Bayesian & Neyman-Pearson detection criteria based models, & applicable for target classification/recognition, realized that **it possesses Explanatory/Interpretability features**, “unlike a black box algorithm”, that represent the proposed attributes of XAI (Explanatory AI - DARPA program [4], see illustration below).



Approach

- The *MaxEnt* machine learning **parametric model** is used for supervised classification and prediction, & serves as the global optimum (sufficient statistics) linearly weighted fusion rule.
- *Unlike Bayes DDF*, *MaxEnt* doesn't require knowledge of prior probabilities of local binary decisions, which are sent to a parallel architecture fusion center, for global predictions
- That is, in real world the performance probabilities of sensors may not be known, *but one can select a statistical model that represents the contextual information “x” of the process and producing an output “y”, and collecting N training data samples of “known features”*. That is binary sensor decisions modeled as **indicator feature functions** $f_n(x,y)=f(x)$.

2

Motivation and Approach (Cont'd)

Approach (Cont'd)

- Having collected *training data set* of N samples, an **empirical** probability distribution, **pe** is formed.
- Subsequently, *the goal is to form a model* probability distribution, **pm**, which satisfies constraints imposed by the **empirical** probability distribution **pe**.
*That is: $E_{pe}[f(x)] = E_{pm}[f(x)]$, and **pm** has to be a proper PDF. The resultant set is ill posed.*
- *Having satisfied the set of ill-posed constraints (by incorporating in the solution the maximum of Shannon entropy, viz. minimum KL divergence, and thereby achieving a uniform distribution for MaxEnt), are solved by optimization methods [1, 2], as part of computing the conditional probability, estimating/learning optimum fusion weights, and allow computation of the global prediction by the model: by forming the maximum of the conditional probability of the MaxEnt classifier.*
- *Than system having learned the sensors performance, can identify context changes; thus becomes **explainable and interpretable**, and provides multi sensors modality and ranked performance.*
- Note: The **Explanatory/Interpretability** features identified and highlighted herein, using the *MaxEnt* model-based machine learning approach, appear analogous to recent sophisticated approaches using model-based deep learning methods, being **Interpretable** (came across by coincidence on March 24, 2021 in a Webinar), detailed in [5] and related references.
Common element: both approaches are *model-based*.

3

About Maximum Entropy (MaxEnt)

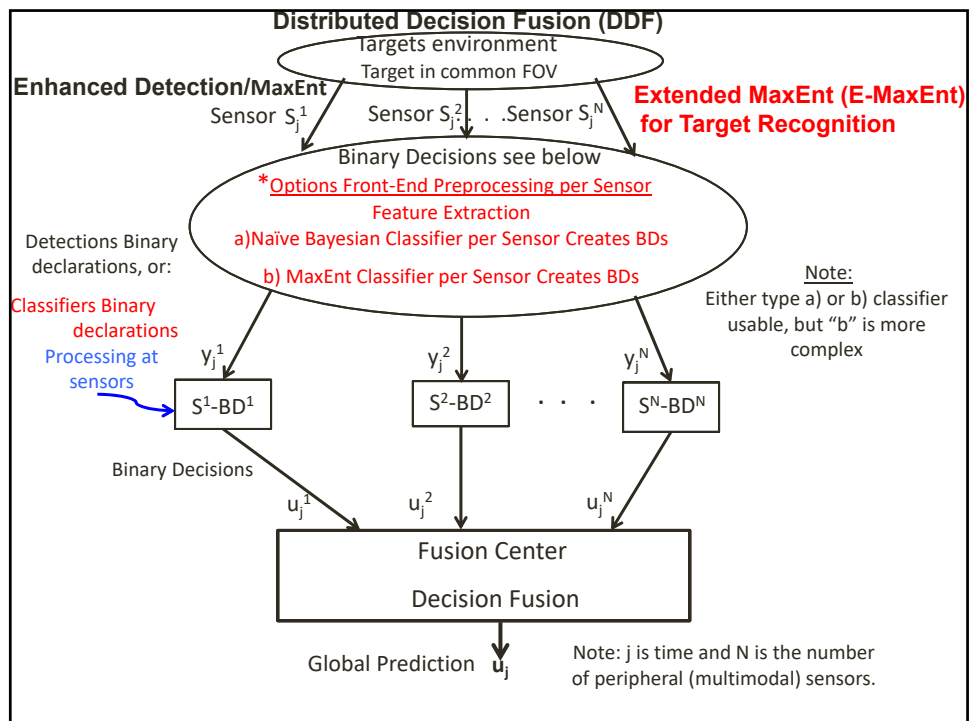
- Why use Maximum entropy-based classifier?
- In real-world applications sensor performance parameters may not be known, and classifiers. e. g., Naïve Bayes:
 - (1) requires exact knowledge of distributions in order to avoid prediction errors,
 - (2) **doesn't provide explanatory/interpretability information alike to MaxEnt.**
- Maximum Entropy → Minimize commitment, choose the most "uniform" distribution, i.e., **one with maximum entropy.**
- **Model** probability estimates should reflect what we know and what we don't know: **"in general, ignorance is preferable to error".**
- *MaxEnt* has been applied to several areas, e.g., word disambiguation it text search, image processing/reconstruction, computer vision, distributed detection [1] and target classification/recognition [2], fault detection, sensor measurements fusion, econometrics, context change detect, surveillance systems, IoT, CPS, sensor failure & degradation notification via **explanatory/interpretability** capability, etc.
- Evolved by correspondence of entropy in statistical mechanics [Jaynes 1957] and Shannon entropy.

4

MaxEnt Application to DDF for DD/Classification/Recognition

- Given the general description of *MaxEnt*, the end goal of the *MaxEnt* parametric model, used for parameter estimation/classification is:
- Having selected a statistical *model*, and associated (selected context attribute) indicator feature functions $f_n(x, y)$, and satisfying ill-posed constraints [1, 2] is to compute the global prediction U_j by the model, which maximizes the conditional probability, (as noted in prior and shown in subsequent slides).
- In reference to the parameters defined in Figure DDF (next slide) for either distributed detection or classification/recognition; symbols are noted below:
- S^i is the i^{th} sensor. Based on the local detection or classification/recognition criteria and threshold at each sensor, it makes a binary decision u_j^i , where, $i=1, \dots, N$; and j is the system time when decisions are made or iteration time/sample numbers in simulation. Note: the values of u_j^i are either 0 or 1.
- The binary decisions are assumed sent synchronized to the fusion center.
- The approach doesn't assume knowledge of actual prior probabilities, and is not dependent on individual sensor declaration scheme
- The optimum weights can be estimated without requiring conditional independence of local decisions.

5



6

MaxEnt Application to DF

- After satisfying constraints [1, 2] and using the binary decisions variables notation, and substituting the feature functions, the resultant expression for the conditional probability can be shown in Equation (x)

$$P(u_j / u_j^1, u_j^2, \dots, u_j^N) = \frac{\exp[\sum_{n=1}^N w_n f_n(u_j^1, u_j^2, \dots, u_j^N)]}{\sum_{x=0}^1 \exp[\sum_{n=1}^N w_n f_n(u_j^1, u_j^2, \dots, u_j^N, u_j = x)]} \quad (x)$$

where, $f_n(u_j^1, u_j^2, \dots, u_j^N, u_j)$ is the n^{th} feature function, w_n is the weight parameter of the n^{th} feature function to be estimated, and x is either zero or one (thus the denominator in (x) is a sum of two terms) indicating the binary declarations of sensors.

- The n^{th} feature function can be further specified as:

$$f_n^k(u_j^1, u_j^2, \dots, u_j^N, u_j = x) = \begin{cases} 1, & \text{if } u_j^k = x; \text{ and } \dots x = k; \text{ o.w.} = 0 \end{cases}$$

where, n is the n^{th} sensor, k corresponds to the sensor declaration state 0 or 1.

- Note: In case of E-MaxEnt for target recognition: - the feature functions depend on sensor type, and target recognition method selected.

7

MaxEnt Application to DF (Cont'd)

- In order to estimate w_n we need to solve the log likelihood equation below, which can be shown minimizes the Kullback-Liebler divergence between the model and the empirical distribution from training samples,

$$L = \sum_{j=1}^{DT} \log P(u_j / u_j^1, u_j^2, \dots, u_j^N)$$

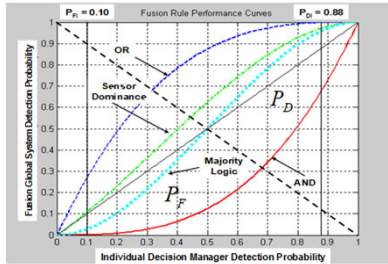
- where, the sum is over the number of training samples, DT , which equals the number of samples in the data set.
Since L is a concave function a global maximum can be found by standard convex optimization methods, such as gradient ascent.
- Used gradient ascent . The solution yields the optimum weights once the algorithm converges, $n=1, \dots, N$.
- Simulations show that the algorithm converges within fewer than fifty iterations. Note that this algorithm is executed for each data sample processed.
- Having learned the optimum weights, we can solve for the goal of the MaxEnt classifier method for DF, i.e., the optimum global predictions, given solution of equation below (yielding either zeros or ones in this case):

$$U_j = \arg \max P(u_j / u_j^1, u_j^2, \dots, u_j^N)$$

8

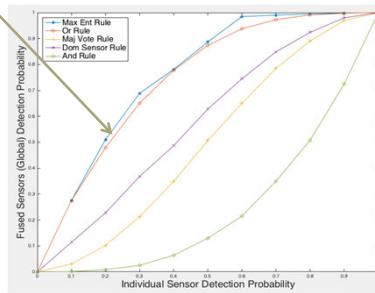
Distributed Decisions Fusion (DDF) for Enhanced Detection - **Explanatory Result** - Bayesian Case

Distributed Bayesian detection fusion concept [6], note: ad-hoc rules needed



- Simulation Example: Bayesian Case via *MaxEnt* – Best
 - No ad-hoc fusion rules needed

MaxEnt



9

Distributed Decisions Fusion (DDF) for Enhanced Detection - **Explanatory Results**

Example: Neyman-Pearson Approach [7, 8] - **Explanatory Result**

The results of the optimum *MaxEnt* weights distributions for the dissimilar sensors case, with different detection performances show, as expected, for the **optimal fusion combination**, that the largest weight attaches to the “**dominant**” sensor with highest P_D .

