Causality-Driven One-Shot Learning for Prostate Cancer Grading from MRI

Gianluca Carloni*[1,2] Eva Pachetti*[1,2] Sara Colantonio[1]

[1]Institute of Information Science and Technologies (ISTI-CNR), Italy [2]University of Pisa, Italy

* These authors contributed equally {gianluca.carloni, eva.pachetti}@isti.cnr.it

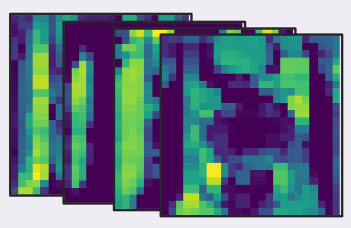
Motivation

- Find a way to exploit intrinsic information: weak **causal signals**
- Work in low-data scenario, as in the medical domain: **Few-Shot Learning**

Methods

1) Extract features

$F^1, F^2, \dots F^k$



2) Compute causality maps

$$P(F^i|F^j) = \frac{P(F^i, F^j)}{P(F^j)}$$

Max

$$P(F^i|F^j) = \frac{(\max_{r,c} F^i_{r,c}) \cdot (\max_{r,c} F^j_{r,c})}{\sum_{r,c} F^j_{r,c}}$$

Lehmer

$$P(F^i|F^j)_p = \frac{LM_p(F^i \times F^j)}{LM_p(F^j)}$$

E.g.	F0	F1	F2	F3
F0	0.51	0.62	0.83	0.31
F1	0.94	0.55	0.81	0.12
F2	0.46	0.61	0.6	0.11
F3	0.85	0.72	0.13	0.50

Conditional asymmetries

$$P(F^0|F^2) > P(F^2|F^0) \implies F^0 \to F^2$$
$$P(F^1|F^0) > P(F^0|F^1) \implies F^1 \to F^0$$
$$P(F^1|F^2) > P(F^2|F^1) \implies F^1 \to F^2$$
$$P(F^3|F^0) > P(F^0|F^3) \implies F^3 \to F^0$$
$$P(F^3|F^1) > P(F^1|F^3) \implies F^3 \to F^1$$
$$P(F^3|F^2) > P(F^2|F^3) \implies F^3 \to F^2$$

