University of Cagliari

Sapienza University of Rome

University of Genova

Pattern Recognition and Application Lab

National PhD in Artificial Intelligence

Trustworthy AI in Video Surveillance: The IMMAGINA Project

Emanuele Ledda, Lorenzo Putzu, Rita Delussu, Giorgio Fumera, Fabio Roli

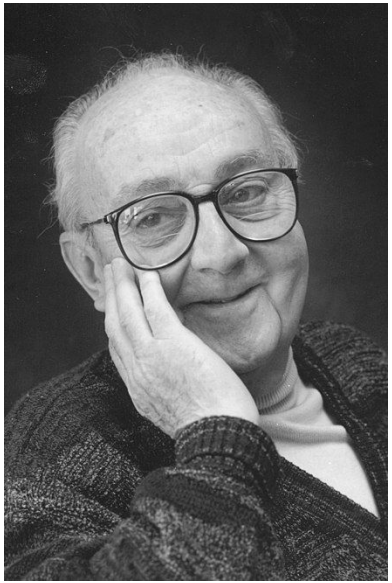
IMMAGINA
IMAGING MANAGEMENT GUIDELINES AND INFORMATICS NETWORK FOR LAW ENFORCEMENT AGENCIES

3rd National Conference on Artificial Intelligence - Pisa, 30/05/2023

Artificial Intelligence and Intelligent Systems
cini National Lab

“All models are wrong but some are useful„

George Box, 1976



Why Are All Models Wrong?

Because the world has many sources of Uncertainty

Intrinsic Randomness = Aleatoric Uncertainty

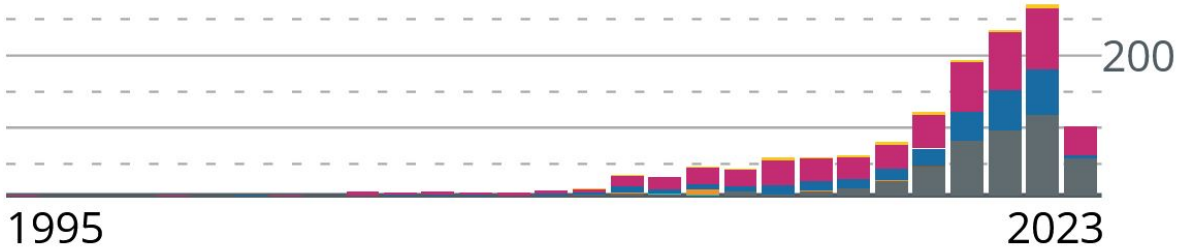
- Exact predictions may not always exist
- From the Latin “Aleator”, i.e. “Dice Player”