For five sensors: $P_{D1} = [.95 .94 .93 .92 .91]$ and $P_F = [.05 .04 .03 .02 .01]$ N-P

MaxEnt fusion results are: $P_D = 0.9983$ and $P_F = 0.0017$ (independent of thresholds)

The corresponding **weight parameters** are:

[3.425782, 3.408893, 2.478245, 2.466412, 2.376458] respectively.

As shown, the best “**dominant**” peripheral sensor gets assigned the largest weight, which explains sensors performance at that time, and degradation or context change among sensors could be detected.

10

Summary

Introduced

- The foundation, application, and efficacy of using *MaxEnt* multimodal sensing machine learning for DDF with applications to DD and target classification/recognition.
- The *MaxEnt* machine learning algorithm capability for providing ***Explanatory/Interpretability features***.
- Examples of Bayesian and Neyman-Pearson approach vs. *MaxEnt* machine learning approach to DD target detection, demonstrating “*MaxEnt*” best performance and *Explanatory/Interpretability features*.
- The systems learns actual sensor detection performances and recognizes context induced changes.
- Based on the proven excellent DDF detection performance of “*MaxEnt*”, one would expect similar result of “*MaxEnt*” for target classification/recognition with *Explanatory/Interpretability features*:
- For Example: finding the presence, viz., recognizing a previously known target of interest, among many similar targets in areas where previously seen targets most likely are, along with other targets, context changes, sensor change or degradation, target location change, with *Explanatory/Interpretability features*, etc.

11

References

1. I. Kadar and K. Hintz, “Adaptive *MaxEnt* Modeling of Distributed Decision Fusion without Knowledge of Prior Probabilities of Local Decisions”, *Fusion 2016, Heidelberg, Germany*.
2. I. Kadar, “A framework for adaptive *MaxEnt* modeling within distributed sensors and decision fusion for robust target detection/recognition”, *Signal Processing, Sensor Fusion and Target Recognition XXVII*, edited Ivan Kadar, Proc. SPIE Vol. 10646, Orlando, FL, April 2018.
3. <https://www.darpa.mil/program/explainable-artificial-intelligence>
4. T. Baltrušaitis, et.al, “Multimodal Machine Learning: A Survey and Taxonomy”, *IEEE Trans. on PAMI*, 2017
5. “Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing”- “Combining accurate model-based algorithms and deep learning: Compact. Interpretable and simple to train data-driven systems” Presentation by Professor Yonina Eldar, yonina.eldar@weizmann.ac.il
SPACE Webinar March 24, 2021
6. Chapter 8: “Overview of Distributed Data Fusion”, *Handbook for Multi-Sensor Data Fusion*, Second Edition: Theory and Practice, Eds: D. Hall, M. Liggins, and J. Llinas, CRC Press, 2009, Florida.
7. S. C. A. Thomopoulos, R. Viswanathan and D. C. Bougoulas. “Optimal decision fusion in multiple sensor systems,” *IEEE Trans/on Aerospace and Electronic Systems*, AES-23(5), pp. 644-653, 1987.
8. R. S. Blum, “Necessary conditions for optimum distributed sensor detectors under the Neyman-Pearson criterion,” *IEEE Trans. Inform. Theory*, vol. 42, pp. 990–994, May, 1996.

12

Frameworks of Maximum Entropy (*MaxEnt*) Based XIA in Multimodal Machine Learning via Distributed Decision Fusion

Ivan Kadar
Interlink Systems Sciences, Inc., Lake Success, NY11042

ABSTRACT

Motivation for the panel presentation was to describe a learning technique: (1) which is unlike a black box, (2) along the proposed principles of the DARPA XAI program [1], and (3) can provide explanatory/interpretable capability. The technique is based on the Maximum Entropy (*MaxEnt*) model-based machine learning approach applied to Distributed Decision Fusion (DDF) for both multimodal sensing [2] target detection and target classification/recognition. The approach was described in previous publications, and demonstrated the *MaxEnt* model-based machine learning DDF achieving the best, realistic detection performance demonstrating the effectiveness of the method. The same effectiveness is expected applying *MaxEnt* model-based machine learning to multimodal sensing for target classification/recognition as described in-part below and references. The model-based aspect is the key element of the algorithm enabling to provide explanatory/interpretable capability.

Keywords: Model-based, Machine Learning, Maximum Entropy, Multimodal Sensing, and Distributed Decision Fusion

1. INTRODUCTION

The Maximum Entropy (*MaxEnt*) information theoretic model parametric framework was introduced in two prior papers by the author [3, 4], both use distributed decision fusion (DDF) without knowledge of prior probabilities of local decisions. Reference [1] demonstrated the effectiveness of the *MaxEnt* fusion center by achieving the best, realistic detection performance with respect to published results of either the Bayesian formulation or the Neyman-Pearson criterion. Reference [4] paper depicted an extension of *MaxEnt* DDF using individual sensor *MaxEnt* classifiers, called *E-MaxEnt*, for targets classification/recognition, and fusing local classifier decisions. Specifically, in *E-MaxEnt* each sensor has a front-end pre-processing system to process unique target attributes extracted for example from observed target imagery, which attributes are stored for reference/learning/comparison in the sensors *MaxEnt* classifiers. Based on the degree of match, each sensor generates local binary decisions that are sent to a *MaxEnt* fusion center, in the usual parallel architecture. No assumptions are made about knowing any local decision rules. At least three sensors are used for taking simultaneous (synchronized) measurements with overlapping FOV overages. Note: The approach depicted in [4] can be applied to any type of sensor i.e., multimodal sensing [2], for target recognition/classification, and is independent on the type of classifier used making the binary decisions.

2. EXPLANATORY/INTERPRETABLE CAPABILITY

Given the above-depicted capability, and the accompanying detailed presentation slides, demonstrates how the algorithm comes up with an answer to explain and interpret outcomes. Basically the *MaxEnt* model-based machine learning approach can compare the established model statistics with the online empirical statistics in the learning process, while guarantying the uniform distribution base in the optimization process, in order to come up the optimum of the conditional probability to make a prediction. If it is a “match”, the algorithm learned the properties of the multimodal sources either in DDF for detection or for target classification/recognition. A succinct few lines summary is presented in the accompanying slides. A detailed description is given below from [3,4]

The first step is to construct a **statistical model, which represents the behavior of a process** by collecting a set of training data from it. For example, in text search use contextual information x in a document, and produce an output y . That is, obtain samples (x_i, y_i) where x_i represent contextual information about words (e.g., “respect”) from the document, and y_i its class (“noun”). This yields an empirical probability distribution, where N is the size of the training data set as shown below. That is one can introduce a function, say an N -dimensional real-valued feature vector function

$f(x,y)$, and depict the above as one feature in the form of an indicator function

$$f_n(x,y) = \{ 1 \text{ if } x \text{ is a 'respect' and } y \text{ is Noun; } 0 \text{ otherwise} \}$$

In the application of *MaxEnt* to DDF, in an analogous manner to the above, the feature functions represent binary declarations from sensors (which are results of hypothesis tests), and $f_n(x,y)$ will depict presence (detected) or absence (not detected) of a target. *The initial training data is generated from a uniform distribution* [0, 1], target present or absent (i.e., detected or not detected), or ATR like application recognized or not recognized the target, is equally likely.

The expected value of $f(x,y)$ with respect to the empirical distribution $\hat{p}(x,y)$ is the statistics needed in the sequel (the notation below will use $f(x)$ and $\hat{p}(x)$ for the empirical distribution): viz., the expected value of the feature vector from the observed probability of x in the training data: $\hat{E}[f] = \sum_{x \in X} \hat{p}(x)f(x)$, that is, the aggregate collected feature vectors constitute testable information for use. The goal is to construct a model distribution p , which satisfies the constraints imposed by the empirical distribution $\hat{p}(x)$, viz.,

$$E[f] = \hat{E}[f] \quad (1)$$

i.e., the expected value needs to be equal to empirical training value, and guarantee that p is proper probability distribution,

$$\sum_{x \in X} p(x) = 1 \quad (2)$$

That is, the classifier is to learn from the training data, and use it to classify (match) incoming information and predict outcome. In general, the training data could also consists of collected random historical data from a particular DF system. This problem is ill posed and many models would satisfy (1) and (2). Invoking the principle of maximum entropy one needs to find a distribution, which, maximizes $H(p)$, where H is Shannon information entropy [3, 4]. The above is a constrained optimization problem, which is solved using the method of Lagrange multipliers, with result yielding the *MaxEnt* exponential classifier [3,4], please see slides for the final resultant equations.

As described above, having an accurate model is the key [5]. Please refer to the detailed slides, references listed for further details. .

SUMMARY

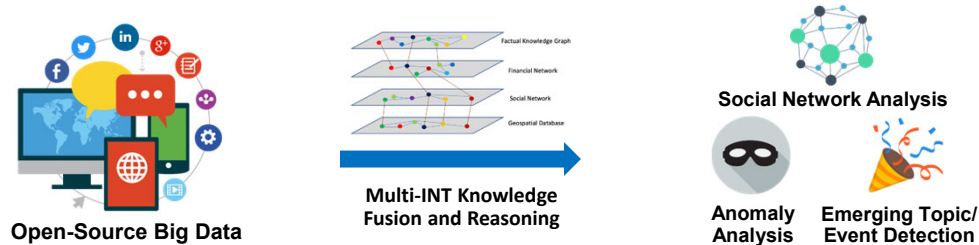
Motivation for the panel presentation was to describe a learning technique: (1) which is unlike a black box, (2) along the proposed principles of the DARPA XAI program [1], and (3) can provide explanatory/interpretable capability. The technique is based on the Maximum Entropy (*MaxEnt*) model-based machine learning approach applied to Distributed Decision Fusion (DDF) for both multimodal sensing target detection and target classification/recognition. The approach was described in previous publications, and demonstrated the *MaxEnt* model-based machine learning DDF achieving the best, realistic detection performance demonstrating the effectiveness of the method. The same effectiveness is expected applying *MaxEnt* model-based machine learning to multimodal sensing for target classification/recognition as described in-part below and references. The *model-based aspect* is the key element of the algorithm enabling to provide explanatory/interpretable capability.

