The DeepHealth Toolkit: a Unified Framework to Boost Biomedical Applications

Michele Cancilla[1], Laura Canalini[1], Federico Bolelli[1], Stefano Allegretti[1], Salvador Carrión[2], Roberto Paredes[2], Jon A. Gómez[2], Simone Leo[3], Marco Enrico Piras[3], Luca Pireddu[3], Asaf Badouh[4], Santiago Marco-Sola[4], Lluc Alvarez[4], Miquel Moreto[4], and Costantino Grana[1]

[1]Università degli Studi di Modena e Reggio Emilia, Italy
[2]Data-intensive Computing Group, CRS4, Italy
[3]Universitat Politècnica de València, Spain
[4]Barcelona Supercomputing Center, Spain

The DeepHealth Project

- Put HPC computing power at the service of biomedical applications
- Increase the productivity of medical personnel and IT professionals
- Offer a unified framework adapted to exploit underlying heterogeneous HPC and Cloud architectures

Duration: 36 months
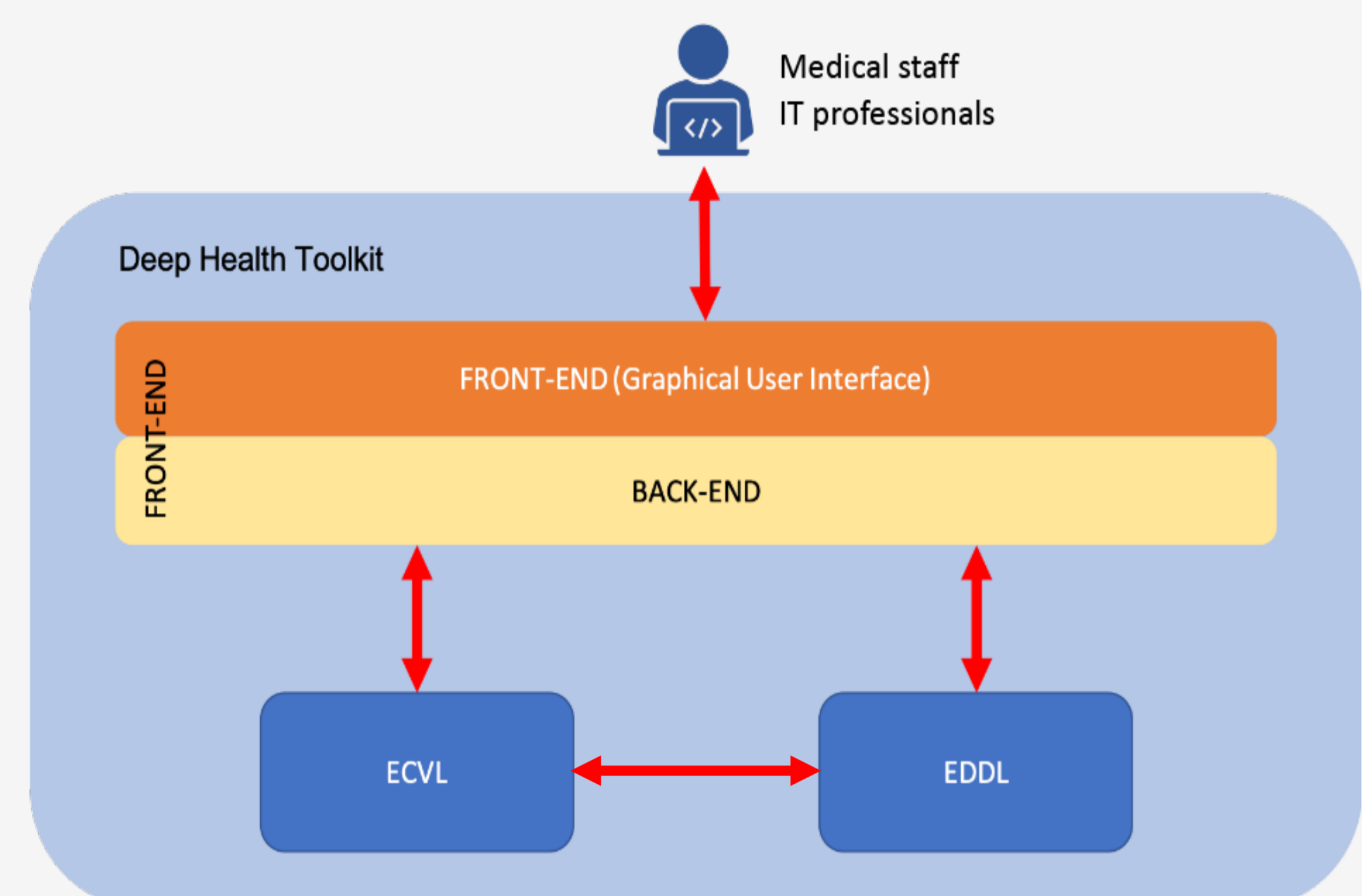**Starting date:** Jan 2019

