Effective Early Polyp Detection from Medical Images with YOLO-V7

F. Iakovidis[1], L. Akritidis[1], P. Bozanis[1]

[1]School of Science and Technology, International Hellenic University

15th International Conference on Information, Intelligence, Systems and Applications (IISA 2024)

Grand Arsenali, Chania, Crete, Greece, 17 – 20 July, 2024

Co-funded by the European Union

ΕΣΠΑ 2021-2027

Partnership Agreement 2021-2027

Support for Internationalization Actions of the International Hellenic University (MIS 6004797)

Polyp Detection

- Polyps constitute a potentially harmful cluster of cells appearing in the colon lining.
- Without early diagnosis and treatment, they can lead to Colorectal cancer (CRC), a quite common and highly lethal type of cancer.
  - the third most lethal cancer type for all ages.
  - the main cause of cancer death for males aged 20–49 and for females aged 40–49.
- Early polyp detection via a colonoscopy screening, can prevent CRC.
- However, traditional colonoscopies exhibit high polyp misdetection rates, leading to late CRC diagnoses.
- The studies for computer-aided detection systems for polyps are highly important.

Object Detection Models

- Divided in two categories:
- Two-stage detectors: the same input image is passed twice through the underlying network.
  - Region-based CNN (R-CNN), Fast R-CNN, Faster R-CNN.
- Single-shot detectors: the same input image is passed only once through the underlying network.
  - Significantly faster and computationally efficient, slightly lower accuracy.
  - YOLO (You Only Look Once) models.

YOLO Object Detectors

- An entire family of single-shot models.
- Renowned for their efficiency and accuracy in real time applications.
- They implement deep learning structures based on Convolutional Neural Networks (CNNs) in order to identify multiple objects within an image.
- In contrast to the traditional classifiers, the YOLO models not only predict class probabilities, but they also output bounding boxes around the recognized objects.
- Over the years, YOLO has evolved through successive updates, each introducing its own performance and architectural enhancements.

Contributions

- In this paper we used a recent model of the YOLO family, YOLOv7.
  - Fast/Accurate object detector in the range of 5-160 fps.
- We combined of YOLOv7 with transfer learning approaches to enhance polyp detection precision & higher inference speed.
- We used non-polyp images, treating them as background in the training process.
  - Achievement: False Positive Rate reduction.
- We also simplified detection by merging the adenomatous and the hyperplastic polyps into a single class.
  - Achievement: Higher accuracy and real-time detection speed.

Dataset

- Dataset: Colonoscopy Polyp Detection and Classification dataset (CPDC).
  - Compared to other polyp detection, CPDC is larger and more diverse.
- In CPDC, each sample image is accompanied by an annotation file with information about the polyp class and the bounding box.
- The dataset includes a collection of training, validation and test polyp images along with their annotations files.
- Training set: 28773 images (1725 of them do not contain polyps).
- Validation set: 4254 images (40 of them do not contain polyps).
- Test set: 4872 images, (153 of them do not contain polyps).
- The original dimensions of the images vary in a range between 384 × 288 and 768 × 576 pixels.

Methodology: Data Processing

- A routine was developed to transform the CPDC annotation files to YOLO-compatible data files. The process:
  - extracted the coordinates of the boxes and their dimensions w_b, h_b,
  - computed the horizontal and the vertical centers of the boxes, C_x and C_y,
  - normalized i) the box dimensions and ii) the coordinates of the centers by dividing them by the image width w and height h, respectively:

$$\hat{w}_b = w_b/w$$

$$\hat{h}_b = h_b/h$$

$$\hat{C}_x = C_x/w$$

$$\hat{C}_y = C_y/h$$

- The YOLO-compatible annotation files stored tuples of the form:

$[\, y \;\; \hat{C}_x \;\; \hat{C}_y \;\; \hat{w}_b \;\; \hat{h}_b \,]$ where y represents the polyp class.