REFERENCES

1. <https://www.darpa.mil/program/explainable-artificial-intelligenceI>.
2. T. Baltrusaitis, et.al, "Multimodal Machine Learning: A Survey and Taxonomy", *IEEE Trans. on PAMI*, 2017
3. Kadar and K. Hintz, "Adaptive *MaxEnt* Modeling of Distributed Decision Fusion without Knowledge of Prior Probabilities of Local Decisions", *Fusion 2016, Heidelberg, Germany*.

4. I. Kadar, "A framework for adaptive *MaxEnt* modeling within distributed sensors and decision fusion for robust target detection/recognition", *Signal Processing, Sensor Fusion and Target Recognition XXVII*, edited Ivan Kadar, Proc. SPIE Vol. 10646, Orlando, FL, April 2018.
5. "Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing"- "Combining accurate model-based algorithms and deep learning: Compact. Interpretable and simple to train data-driven systems" Presentation by Professor Yonina Eldar, yonina.eldar@weizmann.ac.il SPACE Webinar March 24, 2021

Real-time KG-based Trend and Sentiment Analysis System (KG-TSA)



Genshe Chen, Intelligent Fusion Technology, Inc.
gchen@intfusiontech.com

Joint Data Learning at SPIE DCS

1

Background and Motivation

- 
 - **Huge growth in social media, especially online social networking services, has revolutionized the ways people organize, communicate, coordinate, and behavior**
 - Imperative opportunity to develop scalable sentiment analytics to explore the evolving behaviors and predict the intents of the susceptible entities
- 
 - **Key challenge: identify the hidden trend and evolving behaviors and detect the change in sentiment measures using AI/ML-based algorithms**
 - Develop a fundamentally different way to avoid laborious human interpretation and assessment
- 
 - **Social Networks Analysis**
 - Detecting Distributed Covert Networks and their relationships
 - Using target social media connections discover sleeper cells and their activity (e.g. Hezbollah)
- 
 - **Anomaly Detection and Behavior Analysis**
 - Identification of harmful intent; Low SNR signatures (of one or multiple datasets)
 - Detect a front organizations and proxy groups through anomalous activities (e.g. origins of terrorist funding)
- 
 - **Automated entity identification, predicted events, & motivations**
 - Statistically probable event locations and targets
 - The likely harmful action to be taken/hypothetical scenarios
 - List of motivating factors/reasoning (e.g. retaliation for an assassination)

Joint Data Learning

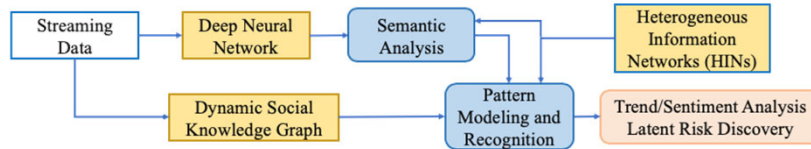
SPIE The international society
for optics and photonics

2

Real-time KG-based Trend and Sentiment Analysis System (KG-TSA)

Technical Approaches:

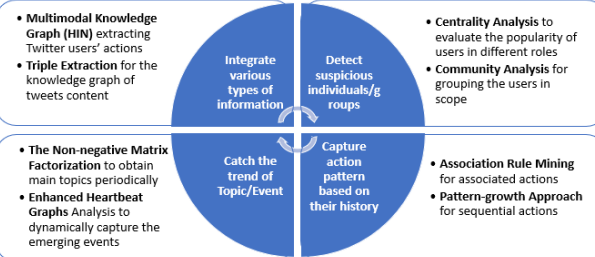
- Dynamic Knowledge Graph for representation and inference, knowledge-aware deep learning for trend and sentiment analysis, and pattern mining for anomaly detection and latent-risk discovery.



System Goal:

- Dynamically provide a solution for intelligence analysis, where no obvious models or signatures are available, to identify potential insider threats. Especially when the risk indicators and warning signs are from low-signal-to-noise events and transactions.

Major Functions VS Key Technical Approaches



Joint Data Learning

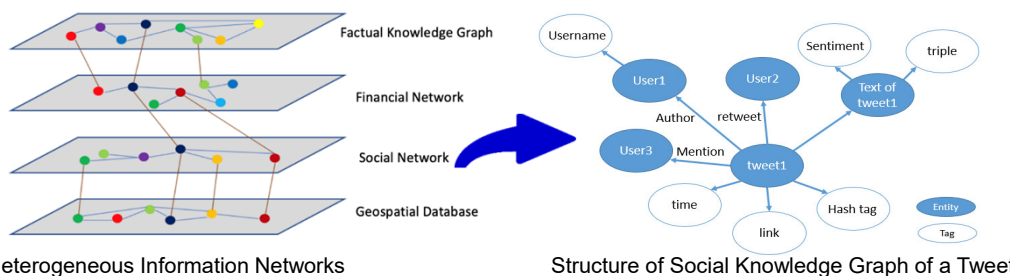
SPIE The international society for optics and photonics

3

Multimodal Knowledge Graph

Multi-Layer Knowledge Graph – Heterogeneous Information Networks (HINs)

- Fuse data from multiple facets for sentiment analysis with multi-layer knowledge graph to connect all components related to an entity and its behavior.
- For each layer, the node denotes the specific concepts associated with an entity. These concepts at different layers are linked together if they are associated with the same entity or location
- Facilitate fusion of observed activities detected by multiple sources with multimodal data types.
- Enable rapid understanding of complex dynamic events associated with the susceptible entities by "connecting the dots" of the risk indicators and warning signs related to hidden activities and transactions.
- This multimodal KG could also be used to profile and discover affected entities or groups of interest by continually monitoring their activities in context related to the changing state of perception



Joint Data Learning

SPIE The international society for optics and photonics

4

Graphic Interface for Social Network and Pattern Analysis

Streaming Tweets by keywords for Suspicious Network (all countries)

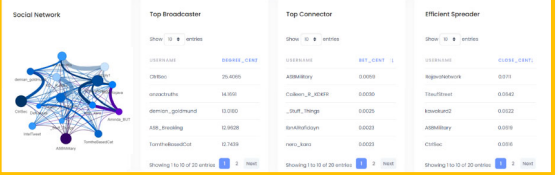
TIME POSTED (UTC)	USERNAME	TEXT
2020/07/20 00:07	Mpuyyeyjss	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	uauyky	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	Mpuyyeyjss	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	uauyky	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	JAMISDSDSD	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE

Showing 1 to 5 of 5,570 entries

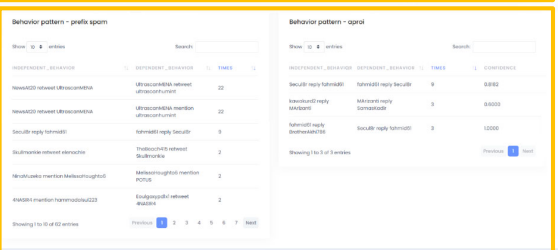
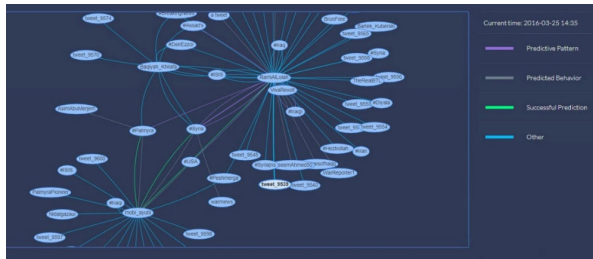
Raw tweet data from Scrapping

dynamically updates by searching based on analysis

Social Network Graph and Centrality Analysis



Multi Modal Knowledge Graph Construction for extracting online action dynamically



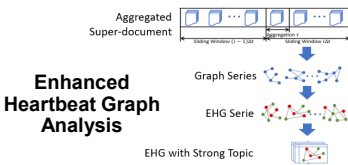
Historical Behavior Patterns

Joint Data Learning

SPIE The international society for optics and photonics

5

Graphic Interface for Emerging Event Detection



Enhanced Heartbeat Graph Analysis

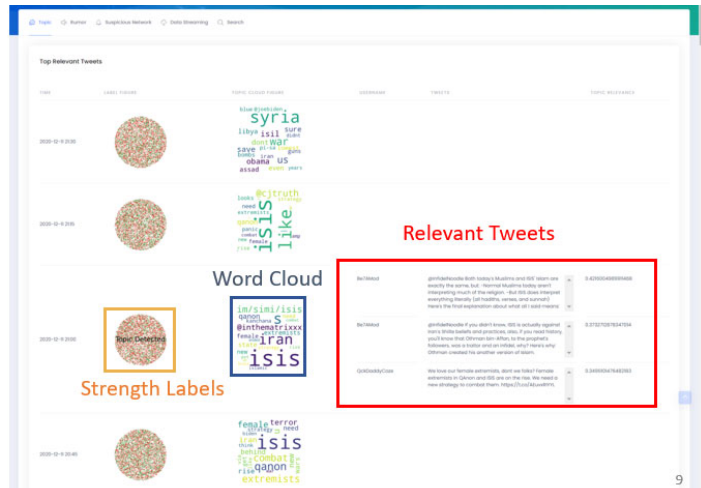
Streaming Tweets by keywords for Suspicious Network (all countries)

