

# Impieghi della computer vision nelle applicazioni industriali

metodologie di progettazione di  
componenti a basso consumo  
energetico



SAPIENZA  
UNIVERSITÀ DI ROMA



Claudia Melis Tonti,  
Dottorato Nazionale in AI presso  
Sapienza  
Supervisore: Irene Amerini





## The vision, perception and learning laboratory



**Prof. Marco Schaerf**  
Full Professor



**Prof. Irene Amerini**  
Associate Professor



**Prof. Roberto Beraldi**  
Associate Professor



**Pietro Manganelli  
Conforti**  
PhD student



**Taiba Majid**  
PhD student



**Emanuele Iacobelli**  
PhD student



**Claudia Melis Tonti**  
PhD student



**Luca Giovannesi**  
PhD student



**Francesco Pro**  
PhD student



**Claudio Schiavella**  
PhD student



**Lorenzo Cirillo**  
PhD student



**Davide Bazzana**  
PhD student



**Giuseppe Daidone**  
PhD student



**Gianmarco Scarano**  
PhD student



**Simone Teglia**  
PhD student



**Michela R. Celsi**  
PhD student

Alcor Lab is one of the laboratory at the Department of Computer, Control and Management Engineering A. Ruberti (DIAG)

# I nostri argomenti di ricerca



Multimedia  
Forensic



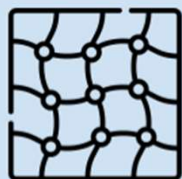
XAI and  
Robustness



Earth  
Observation



Efficient and  
Generalizable  
Models



Computer  
Graphics



Edge  
Computing



Medical  
Imaging



Multimodal  
Computer Vision

# Computer Vision

- Definizione: “La computer vision mira a costruire sistemi autonomi in grado di svolgere alcuni dei compiti che il sistema visivo umano è in grado di compiere.”
- Comprensione di ciò che si vede: estrazione di informazioni di oggetti presenti in immagini, video, ambienti 3D (classificazione)
- Rielaborazione di ciò che si vede per costruire nuove immagini, video, ambienti 3D (generazione)

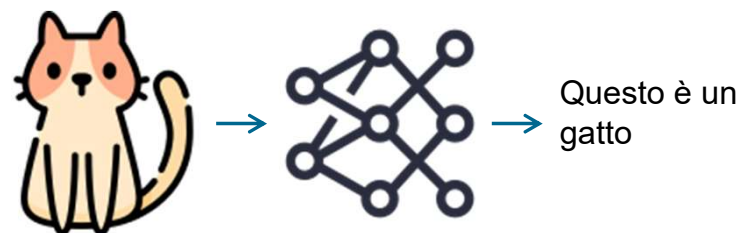
# Evoluzione della Computer Vision

Uso di algoritmi tradizionali matematici per estrazione di informazioni da immagini e video come bordi, angoli e texture (es. Canny edge detection)



↓  
**Deep Learning**  
↓

Uso di modelli di IA per l'estrazione automatica di informazioni e rielaborazione del contenuto.



# Task principali: immagini

- Classificazione

«C'è un gatto nella foto?»



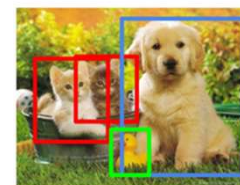
- Localizzazione

«Dov'è il gatto?»



- Rilevamento

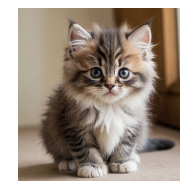
«Che oggetti ci sono nella foto?»



Gatto, Cane,  
Paperella

- Generazione

«Crea l'immagine di un gatto»



