Multimodal colon cancer imaging database for disease modeling

C. Credi[1,2], L. Bocchi[1*], A. Picon[3], C. L. Saratxaga[3], L.F. Sánchez-Peralta[4], J.F. Ortega[4], J. Bote[4], A. Azpeitia Garcia[5], A. Egia[5], B. Glover[6], F.S. Pavone[2,7,8], and R. Cicchi[2,8]

[1] Department of Information Engineering, University of Florence, Florence, Italy
[2] European Laboratory for non-linear spectroscopy, Sesto Fiorentino (FI), Italy
[3] Fundación Tecnalia Research & Innovation, Derio, Spain
[4] Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, Spain
[5] Biobanco Vasco, Fundación Vasca de Innovación e Investigación Sanitarias, Barakaldo, Spain
[6] Imperial College of Science Technology and Medicine, London, UK
[7] Department of Physics, University of Florence, Florence, Italy
[8] National Institute of Optics, National Research Council, Sesto Fiorentino (FI), Italy
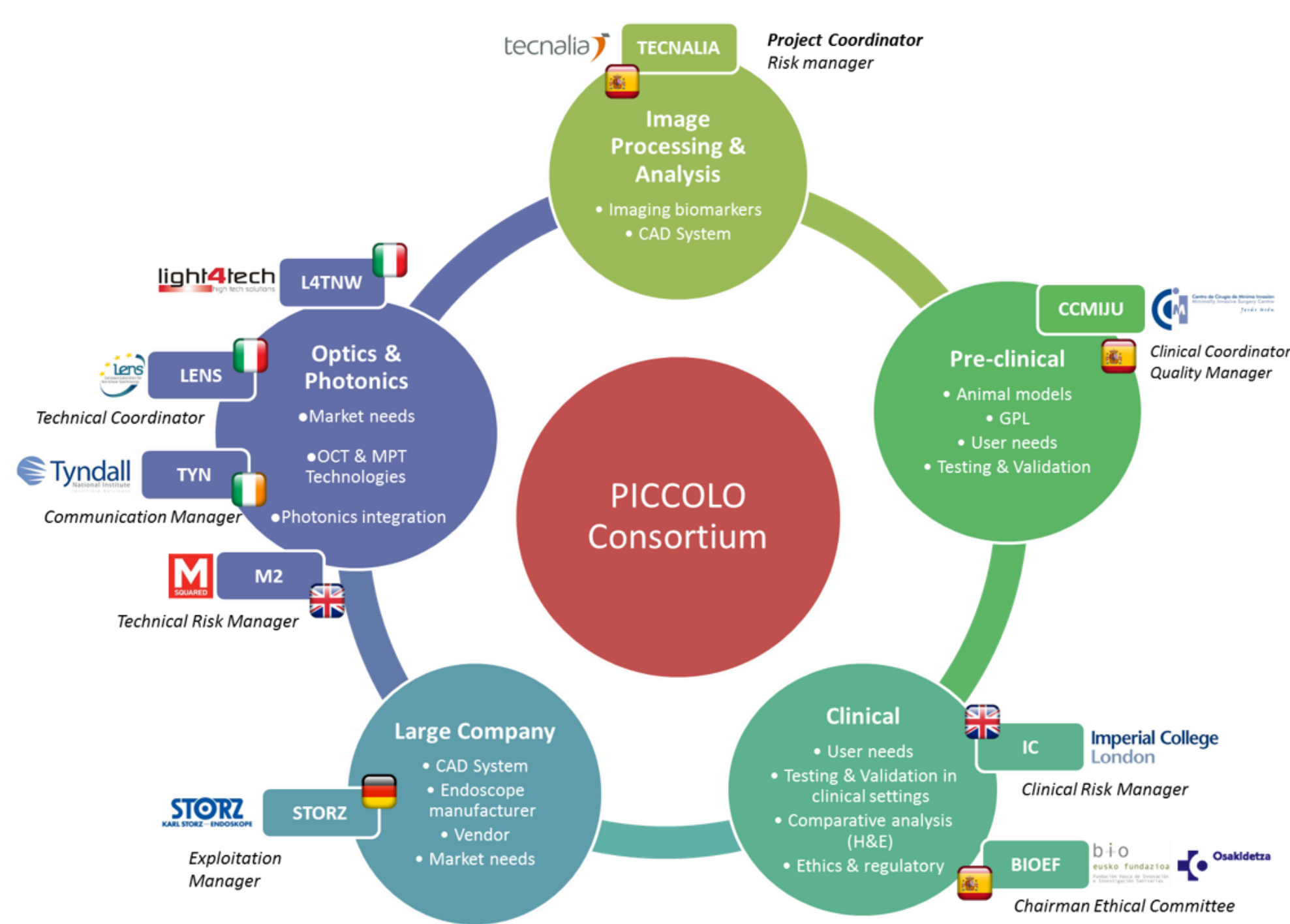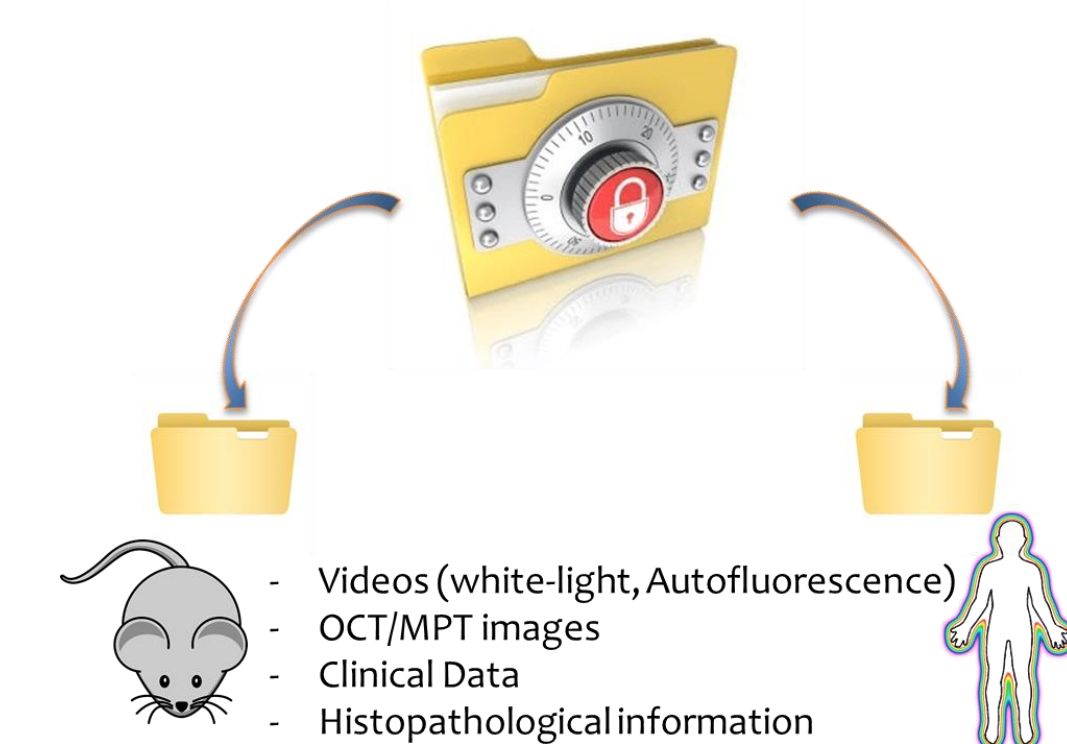
INTRODUCTION