TIME POSTED (UTC)	USERNAME	TEXT
2020/07/20 00:07	Mpuyyeyjss	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	uauyky	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	Mpuyyeyjss	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	uauyky	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE
2020/07/20 00:07	JAMISDSDSD	RT @gndevdewat: gndevdewat PLEASE KEEP IN MIND THAT FROM OUR INTL, REPLICAS SHOW THAT IF WAS THE PROUD BOYS AND REPLICAS THAT CAUSE

Showing 1 to 5 of 5,570 entries

Raw tweet data from Scrapping

Emerging Event with Most Relevant Tweets

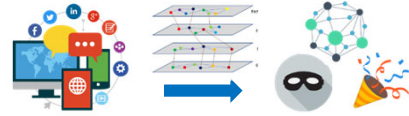


Joint Data Learning

SPIE The international society for optics and photonics

6

Thanks



Qingliang Zhao, Jiaoyue Liu, Nichole Sullivan, Kuochu Chang, John Spina, Erik Blasch, and Genshe Chen "Anomaly detection of unstructured big data via semantic analysis and dynamic knowledge graph construction", Proc. SPIE 11756, Signal Processing, Sensor/Information Fusion, and Target Recognition XXX, 117560N (12 April 2021); <https://doi.org/10.1117/12.2589047>

Genshe Chen, IFT
gchen@intfusiontech.com

Joint Data Learning

SPIE The international society
for optics and photonics

Real-time Knowledge Graph-based Trend and Sentiment Analysis System (KG-TSA)

Qingliang Zhao^a, Jiaoyue Liu^a, Nichole Sullivan^a, Kuochu Chang^b, John Spina^c, Erik Blasch^c, Genshe Chen^{*a}

^aIntelligent Fusion Technology, Inc., 20271 Goldenrod Ln. Germantown, MD 20876;

^bGeorge Mason University, 4400 University Dr. Fairfax, VA 22030;

^cAir Force Research Laboratory, Rome, NY 13441

ABSTRACT

Multi-modal data from social media and micro-blogging provides great opportunity for sentiment analytics, social network analyses, behavior/event prediction, and motivation analysis. To utilize these opportunities automated data collection, exploitation, and dissemination is needed. This paper introduces the core technology Intelligent Fusion Technology, Inc. (IFT) developed for products that perform Multi-Intelligence Knowledge Fusion and Reasoning. This technology is a real-time, knowledge graph-based, trend and sentiment analysis product for automated latent risk discovery. The framework collects real-time streaming data from multiple public data sources. It then interprets and extracts interactions, events, and activities. It uses these extractions to discover relevant patterns, determine, and identify latent risks. This technology performs pattern discovery and activity analysis including centrality analysis, behavior pattern analysis, and emerging topic detection. The framework is implemented in a user-friendly web-application as a working prototype, with real-time streaming data to prove the technical feasibility.

Keywords: data fusion, knowledge graph, open-source big data, semantic analysis, social network analysis, emerging topic detection

1. INTRODUCTION

The past decade has seen a huge growth in social media and micro-blogging usage such as Twitter. This multi-modal data provides opportunity to develop scalable sentiment analytics, explore evolving behaviors, and predict the potential actions of different scales of communities and groups. However, it is not possible to conduct labor-intensive collection, exploitation, and dissemination on such a large scale of data. Therefore, automated entity identification, event prediction, and motivation analysis is needed.

To this end, IFT's developed a Real-time, Knowledge Graph-based, Trend and Sentiment Analysis System (KG-TSA) for automatic and scalable intelligence analysis. Two typical use cases of KG-TSA are (1) Social Networks and Behavior Analysis, and (2) Emerging Event Detection. KG-TSA is the core technology of the ADUSAK and more details can be found in [1].

2. REAL-TIME KG-BASED TREND AND SENTIMENT ANALYSIS SYSTEM

KG-TSA identifies potential inside threats, especially when the risk indicators and warning signs are from low-signal-to-noise data. The system inputs bigdata from social media and other customer provided, private, streaming data sources and outputs trend, sentiment and latent behavior patterns and risks. An overview of KG-TSA is shown in Figure 1.

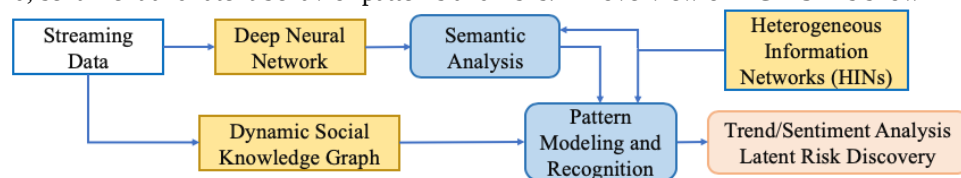


Figure 1. overview of KG-TSA System

First step of the process is streaming data collection from the multi-intelligent sources. Then, the text data is processed by deep neural networks to perform semantic analysis. A social KG is constructed from the raw data as well. Next, the processed data from the different sources are fused in a Heterogeneous Information Network, or HIN. The HIN contains

the processed semantic data and a social Knowledge graph, for pattern modeling. Finally, the results are shown on a graphic user interface.

3. MULTIMODAL KNOWLEDGE GRAPH

The key to our data fusion technology is the Multi-Layer Knowledge Graph. It fuses data from multiple sources via sentiment analysis, with a multi-layer knowledge graph to connect an entity and its behavior. For each layer, the node denotes the specific concepts associated with an entity. These concepts at different layers are linked together if they are associated with the same entity or location. Specifically, in social network layer, a tweet is represented by social network knowledge graph [2] which contains entities such as users and text, relations such as author and retweet, tags such as timestamps and hash tags. The structure of Multi-Layer Knowledge Graph and Social Network Knowledge Graph is shown in Figure 2.

The Multi-Layer Knowledge Graph facilitates fusion of observed activities, detected by multiple sources, with multimodal data types. It enables rapid understanding of complex dynamic events by “connecting the dots” of the risk indicators and warning signs related to hidden activities. This multimodal Knowledge Graph is intended to profile and discover affected entities or groups of interest by continually monitoring their activities in context related to the changing state of perception.

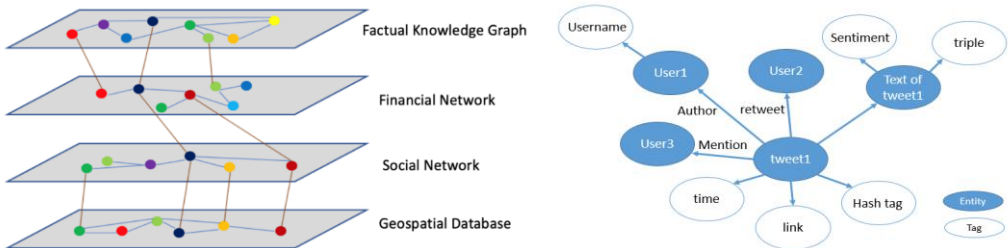


Figure 2. The Multi-Layer Knowledge Graph (Left) and Social Network Knowledge Graph (Right)

The two typical use cases based on Multi-Layer Knowledge Graph, Social Network and Pattern Analysis, are described in Section 4 and 5.

4. GRAPHIC INTERFACE FOR SOCIAL NETWORK AND PATTERN ANALYSIS

The process and graphic user interface (GUI) of Social Network and Pattern Analysis is shown in Figure 3.

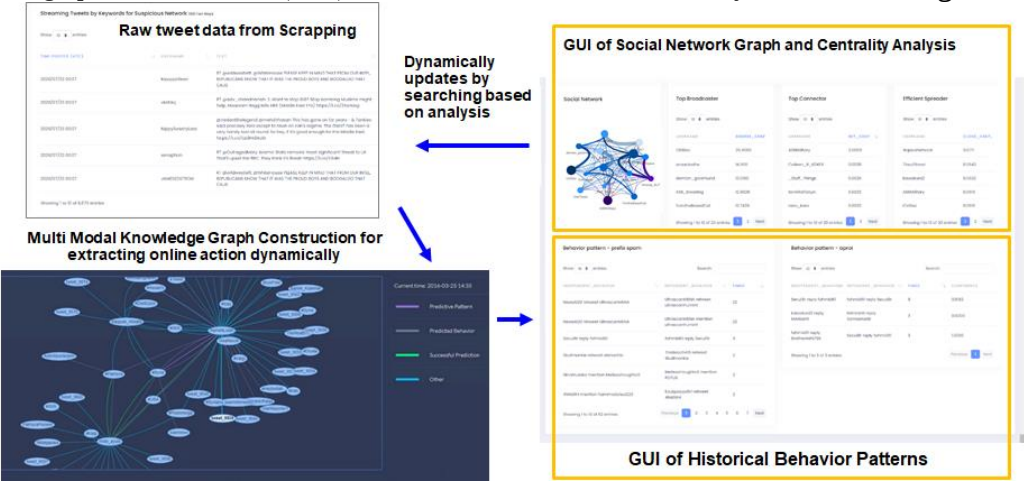


Figure 3. The process and graphic user interface (GUI) of Social Network and Pattern Analysis

In this process, raw data is scrapped using tweet streaming API, then Multi Modal Knowledge Graph is Constructed for extracting people online action dynamically. Results of the Social Network Graph, Centrality Analysis [3], and Historical Behavior Patterns [4] are shown on the Graphic user Interface. And the results feedback to the data collection process, and dynamically update searching scope.

GUI of Social Network Analysis section is a visualized social network graph. In the social network graph, each node represents a user. The larger the node, the more connected that user. Each line represents a connection between two users. In a similar fashion, the thicker the line, the stronger the connection. GUI of Behavior Pattern Section displays the initiator of the independent behavior that exerts influence on the group behavior.

5. EMERGING EVENT DETECTION

Emerging Topic Detection captures the topic of unknown events that are gaining attention and popularity. The process and GUI of Emerging Topic Detection is shown in Figure 4. This approach also starts with raw data construction. Via a series of Enhanced Heartbeat Graphs [5], the top emerging topic is ranked. If the topic strength is strong, it indicates that one or more new events might be gaining attention and popularity. In the GUI of Emerging Topic Detection, Key Word Cloud shows major keywords related to the captured topic. Most Related Tweets of each Topic are screened and displayed to show the contents that are spreading in plain English.