Methodology: Training

- The dataset also includes information about the polyp types: adenomatous, or hyperplastic.
- Here we merged these two types into a single category which allowed a more generalized approach to polyp detection.
- This simplified approach yielded improved performance in terms of both accuracy and real-time detection speed.
- In addition, we feed the model not only with polyp images, but also with non-polyp ones (which the model handles as background).
- Since these background images are treated as negative samples, we achieve a reduction in detection of False Positive samples.
- The absence of polyps in training is almost as important as their presence, as this offers a more realistic training scenario.

Experimental setup (1)

- All the experiments have been conducted on the Google Colab platform. This specific virtual environment offered L4 GPUs with 62.8 GB of available System RAM, 22.5 GB GPU RAM and 201.2 GB of disk space.
- To evaluate detection performance, we employed four well-established measures: Precision, Recall, F-score, and Mean Average Precision.
- We also measured Intersection over Union (IoU), that computes the overlap between the predicted bounding boxes and the actual ones.

Model Hyper-parameters and Training

- We conducted a limited grid search to determine the optimal hyper-parameters for YOLOv7.
  - 415 layers, 37.2 million parameters.
  - Training duration (Colab): 7.5 hours (50 epochs).
- We set the `img` flag, to handle the rescaling of the input images before feeding them to the model.
- Selected image dimensions: 416 × 416
  - larger images (e.g., 640×640), did not yield significant gains in detection performance.
- Batch Size: 16 samples.

TABLE I: Model Hyperparameters.

Parameter	Value	Parameter	Value
lr0	0.01	anchor_t	4.00
lrf	0.1	fl_gamma	0.00
momentum	0.937	hsv_h	0.015
weight_decay	0.0005	hsv_s	0.70
warmup_epochs	3.00	hsv_v	0.40
warmup_momentum	0.80	degrees	0.00
warmup_bias_lr	0.10	translate	0.20
box	0.05	scale	0.90
cls	0.30	fliplr	0.50
cls_pw	1.00	mosaic	1.00
obj	0.70	mixup	0.15
obj_pw	1.00	paste_in	0.15
iou_t	0.20	loss_ota	1.00

Performance Measurements

- The model achieved significant success in polyp detection.
  - Precision=Recall=F1=0.877, mAP@0.5=0.929, mAP@0.5-0.95=0.583.

TABLE II: Polyp detection performance of the employed model in the validation and test sets.

Measure	Validation	Test
Precision	0.914	0.877
Recall	0.813	0.877
F-Score	0.861	0.877
mAP@0.5	0.876	0.929
mAP@0.5-0.95	0.632	0.583

- All 4 measures begin to stabilize between epochs 31-33 and they remain relatively stable afterwards.

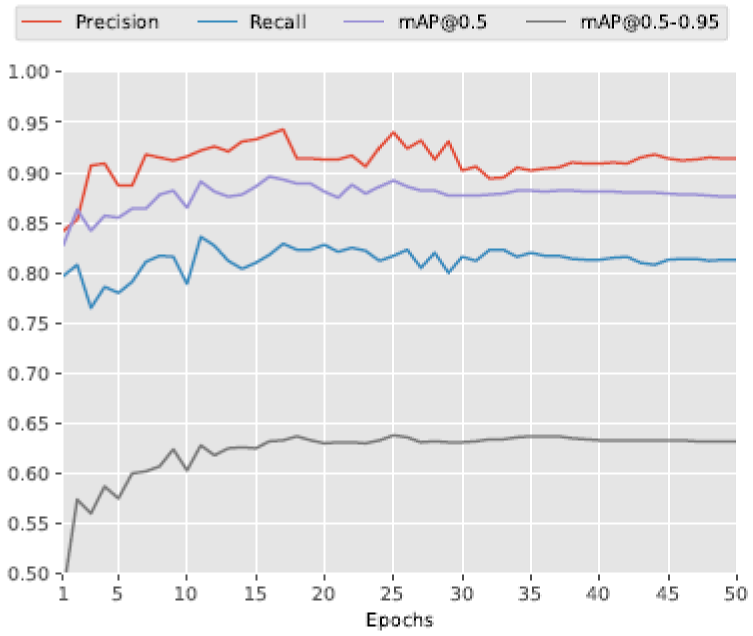


Fig. 1: Fluctuation of values of Precision, Recall, mAP@0.5, mAP@0.5-0.95 during the 50 epochs of YOLOv7 training process.

