# Ph.D. in Information Technology
# Thesis Defenses

**April 16th, 2025**
**At 11:00 a.m.**
**Sala Conferenze Emilio Gatti – Building 20**

**Antonio GIGANTI** – XXXVII Cycle

**DEVELOPMENT OF A FRAMEWORK TO ENHANCE BVOC IMAGING**
Supervisor: Prof. Paolo Bestagini

**Abstract:**

Air pollution remains a major global challenge, particularly in urban areas where high pollutant concentrations negatively impact public health and contribute to climate change. Among the various pollutants, biogenic volatile organic compounds (BVOCs) play a critical role in atmospheric chemistry, influencing the formation of secondary organic aerosols and ground-level ozone, affecting air quality and climate dynamics. Accurately estimating BVOC emissions at high spatial resolution is challenging due to the limitations of satellite observations and computational models. Additionally, forecasting nitrogen dioxide (NO2) concentrations in urban environments is vital for effective air quality management, yet existing models often struggle to capture complex spatiotemporal dependencies.

The thesis aims to address these challenges by proposing novel deep learning (DL) frameworks to tackle two key tasks: (i) improving the spatial resolution of BVOC emission maps through super-resolution (SR) techniques and (ii) developing a robust model for forecasting NO2 concentrations in urban environments using graph neural networks (GNNs).

The first part focuses on SR, a task to generate high-resolution (HR) outputs from low-resolution (LR) inputs. In the context of BVOC emissions, we leverage state-of-the-art neural networks to enhance the spatial detail of numerically simulated and satellite-derived emission maps. Both single-image and multi-image SR (SISR and MISR) tasks are addressed. The MISR model further exploits the interconnections between different BVOC species, allowing for better spatial accuracy by learning shared patterns across multiple emission maps. Additionally, a data transformation strategy is proposed to preprocess the input data, improving the robustness of the SR model. To address the issue of data scarcity, domain adaptation (DA) techniques based on generative adversarial networks (GANs) are employed to bridge the gap between simulated and real-world data. We propose to adapt satellite-derived and numerically simulated emissions by leveraging an unpaired image-to-image translation framework based on CycleGAN. The essential advantage of this approach is that it enables the SR model, initially trained on simulated data, to be applied to satellite observations, which are often of lower resolution and exhibit domain shifts in spatial and dynamic patterns. In this way, the model learns to transfer knowledge from simulations to satellite data, allowing for the generation of HR emission maps while maintaining robustness against domain discrepancies.

The second part of the thesis addresses the problem of forecasting NO2 concentrations in urban areas. Using a GNN-based approach, we automatically learn the spatial relationships between air quality monitoring stations, capturing the spread of pollution across the city. The proposed model further

incorporates historical data and future covariates to improve the accuracy of NO2 predictions. The model was tested on a real-world air quality dataset, demonstrating its ability to outperform traditional forecasting methods in terms of predictive accuracy.

Through an analysis of the state-of-the-art methods, this thesis identifies critical limitations of current SR and forecasting approaches in atmospheric applications. Existing SR methods struggle with non-uniform data distributions and outliers in BVOC emission maps, while many forecasting models fail to capture the spatial dependencies necessary for accurate pollution prediction. This research overcomes these limitations by developing models specifically tailored to the challenges of atmospheric and environmental data.

The results show that the proposed SR framework effectively enhances the spatial resolution of BVOC emission maps, producing accurate HR estimates from coarse data. The domain adaptation strategies ensure the model generalizes well across different inventories, including simulated and satellite-derived emissions. Additionally, the GNN-based NO2 forecasting model demonstrates improved predictive power, providing more accurate air quality forecasts in urban environments.

In summary, this thesis contributes a comprehensive framework that advances the fields of BVOC emission mapping and pollutant forecasting. By leveraging DL and DA techniques, the proposed models enhance atmospheric data analysis and offer practical tools for better environmental monitoring and air quality management.