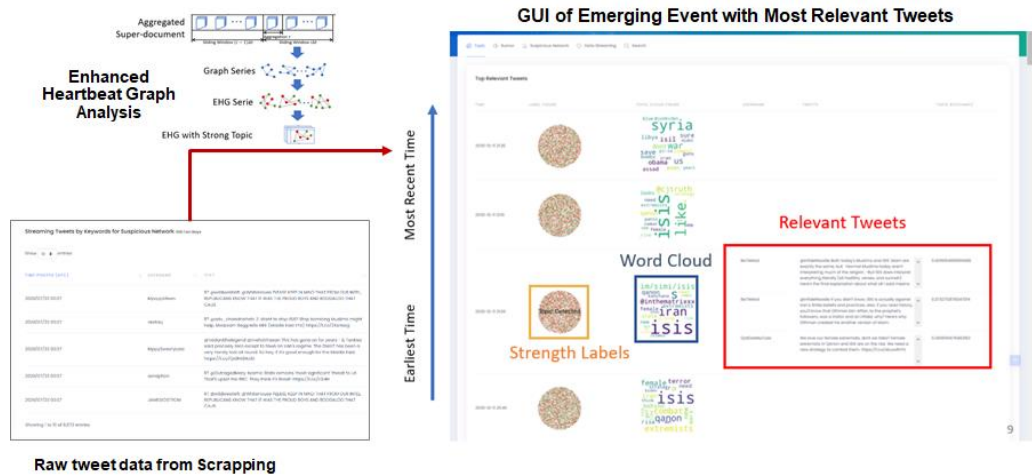


Figure 4. The process and GUI of Emerging Topic Detection

6. SUMMARY

IFT developed a Real-time KG-based Trend and Sentiment Analysis System for automated latent risk discovery. This includes the construction of knowledge graphs by collecting data from big-data sources, knowledge graph based Multi-intelligence fusion by representing factual knowledge and social networks. This technology performs pattern discovery and activity analysis including Centrality Analysis and Behavior Pattern Analysis and Emerging Topic Detection. The framework is implemented in a user-friendly web-application as a working prototype, with real-time streaming data to prove the technical feasibility.

REFERENCES

[1] Q. Zhao, J. Liu, N. Sullivan, K. Chang, J. Spina, E. Blasch, G. Chen, "Anomaly detection of unstructured big data via semantic analysis and dynamic knowledge graph construction," Proc. SPIE 11756, 117560N (2021) .
[2] R. Clancy, I. F. Ilyas and J. Lin, "Scalable Knowledge Graph Construction from Text Collections," Proc. the Second Workshop on Fact Extraction and VERification, 39-46, (2019).
[3] L. C. Freeman, "A set of measures of centrality based upon betweenness," Sociometry, vol. 40, no. 1, 35-41 (1977)
[4] R. Agrawal, T. Imielinski, A. Swami, "Mining association rules between sets of items in large databases," ACM SIGMOD International Conference on Management of Data, 207-216 (1993)
[5] Z. Saeed, R. Abbasi, I. Razzak, O. Maqbool, A. Sadaf and G. Xu, "Enhanced Heartbeat Graph for emerging event detection on Twitter using time series networks," Expert Systems with Applications, vol. 136, 115-132 (2019)

JOINT DATA LEARNING: **Robust Out of Distribution Radio Frequency Target Recognition**

Panel Discussion
SPIE DCS 2021

Dr. Uttam K. Majumder,
Senior Electronics Engineer
Air Force Research Lab
AFRL PA Number: AFRL-2020-0519

1

Research Challenge:

- How to improve In-library/In-Distribution (ID) Target Detection and Classification performance when we expect similar targets in the scene i.e., separating ID targets from out-of-library / Out-of-distribution (OOD) targets/confuser?
- ***Hypothesis:***
 - Develop a classifier by “Joint Data Learning (JDL)” Approach
 - Develop a Deep Learning (DL) Model using ID data as well as OOD data
 - For Robust Radio Frequency Target Recognition, we used both Expected Targets (ID) and unexpected targets (OOD) data to train a Deep Neural Networks
 - JDL approach yields better detection and classification Performance
- **Our Goal**
 - Co-design a Deep Learning-based SAR-ATR system for high ID accuracy and robust OOD detection

2

ii

In Distribution (ID) Data: Synthetic and Measured Paired Labeled Experiment (SAMPLE) Dataset

- **Paired** (real, synth) SAR signatures for MSTAR targets

- **Public Dataset Configs**

- **Azimuth Range: $10^{\circ} - 80^{\circ}$**
- **Elevation Range: $14^{\circ} - 17^{\circ}$**
- **10 data classes**

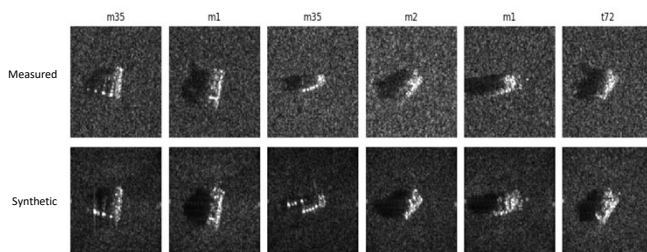


TABLE I
DETAILS OF THE SAMPLE DATASET

Class #	Class Name	# Train	# Test	Total
0	2S1	116	58	174
1	BMP2	55	52	107
2	BTR70	43	49	92
3	M1	78	51	129
4	M2	75	53	128
5	M35	76	53	129
6	M548	75	53	128
7	M60	116	60	176
8	T72	56	52	108
9	ZSU23	116	58	174
Totals		806	539	1345

Parameter	Value	Parameter	Value
Range resolution	0.30 m	Bandwidth	591 MHz
Range pixel spacing	0.20 m	Center frequency	9.6 GHz
Range extent	25.8 m	Image size	128 × 128
Cross-range resolution	0.30 m	Polarization	HH
Cross-range pixel spacing	0.20 m	Elevations	14°-18°
Cross-range extent	25.8 m	Taylor weighting	-35 dB

3

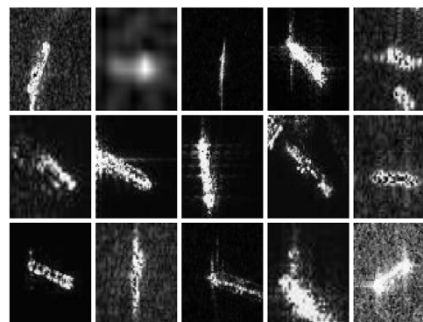
Out of Distribution (OOD) Data: Ship Detection Dataset

- Originally a SAR ship detection dataset covering a variety of sensors and imaging modes
- 59,535 total ships from 43,819 full frame images
- Represents a diverse mix of SAR data from different sensors, imaging modes, polarizations and resolutions for large ships on the open water
 - **Note, this data has no direct relationship to the SAMPLE/MSTAR data and is unlabeled**

Table 1. Detailed information for original synthetic aperture radar (SAR) imagery.

Sensor	Imaging Mode	Resolution Rg. × Az. (m)	Swath (km)	Incident Angle (°)	Polarization	Number of Images
GF-3	UFS	3 × 3	30	20~50	Single	12
GF-3	FSI	5 × 5	50	19~50	Dual	10
GF-3	QPSI	8 × 8	30	20~41	Full	5
GF-3	FSII	10 × 10	100	19~50	Dual	15
GF-3	QPSII	25 × 25	40	20~38	Full	5
Sentinel-1	SM	1.7 × 4.3 to 3.6 × 4.9	80	20~45	Dual	49
Sentinel-1 ¹	IW	20 × 22	250	29~46	Dual	10

¹ Only single cross-polarization SAR images are used to construct the dataset. Ultrafine Strip-Map (UFS), Fine Strip-Map 1 (FSI), Full Polarization 1 (QPSI), Full Polarization 2 (QPSII), S3 Strip-Map (SM), and Fine Strip-Map 2 (FSII) imaging modes, respectively.



4

iii

Our OOD Detection Workflow

- **Step 1:** Train classifier to be accurate on ID training dataset
 - ✓ Training is performed “offline” with no knowledge of the expected OOD samples
- **Step 2:** Instantiate an OOD detector for the trained classifier that produces an “OOD score” for each testing input sample
 - ✓ Make ID/OOD decision based on thresholding this OOD score
 - ✓ Threshold set based on tolerance for errors

5

Detection Results on Holdout OOD Dataset

TABLE I
DETECTION RESULTS FOR DIFFERENT MODEL TRAINING ALGORITHMS AND COMBINATIONS OF K AND J (NOTATION = AUROC / TNR@95TPR)

Training Method	$K = 1$			$K = 0.5$			$K = 0$		
	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$	$J = 3$
standard	92.5 / 75.2	91.4 / 72.2	92.3 / 74.4	88.9 / 60.9	87.7 / 56.8	89.3 / 59.8	72.8 / 26.5	71.7 / 22.9	72.4 / 23.4
lblsm	97.6 / 88.3	97.7 / 87.5	97.8 / 87.8	95.2 / 73.1	94.5 / 71.4	95.1 / 73.5	76.1 / 22.9	77.3 / 24.2	76.9 / 24.0
AT ($\epsilon = 2$)	94.6 / 81.3	92.3 / 75.6	91.6 / 74.0	90.2 / 68.8	89.2 / 65.0	90.7 / 66.4	74.7 / 30.4	72.5 / 22.8	73.5 / 24.8
AT ($\epsilon = 8$)	96.0 / 84.5	94.0 / 81.6	93.8 / 80.4	92.8 / 77.1	91.7 / 74.1	92.5 / 74.6	73.4 / 30.7	71.8 / 25.8	73.0 / 27.9
OE	97.0 / 86.2	96.2 / 82.3	96.3 / 83.2	93.5 / 70.7	93.1 / 68.0	93.1 / 71.1	80.6 / 31.6	79.1 / 32.7	80.5 / 34.3
AdvOE ($\epsilon = 2$)	98.6 / 93.5	98.0 / 90.5	98.2 / 91.4	96.4 / 83.0	95.8 / 80.7	96.1 / 82.0	82.1 / 39.6	81.0 / 36.6	81.6 / 39.2
AdvOE ($\epsilon = 8$)	99.2 / 96.9	99.0 / 95.6	99.0 / 95.8	97.5 / 89.1	97.8 / 89.4	97.7 / 90.1	82.9 / 41.6	81.3 / 39.9	83.7 / 45.2