Lack of Knowledge = Epistemic Uncertainty

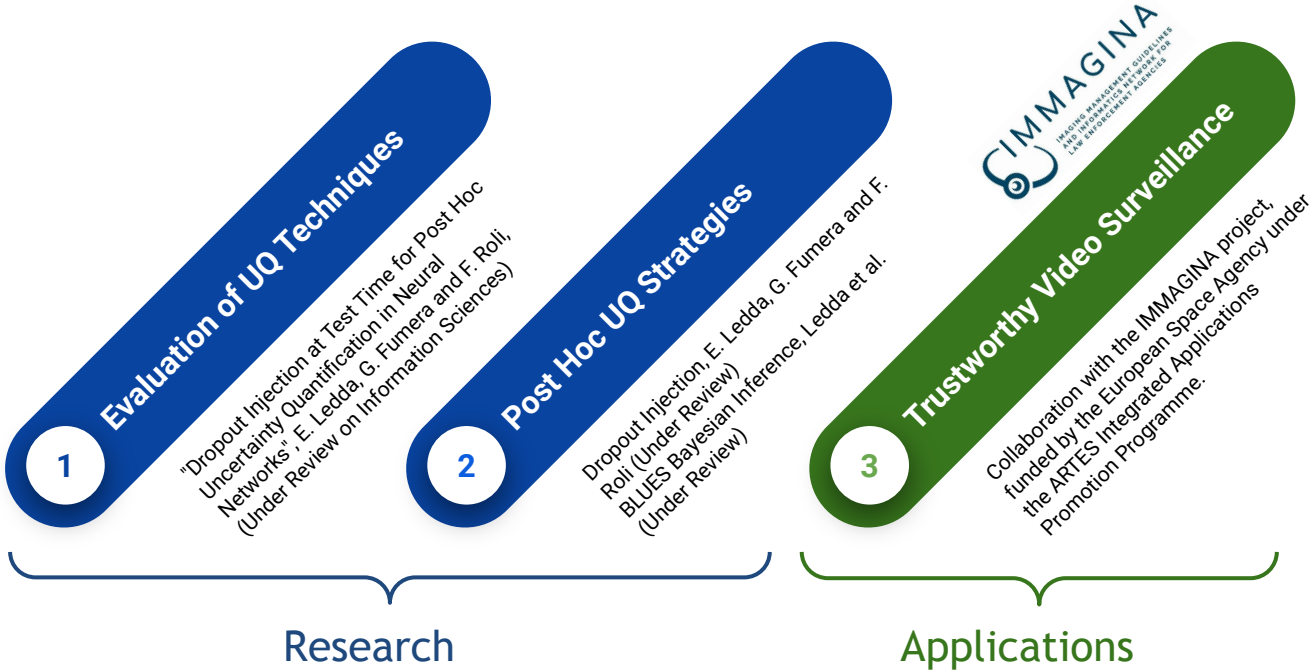
- We do not have perfect knowledge
- from the greek “Episteme”, i.e. “Knowledge”

Trustworthiness from the Uncertainty Perspective

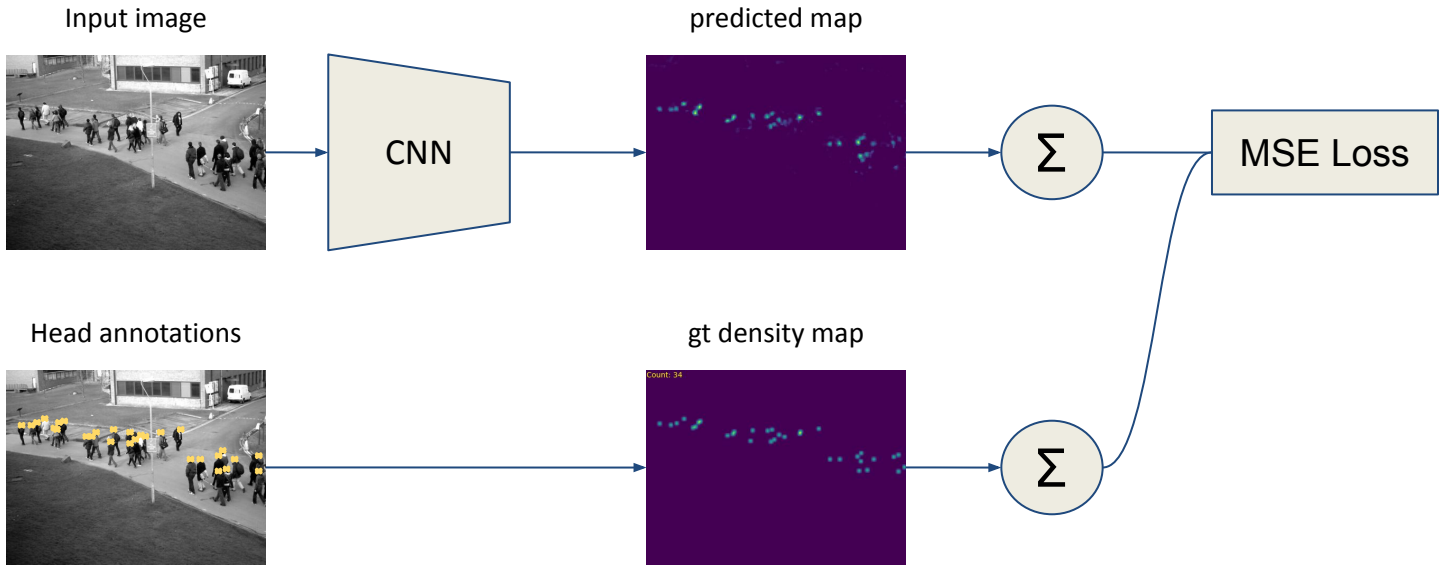
- Trustworthy AI as an AI which **“knows when it does not know”**
- Let people know when to Trust a model’s prediction: hence, when a model is wrong and when it is useful
- Challenging, still rising topic in AI



