Semantic Object Segmentation in Cultural Sites using Real and Synthetic Data

Francesco Ragusa[1,2], Daniele Di Mauro[1], Alfio Palermo[1], Antonino Furnari[1], Giovanni Maria Farinella[1]

[1]FPV@IPLAB - Department of Mathematics and Computer Science, University of Catania, Italy
[2]XDG - XENIA Progetti s.r.l., Acicastello, Catania, Italy

Motivations

Wearable devices equipped with a camera, such as smart glasses, can be used in cultural sites to develop user-centered applications.

We consider here the problem of object semantic segmentation. Since collecting and labeling large datasets of real images is expensive and time consuming, we investigate whether the use of synthetic images can be useful to achieve good segmentation performance on real data.

Ramadhanakbr, CC BY-SA 4.0, via Wikimedia Commons

Bellomo Palace Regional Gallery

The cultural site where we collected the data is the Bellomo Palace Regional Gallery. The origins of the palace have been traced to the 12th century, the time of Hohenstaufen rule of Sicily.

The museum houses a collection of Sicilian art (paintings, sculpture and decorative arts) from Syracuse and the region from several centuries. The arguably most noteworthy item in the collection is an Annunciation by Antonello da Messina

Dataset: Numbers

DETAILS ABOUT THE PROPOSED DATASET, INCLUDING THE NUMBER OF REAL AND SYNTHETIC IMAGES

	Resolution	Artworks	Environments	Masks	Training Images	Val. Images	Test Images	All Images
Real	1280x720	24	11	5624	4740 (85%)	170 (3%)	678 (12%)	5580
Synthetic	1280x720	24	11	24000	12000 (50%)	1200 (5%)	10800 (45%)	24000

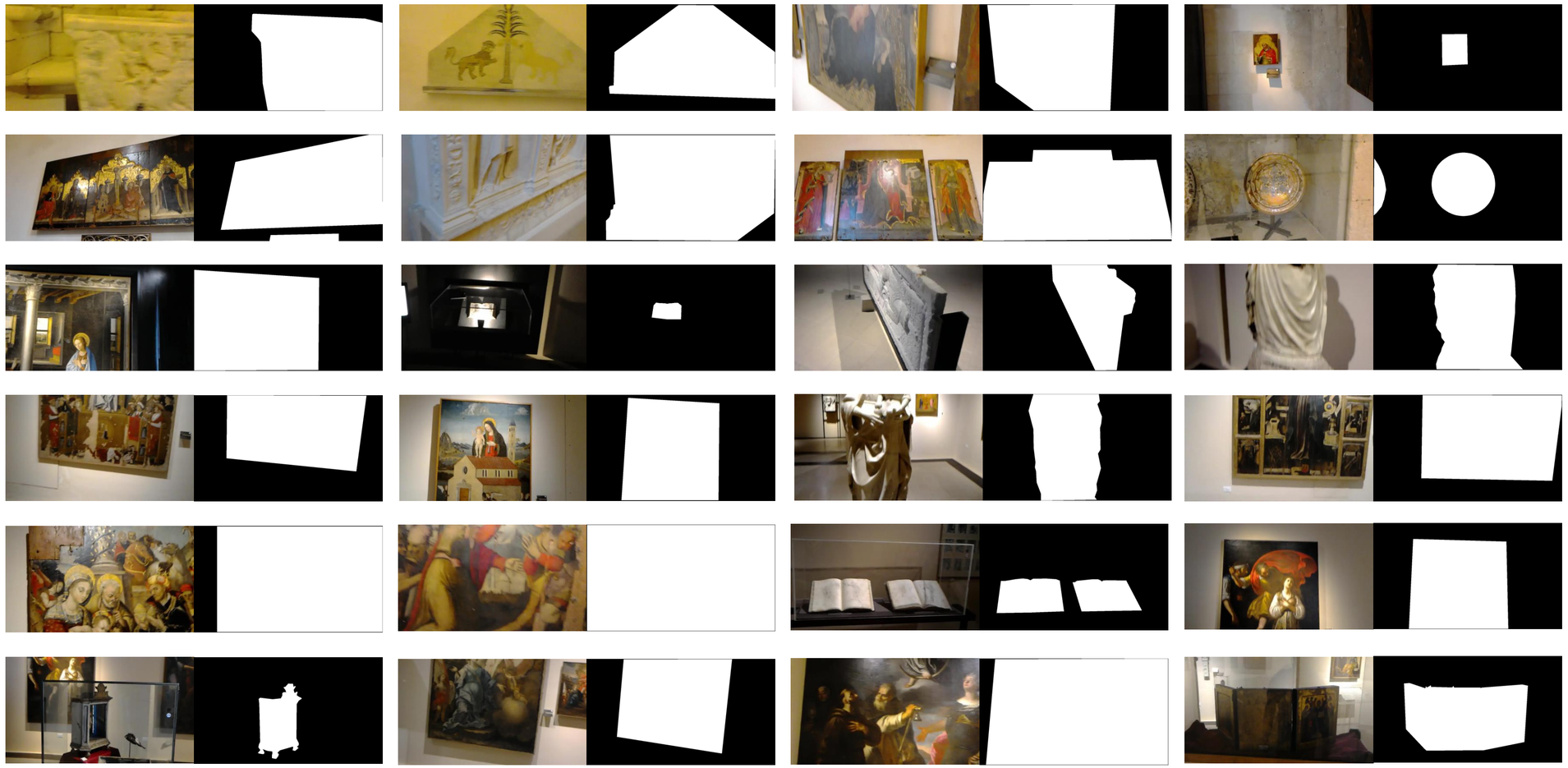
CLASSES WITH NUMBER OF ANNOTATION MASKS. EACH POINT OF INTEREST IS DENOTED BY AN ID IN THE FORM X.Y (E.G., 2.1) WHERE "X" DENOTES THE ENVIRONMENT IN WHICH THE ARTWORK IS LOCATED AND "Y" IDENTIFIES THE ARTWORK