6

liii

Summary on Robust OOD Detection based on Joint Data Learning Approach

Model training procedure matters very much

- Even if no OE dataset is available, Adversarial Training (AT) is better than standard training in all cases
 - General/Vanilla OE is not great, and in most cases marginally outperforms the plain AT models
 - AdvOE is significantly better than all other training methods, and when $\text{eps}=8$ it is the top performer across all K and J configs tested
-
- $K=0$ (i.e., 100% synthetic training data) is much harder than $K=\{0.5, 1\}$
 - AdvOE has near-perfect detection at $K=1$

N. Inkawhich, E. Davis, M. J. Inkawhich, U. Majumder and Y. Chen, "Training SAR-ATR Models for Reliable Operation in Open-World Environments," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, doi: 10.1109/JSTARS.2021.3068944.

N. Inkawhich, E. Davis, M. J. Inkawhich, U. Majumder and Y. Chen, E. Tripp, "Bridging a Gap in SAR-ATR: Training on Fully Synthetic and Testing on Measured Data," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 2942-2955, 2021,

Joint Data Learning Approach for Robust Out of Distribution Radio Frequency Target Recognition

Uttam K Majumder and Nathan Inkawhich

Introduction

It is often the case that a single data source is used for training deep neural networks (DNN) to classify targets. In this case, DNN model is very much capable of classifying known targets (in-library). However, unknown targets, in particular, out of library (OOL) targets similar to the in-library targets can be falsely classified as in-library targets if the DNN has not been trained with additional data sources beyond the in-library data. Most synthetic aperture radar automatic target recognition (SAR-ATR) systems are designed to achieve maximum accuracy for a limited set of in-library classes, yet ignore the implications of encountering novel data classes during deployment. Even worse, the standard deep learning training objectives fundamentally inherit a closed world assumption, and provide no guidance for how to handle out-of-library (OOL) data. In this work, we develop a novel training procedure called Adversarial Outlier Exposure (AdvOE) to co-design the ATR system for accuracy and OOL detection. Our method introduces a large, diverse and unlabeled auxiliary training dataset containing samples from the OOL set. The AdvOE objective encourages a Deep Neural Network to learn robust features of the in-library training data, while also promoting maximum entropy predictions for adversarial perturbed versions of the OOL data. We experiment with the recent SAMPLE (Synthetic and Measured Paired Labeled Experiment (SAMPLE)) dataset, and find our method nearly doubles OOL detection performance over the baseline in some settings, and excels when using only synthetic training data are used. As compared to several other advanced ATR training techniques, AdvOE also affords significant improvements in both classification and detection statistics. For joint data learning approach, we used SAMPLE dataset for in-library targets classes and SAR ship detection dataset for out-of-library targets.

In-Library / In-Distribution Data

SAMPLE dataset contains (measured, synthetic) pairs from the ten Moving and Stationary Target Acquisition and Recognition (MSTAR) target classes. In total, there are 806 training and 539 test pairs. The measured components were taken directly from the MSTAR public release dataset while the corresponding synthetic targets were constructed via electromagnetic signature prediction from meticulous Computer Aided Design (CAD) models. There are two key parameters associated with this dataset and its experiments that are referenced throughout this work. K sets the fraction of measured training data, while the test data is always 100% measured ($0 \leq K \leq 1$). For example, if $K = 0.75$ then 604/806 training pairs are represented by their measured component while the remaining 202/806 are represented by their synthetic component. As a special case, if $K = 0$ then 100% of the training samples are synthetic. To manufacture the OOL problem, parameter J is introduced to set the number of classes that are held-out from the

training set ($1 \leq J \leq 8$). So, the classifiers are trained on the remaining $10 - J$ classes, and at evaluation time the test data for the held-out J classes is considered OOL while the test data for the other $10 - J$ classes is considered in-distribution (In-library).

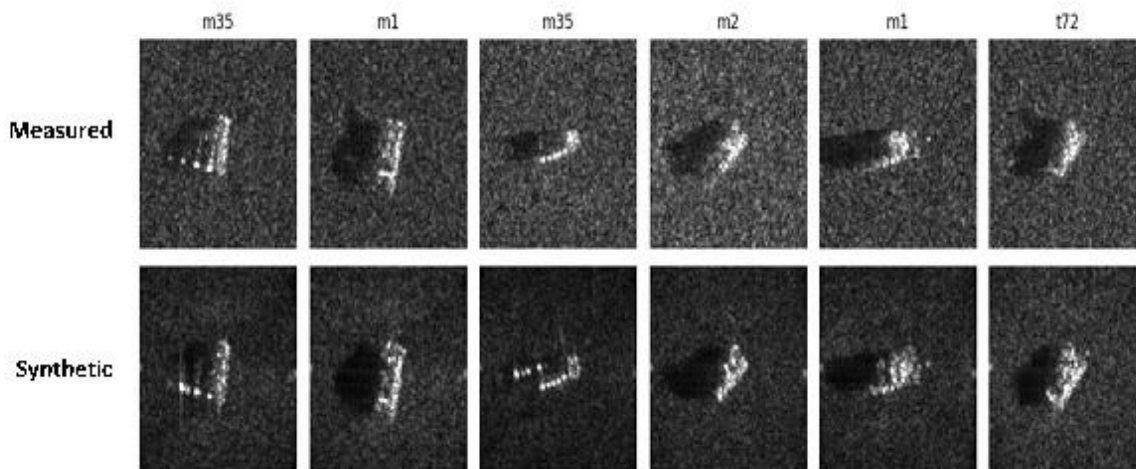


Figure 1: SAMPLE Data (Targets)

TABLE I
DETAILS OF THE SAMPLE DATASET

Class #	Class Name	# Train	# Test	Total
0	2S1	116	58	174
1	BMP2	55	52	107
2	BTR70	43	49	92
3	M1	78	51	129
4	M2	75	53	128
5	M35	76	53	129
6	M548	75	53	128
7	M60	116	60	176
8	T72	56	52	108
9	ZSU23	116	58	174
Totals		806	539	1345

Out of Library (OOL) Ship Detection Dataset

The SAR-Ship-Dataset was originally proposed as a large object detection dataset that covers a variety of sensors, imaging modes, resolutions, and polarizations (with no relationship to the SAMPLE dataset). In total, there are 59,535 ships from 43,819 full frame images. For our purposes, each ship was chipped from the full images using the ground truth box information.

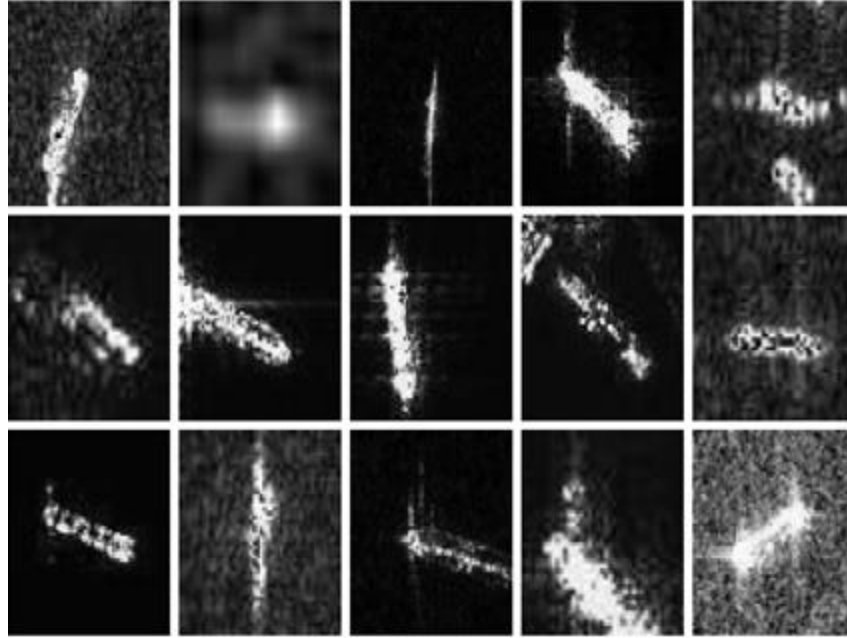


Figure 2: Ship Detection dataset

Table 2 Detailed information for original synthetic aperture radar (SAR) imagery.

Sensor	Imaging Mode	Resolution Rg. $\times$ Az. (m)	Swath (km)	Incident Angle ($^{\circ}$)	Polarization	Number of Images
GF-3	UFS	3×3	30	20~50	Single	12
GF-3	FSI	5×5	50	19~50	Dual	10
GF-3	QPSI	8×8	30	20~41	Full	5
GF-3	FSII	10×10	100	19~50	Dual	15
GF-3	QPSII	25×25	40	20~38	Full	5
Sentinel-1	SM	1.7×4.3 to 3.6×4.9	80	20~45	Dual	49
Sentinel-1 ¹	IW	20×22	250	29~46	Dual	10

¹ Only single cross-polarization SAR images are used to construct the dataset. Ultrafine Strip-Map (UFS), Fine Strip-Map 1 (FSI), Full Polarization 1 (QPSI), Full Polarization 2 (QPSII), S3 Strip-Map (SM), and Fine Strip-Map 2 (FSII) imaging modes, respectively.