**Daniele Ugo LEONZIO**– XXXVII Cycle

**DATA DRIVEN TECHNIQUES FOR LEAK DETECTION IN WATER DISTRIBUTION NETWORK**

Supervisor: Prof. Marco Marcon

This thesis presents innovative data-driven techniques for leak detection in WaterDistribution Network (WDN), which are crucial infrastructure components responsible for providing potable water to communities around the world. As urban populations continue to grow and existing infrastructures age, addressing water losses due to leaks has become an increasingly pressing concern. It is estimated that leaks account for up to 30% of the total water supplied in metropolitan areas, resulting in significant financial losses, environmental degradation, and potential safety hazards. The World Bank estimates that up to 126 billion cubic meters of Non Revenue Water (NRW) are lost annually due to leaks, highlighting the critical need for effective leak detection methods. Consequently, the demand for timely and accurate leak detection solutions is paramount, not only to minimize resource wastage but also to reduce repair costs, ensure service quality, and prevent risks associated with water contamination. Traditional leak detection methods, while useful, often suffer from limitations such as false alarms and inaccurate localization, leading to prolonged repair times and increased operational costs. To overcome these challenges, modern electronic and information technologies have enabled the development of advanced sensing systems that leverage data-driven approaches.

This work focuses on harnessing Machine Learning (ML) and Deep Learning (DL) algorithms to enhance leak detection capabilities within WDN. The first section of this thesis addresses the challenge of detecting and localizing water leaks using time-series data. We propose a series of

approaches based on autoencoder models, trained on data collected under normal operating conditions without leaks. By establishing a baseline of typical pressure measurements, the models are able to identify anomalous patterns indicative of leaks. Additionally, these models incorporate Graph Signal Processing (GSP) techniques, enabling the reconstruction of missing or corrupted data caused by sensor failures. By modeling the WDN as a graph, we can exploit the spatial relationships among sensors to accurately interpolate and reconstruct lost signals, thereby improving the system's resilience against sensor malfunctions. This robust approach not only enhances detection accuracy but also ensures operational efficiency, enabling the leak detection system to function reliably even in the presence of missing data. The second focus of this thesis is on leak detection using vibroacoustic sensors. When a leak occurs, it generates unique acoustic signatures characterized by changes in noise levels and pressure wave patterns that propagate through the pipeline. This work investigates the effectiveness of various ML algorithms, including Random Forests and Gradient Boosting, in classifying these signals and distinguishing between leak-related and normal operational sounds. Furthermore, we explore the potential of advanced deep learning techniques, specifically complex-valued Convolutional Neural Network (CNN), which are adept at processing the intricate phase and amplitude information inherent in vibroacoustic signals. The integration of multiple sensor types, including hydrophones and accelerometers, is examined as a means of enhancing leak classification performance through sensor data fusion techniques. This multi-sensor approach allows for cross-validation of signals and mitigates the limitations associated with individual sensor types, ultimately improving detection accuracy and robustness.

Finally, this thesis delves into the challenge of domain adaptation, which is critical for ensuring the transferability of models trained on controlled, laboratory-scale data to real-world WDN characterized by complex and noisy conditions. We evaluate various domain adaptation techniques tailored for leak detection tasks, enabling our models to function effectively in operational environments. This aspect of our research assesses the viability of domain adaptation in improving model performance and generalization capabilities, ultimately aiming to provide comprehensive solutions for effective leak management within WDN. The findings presented in this work contribute significantly to the advancement of leak detection methodologies, demonstrating their potential to enhance the management and sustainability of water resources. By addressing the multifaceted challenges of leak detection, localization, and classification, this thesis offers valuable insights and practical solutions for mitigating the impact of water losses, thereby supporting the ongoing efforts to secure water resources for future generations.

**PhD Committee**

Prof. Stefano Tubaro, **Politecnico di Milano**

Prof. Enrico Magli, **Politecnico di Torino**

Prof. Simone Milani, **Università degli Studi di Padova**