Nowadays, the possibility to have digital databases organized to collect extensive amount of data on biological samples it is of great relevance for translational research. Digital biobanks represent a powerful tool that should gather and provide the access to biological data for research purposes. In the framework of the H2020 PICCOLO project, a digital image database of healthy colon, neoplastic and hyperplastic lesions for algorithm training will be produced. Such database should be organized i) to centralize images acquired by dislocated institutions; ii) to combine images with clinical data; iii) to share the information within the consortium while avoiding loss of data or risk of mislabeling, iv) to comply with EU data protection. To this end, standardization and synchronization of acquisition protocols allow secure data management as also privacy issues are addressed. Finally, optical coherence tomography (OCT) and multiphoton tomography (MPT) images are being collected through the endoscopic tool developed within the project, together with wide-field and NBI images from clinical colonoscopy and histopathological images and report, taken as benchmark. OCT/MPT imaging library will boost for biomarkers discovery, analysis of the OCT/MPT technology potential for colon cancer diagnosis, analysis on clinical prognosis related to the OCT/MPT biomarkers, and lesion grading studies among others.
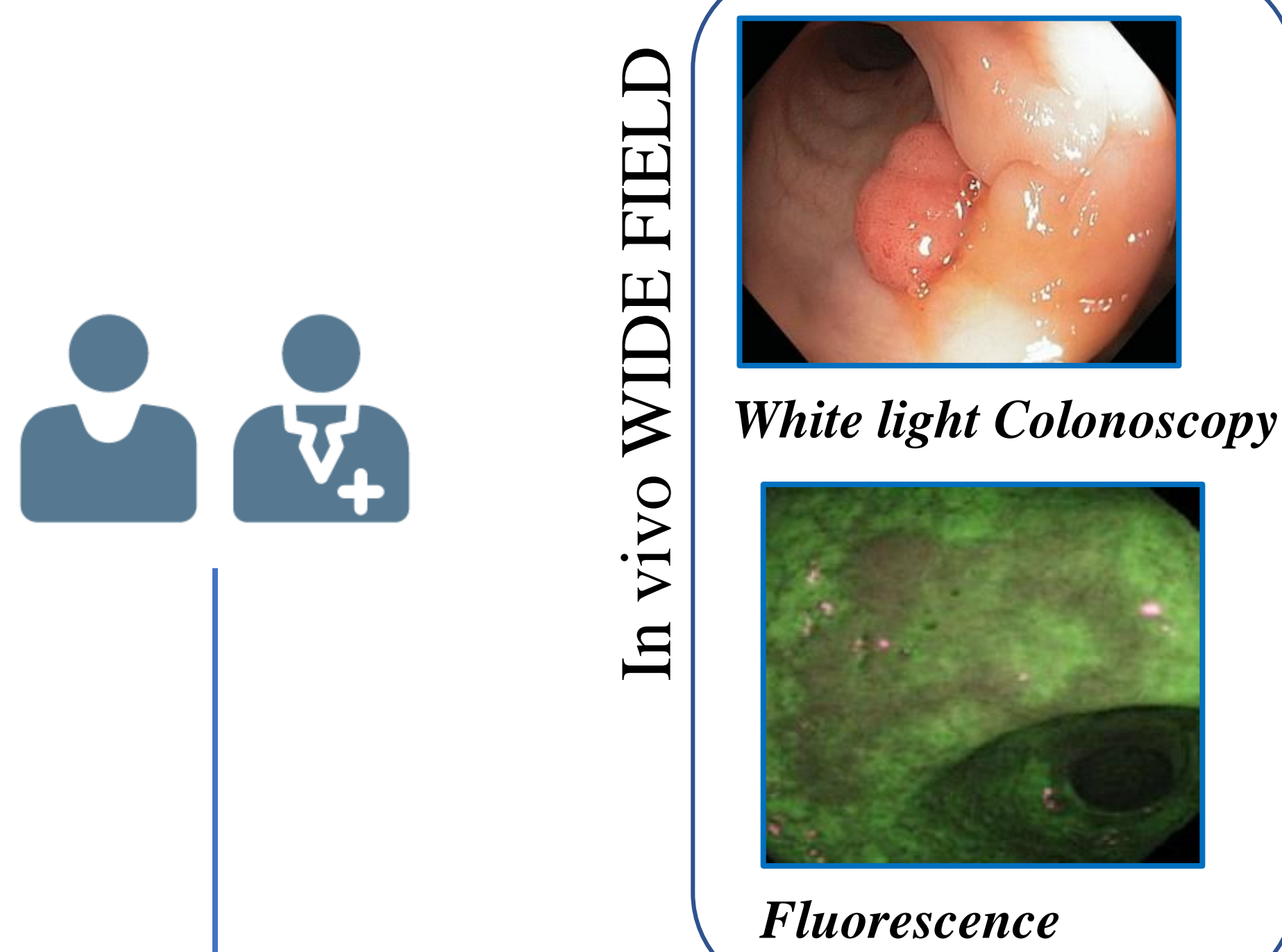
DATABASE