Joint Learning Algorithmic Approach and Results

Following steps summarizes overall algorithm implementation steps:

- **Step 1:** Train classifier to be accurate on ID /In-library training dataset
 - ✓ Training is performed “offline” with no knowledge of the expected OOL samples
- **Step 2:** Instantiate an OOL detector for the trained classifier that produces an “OOl score” for each testing input sample
 - ✓ Make in-library /OOl decision based on thresholding this OOL score
 - ✓ Threshold set based on tolerance for errors

Training Method	$K = 1$			$K = 0.5$			$K = 0$		
	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$	$J = 3$	$J = 1$	$J = 2$	$J = 3$
standard	92.5 / 75.2	91.4 / 72.2	92.3 / 74.4	88.9 / 60.9	87.7 / 56.8	89.3 / 59.8	72.8 / 26.5	71.7 / 22.9	72.4 / 23.4
lblsm	97.6 / 88.3	97.7 / 87.5	97.8 / 87.8	95.2 / 73.1	94.5 / 71.4	95.1 / 73.5	76.1 / 22.9	77.3 / 24.2	76.9 / 24.0
AT ($\epsilon = 2$)	94.6 / 81.3	92.3 / 75.6	91.6 / 74.0	90.2 / 68.8	89.2 / 65.0	90.7 / 66.4	74.7 / 30.4	72.5 / 22.8	73.5 / 24.8
AT ($\epsilon = 8$)	96.0 / 84.5	94.0 / 81.6	93.8 / 80.4	92.8 / 77.1	91.7 / 74.1	92.5 / 74.6	73.4 / 30.7	71.8 / 25.8	73.0 / 27.9
OE	97.0 / 86.2	96.2 / 82.3	96.3 / 83.2	93.5 / 70.7	93.1 / 68.0	93.1 / 71.1	80.6 / 31.6	79.1 / 32.7	80.5 / 34.3
AdvOE ($\epsilon = 2$)	98.6 / 93.5	98.0 / 90.5	98.2 / 91.4	96.4 / 83.0	95.8 / 80.7	96.1 / 82.0	82.1 / 39.6	81.0 / 36.6	81.6 / 39.2
AdvOE ($\epsilon = 8$)	99.2 / 96.9	99.0 / 95.6	99.0 / 95.8	97.5 / 89.1	97.8 / 89.4	97.7 / 90.1	82.9 / 41.6	81.3 / 39.9	83.7 / 45.2

Figure 3: Summary of detection and classification results with joint data learning (JDL) approach. When $K=0$, (i.e., 100% synthetic data were used, JDL approach achieve better accuracy than the standard non-JDL approach.

Conclusion

In an effort to co-design SAR-ATR models for both improved accuracy on in-distribution targets and out-of-library target detection, we develop the Adversarial Outlier Exposure (AdvOE) training using two different datasets. The AdvOE method simultaneously learns the robust features of the in-distribution data for high classification performance, while also learning to exclude the unknown classes different from the in-library.

References

- [1] N. Inkawhich, E. Davis, U. Majumder, C. Capraro, and Y. Chen, “Advanced techniques for robust sar atr: Mitigating noise and phase errors,” in IEEE International Radar Conference (RADAR), 2020.
- [2] S. Chen, H. Wang, F. Xu, and Y. Jin, “Target classification using the deep convolutional networks for sar images,” IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4806–4817, 2016.
- [3] N. Inkawhich, M. Inkawhich, E. Davis, U. Majumder, C. Capraro, and Y. Chen, “Bridging a gap in sar-atr: Training on fully synthetic and testing on measured data,” Preprint Under Review, 2020.
- [4] E. Zelnio and A. Pavy, “Open set SAR target classification,” in SPIE Algorithms for Synthetic Aperture Radar Imagery XXVI, 2019.
- [5] B. Lewis, T. Scarnati, E. Sudkamp, J. Nehrbass, S. Rosencrantz, and E. Zelnio, “A SAR dataset for ATR development: the Synthetic and Measured Paired Labeled Experiment (SAMPLE),” in SPIE Algorithms for Synthetic Aperture Radar Imagery XXVI, 2019.
- [6] U. K. Majumder, E. P. Blasch, and D. A. Garren, Deep Learning for Radar and Communications Automatic Target Recognition. Artech House, 2020.

- [7] Y. Wang, C. Wang, H. Zhang, Y. Dong, and S. Wei, “A sar dataset of ship detection for deep learning under complex backgrounds,” *Remote Sensing*, vol. 11, no. 7, p. 765, Mar 2019
- [8] D. Hendrycks, M. Mazeika, and T. G. Dietterich, “Deep anomaly detection with outlier exposure,” in *International Conference on Learning Representations (ICLR)*, 2019.
- [9] T. Ross, S. Worrell, V. Velten, J. Mossing, and M. Bryant, “Standard sar atr evaluation experiments using the mstar public release data set,” in *SPIE Conference Algorithms for Synthetic Aperture Radar Imagery V*, 1998. JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015

EXPLORING SYNERGIES BETWEEN ARTIFICIAL INTELLIGENCE, SYSTEMS ENGINEERING, AND INFORMATION FUSION

Dr. Ali K. Raz
Assistant Professor Systems Engineering
Assistant Director of Intelligent Systems and Integration at C4I and Cyber Center
George Mason University
araz@gmu.edu



1

RESEARCH MOTIVATION AND BACKGROUND

Future operational concepts of in both civil and defense domains increasingly demand *integration and interoperability* of multiple *intelligent systems*



<https://utm.arc.nasa.gov/index.shtml>



<https://www.darpa.mil/program/collaborative-operations-in-denied-environment>

Intelligent Systems require:
*Artificial Intelligence and
Machine Learning*

Integration requires:
*Systems Engineering and
Systems of Systems Engineering*

Interoperability requires:
*Processing and Fusion of
Heterogeneous Information*

GEORGE MASON UNIVERSITY

2

lx

SYSTEMS ENGINEERING VIEW OF AI

- INCOSE's SE Definition:

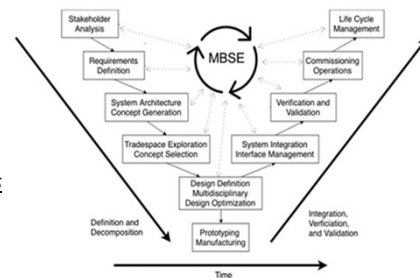
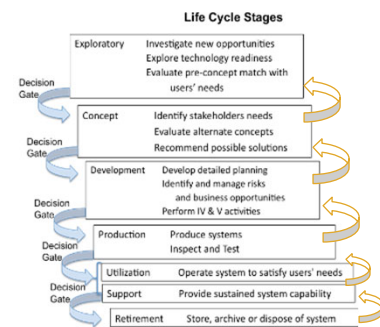
"a transdisciplinary and integrative approach to enable the successful realization, use, and retirement of engineered systems, using systems principles and concepts, and scientific, technological, and management methods." (INCOSE SE State of Discipline)

- SE4AI:

- Utilize systems principles and concepts for engineering AI solutions
- Examine, develop, and build AI with the SE life-cycle stages

- AI4SE:

- Make AI an integral part of scientific and technological methods for engineered systems
- Utilize AI during the life-stages of system development



GEORGE MASON UNIVERSITY

3

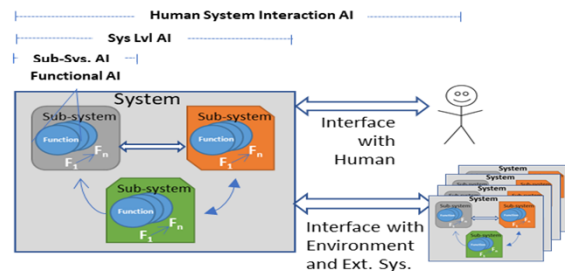
WHERE IS AI/ML IN SYSTEMS?

- SE4AI:

- AI and ML can manifest throughout a system's hierarchical decomposition
 - Functional Level
 - Sub-System Level
 - System-Level
 - Human and external System Interaction
- Different levels within a system could have different AI/ML types
 - SEs need to integrate across the various types
- Imperative for T&E frameworks for AI systems

- AI4SE:

- How can AI/ML techniques help design and engineer better systems
 - Identify interactions and characterize emergent behavior in systems

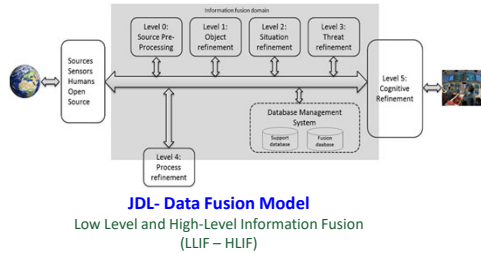


GEORGE MASON UNIVERSITY

4

lxi

AI4SE CASE STUDY: CHARACTERIZING INFORMATION FUSION (IF) SYSTEM DESIGN WITH SE AND AI



Status-Quo: Independent Design and Evaluation of Fusion System Elements

- ✓ Sensors Design and Measurement Quality
- ✓ State-Estimation and Track Fusion Algorithms
- ✓ Association and Correlation Approaches
- ✓ Gap between LLIF and HLIF

- Information Fusion (IF) Systems are complex systems where the capability emerges from interaction of multiple elements
- Various Design Considerations of IF Systems
 - Information quality, transmission, communication, association, correlation, fusion architectures etc.
 - Low-Level and High-Level Information Fusion processes and algorithms

Complex Systems Design and Evaluation Requirements

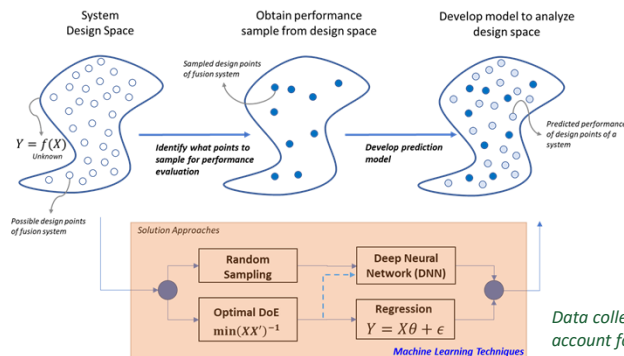
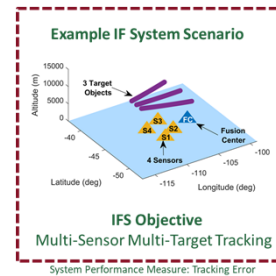
- ❑ How to characterize the collective system level behavior?
- ❑ How does variation in different elements impact the end-system performance?
- ❑ How to select the most promising combination of element design/performance?

