# EUVIP2021 Tutorial Proposal

**Title:** Explainable and Robust Machine Learning for Images




**By:** Ghassan AlRegib, Mohit Prabhushankar, Gukyeong Kwon, and Jinsol Lee

Georgia Institute of Technology

**Overview, objectives, and relevance of the tutorial**

In recent years, artificial intelligence systems achieved state-of-the-art performances in image classification tasks. Specifically, classification algorithms surpassed top-5 human error rate of 5.1% on ImageNet. Even though these advancements are promising, images in these datasets do not cover diverse real-world scenarios. Distortions may include perceptually unpleasant camera related issues like blur, motion blur, overexposure, underexposure, and noise. Moreover, environmental conditions such as rain, snow, and frost can affect the field of view. These non-ideal conditions impact the performance of artificial intelligence (AI) algorithms. Furthermore, gaining an insight into understanding the decision made by an AI algorithm under such scenarios is crucial in building robust AI models. Over the past a few years, there has been quite progress in AI Explainability. As an example, Grad-CAM has been widely used to visually justify the decision made by a classification network by answering 'Why P?' where P is its prediction. We add context and relevance to this question by answering `Why P, rather than Q?' where Q is some contrast prediction. In some cases, such context can be more descriptive for interpretability. For instance, in autonomous driving applications that recognize traffic signs, knowing why a particular traffic sign was chosen over another is informative in contexts of analyzing decisions in case of accidents. This has a broader impact in applications such as medical and subsurface image analysis. Other modalities of the considered question include `Why P, rather than P?' and `Why P, rather than all other classes?'. We show that explicitly tying the inference to these questions makes the neural networks more robust while also providing justifications that are contextually relevant. Specifically, we analyze robustness in the domain of recognition, out-of-distribution detection, novelty detection, open-set recognition, and Image Quality Assessment.

The tutorial will have three major parts:

- Part 1: Explainability and Reasoning
- Part 2: Recognition under domain shift, Out-of-distribution and Novelty Detection
- Part 3: Image Quality Assessment and Saliency Detection

Outline:

- Overview of algorithms robustness for image processing and analysis application
- Overview of Explainability in AI and ML
- A brief literature review of the recent works in both robustness and explainability in AI and ML for image processing and analysis applications
- Activations-based versus gradients-based methodologies
    - Intuition behind gradients-based techniques
- Utilizing gradients to explain networks behavior and empower them for better generalizability
    - The effectiveness of contrastive reasoning to provide better representation space than Inductive reasoning
    - Understanding the purview of neural networks
- Applications to be discussed:
    - Recognition
    - Domain Adaptation
    - Open-set recognition
    - Out-of-distribution Detection
    - Novelty Detection
    - Adversarial Image Detection
    - Image Quality Assessment (IQA)
    - Saliency Detection
- Overview of related papers and their available codes and datasets
- Conclusions and take-away messages

---

**Biographical Sketches:**

*Ghassan AlRegib* (alregib@gatech.edu) is currently Professor in the School of Electrical and Computer Engineering at the Georgia Institute of Technology. He was a recipient of the ECE Outstanding Graduate Teaching Award in 2001 and both the CSIP Research and the CSIP Service Awards in 2003, the ECE Outstanding Junior Faculty Member Award, in 2008, and the 2017 Denning Faculty Award for Global Engagement. His research group, the Omni Lab for Intelligent Visual Engineering and Science (OLIVES) works on research projects related to machine learning, image and video processing, image and video understanding, seismic interpretation, machine learning for ophthalmology, and video analytics. He has participated in several service activities within the IEEE. He served as the TP co-Chair for ICIP 2020.

*Mohit Prabhushankar* (mohit.p@gatech.edu) received the B.S Degree in Electronics and Communications Engineering from PES Institute of Technology, Bangalore, India in 2014 and the M.S Degree in Electrical Engineering from Georgia Institute of Technology, Atlanta, GA, USA, in 2015. Since then, he has been a Ph.D student in the Omni Lab for Intelligent Visual Engineering and Science (OLIVES) lab, headed by Ghassan AlRegib, working in the fields of image processing and machine learning. He is the recipient of the Best Paper award at ICIP 2019 and Top Viewed Special Session Paper Award at ICIP 2020. He has served as a Teaching Fellow at Georgia Tech since 2020.



*Gukyeong Kwon* (gukyeong.kwon@gatech.edu) is a Ph.D. candidate in the School of Electrical and Computer Engineering (ECE) at the Georgia Institute of Technology where he received his M.S. degree in 2018. He is a co-recipient of the Finalist of the World's First 10K Best Paper Award at the IEEE International Conference on Multimedia and Expo in 2017, the Best Paper Award at the IEEE International Conference on Image Processing in 2019, and the Top Viewed Special Session Paper Award at the IEEE International Conference on Image Processing in 2020. His research has primarily focused on the robustness of machine learning, multimodal representation learning, and learning with limited data.

*Jinsol Lee* (jinsol.lee@gatech.edu) received the B.S. degree in Electrical Engineering at Purdue University in 2016 and the M.S. degree in Electrical and Computer Engineering at Georgia Institute of Technology in 2017, where she is pursuing her Ph.D. degree. Her research interests include machine learning, computer vision, and image processing. Her research mainly focuses on the robustness of machine learning models.

**Webpage : www.ghassanalregib.info and https://github.com/olivesgatech**