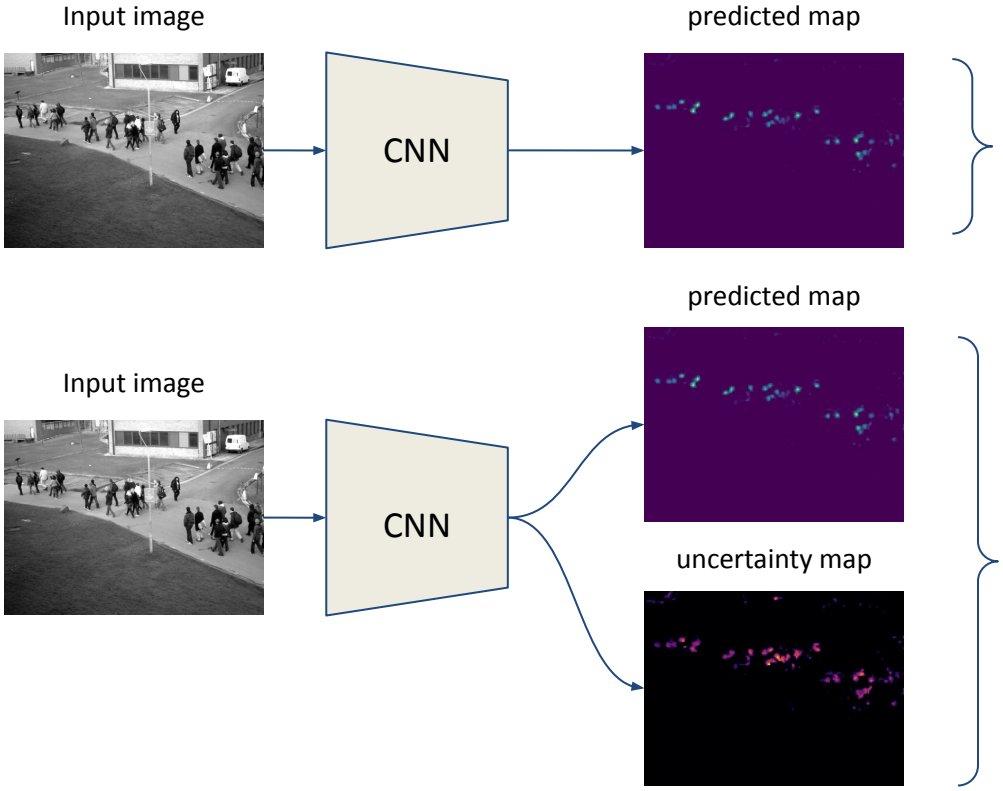
Our Line of Research and Application



Crowd Counting and Density Estimation



Uncertainty Quantification for Crowd Counting

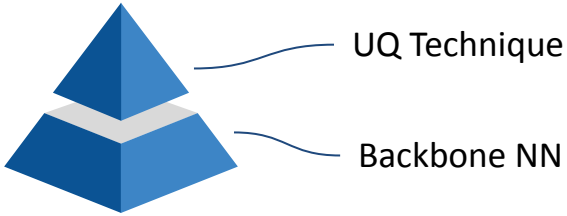


- **Deterministic** model and output
- **Point predictions** (e.g. count 31)

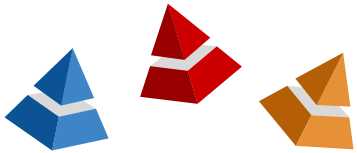
- **Probabilistic** model and output
- **Confidence interval** on the prediction (e.g. count 31 ± 3)
- Working as a **post hoc UQ** technique