GEORGE MASON UNIVERSITY

5

CASE STUDY EXAMPLE AND APPROACH (SUMMARIZED)

- Design Considerations for Simple IF System Scenario
 - System Architecture: Centralized, Distributed, Hierarchical
 - Measurement quality and frequency of different sensors
 - Measurement and track fusion algorithm
 - Placement of sensors, pairing of sensors in distributed architectures
 -



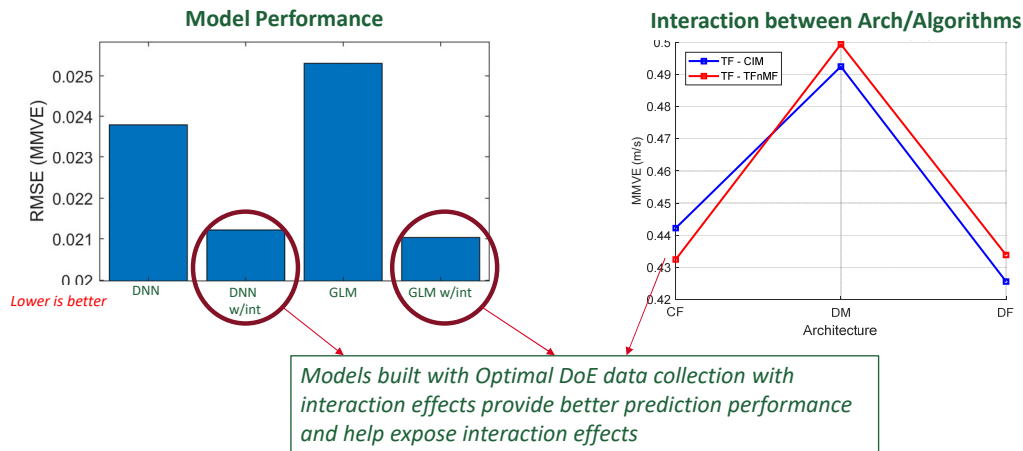
Data collection with optimal DoE helps account for interactions in the model

GEORGE MASON UNIVERSITY

6

lxii

CASE STUDY RESULTS



- **AI Implications:** Deploying AI/ML techniques for IF system evaluation can help better understand system performance and investigate competing design options
- **SE Implications:** It is not just data, but how data is collected can have impact on AI/ML performance

GEORGE MASON UNIVERSITY

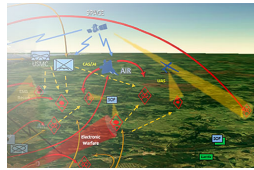
7

AI, SE, AND IF ROLES FOR COLLABORATIVE AUTONOMOUS SYSTEMS

- **Future Vision: Collaborative Autonomous Systems**
 - AI4SE and SE4AI both need to advance together for transitioning AI into operational systems



Urban Air Mobility



Multi-Domain C2



Human Space Exploration



Intelligent Transportation

- **Exploring AI, SE, IF synergy**
 - AI/ML helps improve SE and IF processes
 - SE helps with better design of AI and IF systems
 - IF helps enable interoperability in distributed systems

GEORGE MASON UNIVERSITY

8

lxiii



Exploring Synergies between Artificial Intelligence, Systems Engineering, and Information Fusion

Ali K. Raz

Systems Engineering and Operations Research, George Mason University
4400 University Dr., MS 4A6, Fairfax, VA 22030-4444 USA

INVITED PANEL DISCUSSION POSITION PAPER

Future operational concepts of complex system missions increasingly demand integration and interoperability of multiple systems some of which are expected to be intelligent systems. For example, consider the Unmanned Aircraft System (UAS) integration in the National Airspace [1] or the multi-domain command and control system that integrates the air, space, and cyber domains for enhanced mission capabilities. A common theme among both these examples is the integration and interoperability of multiple conventional and intelligent systems. Inculcating intelligence in system requires incorporating artificial intelligence (AI) and machine learning (ML) techniques as part of a system functionality. Integration of multiple heterogeneous systems requires utilizing tools and techniques of system engineering discipline, while engineering effective interoperability will demand an ability for processing and fusion of heterogeneous data and information. In this panel presentation, I will highlight the importance of exploring synergies between systems engineering, AI/ML, and information fusion which are often viewed as disparate and independent disciplines but yet will need to work together to enable operational concepts where independent intelligent systems can integrate and interoperate.

Synergy Between AI/ML and Systems Engineering:

The International Council on Systems Engineering (INCOSE) defines systems engineering as “a transdisciplinary and integrative approach to enable the successful realization, use, and retirement of engineered systems, using systems principles and concepts, and scientific, technological, and management methods” [2]. Evolving capabilities of systems and the resulting design and engineering complexities demand a corresponding evolution of systems principles and technological methods that need to be incorporated into the transdisciplinary nature of systems engineering. In this regard, the technological advances in AI/ML have become particularly important for the future of systems engineering as they provide a promising application for both the engineered system and the evolution of systems principles and concepts. One the one hand, operational concepts of future engineered systems, such as self-driving cars, demand an increasing level of autonomy which can only be successfully realized by the systems engineering of AI. On the other hand, the growing complexity of these future systems paired with a plethora of design considerations, intertwined operational requirements, and emergent behavior, demands the use of AI/ML for systems engineering of these future systems. Hence, the quest for exploring synergies between systems engineering and artificial intelligence becomes a dichotomous exercise of two parallel paths: one investigating systems engineering for artificial intelligence (AI4SE) and the other studying artificial intelligence for systems engineering (SE4AI). In this presentation, I will cover both of these paths to examine complementary opportunities the two disciplines can offer each other.

Synergy Between Information Fusion and Systems Engineering:

Developing real-time situational awareness of unfolding scenarios requires rapid processing of heterogeneous data from geographically distributed sensing, information processing, information fusing, and decision-making systems. These systems, such as radars, fusers, command and control centers, including human decision-makers and autonomous systems, all contribute varying levels of fundamental capabilities to situational awareness development. The overall situational awareness mission, nevertheless, is only realized from a synergistic orchestration of these individual yet interdependent capabilities. In this regard, these individual systems—although complex and esoteric in nature—become parts of a larger engineered system. Referring to the operational concepts of integrating UAS in National Airspace or the multi-domain C2 discussed earlier, developing situational awareness—which is a key information fusion problem [3]—is required for interoperability of systems, while engineering a system with multiple information producing and fusing elements becomes a system engineering problem [2].

Tracing back to the historical origins of data fusion and situation awareness to the JDL Data Fusion Process model, the different parts and elements of information fusion system are traditionally individually designed and developed with limited to no attention paid to the various interdependencies that inherently arise in the synergistic orchestration of the situational awareness mission [4]. For example, the interactions and interdependence between different JDL levels, most noticeably between low-level and high-level JDL levels, is rarely addressed by methodical design processes in the information fusion literature. Development of such methodical design processes that elucidate complex interactions in distributed heterogeneous systems is an active and dedicated area of research in systems engineering [5]. In this presentation, I will examine the need for system engineering of information fusion system.

Synergy Between AI/ML, Systems Engineering, and Information Fusion:

In this presentation, I will cover a case study for design of information fusion system using system engineering techniques aided by the application of AI/ML. Information fusion system (IFS) is a complex system consisting of various interdependent elements that are typically designed, evaluated, and integrated independently [6]. However, siloed performance evaluation does not equate to a system-level performance in complex systems. Since the IFS capability results from integrated behavior of these elements, characterizing interactions becomes critical for engineering an IFS. In the case study, which is based on [7], I will present machine learning techniques (deep neural networks and general linear models) to build predictive performance model of the IFS using Optimal Design of Experiments. Application of ML methods complements the systems engineering approach and leads to a comprehensive understanding of IFS design and performance.

REFERENCES

- [1] “Unmanned Aircraft System Traffic Management (UTM).” https://www.faa.gov/uas/research_development/traffic_management/ (accessed Jun. 20, 2020).
- [2] INCOSE, *INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities*, 4 edition. Hoboken, New Jersey: Wiley, 2015.
- [3] James Llinas and David L. Hall, “Multisensor Data Fusion,” in *Handbook of Multisensor Data Fusion*, 0 vols., CRC Press, 2008, pp. 1–14. Accessed: Sep. 01, 2016. [Online]. Available: <http://dx.doi.org/10.1201/9781420053098.ch1>
- [4] E. Blasch, É. Bosse, and D. Lambert, *High-Level Information Fusion Management and System Design*, 1st ed. Artech House, 2012.
- [5] A. K. Raz, C. R. Kenley, and D. A. DeLaurentis, “A System-of-Systems perspective for information fusion system design and evaluation,” *Inf. Fusion*, vol. 35, pp. 148–165, May 2017, doi: 10.1016/j.inffus.2016.10.002.
- [6] A. K. Raz, J. Llinas, R. Mittu, and W. Lawless, “Engineering for Emergence in Information Fusion Systems: A Review of Some Challenges,” in *2019 22th International Conference on Information Fusion (FUSION)*, Jul. 2019, pp. 1–8.
- [7] A. K. Raz, P. Wood, L. Mockus, D. A. DeLaurentis, and J. Llinas, “Identifying Interactions for Information Fusion System Design Using Machine Learning Techniques,” in *2018 21st International Conference on Information Fusion (FUSION)*, Jul. 2018, pp. 226–233. doi: 10.23919/ICIF.2018.8455429.