In order to properly store, centralize and share the extensive amount of data that will be generated within PICCOLO project, a Network-Attached Storage (NAS) has been selected to storage device exploitable for PICCOLO database generation. NAS systems enable to provide access to the same files through standard Ethernet connection to multiple heterogeneous computer clients, a fundamental aspect within the framework of PICCOLO project

Type of samples		N°	
Genetically modified murine models – murine model of hyperplasia (healthy and diseased) (neoplasia and hyperplasia)	Ex-vivo	≈ 200	OCT/MPT imaging + histological data + colonoscopy
	In-vivo	≈ 340	
Human samples (healthy and diseased) – ex vivo		≈ 400	

- Videos (white-light, Autofluorescence)
- OCT/MPT images
- Clinical Data
- Histopathological information

DETECTION

In vivo WIDE FIELD

White light Colonoscopy

Fluorescence

CHARACTERIZATION

Ex vivo on resection

OCT Images: a) Cross-section, c) 3D

500 μm

a) Histological hematoxilin-eosin
i) MPT images

IMAGE ANALYSIS

Training

Hyperplastic and neoplastic semantic segmentation for:
- *Margin assessment*
- *Infiltration analysis*
- *Localization*
- *Classification*

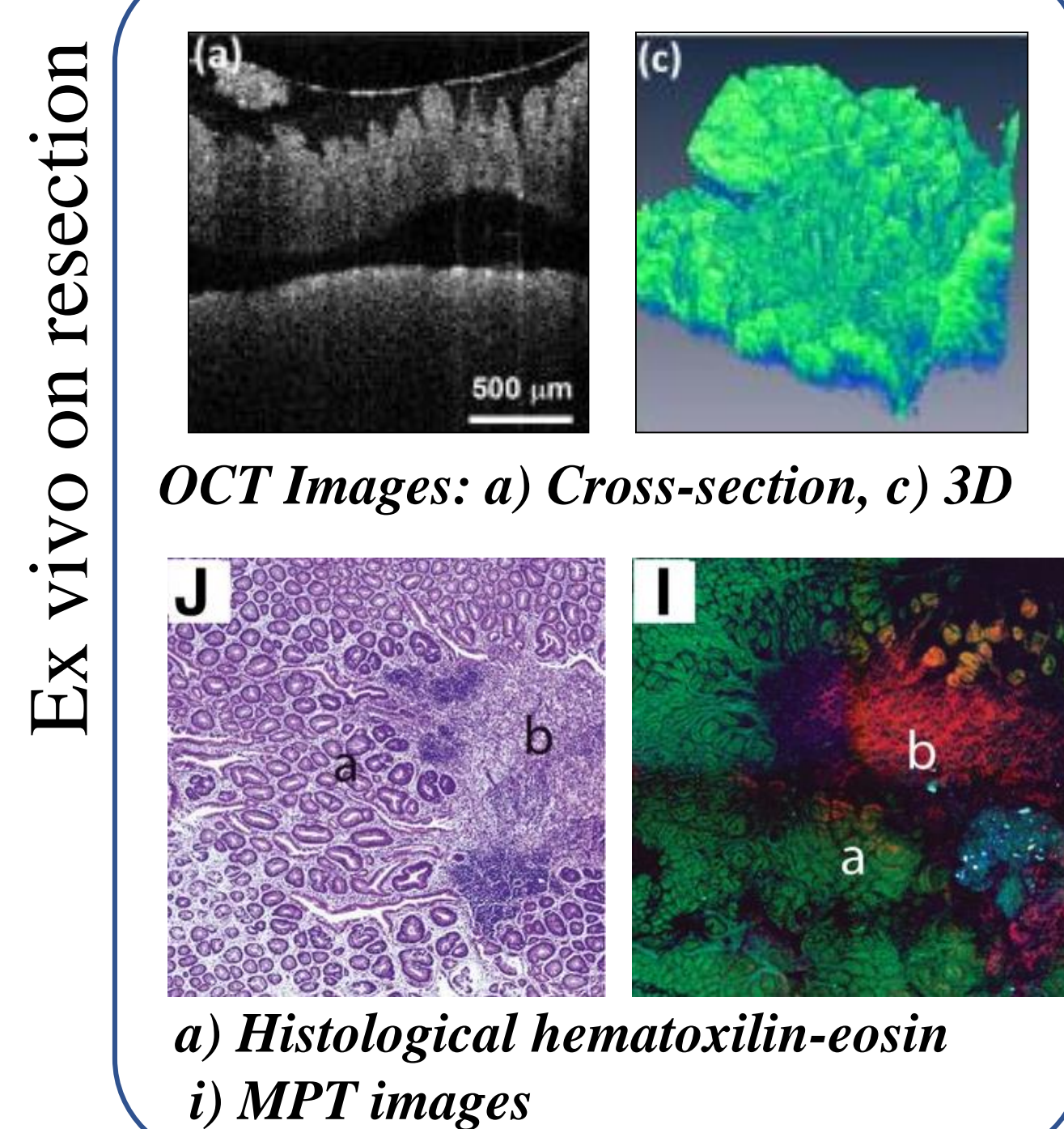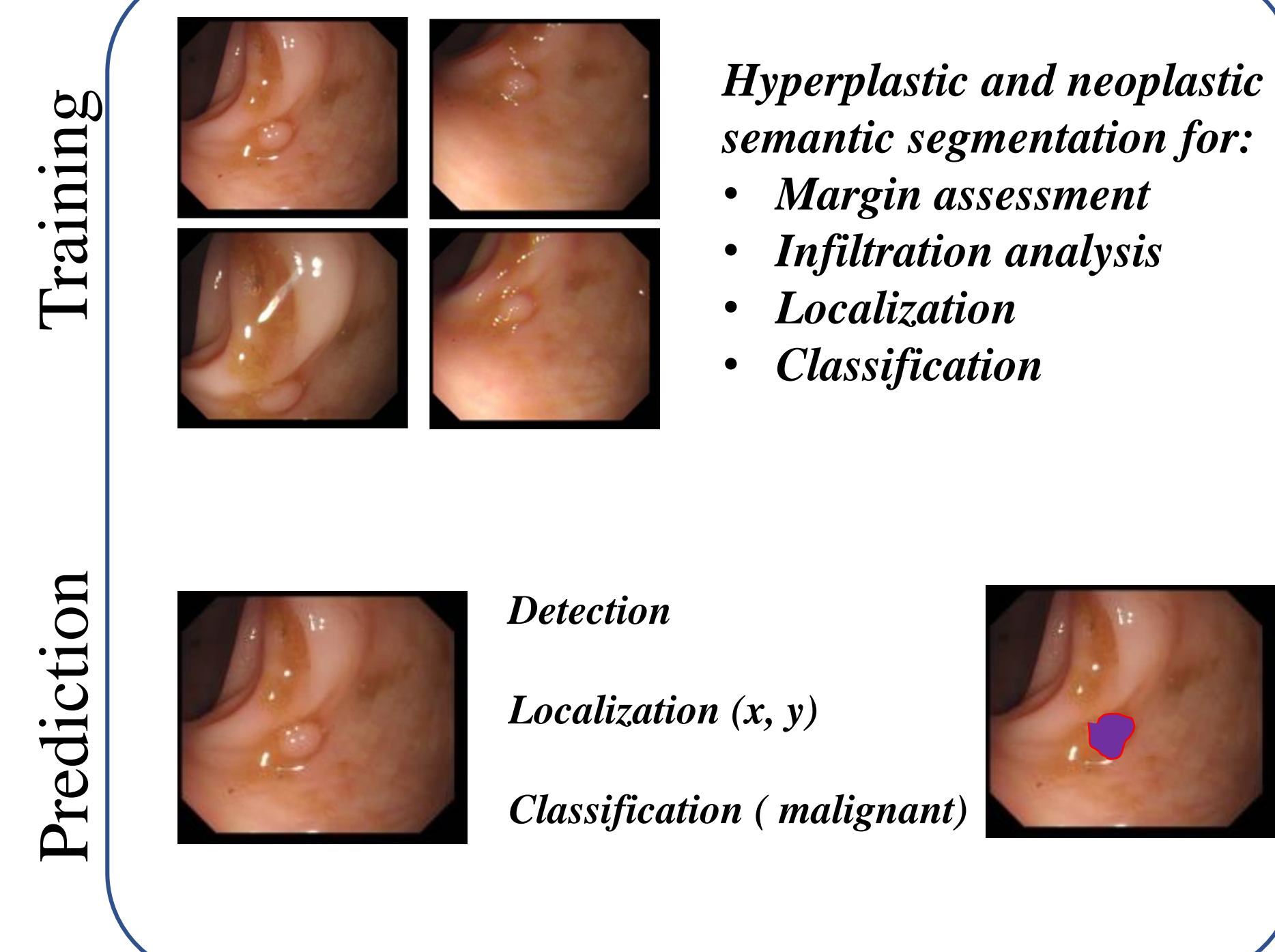Prediction

