Few Shot Object Recognition summary

E. Patsiouras, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 2.5

Few-shot Image Learning

- The goal of few-shot image learning is to utilize a very small amount of training examples in order to train a machine learning model to recognize a given number of image classes.

- Human visual system is able to effortlessly learn novel concepts from only a few examples.

- Applying the same mechanism to deep learning visual recognitions systems is a much more difficult task, having a wide range of real-world visual recognition applications.

Few-shot Image Learning

Bottlenecks of a typical many-shot based learning model:

- Hundreds or thousands of training examples for each image class are needed.
- Gathering enough training image examples for specific classes can be rather difficult (e.g. data accessibility issues).
- Manual labeling thousands of training examples can be prohibitive.
- A large number of training samples may lead to many gradient-based training iterations hence imposing further computational costs.
- Re-training such a model on new classes requires gathering/labeling sufficient data for every new class start a new training cycle.

Few-shot Image Learning

- Few-shot learning methods are typically designed to provide adequate re-training for new classes given only a few sample images from each one. For example, in few-shot object recognition, we wish to develop a learning model that is able to accurately recognize and classify unseen objects (meaning new classes) using only 1-5 training examples per new object.

- Few-shot learning is usually studied using the ***n-way k-shot*** classification scheme, i.e. we aim to discriminate between **n** classes using only **k** examples of each.

- Typical examples used for performance measurement is the **5-way 1-shot** and **5-way 5-shot** schemes in which we aim to discriminate between 5 novel classes using 1 and 5 examples respectively from each one.

Few-shot Image Learning

Few-shot learning methods can be roughly categorized into 4 classes:

1. **Data augmentation methods**: Provided only a limited number of training examples for some image classes data augmentation techniques can be invoked in order to increase the amount of existing data.

2. **Metric learning-based methods**: Metric learning-based approaches attempt to classify a test example from an unknown class by comparing (based on a distance metric) it to every labeled training example.

3. **Parameter generation methods**: Generate model parameters, i.e. classification weights, for new image classes provided only a few available training data of them.

4. **Meta-learning methods:** Given a few training examples of a new task a meta-learner tries to quickly learn how to "solve" this new task.

Few-shot Image Recognition for UAV Sports Cinematography

- In the context of intelligent UAV sports cinematography, we wish to recognize not only general image classes (*base* classes), e.g. cyclists, runners, boat rowers, but also specific athletes (*novel* classes), e.g. known champions or athletes that lead a particular race, that have distinguished exterior features.

- In our application scenario we wish to recognize a leader ranking athlete (*novel* class) using only few samples by retraining a CNN model that recognizes cyclist athletes (*base* class) [1].

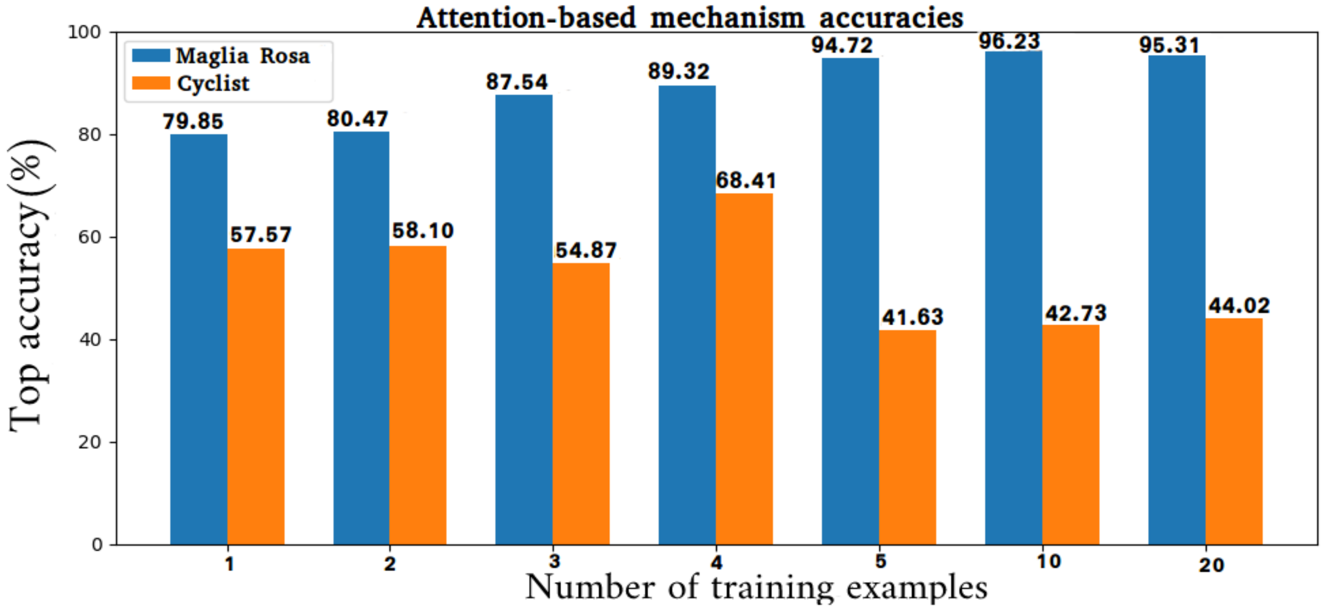
Figure 1.

Figure 2.

[1] E. Patsiouras, A. Tefas, I. Pitas. “Few-shot image recognition for UAV sports cinematography”. *Unpublished*.

Few-shot Image Recognition for UAV Sports Cinematography

For our application scenario our goal is to accurately recognize both base and novel classes, i.e. "cyclist" and "maglia rosa" classes respectively, in a unified and dynamic manner. Baseline results of [2]:



Poor results for the accurate recognition of both classes. This is caused by the fact that the novel class ("maglia rosa") is a subclass of the base class ("cyclist").

[2] S. Gidaris and N. Komodakis. "Dynamic few-shot visual learning without forgetting". In *Proceedings of the IEEE Conference of Computer Vision and Pattern Recognition*, pages 4367-4375, 2018.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr