Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

GRAIL

Gamma-Ray Imaging with deep Learning

Alberto Garinei – Department of Engineering Science, Guglielmo Marconi University

Spoke 3 :

Fabio Gargano - INFN Bari

Spoke 3 Progetti Bandi a Cascata, 09/01, 2025

Project Overview

The "Gamma-RAy Imaging with deep Learning" (GRAIL) project aims to transform calorimetric data analysis by integrating advanced deep learning techniques implemented directly on edge within the satellite. This approach minimizes computational resource usage while addressing the critical challenge of transmitting large volumes of data from space to Earth.

The primary focus is on improving particle identification and event reconstruction for low-energy particles in space-based calorimeters, ensuring high efficiency and precision in resource-constrained environments.

Technical Objectives

1. **Data Structure**

   Identify the most efficient data structure for calorimetric representation to improve analysis precision and scalability.

2. **Event Reconstruction and Particle Identification**

   Enhance the reconstruction of particle trajectories and identification of particle types across a wide range of energies.

3. **Computational Efficiency**

   Optimize algorithms for **real-time on-edge predictions** directly on satellite hardware, ensuring efficient processing under strict resource constraints.

Methodologies and Solutions

1. **Data Structure Methods**:

- Transition from image-based data representations to **point cloud-based** data, which provide a more flexible and precise description of calorimetric particle showers.
- Develop **Graph Neural Networks** (GNN) to represent the complex structure of calorimetric showers and enhance energy estimation precision.
- Incorporate advanced architectures **autoencoders** to enhance feature extraction and representation, ensuring scalability and adaptability for high-dimensional data.

2. **Event Reconstruction and Particle Identification**:

- Use models like Convolutional Neural Networks (**CNN**), Recurrent Neural Networks (**RNN**) and/or **transformers** to analyze temporal and spatial sequences of calorimetric events, identifying energy, particle trajectories, and initial interactions.

Methodologies and Solutions

3. **Computational Efficiency** :

- Optimize algorithms for deployment **on-edge**, enabling predictions and data processing directly on satellites to minimize bandwidth usage and improve real-time decision-making capabilities.
- Evaluate and test **quantum hybrid algorithms**, which combine quantum and classical computing paradigms to enhance the efficiency of high-dimensional data processing tasks, such as shower reconstruction and feature extraction.
- Tailor the **implementation to space-grade hardware** constraints, such as limited power and processing capabilities, ensuring sustainability and reliability in extreme environments.

Technical Objectives, Methodologies and Solutions

AI techniques in calorimetric data analysis have been successfully applied in various missions and experiments, such as:

- Machine learning algorithms to classify particle interactions and enhance event reconstruction.
- Neural networks to improve accuracy in identifying and characterizing particle showers.

↓

Most of the existing research has focused on **preliminary evaluations** of data efficiency or model precision. However, these studies have been conducted with access to substantial computational resources available in **ground-based environments**, whereas our approach leverages the **limited resources available** directly on the satellite, emphasizing the need for **highly optimized and efficient** algorithms.

Why on edge?

Challenges of Data Transmission to Earth:

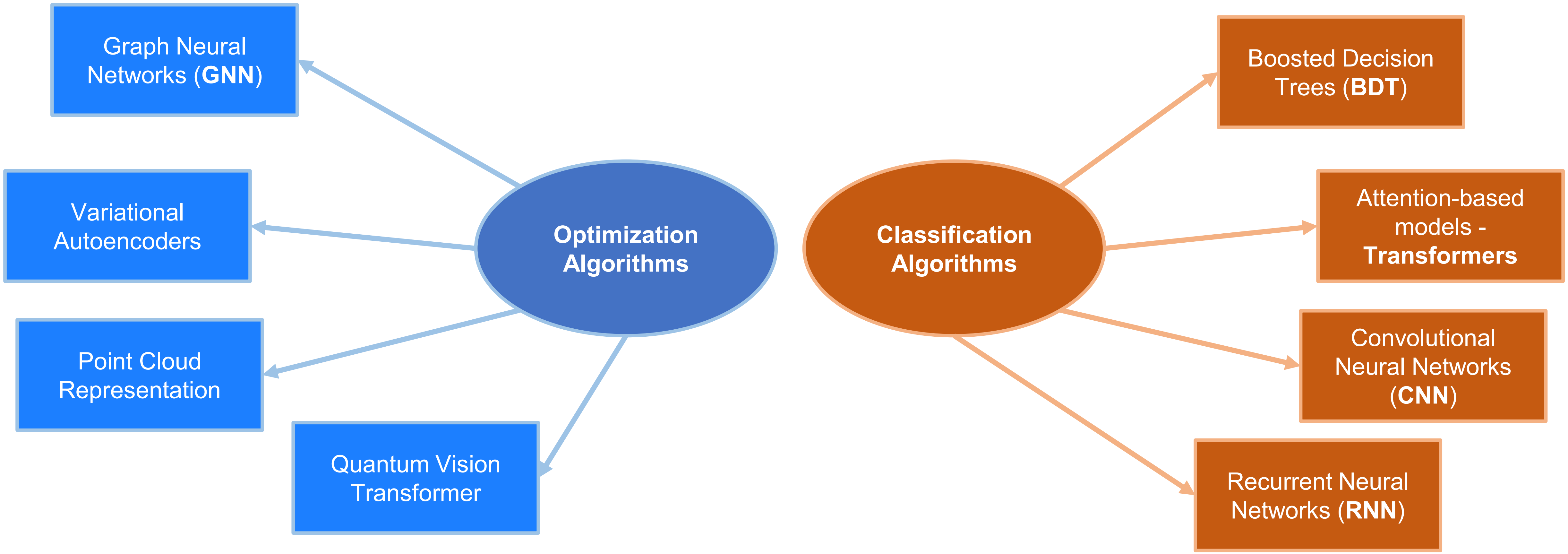
- Calorimeters generate high-resolution 3D data, often megabytes per event, accumulating **terabytes per day**.
- **Limited satellite bandwidth** cannot sustain such volumes.
- **Restricted time** windows for data transfer to ground stations.
- **Delayed data** processing **inhibits real-time responses** for critical events.