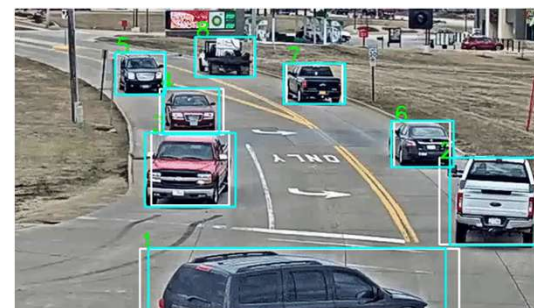


# Task principali: video

- Riconoscimento dell'azione
- Tracciamento oggetti
- Generazione/Deep Fake

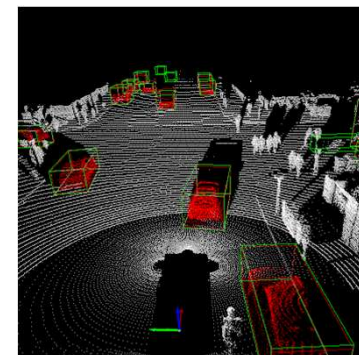
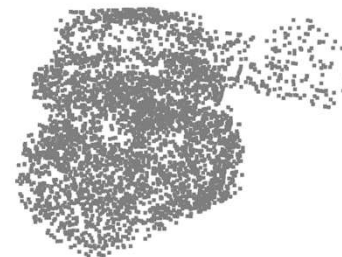


“La persona nel video  
sta correndo”



# Task principali: 3D

- Rappresentazione
- Localizzazione
- Generazione del movimento





# Il problema dell'overload energetico

ChatGPT e Gemini

- Comprensione del testo
- Generazione del testo
- Classificazione e rilevamento di immagini
- Generazione di immagini
- Generazione di video

# Il problema dell'overload energetico

## Training

- Convoluzione: filtrare le informazioni dell'input
- Metodo: si usa una finestra che funziona come «lente di ingrandimento» su determinati punti dell'input. La finestra «scorre» nell'input passo dopo passo, catturando il contesto da cui poi estrae l'informazione

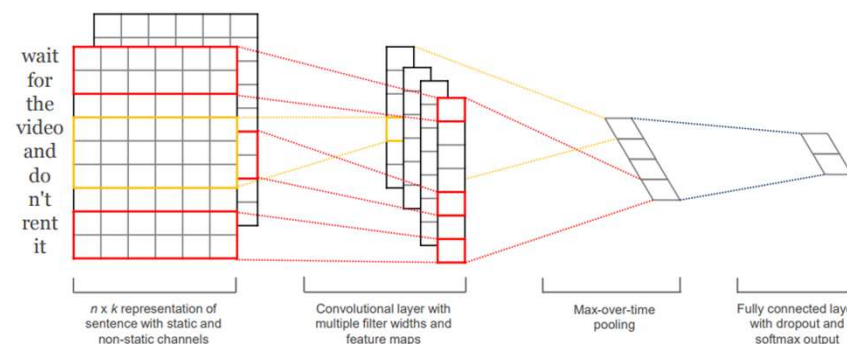


Figure 1: Model architecture with two channels for an example sentence.

Kim, Yoon. "Convolutional Neural Networks for Sentence Classification." *Conference on Empirical Methods in Natural Language Processing* (2014).