ID	Class	Annotations
2.1	Acquasantiera	244
2.3	LastraconLeoni	248
3.1	MadonnainTrono	237
3.2	FrammentoS.Leo	186
4.1	MadonnainTrono	245
4.2	MonumentoE.d'Aragona	222
4.3	Trasf.Cristo	233
4.4	Piatti	208

ID	Class	Annotations
5.1	Annunciazione	303
5.2	LibroD'OreMin	253
5.3	LastraG.Cabastida	307
5.4	MadonnadelCard.	223
7.1	DisputaS.Tomm.	200
7.2	TraslazioneS.Casa	279
7.3	MadonnacolBam.	231
8.1	ImmacolataConc.	245

ID	Class	Annotations
9.1	AdorazionedeiMagi	230
9.2	S.ElenaCost.eMadon.	247
9.3	TaccuinidiDisegni	212
10.1	MartirioS.Lucia	196
10.2	VoltodiCristo	210
11.1	Miracolodi.S.Orsola	250
11.2	Immacolata	219
21.1	StoriedellaGenesi	196

Dataset: Real Samples



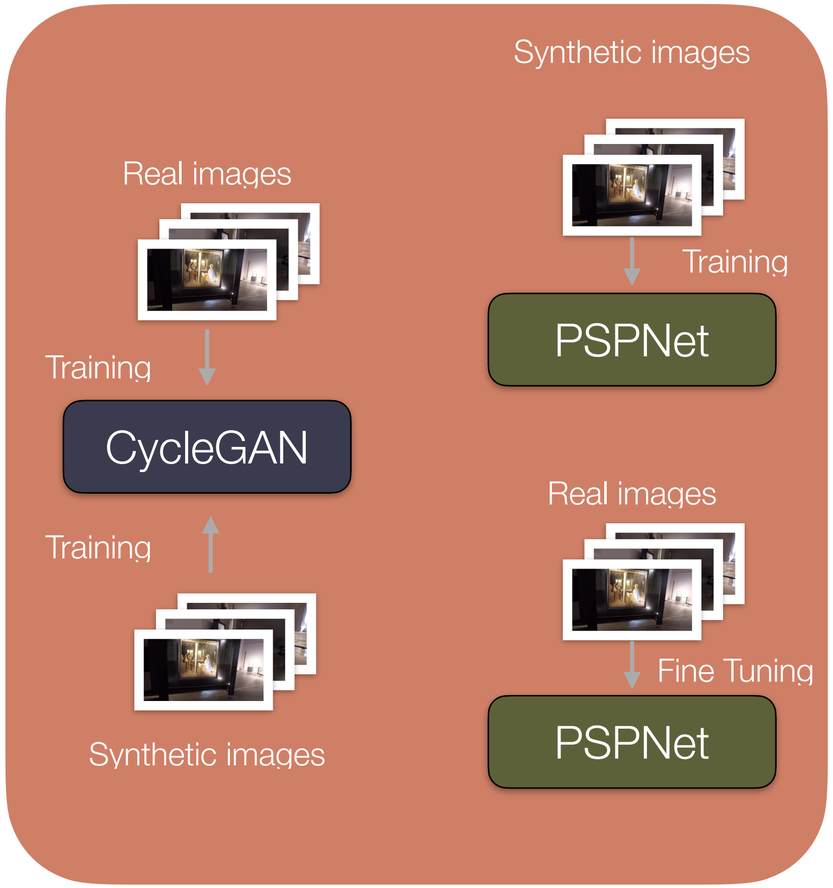
Dataset: Synthetic Samples

Baselines

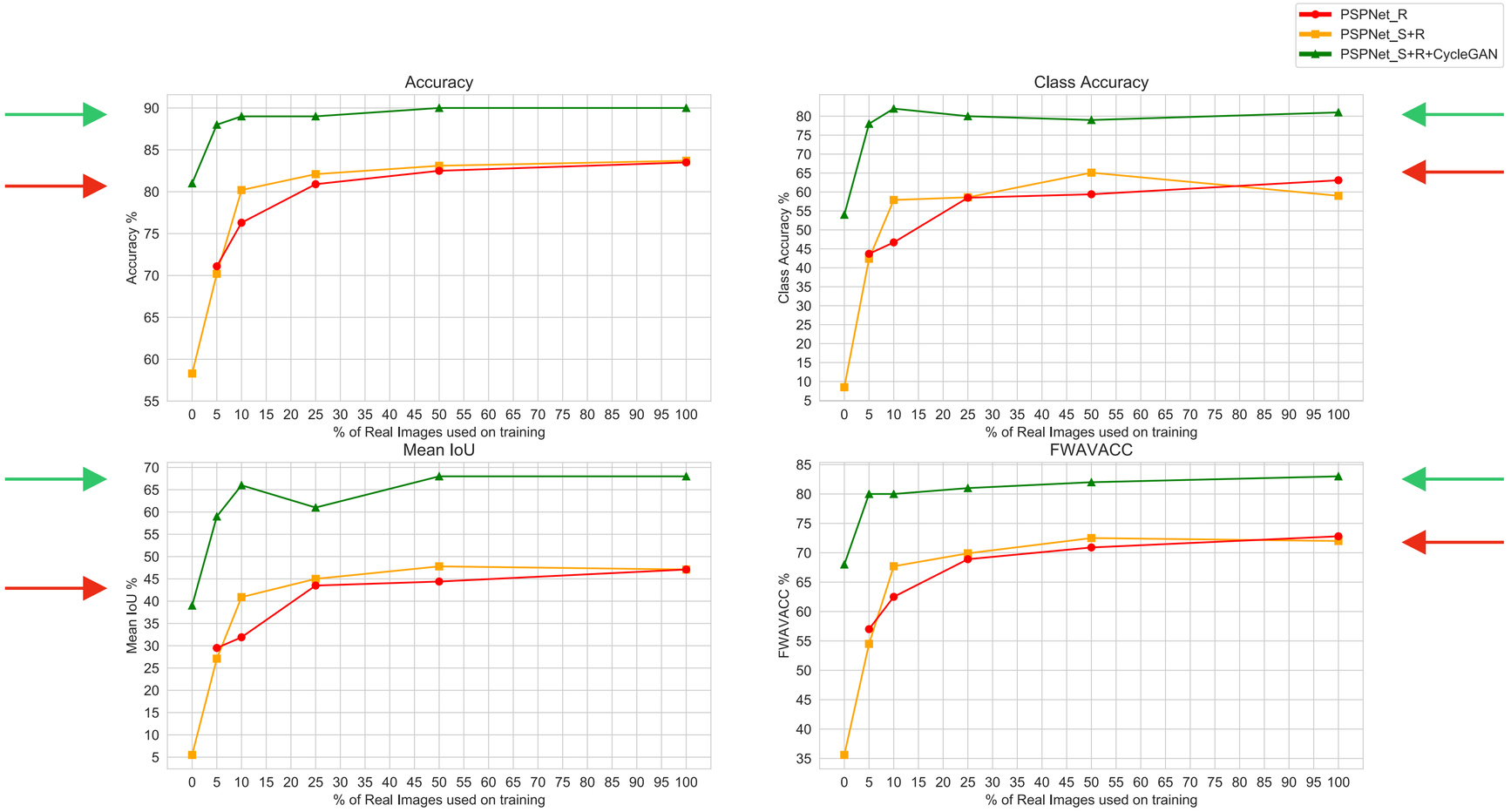
PSPNet_R

PSPNet_S+R

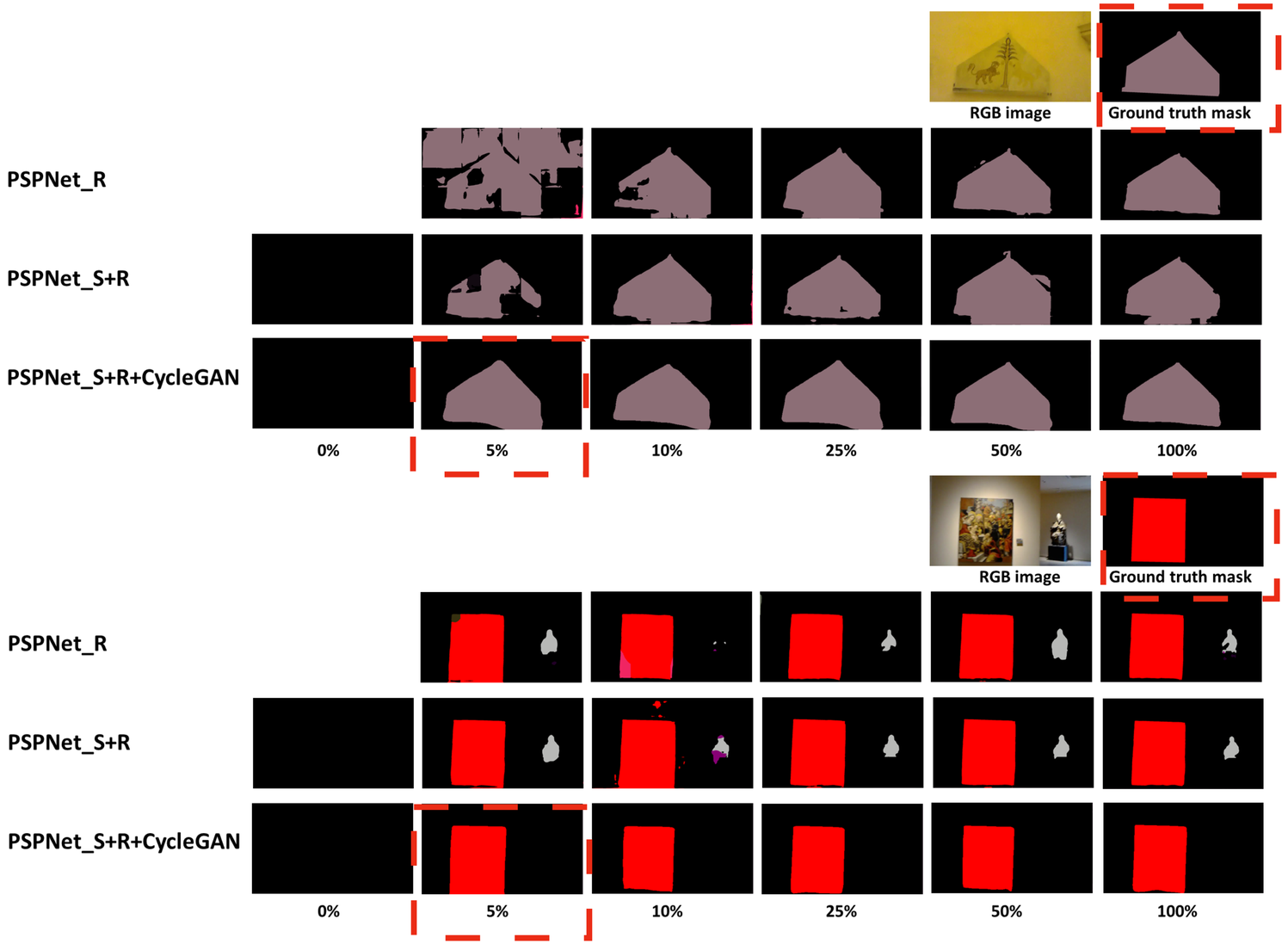
PSPNet_S+R+CycleGAN



Quantitative results: Graphs



Qualitative Results



Conclusions

- We have considered the problem of object segmentation in cultural sites.
- Starting from the assumption that manually labeling images with semantic masks is expensive and time-consuming, we have studied whether the availability of large amounts of synthetic images can allow to improve performance on real images.
- Results highlight that synthetic images can be beneficial to improve performance on real data, especially when coupled with image-to-image translation techniques, to reduce the domain shift arising from the two different data sources.
- For Further information on this work and to download the dataset https://iplab.dmi.unict.it/EGO-CH-OBJ-SEG/ and https://iplab.dmi.unict.it/fpv/