Budget 14.642.366 €
EU funding 12.774.824 €

22 partners from **9 countries**
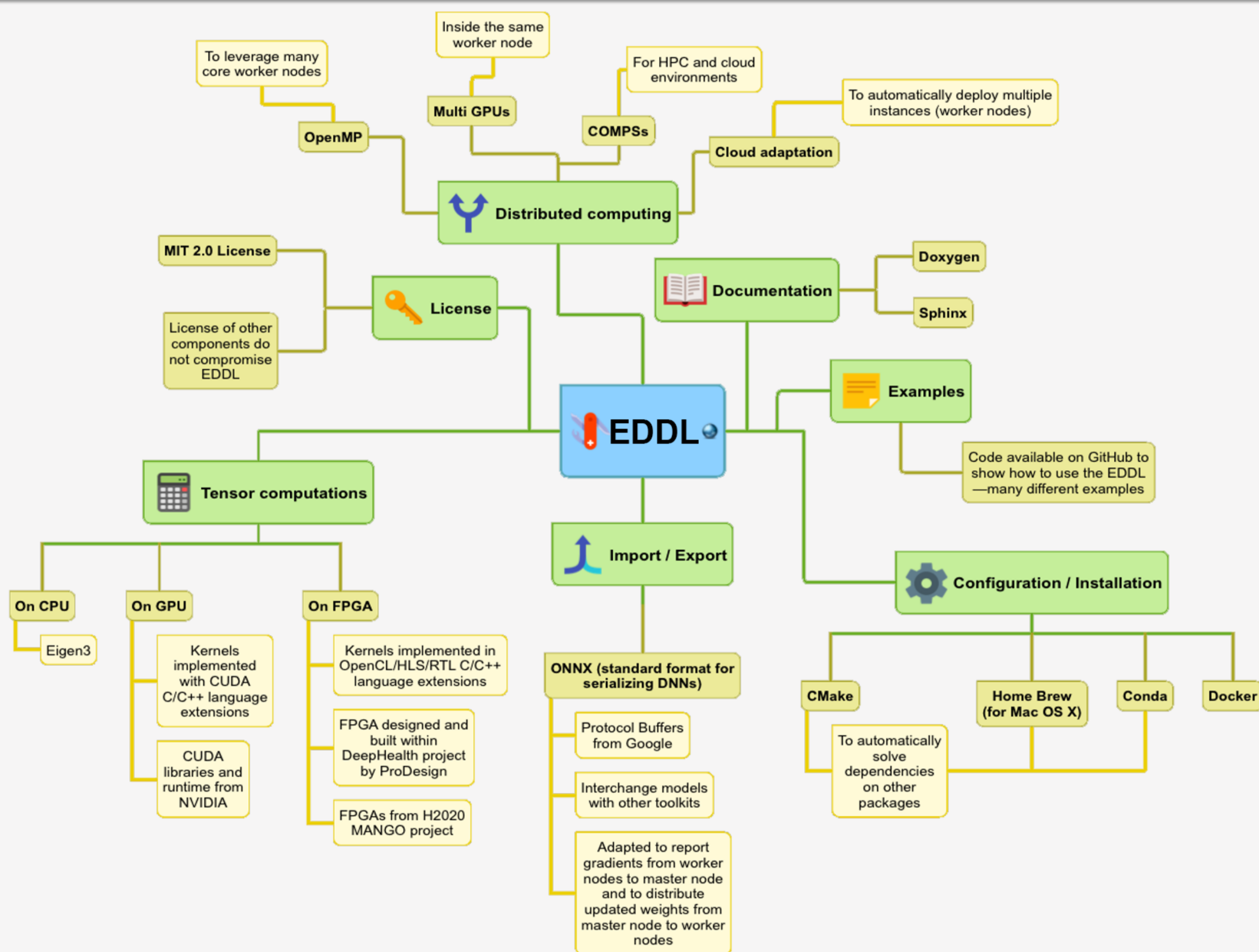Research centers, Health organizations, large industries and SMEs