Visual Performance Inspection (Successful Detection)

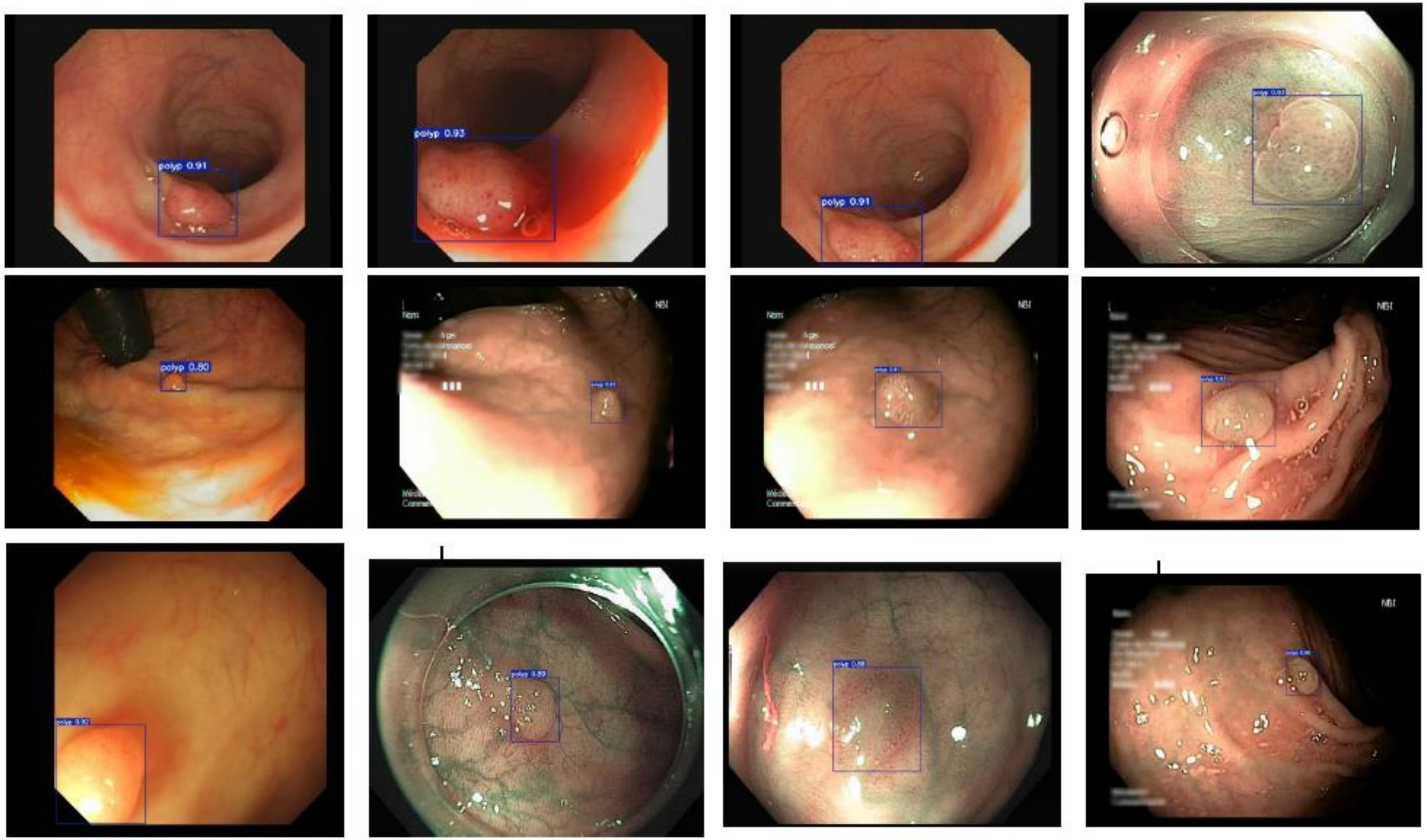


Fig. 2: Examples of successful polyp detection in 12 images of the Colonoscopy Polyp Detection and Classification dataset.

Visual Performance Inspection (Failed Detection)