Detection

Localization (x, y)

Classification (malignant)

Database generation was approved by the Ethics and Investigation committee and a risk analysis was performed to ensure protection of personal data as requested by General Data Protection Regulation (GDPR). Regulation (EU) 2016/679

Security codes are released from the NAS administrators to consortium users upon signing a Consortium Agreement
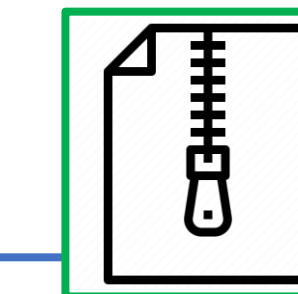
Lossless image compression formats: Digital Imaging and Communication in Medicine (DICOM), JavaScript Object Notation (JSON) Open Microscopy Environment.

CONCLUSION

The exhaustive animal and human colon imaging database created within the PICCOLO project will represent a powerful tool for the biomedical research community i) for the identification and validation of the imaging biomarkers; ii) for the development of tumoral models; iii) to foster and standardize the use of OCT/MPT techniques for colonoscopy.

REFERENCES

- PICCOLO H2020 Project, Multimodal highly sensitive PhotonICs endoscope for improved in-vivo COLOn Cancer diagnosis and clinical decision support. 2017-2020.
- Available at: www.piccolo-project.eu This work is supported by PICCOLO project. This project has received funding from the European Union's Horizon2020 Research & Innovation Programme under Grant Agreement No. 732111. This information reflects only the institutions view, so the Agencies and the Commission are not responsible for any use that may be made of the Information it contains.