Benefits of On-Edge Predictions:

- **Drastic Data Reduction**: Raw data (~MB per event) replaced with synthetic results (~KB per event). Reduces transmission volume by orders of magnitude.
- **Energy Efficiency**: Conserves energy by minimizing data transmission.
- **Real-Time Decision Making**: Immediate processing prioritizes critical events → Adaptive strategies improve mission efficiency.

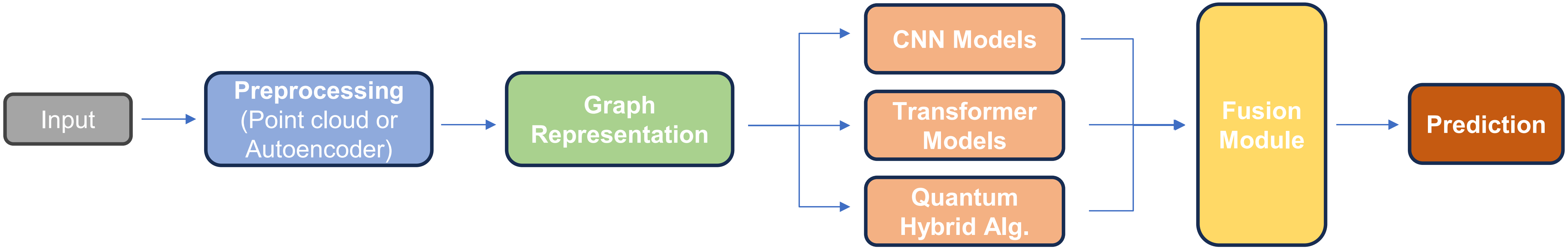
Technical Objectives, Methodologies and Solutions

Deep Learning algorithms to process and analyze spatial, temporal and energetic data from calorimeters.



Technical Objectives, Methodologies and Solutions

Rappresentation of the Process



Involved Staff and new recruitments

Involved Staff:

- Prof. Alberto Garinei: Full Professor of Mechanical and Thermal Measurements
- Prof. Francesca Fallucchi: Associate Professor of Computer Science
- Prof. Matteo Martini: Associate Professor of Experimental Physics

Additional staff:

Three Associate Professors of the Department of Engineering Sciences

Staff	Level	Additional staff
ALBERTO GARINEI	Full Professor	N
MATTEO MARTINI	Associate Professor	N
FRANCESCA FALLUCCHI	Associate Professor	N
SABINO MEOLA	Associate Professor	Y
FABRIZIO ZUCCARI	Researcher	Y
ADRIANO RANTIANGELI	Researcher	Y

Involved Staff and new recruitments

Preliminary budget: personnel and licensing costs

Updated budget: only personnel costs

Voce di costo	Anno 2024	Anno 2025	Totale
Attività A: WP1 Studio e Ideazione Algoritmi			
Spese di personale	10.000,00 €	55.122,00 €	65.122,00 €
Attività B: WP2 Porting degli Algoritmi su GPU			
Spese di personale		65.122,00 €	65.122,00 €
Costi per materiali, attrezzature e licenze			- €
Totali costi	10.000,00 €	120.244,00 €	130.244,00 €
Altri costi indiretti (specificare)			
Attività A: WP1 Studio e Ideazione Algoritmi			
Costi indiretti, determinati forfettariamente e pari al 15% dei costi diretti ammissibili per il personale	1.500,00 €	8.268,30 €	9.768,30 €
Attività B: WP2 Porting degli Algoritmi su GPU			
Costi indiretti, determinati forfettariamente e pari al 15% dei costi diretti ammissibili per il personale	- €	9.768,30 €	9.768,30 €
Totali costi complessivi			149.780,60 €

Timescale, Milestones, SAL

Work package title	#	Sub-Work package title	dic-24	gen-25	feb-25	mar-25	apr-25	mag-25	giu-25	lug-25	ago-25	set-25	ott-25	nov-25
					SAL1			SAL2			SAL3			SAL4
WP1 - Studio e Ideazione Algoritmi	WP1.1	1 - Analisi Preliminare e Definizione dei Requisiti												
	WP1.2	2. Sviluppo e Valutazione di Prototipi di Algoritmi												
	WP1.3	3. Ottimizzazione e Selezione degli Algoritmi												
WP2 Porting degli Algoritmi su GPU	WP2.1	1 - Adattamento degli Algoritmi per l'Esecuzione su GPU												
	WP2.2	2 - Testing e Profiling su Piattaforme CUDA												
	WP2.3	3. Ottimizzazione delle Prestazioni e Scalabilità												

SAL 1 – M3: end of WP1.1 - Algorithm dataset and system requirements

SAL 2 – M6: start of WP1.3 - First algorithm prototype

SAL 3 – M9: end of WP1.2 - Definitive algorithm prototype

SAL 4 – M12: end of the project - Validated and optimized algorithm

Thanks!