# Il problema dell'overload energetico

## Training

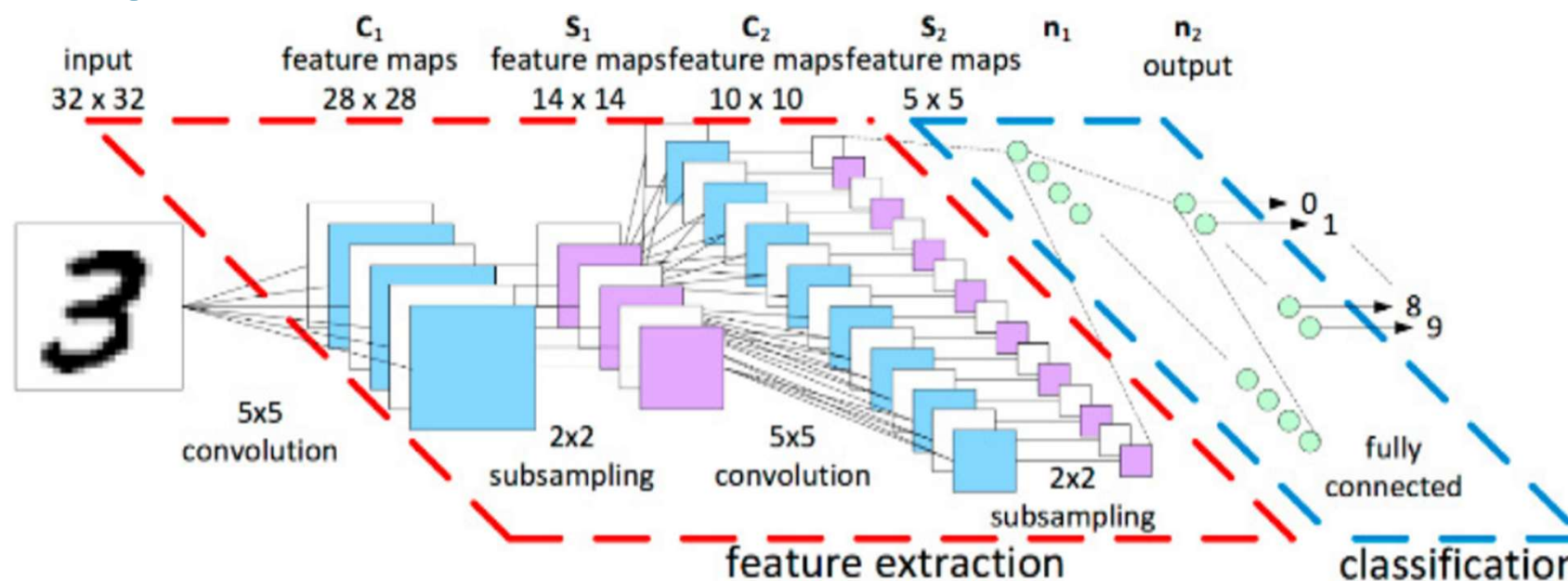
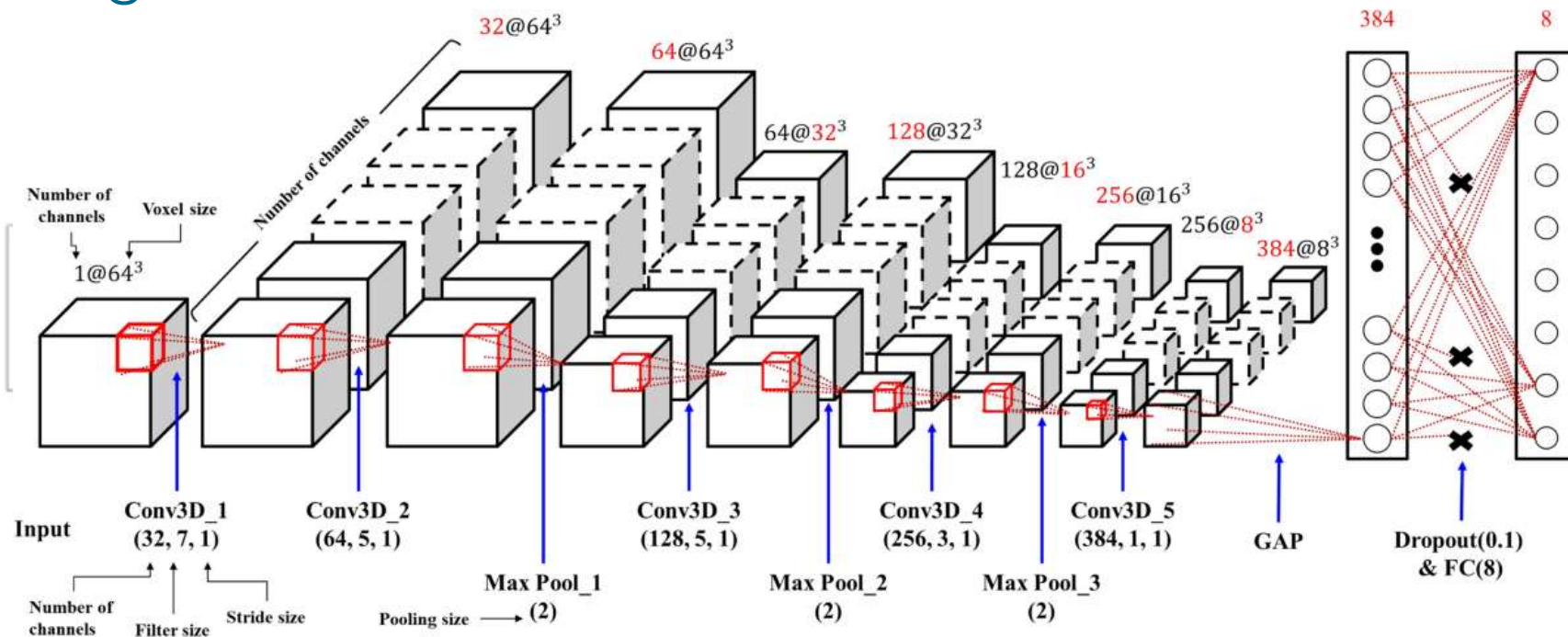