The DeepHealth Toolkit

- Three main components:
  - EDDL – European Distributed Deep Learning Library
  - ECVL – European Computer Vision Library
  - Front-end – RESTful web service + web-based GUI
- Toolkit open-source and available at github.com/deephealthproject

Medical staff
IT professionals

Deep Health Toolkit

FRONT-END (Graphical User Interface)

BACK-END

ECVL ⟷ EDDL

EDDL

Inside the same worker node
To leverage many core worker nodes
Multi GPUs
For HPC and cloud environments
To automatically deploy multiple instances (worker nodes)
OpenMP
COMPSs
Cloud adaptation
MIT 2.0 License
Distributed computing
Doxygen
Documentation
License of other components do not compromise EDDL
License
Sphinx
Examples
EDDL
Code available on GitHub to show how to use the EDDL —many different examples
Tensor computations
Import / Export
Configuration / Installation
On CPU | On GPU | On FPGA
Eigen3
Kernels implemented with CUDA C/C++ language extensions
Kernels implemented in OpenCL/HLS/RTL C/C++ language extensions
ONNX (standard format for serializing DNNs)
CMake | Home Brew (for Mac OS X) | Conda | Docker
CUDA libraries and runtime from NVIDIA
FPGA designed and built within DeepHealth project by ProDesign
Protocol Buffers from Google
To automatically solve dependencies on other packages
FPGAs from H2020 MANGO project
Interchange models with other toolkits
Adapted to report gradients from worker nodes to master node and to distribute updated weights from master node to worker nodes