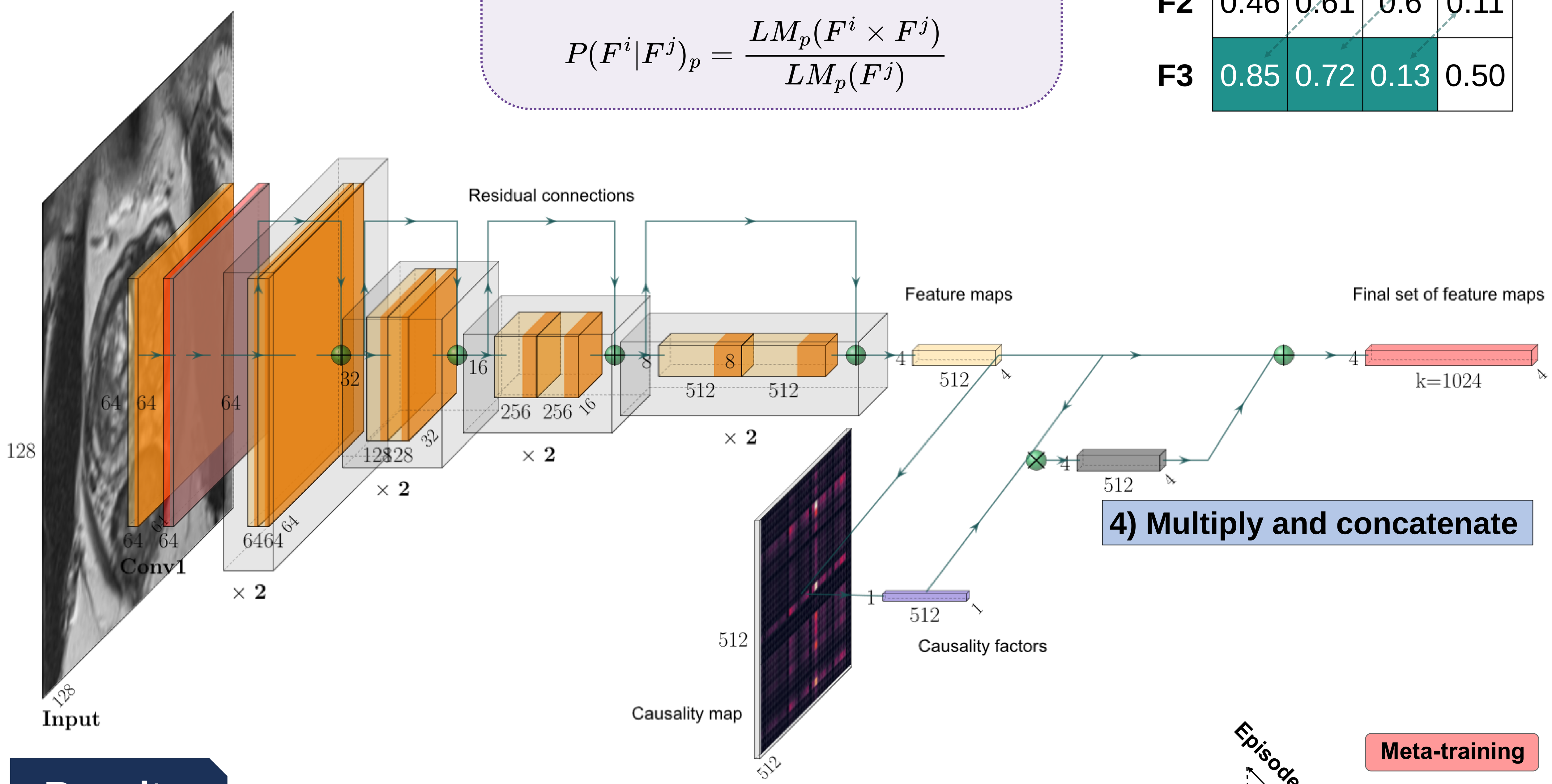
3) Determine causality factors

mulcat

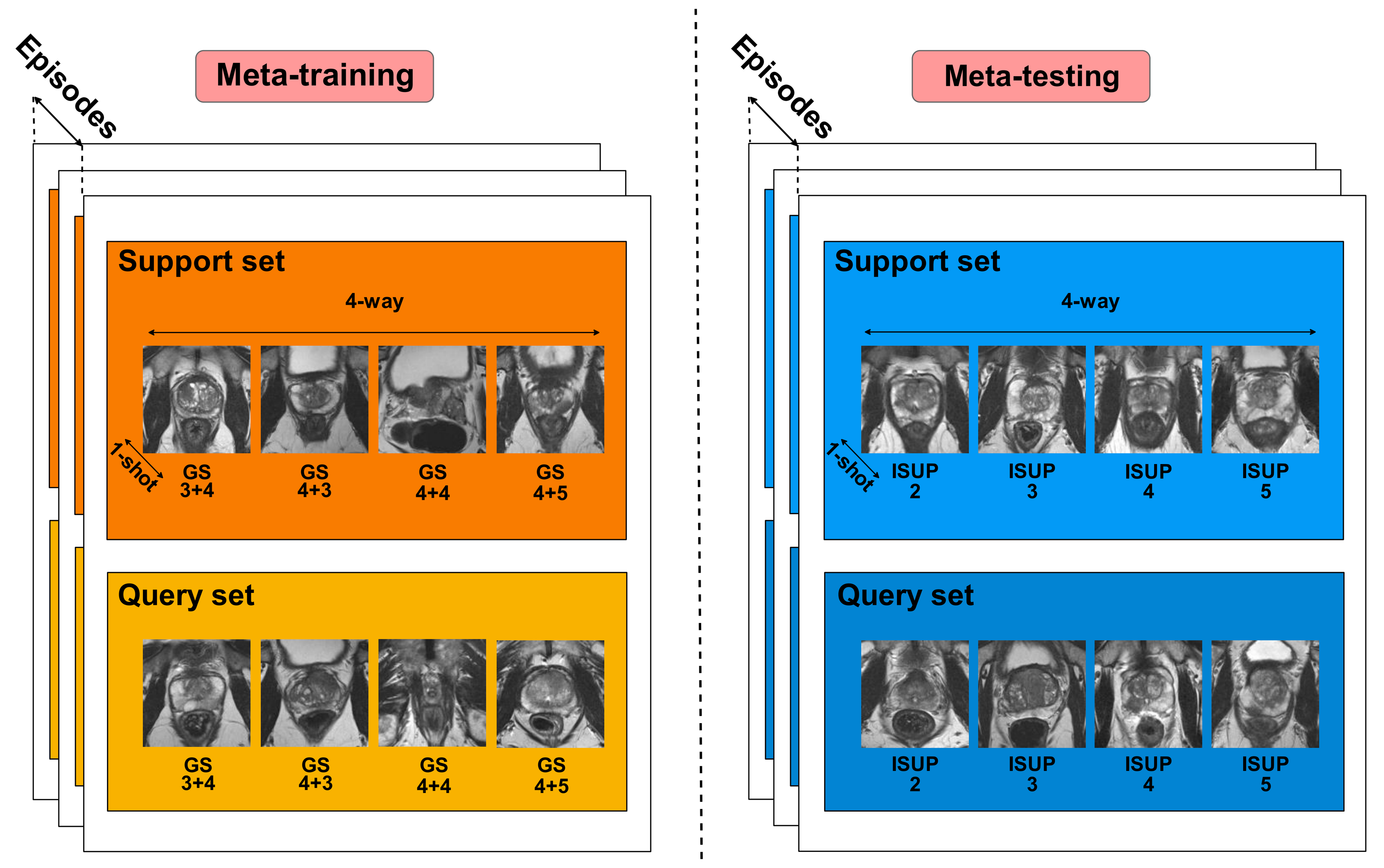
$$F^0 : 1 - 2 = \mathbf{-1}$$
$$F^1 : 2 - 1 = \mathbf{1}$$
$$F^2 : 0 - 3 = \mathbf{-3}$$
$$F^3 : 3 - 0 = \mathbf{3}$$