Fig. 1 Internal Layers of CNNs

Neha Sharma, Vibhor Jain, Anju Mishra, An Analysis Of Convolutional Neural Networks For Image Classification, Procedia Computer Science, Volume 132, 2018, Pages 377-384

# Il problema dell'overload energetico

## Training

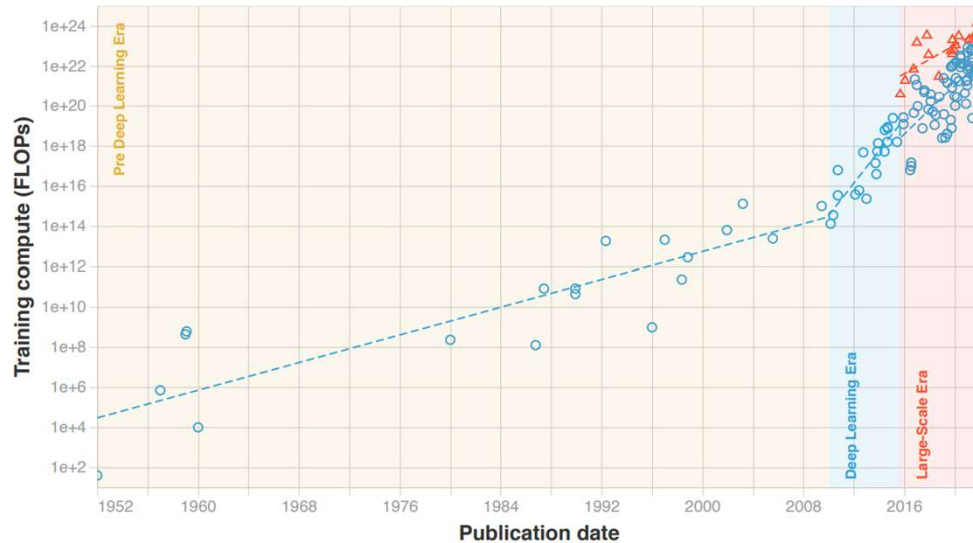


Lee, J., Lee, H. & Mun, D. 3D convolutional neural network for machining feature recognition with gradient-based visual explanations from 3D CAD models. Sci Rep 12, 14864 (2022).

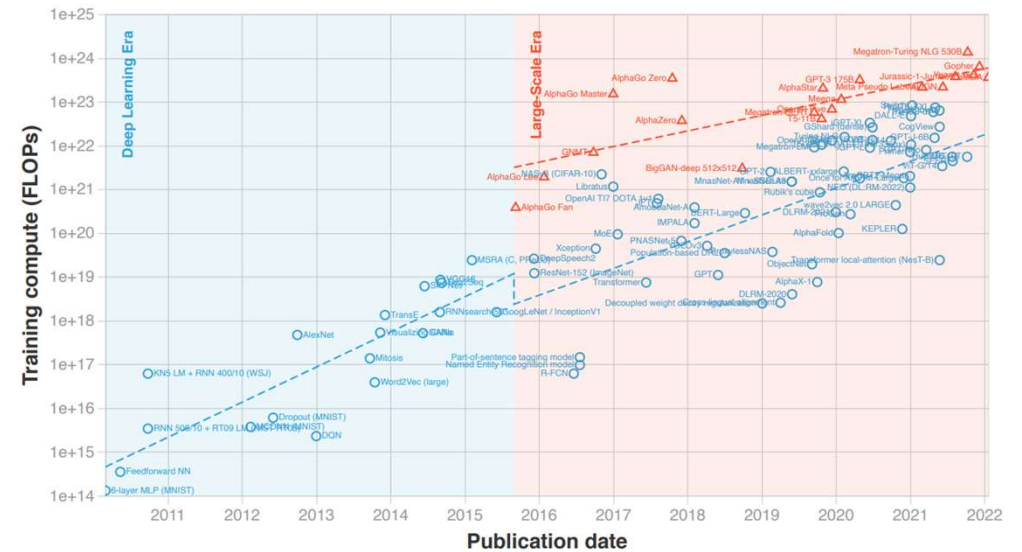
# Il problema dell'overload energetico

## Training

Training compute (FLOPs) of milestone Machine Learning systems over time  
n = 121



Training compute (FLOPs) of milestone Machine Learning systems over time  
n = 102



Sevilla, Jaime et al. "Compute Trends Across Three Eras of Machine Learning." 2022 International Joint Conference on Neural Networks (IJCNN) (2022): 1-8.