Post Hoc vs Ad Hoc Uncertainty Quantification

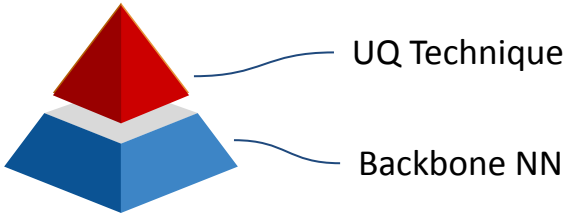
Ad Hoc Techniques



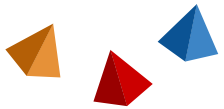
- Integrated inside the network at **training time**
- Changing UQ technique **requires re-train** the backbone



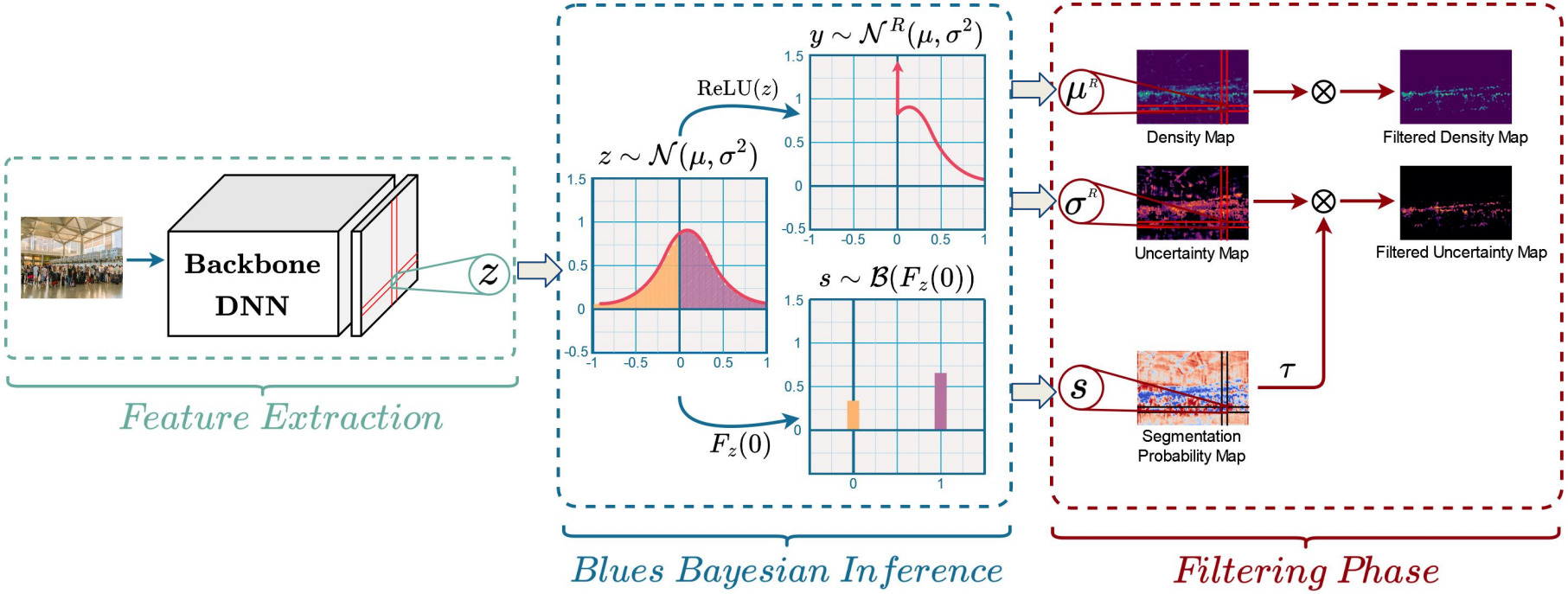
Post Hoc Techniques



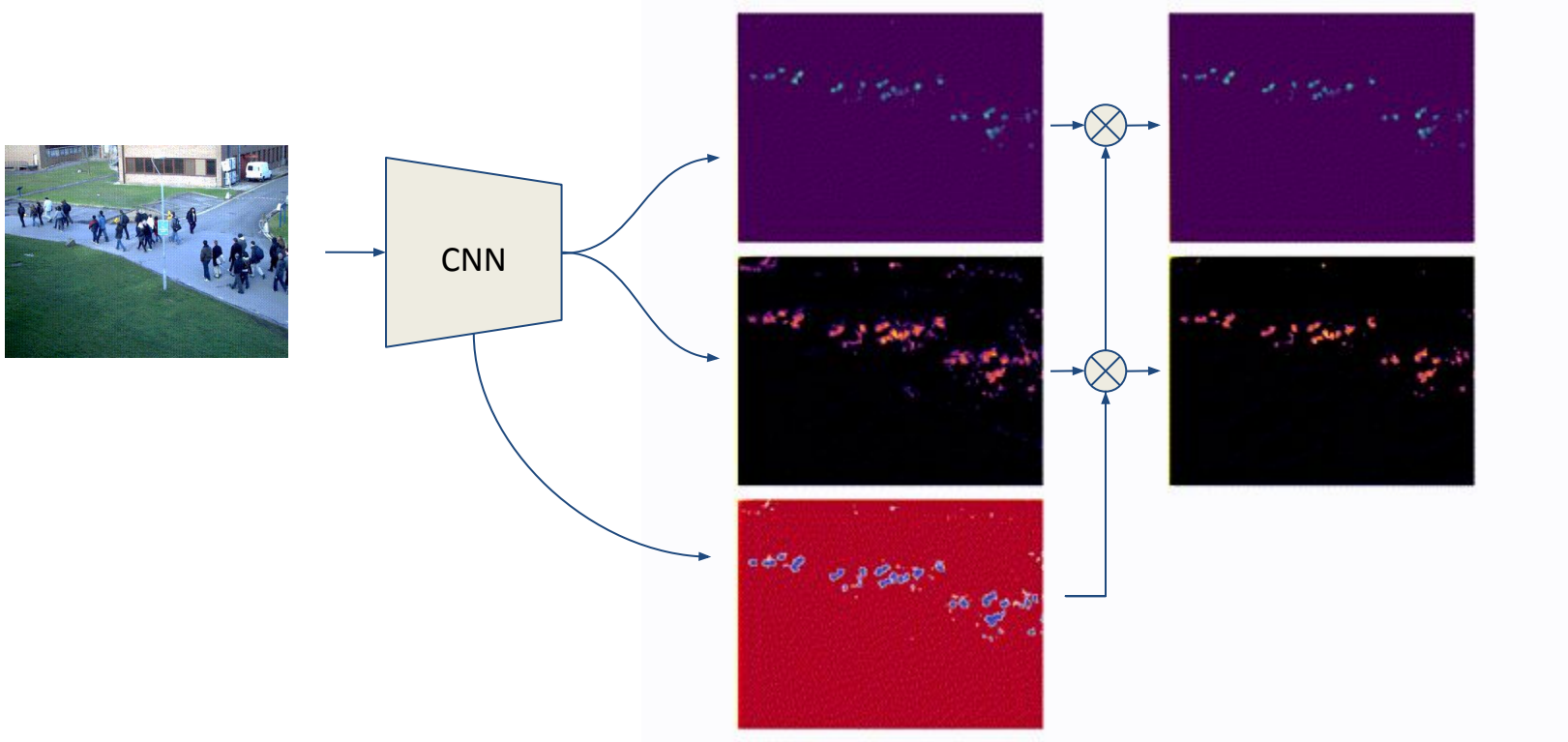
- Integrated on an **already trained** neural network
- Changing UQ technique **does not require to re-train** the backbone
- Act as a **plug-in**



Steps Ahead: Before-ReLU-Estimates Bayesian Inference



Applicative Results



University of Cagliari

Sapienza University of Rome

University of Genova

Pattern Recognition and Application Lab

National PhD in Artificial Intelligence

Thank you for the attention!

Emanuele Ledda

National PhD Student in Artificial Intelligence
Università degli Studi di Roma La Sapienza
emanuele.ledda@uniroma1.it