ECVL

- Integrates existing state-of-the-art Computer Vision and Image Processing libraries
- Supports multiple medical imaging formats (NIfTI, DICOM, TIFF, whole-slide)
- Core functionalities implemented for both 2D images and 3D volumes:
  - Reading and writing
  - Processing
  - Visualizing
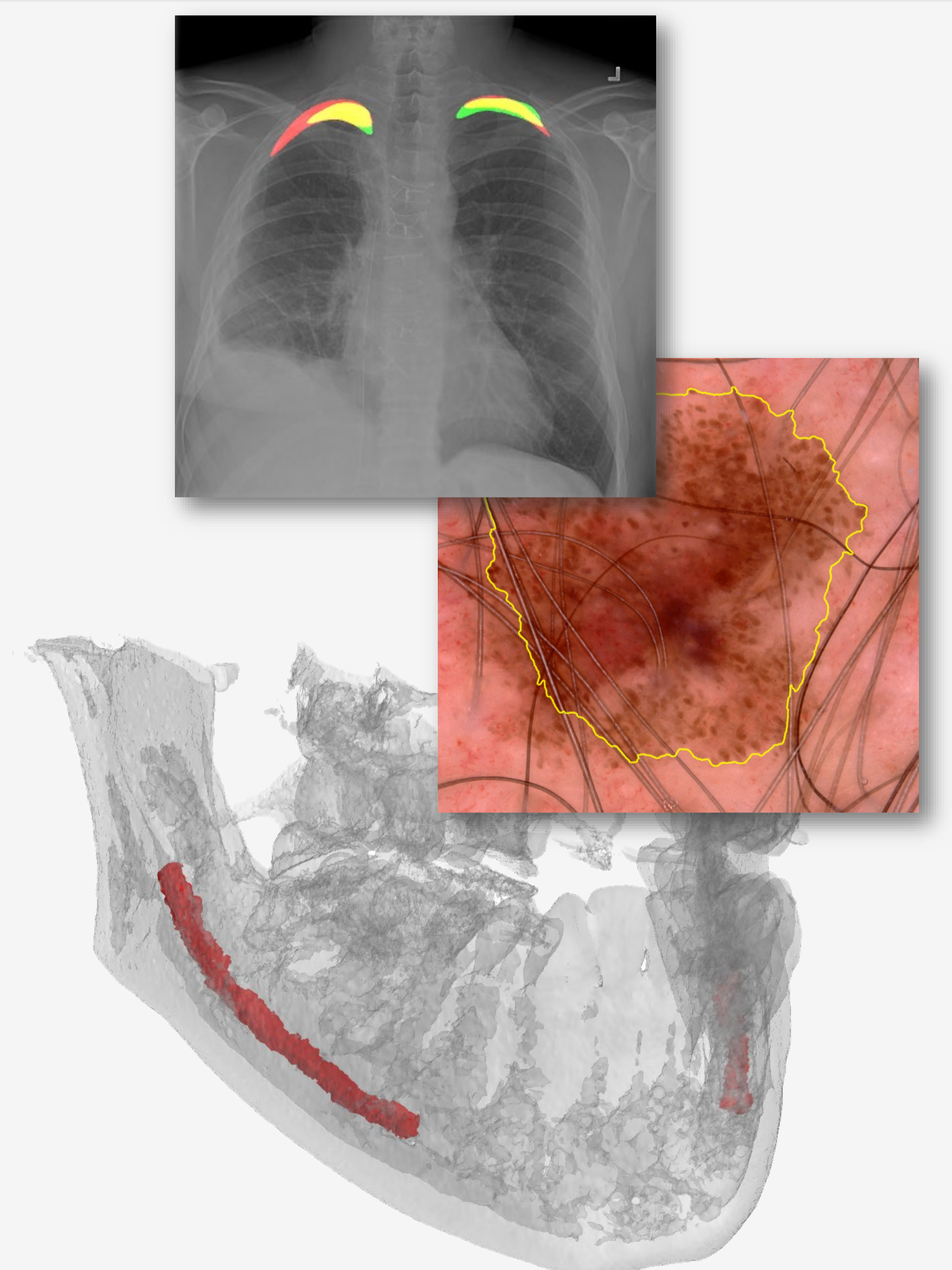
Python APIs

- PyEDDL and PyECVL have been designed for binding Python code to existing C++ code
- Seamless conversion between EDDL Tensor or ECVL Image objects and NumPy arrays

EDDL → PyEDDL

NumPy

ECVL → PyECVL

Benchmarking

GPU TIME PER EPOCH (s)

	VGG16	VGG19	RESNET18	RESNET50
EDDL	204	260	49	132
Keras	68	81	26	52
PyTorch	77	126	60	92

TEST ACCURACY (%)

	VGG16	VGG19	RESNET18	RESNET50
EDDL	76,4	61	67,3	61,9
Keras	71,7	59,9	67,6	61,3
PyTorch	76,2	59,7	66,4	63,1

CONNECTED COMPONENTS LABELING (ms)

	3DPES	FINGERPRINTS	MIRFLICKR
ECVL_Spaghetti	0,88	0,34	0,49
OpenCV_SAUF	1,14	0,45	0,64
OpenCV_BBDT	0,97	0,37	0,54