# Il problema dell'overload energetico

## Inferenza

task	inference energy (kWh)	
	mean	std
text classification	0.002	0.001
extractive QA	0.003	0.001
masked language modeling	0.003	0.001
token classification	0.004	0.002
image classification	0.007	0.001
object detection	0.038	0.02
text generation	0.047	0.03
summarization	0.049	0.01
image captioning	0.063	0.02
image generation	2.907	3.31

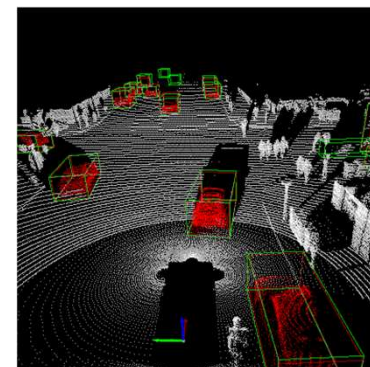
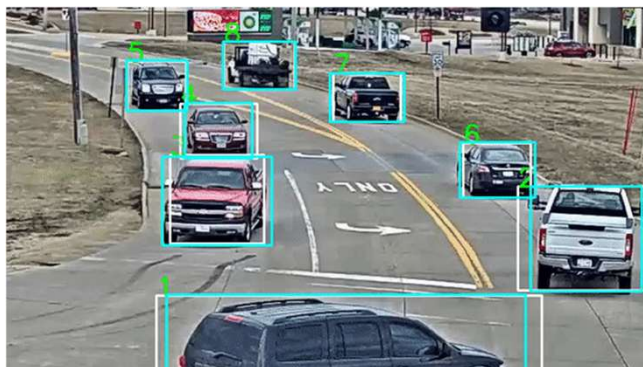
Table 2. Mean and standard deviation of energy per 1,000 queries for the ten tasks examined in our analysis.

Luccioni, Sasha et al. "Power Hungry Processing: Watts Driving the Cost of AI Deployment?" *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (2023): n. pag.



# Il problema dell'overload energetico

Il dataset

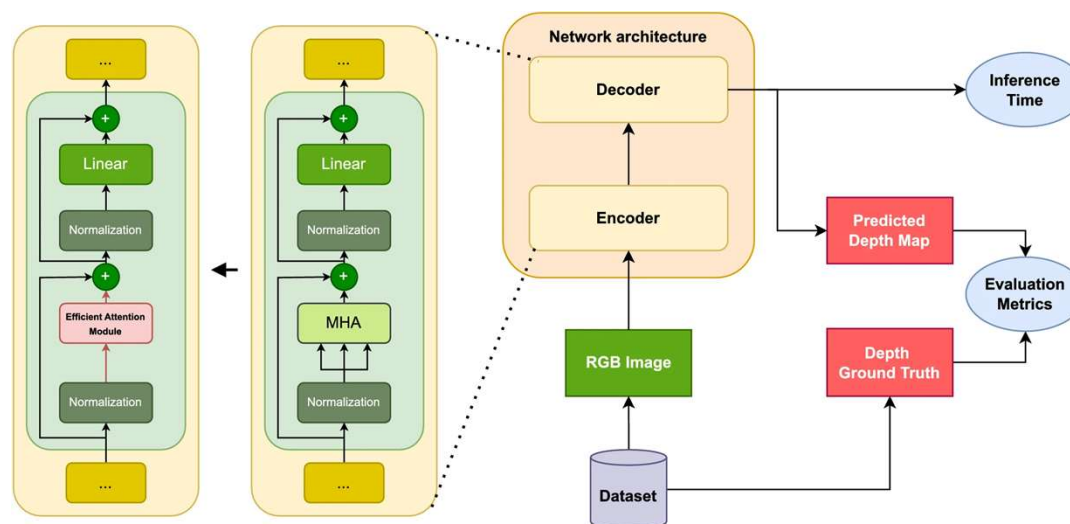


# Come lo gestiamo noi?

- Metodi efficienti per l'inferenza:
  - Riduzione degli sprechi sull'architettura
    - Eliminazione di layer non necessari
    - Semplificazione delle operazioni
    - Distillazione
- Compressione del dataset