mulcatbool

$$F^0 : (1-2) \overset{?}{>} 0 = \mathbf{0}$$
$$F^1 : (2-1) \overset{?}{>} 0 = \mathbf{1}$$
$$F^2 : (0-3) \overset{?}{>} 0 = \mathbf{0}$$
$$F^3 : (3-0) \overset{?}{>} 0 = \mathbf{1}$$



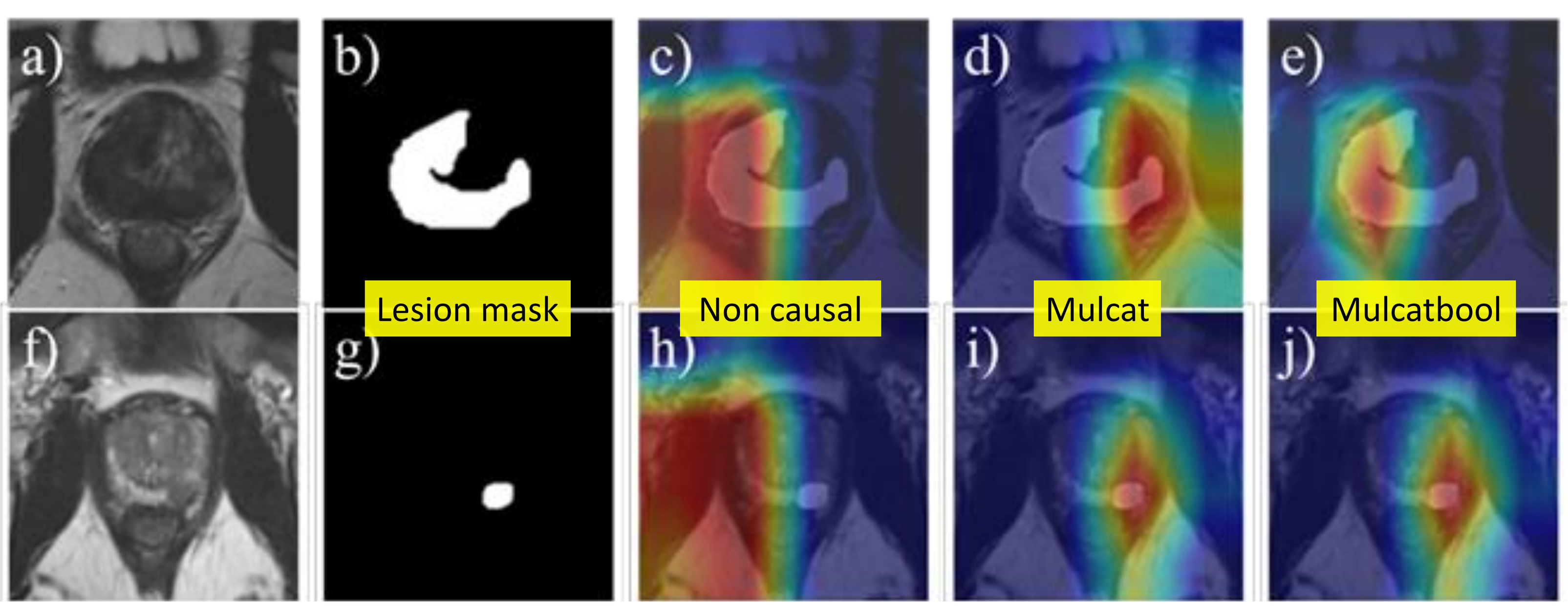
4) Multiply and concatenate



Results

Setting	2-way 1-shot	4-way 1-shot	4-way 1-shot*
Main experiments			
Non causality-driven	0.539 (0.141)	0.585 (0.068)	0.586 (0.118)
Causality-driven mulcat	0.550 (0.144)	0.611 (0.069)	0.712 (0.118)
Causality-driven mulcatbool	**0.556** (0.141)	**0.614** (0.067)	**0.713** (0.119)
Ablation study			
Ablation mulcat	0.535 (0.143)	0.557 (0.063)	0.557 (0.111)
Ablation mulcatbool	0.540 (0.139)	0.571 (0.068)	0.612 (0.119)

*: trained to distinguish four classes (ISUP 2 – 5), but the AUROC is computed between ISUP 2 vs. rest.



Conclusion

Integrating causality into a CNN model can:

- **Enhance** performance (classification metrics)
- Help the model focus more accurately on the critical regions of the image, leading to more **reliable** and **robust** predictions.