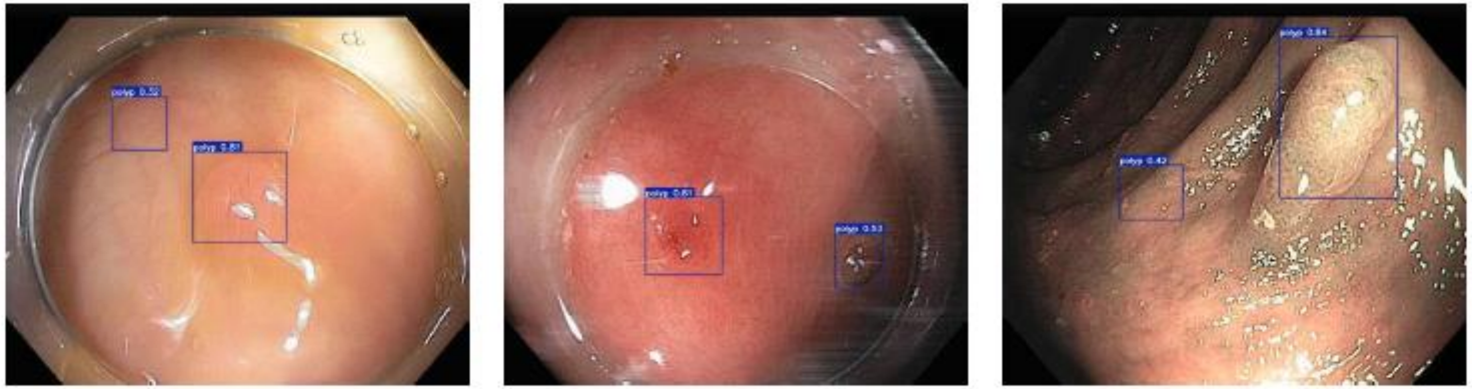
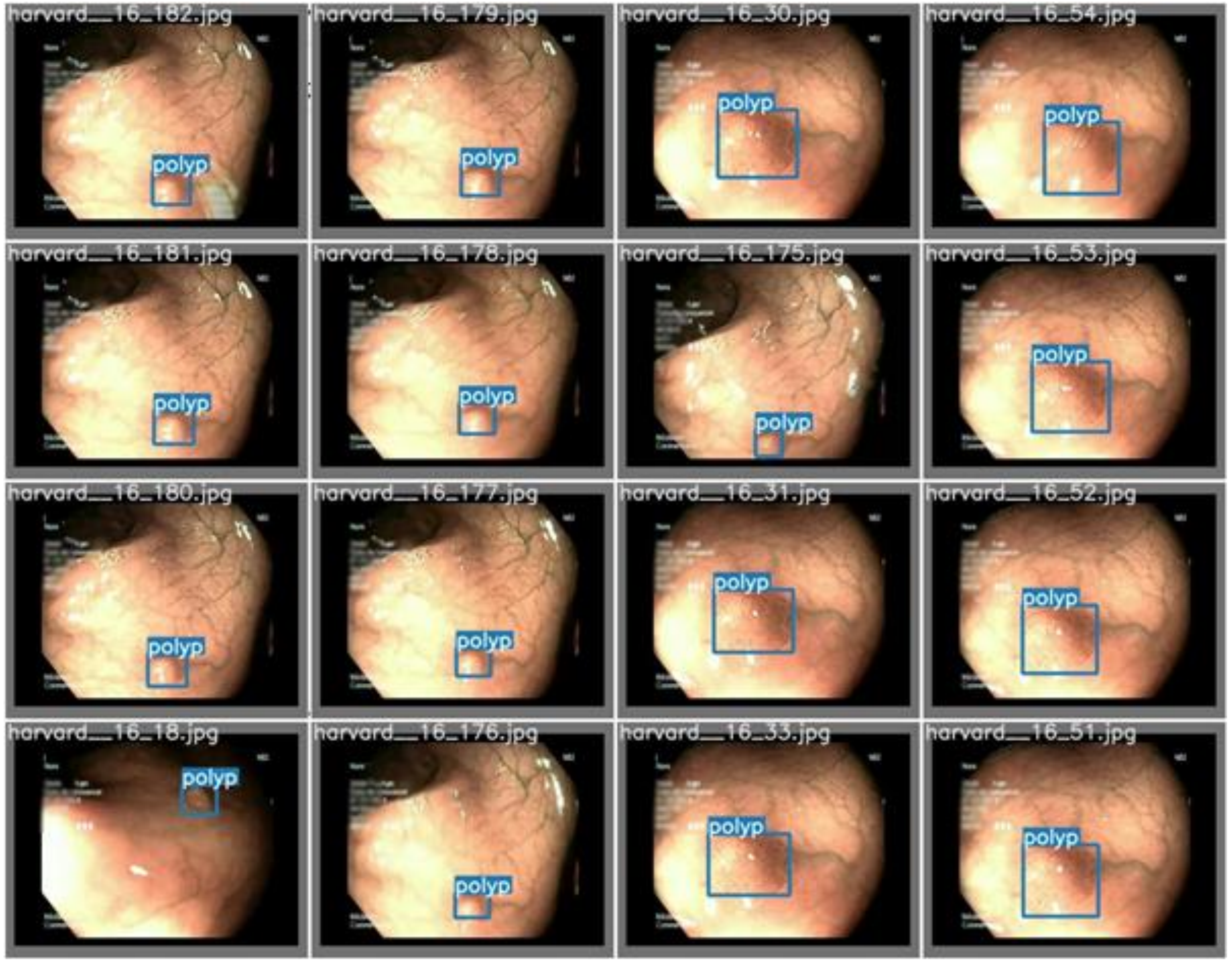
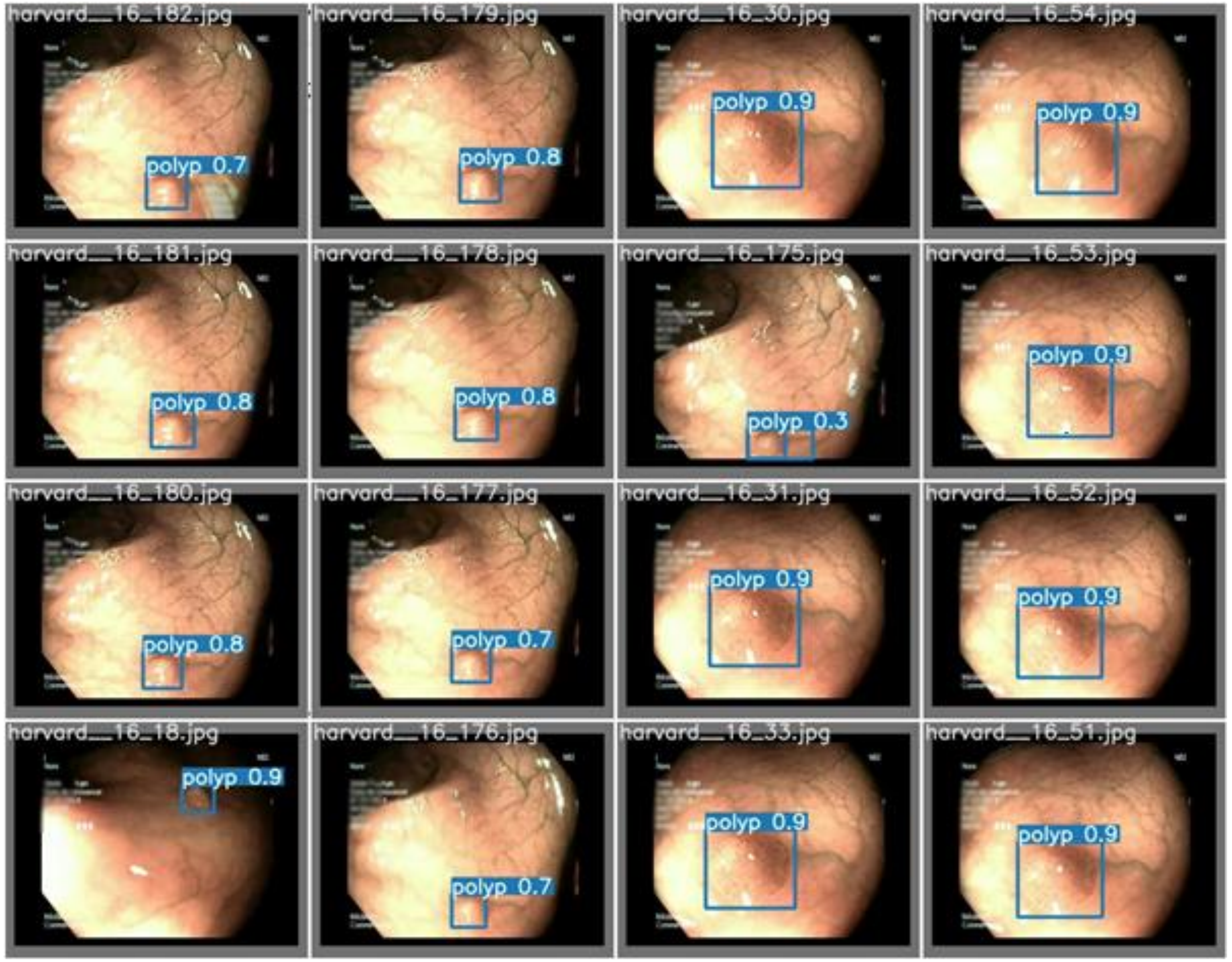


Fig. 3: Examples of failed polyp detection in 3 images of the Colonoscopy Polyp Detection and Classification dataset.

Visual Performance Inspection (Random Batch)



The bounding boxes indicate the ground truth position



The bounding boxes indicate the position predicted by our model, along with the confidence score.

Conclusions & Future Work

- We studied effective data preprocessing techniques for early polyp detection with the YOLOv7 model.
- We transformed and normalized the annotations off the original dataset into a specific format, so that the model could receive its inputs properly.
- We achieved very satisfactory results in very high detection speeds.
- Our future work includes:
- Using image segmentation techniques to improve detection performance.
- Using image augmentation methods for noise reduction and confrontation of imbalance.
- Experiments with different architectures and deep learning techniques. (e.g. R-CNN, or even different YOLO versions).
- Merging multiple public datasets for both polyp and non-polyp images.

Thank you for watching

I would be happy to answer your questions.

Please send them to lakritidis@ihu.gr

Co-funded by the European Union

ΕΣΠΑ 2021-2027

Support for Internationalization Actions of the International Hellenic University (MIS 6004797)

Co-financed by Greece and the European Union (European Social Fund-SF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning 2014-2020" in the context of the project "Support for International Actions of the International Hellenic University" (MIS 6004797)