# Le nostre ricerche a riguardo

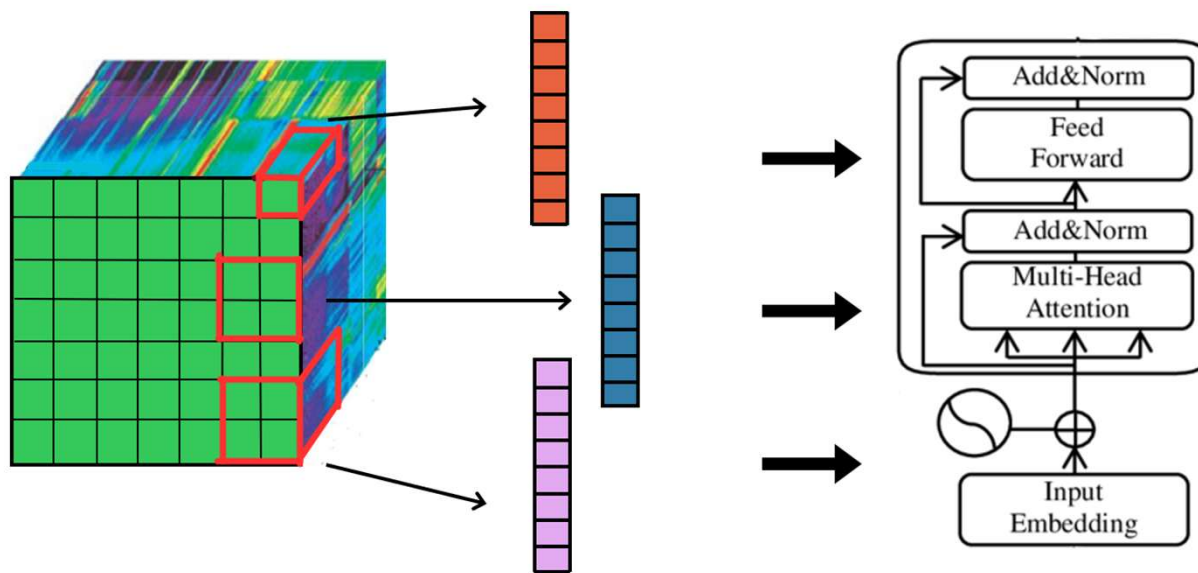
Efficient attention ViT for MDE on resource-limited hardware



Schiavella, C., Cirillo, L., Papa, L. et al. Efficient attention vision transformers for monocular depth estimation on resource-limited hardware. Nature Scientific Reports, 2025

# Le nostre ricerche a riguardo

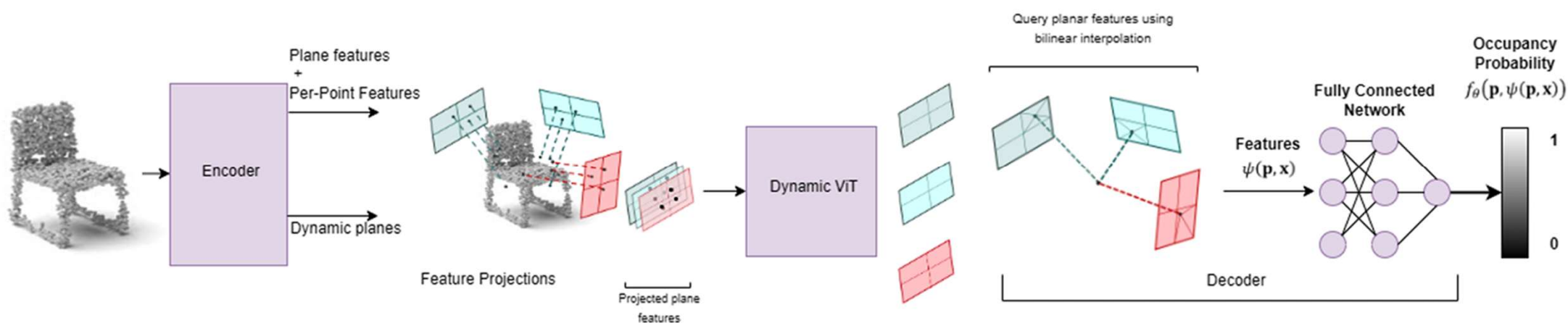
## Geometric Algebra Transformer



[4] Fabi., Schiavella, Amerini. Geometric Algebra Transformer with Spectral-Spatial-Volumetric Feature Extraction for Hyperspectral Image Classification. Pattern Recognition Letters. 2025

