University of Messina, Italy
Department of Mathematical and Computer Sciences,
Physical Sciences and Earth Sciences

Subject: Internship Proposal

| | |
|---|---|
| *ID* | PTI_Daniele Ravi_16/07/2025 12.09.29 |
| *Data* | 16/07/2025 12.09.29 |

**Project Supervisor**

| | |
|---|---|
| *Surname* | Daniele |
| *Name* | Ravi |
| *Department* | 20/11/1983 |
| *Laboratory* | 511 |
| *E-mail* | daniele.ravi@unime.it |
| *Phone number* | 3717605146 |

**Project Co-Supervisor**

| | |
|---|---|
| *Surname* | Fazio |
| *Name* | Maria |
| *Job Position* | Associate Professor |
| *Department* | MIFT |

| | |
|---|---|
| *Laboratory* | 511 |
| *E-mail* | maria.fazio@unime.it |
| *Phone number* | |

**Project details**

| | |
|---|---|
| *Title* | Deep Learning for Multimodal Fusion of Genomic and Imaging Data Using Transformers and LLMs |

*Detailed description:* Context and Motivation
Modern precision medicine is increasingly powered by multimodal data integration, combining genomic data (such as DNA sequencing) with medical imaging (e.g., MRI, CT scans, histopathology) for more accurate diagnosis, prognosis, and patient stratification. Effectively merging these inherently different data types—each with their own structure, dimensionality, and information content—is a central challenge. Recent advances in transformer architectures and large language models (LLMs) have shown remarkable performance in building joint representations from diverse modalities. Incorporating these models enables unified embedding spaces that capture meaningful relationships between genetics and imaging, thereby boosting predictive power in disease modeling and personalized medicine.

Internship Objectives
Literature review: Systematically survey the latest research on multimodal data fusion in biomedical AI, with a particular focus on the use of transformer models and LLMs for handling genomic and imaging data streams.

Data preparation: Preprocess and harmonize imaging and genomic datasets, ensuring compatibility for downstream embedding.

Development and comparison of fusion models:

Implement transformer-based architectures (such as Multimodal Transformers or Vision Transformers for imaging, and LLMs for sequence or tabular genomic data) to extract rich, aligned latent representations from each data type.

Experiment with joint embedding techniques, where LLMs are used to embed text-based

genomic profiles and transformers process image data, followed by fusion strategies (early/intermediate/late fusion) in a unified network.

Evaluation: Benchmark the models on clinically meaningful tasks such as survival prediction, disease classification, or patient risk profiling, comparing transformer/LLM-based approaches with classical multimodal fusion methods.

Expected Outcomes
Comprehensive Literature Review
A structured critical review of the state-of-the-art in multimodal biomedical AI with a focus on transformer architectures and LLMs for combining DNA/genomic and imaging data.

Analize/Develop deep learning pipeline processing/embedding imaging and DNA data using transformer/LLM-based embeddings.

Comparative analysis of different fusion strategies and demonstration of improved clinical/biological prediction.

| | |
|---|---|
| *Duration (month – max 12)* | 6 |
| *Duration (hours)* | 150 |
| *Open positions* | 1 |

**Internship Skills**

*Technical requirements:* • Background in machine learning and deep learning (Python, PyTorch or TensorFlow).
- Experience with data preprocessing and integration.
- Interest in clinical research and good communication skills

| | |
|---|---|
| *Other skills* | |