# Le nostre ricerche a riguardo

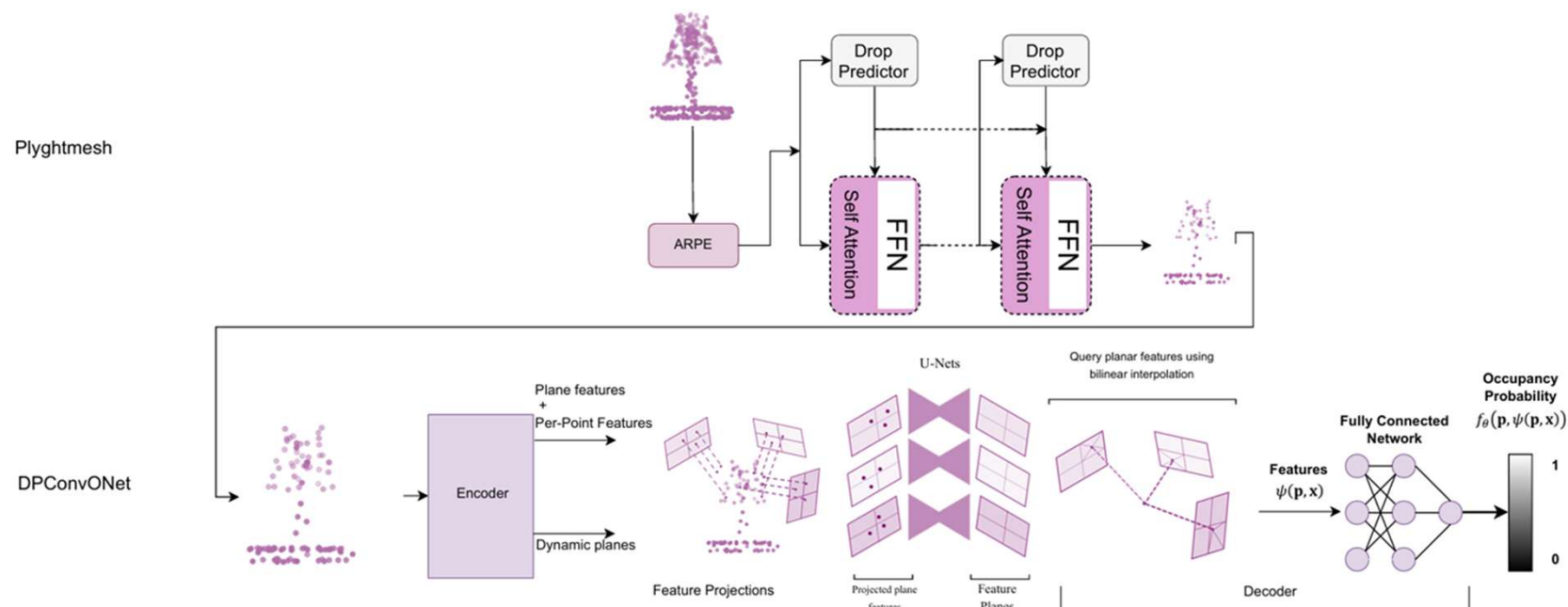
## 3D mesh reconstruction



Tonti, Claudia, and Irene Amerini. "Lightweight transformer occupancy networks for 3d virtual object reconstruction." Lightweight Transformer Occupancy Networks for 3D Virtual Object Reconstruction. SciTePress, 2025. 408-414.

# Le nostre ricerche a riguardo

## Point Cloud Compression





# Contatti

## Alcor Lab



**WEBSITE**

<https://alcorlab.diag.uniroma1.it/>



**EMAIL**

[alcor@diag.uniroma1.it](mailto:alcor@diag.uniroma1.it)

## Personal contacts



**Claudia Melis Tonti**

[melistonti@diag.uniroma1.it](mailto:melistonti@diag.uniroma1.it)



**SAPIENZA**  
UNIVERSITÀ DI ROMA